INCOME DISTRIBUTION
AND
THE SOCIAL WELFARE FUNCTION

by

Sherman Robinson
Princeton University

Research Program in Economic Development
Woodrow Wilson School
Princeton University
Princeton, New Jersey

Discussion Paper 552
February 1975

I wish to thank I. Adelman, L. Joy and R. Webb for helpful comments on an earlier draft. They, of course, cannot be held responsible for persisting errors.

Note: Discussion Papers of the Research Program in Economic Development are preliminary material circulated to stimulate discussion and critical comment. Please do not refer to discussion papers without permission of the author.
1. INTRODUCTION

The modern (i.e., post Pigou) tradition in welfare economics has sought to separate questions of income distribution from other welfare criteria. The "equity versus efficiency" dichotomy is widely accepted, partially because economists feel that they can speak with some authority about efficiency but have no more expertise on questions of equity than any other citizens. Analytically, an economist studies efficiency -- Pareto optimality -- because in neoclassical welfare economics efficient points are always better than inefficient ones. He then calls upon some externally provided value judgments or social welfare function to choose the "best" efficient points and so settle the equity problem. Most practicing economists involved in benefit cost analysis or applied welfare economics (which should probably be defined to include all policy analysis) do not usually bother with the second step. It is enough simply to increase efficiency.

The purpose of this article is to examine some alternative reasonable sets of assumptions about how income distribution affects social welfare and consider their effect on the equity-efficiency dichotomy and so on the postulates of applied welfare economics generally. The framework used for considering welfare value judgments is that of a smooth "well-behaved" social welfare function. The use of such a function is unnecessarily restrictive in that it would be possible to recast the entire discussion in terms of "social ordering" and "choice sets." It is, however, very convenient and makes the main points of the argument much more readily understandable.
There is no discussion of where the social welfare function comes from, a problem that has preoccupied a number of theorists. The function is assumed to embody societal value judgments and so provide a valid framework for evaluating social policy alternatives. The intent here is to examine the nature and implications of the value judgments embodied in the function, but not to examine how they arose in the first place.

There will be no explicit examination of prices and the role of markets. It will simply be assumed that some sort of competitive market system is at work which generates equilibrium prices at all times. An important issue in both theoretical and applied work is that equilibrium prices are not independent of the distribution of income. This entire question is simply ignored in the analysis below. Incomes are assumed to be measured in real terms and to have been properly corrected for price changes.

In the analysis below, I shall first briefly set out virtually the simplest set of basic assumptions -- both normative and positive -- underlying the traditional welfare paradigm and then discuss how questions of income distribution have been considered within this framework. Second, I consider reasonable alternate assumptions and discuss their implications for the equity-efficiency dichotomy. Third, I discuss additional conditions relating to "distributional neutrality" which are required to maintain the usefulness (in welfare terms) of the motion of a Pareto Optimum. Finally, I suggest some ways in which income distribution can be included in social welfare functions, and at what cost in terms of the traditional assumptions.
2. THE TRADITIONAL APPROACH TO INCOME DISTRIBUTION

First, assume that society is made up of \( n \) individuals each with a "well behaved" utility function which is only a function of that individual's income:

\[
U_i(Y_i) \quad i = 1, \ldots, n
\]

A person is made "better off" by increasing his utility index and that, in this formulation, can be done only by increasing his income. Whether the index is cardinal or ordinal is immaterial. For convenience, "well behaved" means a twice differentiable convex function. One usually also assumes that individuals are utility maximizers and it is convenient to use income \((Y_i)\) rather than the actual consumption bundle.

There are two more serious assumptions involved in (1). The first is that only income affects utility, that we are dealing with "economic utility functions." One can assume that non-economic factors affect utility, but they must be seen as exogenous parameters, independent of income.

The second assumption is that only one's own income matters. My utility is not affected by another person's income or utility. This assumption is much discussed in the literature on externalities and will be discussed below in terms of income distribution. One need only note here that it is a factual, not a normative, assumption and so is in principle amenable to empirical testing.

The utility function in (1) is assumed to be a monotonically increasing function of \( Y_i \). That is:
\[ \frac{\partial U_i}{\partial y_i} > 0 \]

for all \( y_i > 0, \ i = 1, \ldots, n. \)

This formulation incorporates the usual assumptions that "more is better than less" and that a person is never satiated. Given (1), (2) implies that the only way a person can be made better off is to give him more income.

Under various assumptions about the workings of the economy, one can show that the economic system will reach an equilibrium at a Pareto optimal point; that is, a position from which the economy cannot move without making someone worse off. It can further be shown that if the economy is not at a Pareto optimal point, then it is possible to make a move so that at least one person is made better off and no one is made worse off. Such a move defines a Pareto improvement and if a potential move represents a Pareto improvement, it will be said to satisfy the Pareto criterion. The notion that a Pareto improvement always makes society "better off" is perhaps the most fundamental value judgment in welfare economics. It really underlies the economic definition of efficiency.

Of course, the economy may be at equilibrium at an infinite number of Pareto optimal points and there is no way to choose among them on the basis of the Pareto criterion. Welfare economists have postulated an externally supplied social welfare function to evaluate different Pareto optima. Following Graaf [1957], such a function is called a Bergson social welfare function and has, as arguments, the cardinal utilities of each individual in the society and nothing else. In addition, the Bergson function is a Utility-Paretian social welfare function which is defined
as a social welfare function with the property that a Pareto improvement always increases social welfare.

These properties can be set forth as follows. Define the social welfare function as:

\[ W(U_1, U_2, \ldots, U_n) \]

where

\[ \frac{\partial W}{\partial U_i} > 0 \quad \text{for all } i \]

Take the total derivative of (3):

\[ dW = \sum_i \frac{\partial W}{\partial U_i} dU_i \]

The definition of a Utility-Paretian social welfare function can be stated in terms of the total derivative. The function is Utility-Paretian if:

\[ dU_i > 0 \text{ and } dU_j = 0 \text{ for all } j \neq i \text{ implies } dW > 0 \text{, for all } i. \]

Clearly, the Bergson welfare function (3) is Utility-Paretian since

\[ \frac{\partial W}{\partial U_i} > 0 \text{ for all } i. \]

Note that the formal definition of Utility-Paretian is equivalent to saying that a Pareto improvement always increases social welfare.

Taking the total differential of \( U_i \) with respect to income, equation (5) can be rewritten:

\[ dW = \sum_i \frac{\partial W}{\partial U_i} \frac{\partial U_i}{\partial Y_i} dY_i \]

Define a social welfare function as Income-Paretian if \( dY_i > 0 \) and \( dY_j = 0 \) for all \( j \neq i \) implies \( dW > 0 \) for all \( i \). Given (2), the Bergson social welfare function is also Income-Paretian. That is, making one person richer (as opposed to "better off") without making any one else poorer always improves social welfare. Given the assumptions (1) - (4), the welfare function is both Income- and Utility-Paretian and can be simply called Paretian. Being "better off" and "richer" are fully equivalent,
both for individuals and society.

The distinction between the Income-Paretian and Utility-Paretian properties of a social welfare function are often ignored. Given the assumptions made so far, ignoring the distinction does not cause any problems. However, as will be shown below, the distinction is important when the paradigm is extended. The notions of a Pareto improvement and the corresponding Pareto criterion are usually defined in terms of utility but applied in terms of income. Below, when I mean to refer to income changes, I shall refer to an "Income-Pareto improvement" (or criterion). Without the qualification, the term is defined in terms of utility.

The assumptions given in (1) - (4) represent perhaps the simplest set of assumptions which underly the basic paradigm of modern welfare economics. Given these assumptions, the welfare analysis of income distribution is really limited to discussing the shape of the social welfare function and is purely a matter of value judgments. Furthermore, since the social welfare function is Pareto, it is possible to separate questions of income distribution (equity) from questions of efficiency. Regardless of questions of equity, it is always better to be more efficient since improving efficiency is always a Pareto improvement and so increases social welfare. Note here that improving efficiency is an ex post concept -- society is not better off unless someone is made better off or richer in fact (not potentially) without in fact making anyone else worse off.
3. INCOME DISTRIBUTION AND THE TRADITIONAL PARADIGM

It has been argued since Marshall and Pigou that questions of income distribution can be considered within the traditional framework by making a further factual assumption; namely, that all individuals have a diminishing marginal utility of income. That is:

$$\frac{\partial^2 U_i}{\partial Y_i^2} < 0 \text{ for all } i$$

(7)

If one were a Benthamite utilitarian and so wished to maximize the total number of "utils" in society, then it is obvious that social welfare will be increased by taking from the rich who have a low marginal utility of income and giving to the poor. Note, however, that this policy will not necessarily lead to perfect equality. Making transfers so as to equalize everyone's marginal utility of income will not equalize total utilities (unless, for example, everyone had the same utility function).

An important point to note is that the factual statement embodied in (7) is not, alone, enough to enable one to make statements about the effect of distribution on social welfare. One must, in addition, specify the social welfare function. For example, consider a social welfare function as the weighted sum of individual utilities:

$$W = \sum_{i=1}^{n} B_i U_i$$

(8)

If $W$ is to be Paretoian, it must be true that $B_i > 0$. The weights may, however, be different and could be specified as functions of variables such as income. For a simple Benthamite welfare function (with cardinal utility functions) one would probably assume that $B_i = 1$ for all $i$. However, it is clearly possible to specify that some of the $B_i$ are increasing functions of income in such a way that the weights increase faster than
the marginal utility of income declines. In such a case, increasing inequality will increase social welfare even if there is declining marginal utility of income.

Even if one were willing to specify a Benthamite social welfare function, this approach seems a rather fragile way to incorporate equity judgments into the welfare function. In common with simple utilitarian philosophy, the approach can easily have unsatisfactory implications. For example, assume there is one fanatic miser in the society for whom the marginal utility of income increases very rapidly with income. To equate marginal utilities, one might have to give him most or all of aggregate income. It would seem to make more sense to specify the social welfare function in such a way that it directly embodies distributional judgments unambivalently. For example, one might make the \( B_i \) weights in (6) rapidly decreasing functions of income \( Y_i \).

4. DISTRIBUTION EXTERNALITIES

There is a long literature on consumption externalities where a person's utility is affected by someone else's consumption or utility. There is also some work done considering income distribution as an externality. Within this framework, one can speak of a Pareto optimal income distribution. It is worthwhile to consider the effect of the existence of such externalities on the traditional paradigm, particularly on the nature of a Pareto improvement.

Replace assumption (1) with

(9) \[ U_i(Y_i, D_i) \]

where \( D_i = D_i(Y_i, \ldots, Y_n) \). \( D_i \) is some distribution statistic (or a vector of them) and need not be specified further. For convenience,
assume that as $D_i$ increases, the distribution becomes more unequal — it is an increasing measure of inequality (e.g., the Gini coefficient).

It is clear that equation (5) giving the total derivative $dW$ with respect to changes in individual utilities $dU_i$ remains unchanged. Thus the welfare function is still Utility-Paretian. However, the total derivative with respect to changes in income is different. It is now:

$$dW = \sum_i \frac{\partial W}{\partial U_i} \frac{\partial U_i}{\partial y_i} dy_i + \sum_i \frac{\partial W}{\partial D_i} \frac{\partial U_i}{\partial D_i} dD_i$$

where

$$dD = \sum_j \frac{\partial D_j}{\partial y_j} dy_j$$

The first term in equation (10) gives the change in social welfare of a set of income changes, ignoring any changes in distribution. The second term gives the effect of distribution changes and neither its sign nor its magnitude is known. Thus, the social welfare function is no longer guaranteed to be Income-Paretian.

One can use the second term in (10) to define conditions representing two notions of a "distributionally neutral" set of income changes. They are:

$$dD_i = 0 \text{ for all } i, \ "data\ neutrality"$$

$$\sum_i \frac{\partial W}{\partial D_i} \frac{\partial U_i}{\partial D_i} dD_i = 0, \ "value\ neutrality"$$

Distributionally neutral in the sense of data neutrality implies that all the distribution statistics that people care about remain constant as incomes change. Neutral in the sense of value neutrality implies that the income distribution changes are offsetting in their effect on social welfare. Data neutrality is stronger in that it implies value neutrality.
However, using (12B) requires the complete specification of the social welfare function while (12A) requires only measuring the distribution statistics that "matter" to people.

Under the externality approach, the social welfare function is still Utility-Paretian. However, it is no longer true that making one person richer (as opposed to making him happier) without making anyone else poorer is an unequivocal improvement. The existence of the distribution term in equation (10) implies that the Pareto criterion cannot be extended to increases in income as opposed to increases in utility.

The income Pareto criterion will work, however, if the changes in income are distributionally neutral in either of the senses discussed above. This question of distributional neutrality really underlies the recent controversy about including distributional weights in assessing the benefits in benefit-cost analysis. Those who argue that income distribution should be ignored - or at least not incorporated into the analysis - when doing benefit-cost analysis are implicitly either: (1) assuming the separation of efficiency and equity criteria is valid, (2) assuming that the changes are distributionally neutral, (3) or making a strong value judgment that distribution "ought" to be ignored. (This latter approach will be discussed further in section 6 below.)

The first and second possible assumptions are really equivalent. The classical dichotomy between efficiency and equity holds only when changes are distributionally neutral since only for this restricted set of possible changes is the social welfare function guaranteed to be both Utility- and Income-Paretian.
The failure of the classical dichotomy when income changes are not distributionally neutral has been fought about for years. The entire literature on compensation tests has revolved around the issue. The conclusion of this controversy seems to be that no matter how much ingenuity one uses in devising fictional intermediate points, the initial and final points cannot be compared according to any income variant of the Pareto criterion (or compensation test) when the change is not distributionally neutral. 8

If the income changes being considered are not distributionally neutral -- and for many projects analyzed by means of benefit-cost analysis, the changes are not neutral -- then it simply is not possible to base the analysis on any notion of a Pareto improvement in terms of income without specifically analyzing the effect of the redistribution on social welfare. It seems pointless to insist on doing a benefit cost analysis based on the separability of the two criteria, and so behave as if the changes are distributionally neutral when, in fact, they are not. Such an analysis seems obviously inferior to, and is certainly in no sense more scientific than, an honest attempt to include income distribution criteria directly into the analysis.

Since the income Pareto criterion test is valid for distributionally neutral policies, one way to apply benefit-cost analysis appropriately would be to insist that, in fact, redistributions are made so that the final change is distributionally neutral. Then, regardless of how distribution affects the social welfare function, the equity-efficiency dichotomy is valid and benefit-cost analysis works. It is a separable question how much better in terms of social welfare one can do by further
redistribution. Of course, maximizing social welfare subject to the constraint that changes be distributionally neutral will, if the constraint is binding, lead to a solution with lower social welfare than would arise from maximization not subject to the constraint.

The question of distributional neutrality is really an empirical or factual question on two levels. First, it may be that people in fact do not care about income distribution. If so, then \( \partial U_j / \partial D_j = 0 \) and all policies are distributionally neutral. Second, even if they do care, the changes in income being considered may still be distributionally neutral in that they either do not change the relevant distribution statistics or change them in offsetting ways.

Without knowing the exact effects of distribution on social welfare, one might seek to achieve distributional neutrality in the sense of data neutrality. Even without knowing all the relevant statistics, one can seek approximate neutrality. I would be willing to argue, for example, that a first approximation to a distributionally neutral change would be achieved if the gainers, in fact, compensate the losers. One might attach this as an additional condition that policy makers and those who do benefit-cost analysis must satisfy if they wish to avoid analyzing explicit distributional effects of policies. Note that this implies that poor gainers must compensate rich losers if necessary. I suspect that such poor-gainer rich-loser situations, while possible, are very rare in practice. Some of the implications of distributional neutrality will be discussed in detail in the next section.
To summarize, the existence of distributional externalities invalidates the traditional separation of equity and efficiency criteria. The social welfare function is still Utility-Paretian, but is no longer necessarily Income-Paretian. Making one person richer without making anyone else poorer does not necessarily increase social welfare. Traditional benefit-cost analysis is valid only when the social welfare function is Income-Paretian. Compensation tests are simply not valid unless, perhaps, the compensation is actually paid. Note that the problem is not one of values - the only normative assumption made is that social welfare is an increasing function of individual utilities.

There are two ways out of the dilemma. First, one can insist that any changes be distributionally neutral and so do constrained maximization of the social welfare function. Since the social welfare function is income Pareto for distributionally neutral changes, then benefit-cost analysis is applicable. Second, one can accept that the equity-efficiency separation is invalid and try to incorporate equity criteria directly into the analysis.

In taking the first way out of the dilemma, it would seem easiest to try to achieve neutrality in the first sense (that of not changing the distribution statistics) since this approach does not require explicit knowledge of the social welfare function. If one can specify how distribution affects individual utility and social welfare, then it would seem that one could do better by abandoning the traditional dichotomy and use the social welfare function directly (and so derive distribution weights).
5. DISTRIBUTIONAL NEUTRALITY

In this section we examine the implications of data neutrality (equation 12A). First, assume that all the distributional statistics $D_i$ are a function only of relative income, implying that a proportional change in all incomes leaves the statistics unchanged. Formally, they are homogeneous of degree zero in all incomes. Clearly, any proportional changes in all incomes is distributionally neutral in this case. One can write $dY_j = kY_j$ where $k$ is the proportional change:

$$dD_i = \sum_j \frac{\partial D_i}{\partial Y_j} dY_j = k \sum_j \frac{\partial D_i}{\partial Y_j} Y_j = 0$$

by Euler's equation for homogeneous functions.

The conditions for data neutrality can be seen as a set of homogeneous simultaneous equations in $dY_j$:

$$\frac{\partial D_1}{\partial Y_1} dY_1 + \frac{\partial D_1}{\partial Y_2} dY_2 + \ldots + \frac{\partial D_1}{\partial Y_n} dY_n = 0$$

$$\frac{\partial D_n}{\partial Y_1} dY_1 + \frac{\partial D_n}{\partial Y_2} dY_2 + \ldots + \frac{\partial D_n}{\partial Y_n} dY_n = 0$$

or

$$Jd = 0$$

where $J$ is an $n \times n$ matrix with $J_{ij} = \frac{\partial D_i}{\partial Y_j}$ and $d_j = dY_j$.

The condition for there to be a non-trivial solution to this system is that the determinant of $J$, the Jacobian, equal zero. If the Jacobian does not vanish, then there does not exist a distributionally neutral set of income changes (in the sense of data neutrality).
One cannot, in general, assume that the Jacobian vanishes and it is certainly possible to construct examples where it does not vanish -- for example, everyone has a different distribution statistic and at least one of them is affected by proportional changes in all incomes. However, it is reasonable to assume that a number of people consider the same distribution statistic as relevant so that $D_i = D_j$ for some $i \neq j$. People in the same social class, ethnic group, or income class might well so indentify themselves with their group that the income distribution statistic they consider relevant is one that reflects the relative position of their group in the income distribution rather than their personal position. Indeed, one might well define a social class as being comprised of those individuals who share the same concern about their collective position in the income distribution.

One point is worth noting about this view of a person as being concerned not with his personal position in the distribution but his group's position. It is at first glance contrary to the utilitarian (and neo-classical) view that every individual pursues only his personal self-interest. However, in a society where economic policy and the resulting distribution are strongly affected by political forces, and where those political forces are influenced (or even defined) by various "pressure groups" or political constituencies, than it is clearly in a person's self-interest to associate himself with a group and to care about and pursue the collective interests of the group.

One need not assume that people do not care about the intra-group distribution but that for many purposes it is reasonable to consider only changes in the inter-group distribution and to assume that the intra-group distribution is optimal, or adjusts itself exogenously.
The result of these assumptions is that various of the J
matrix are identical and so it is not of full rank. If there are m such
if interest-groups, then the rank of J is m (or m-1, in addition, all the
distributional statistics are unaffected by proportional changes.) At
first glance, it seems that one is free to set any n-m income changes and
solve for the remaining m changes so that the set is distributionally
neutral. However, this is not proper since any such solution involves
changing the intra-group distribution in an arbitrary way. One must
restrict the analysis to changes that do not affect the intra-group
distributions -- for example, reduce the \( dY_j \) variables to the n
mean incomes.

By considering interest-groups defined as people concerned
with the sum distribution, one can reduce the completely intractable
problem of considering n different distribution statistics where n might
be millions to a much reduced problem of considering m interest groups
where m might be ten, or even less. For many cases, both in planning
and in benefit-cost analysis, it might well be reasonable to assume
that the intra-group distributions will be optimal and so we need only
concern ourselves with the inter-group distribution, comparing mean group-
incomes. We are then left with a J matrix of partial derivatives which
is m by m and also m \( dY_j \) variables where \( dY_j \) is now defined to be the
change in group mean incomes.

If the interest groups are truly different, one would assume
that the resulting J matrix would be either (1) of full rank or (2) of
rank m-1 if all the distribution statistics are a function only of relative
incomes. If it is of full rank, then there exists no set of income changes
that achieves data neutrality and any policies will result in distributional
trade-offs that must be considered explicitly. If J is of rank \( m-1 \),
then it is always possible to find a set of distributionally neutral
income changes, namely proportional changes. Whether or not the
appropriate distribution statistics are a function only of relative
incomes is really an empirical question. It seems reasonable to assume
that for relatively small changes in absolute income, only changes in the
relative distribution are important -- for large changes, it seems less
tenable.

There are two ways changes in the absolute income level might
affect how people view distribution. First, it might affect how they
feel about the relative distribution, but not change their basic feeling
that it is the relative distribution that matters. In this case, it
affects how the \( D_j \) variables affect utility, but still implies that a
statistic measuring the relative distribution is relevant. 10 Distributional
neutrality is still feasible. Second, proportional changes in group incomes
implies a widening absolute difference between groups. As average income
increases, the widening absolute gap will itself become important and so
people will no longer consider only statistics measuring the relative
distribution. In this case, the J matrix will be of full rank and distri-
butional neutrality in the sense of data neutrality is no longer possible.
The trade-offs must be considered explicitly. If one still wishes to
achieve distributional neutrality, it must be in the sense of value
neutrality (equation 12B).
6. UTILITY, DISTRIBUTIONAL EXTERNALITIES, AND SOCIAL WELFARE

The exact definition of the distribution statistics which are arguments in the utility functions in the previous sections was left purposely vague. The effects of the existence of distributional externalities on the traditional welfare analysis do not depend on the exact form the externalities take. There are, however, a number of important questions relating to the form of the externalities.

Usually, distributional externalities are characterized by terms such as "charity" or "envy." A person's utility is affected by being aware of people who are richer ("envy") or poorer ("charity"). Presumably, the appropriate distribution statistics which would reflect such feelings would measure an individual's relative income ranking compared to those richer and/or poorer. For example, the ratio of one's income to the income of the other people or groups (or perhaps the absolute difference) would seem the most relevant statistic.

The charity notion of distributional externalities has led to a discussion of Pareto optimal redistribution through charitable organizations. The approach has been to treat this particular distributional externality as analogous to any other externality and to seek market mechanisms that incorporate it into the rest of the economic system. The existence of organized charities is seen as proof that the distributional externality is not really an externality at all. Presumably, one could view burglary as the analogous institutional mechanism for incorporating "envy" into the economic system.

The most extreme variant of the charity/envy approach is to argue that policy makers need not worry about income distribution since
market mechanisms will guarantee that the second round distribution (after charitable transfers are made) will be Pareto optimal. As has been pointed out by Musgrave, this view is incorrect. The second round distribution will not be independent of the first round, pre-transfer distribution. There will be many Pareto optimal distributions and we are still left with the problem of choosing among them. The Income-Pareto criterion still is not applicable unless the income changes are distributionally neutral.

The charity/envy approach is not the only reasonable way to view distributional externalities. Thurow (1971) argues that income distribution should be seen as a pure social good. The income distribution is viewed as one of the attributes of a society and everyone in the society partakes equally of whatever "benefits" the distribution engenders. This approach would lead to the inclusion of the same measures of the overall income distribution being included in each person's utility function. Thus one would write $U_i (Y_i, D)$ where $D$ is one (or more) measures characterizing the overall distribution.

It should be noted that the two approaches are in no way mutually exclusive. Indeed, they both can be included in the general framework presented above by merely defining $D_i$ as a vector of distributional statistics, some particular to the individual, some general. The implications of either approach for traditional welfare economics are the same. The two approaches do, of course, have different implications for exactly what constitutes a distributionally neutral set of income changes.

As well as having different implications for the appropriate distribution statistics to include as arguments in utility functions, the two approaches also, implicitly, involve different assumptions about
people's attitudes and motivations. The charity/envy approach represents only a minor extension of the usual utility maximizing homo-economicus who seeks only to improve his personal well-being. Instead of caring only about his own creature comforts, he also feels envious towards those richer than he and/or charitable towards those poorer. While this view of man is consistent with simplistic utilitarianism, it seems wildly at odds with the view that people are citizens who also care about the nature of the society in which they live. The social good approach permits a much broader and more realistic view of what motivates individuals. Individuals as citizens and voters clearly care about the aggregate distribution and such attitudes are clearly reflected in social policy. In the discussion of distributional neutrality, groups were defined with similar attitudes about income distribution. That approach, in which it is the inter-group distribution that "matters," can be seen as a compromise between the view of people as purely self-interested and the view of people caring about the overall distribution.

So far, very little has been said about the nature of the social welfare function. It has been assumed to be an increasing function of individual utilities and to include no other arguments. A question that has haunted welfare economics is "What is the source of the value judgments embodied in the social welfare function?" There has been an active literature exploring the conditions necessary for some kind of "democratic" rules of decision making to yield a well-defined social welfare function (or, more generally, a well-defined social ordering). The usual view in welfare economics has been to assume that such judgments are provided exogenously -- which is a way of avoiding the question. Given the
difficulties inherent in deriving a social welfare function from some more basic postulates, welfare economists have naturally sought to base their recommendations on simple and "widely acceptable" assumptions about the underlying social welfare function.

Under the traditional assumptions, the social welfare function has been shown to be both Utility-Paretian and Income-Paretian. The value judgments implicit in this formulation — that society is better off if one person is made better off without making anyone else worse off — seems both reasonable and "widely accepted." Unfortunately, the Paretian value judgment alone is not sufficient to make judgments about income distribution; it only permits efficiency criteria to be valid regardless of the distributional effects. If distributional externalities are present then the social welfare function is still Utility-Paretian but not Income-Paretian and only by restricting the analysis to distributionally neutral changes can the traditional equity-efficiency separation be saved. Saving the traditional separation still does not enable welfare economists to say anything about distribution — it only recaptures lost ground at additional cost and does not add anything. Distributional value judgments are still required.

Analytically, the theoretically most appropriate way to include distributional judgments would be to explore various analytic forms of the social welfare function given in (3) and to specify exactly the effect of the distributional externalities on individual utility. If people feel strongly enough about distribution statistics, then even a social welfare function that is not inherently egalitarian might yield an equal distribution when externalities are taken into account. Thus, maximizing a utilitarian additive social welfare function — e.g., equation
(8) -- could easily yield perfect equality if everyone views equality as desirable, even without assuming declining marginal utility of income.

The main problem with any approach that requires the exact specification of utility functions is that it requires a great deal of knowledge -- or a great deal of courage. In practice, economists have tended to equate utility and income and have sought to maximize the present discounted value of income (or consumption) streams. There are many examples of this approach in the literature on benefit-cost analysis, optimal growth models, and planning models. The usual justification for the practical approach has been that not only are distributional value judgments exogenously provided, but also the actual redistribution is somehow handled independently of economic institutions. Economists need not worry about distribution since it is handled "elsewhere" in the system, presumably through magical, costless, lump sum transfers.

While the practical approach is obviously naive, it is, nevertheless, practical. It is not really very helpful to argue that the approach is incorrect and then suggest a solution that cannot be implemented. There is, however, at least one possible solution that can be implemented. Retain the simple utility-equals-income view but include distributional statistics directly as arguments in the social welfare function. There are serious costs associated with this solution which will be discussed below, but it has the obvious advantages that (1) it is practical and (2) it forces the welfare economist to make his distributional value judgments explicit. Also, as will be discussed below, it can be justified on theoretical as well as practical grounds.
7. DISTRIBUTION AND THE SOCIAL WELFARE FUNCTION

The inclusion of distributional statistics directly in the social welfare function at first seems only a minor extension of the traditional approach. The properties of the utility function are the same and do not involve any externalities:

\[ U_i(Y_i) \]

\[ \frac{\partial U_i}{\partial Y_i} > 0 \text{ for all } i \text{ and } Y_i > 0 \]

The social welfare function now becomes:

\[ W(U_1, \ldots, U_n, D) \]

where

\[ D = D(Y_1, \ldots, Y_n) \]

The variable \( D \) is a distribution statistic (or a vector of them, if desired) and so is a function of incomes.

Condition (4) still holds - the social welfare is an increasing function of individual utilities:

\[ \frac{\partial W}{\partial U_i} > 0 \]

However, (5) becomes:

\[ dW = \sum_i \frac{\partial W}{\partial U_i} \: dU_i + \frac{\partial W}{\partial D} \: dD \]

and (5) becomes:

\[ dW = \sum_i \frac{\partial W}{\partial U_i} \frac{\partial U_i}{\partial Y_i} \: dY_i + \frac{\partial W}{\partial D} \: dD \]

where:

\[ dD = \sum_i \frac{\partial D}{\partial Y_i} \: dY_i \]
The social welfare function is certainly not income Paretian since the
sign of \( dW \) involves a distributional term \( (\partial W/\partial D) \, dD \), whose magnitude
and sign are unknown. The social welfare function would still be Utility-
Paretian if \( dU_i \) could be changed independently of \( dD \). However,
\[ dD = \sum_i \frac{\partial D}{\partial Y_i} \, dY_i. \]
If we make the \( i \)'th person better off without making
anyone else worse off, then that implies we make the \( i \)'th person richer,
since from (1), the only way to make person \( i \) better off is to make him
richer. The only way to guarantee that the social welfare function is
Utility-Paretian is to assume either \( \partial W/\partial D = 0 \) or \( \partial D/\partial Y_i = 0 \) for all \( i \).
Neither of these conditions is reasonable, so there is a potential con-
flict between this formulation of the social welfare function and the
fundamental Paretian value judgment. It is possible to make someone
better off without making anyone else worse off and yet have social
welfare decline. The social welfare function is not Paretian.

As was done in the externality case, one can define a distribu-
tionally neutral set of income changes such that \( dD = 0 \). For this
restricted set of possible income changes, the social welfare function
is both Income-Paretian and Utility-Paretian. Again, one can duck the
question of income distribution by restricting oneself to changes that
do not affect the distribution.

The fact that the social welfare function is no longer
necessarily Utility-Paretian sounds rather serious. One is being asked
to give up the fundamental utilitarian value judgment underlying modern
welfare economics. One answer is to argue, as was done at the end of
the last section, that including distribution in the social welfare
function is a convenience only, an approximation to the theoretically
more appropriate approach of fully specifying the distributional
externalities. One knows that it is incorrect to ignore distributional externalities, but it is far more practical to approximate their effect by including distribution statistics directly in the social welfare function than it is to specify analytically the form of all the individual utility functions. When, for example, an increase in social welfare is brought about by a change in income distribution, one assumes that the actual mechanism at work is that the change in distribution affects the utilities of "enough" people so that social welfare is increased.

When using this approximation argument, the sources of the distributional value judgments are assumed to be a well-behaved Paretian social welfare function and the existence of distributional externalities. One is thus not free to impose any distributional value judgments directly but is constrained by the existence of distributional externalities. The distributional value judgments must still be consistent with a Utility-Paretian social welfare function even though the welfare function used as an approximation will not itself by Paretian. Fundamentally, the approach is still rooted in utilitarianism.

Rather than justify this social welfare function approach as an approximation, it is possible to justify it normatively. One might argue that in an individualistic society it is a justifiable normative assumption to define individual utility selfishly and so exclude distributional externalities. The social welfare function can be seen as reflecting a social contract among the members of society. The social contract clearly involves distributional judgments -- indeed, along with judgments about political and social justice, such distributional judgments form the basis of the social contract.
There is, however, no particular reason to assume that the Paretian value judgment should be reflected in the social contract. It is possible to develop a social contract approach to distributional judgments that is quite distinct from the utilitarian approach that underlies the Paretian value judgment. In any case, there seems no strong reason to assume that the Paretian value judgment is the most fundamental of all, more fundamental than any equity judgments. Indeed, such a view seems at odds with much political theory and political rhetoric.

If one accepts the normative (rather than the approximation) justification for including distribution statistics directly into the social welfare function, then one is still faced with the problem of specifying the source of the value judgments — the social contract underlying equity judgments. The problem is really the same as that facing the traditional welfare economists, except that it cannot be ducked — and one is no longer restricted to Paretian social welfare functions.

One can specify many social contracts and corresponding social welfare functions — one for each point in the political spectrum. Indeed, one can specify a "conservative" social contract that re-establishes the traditional dichotomy. First, as a value judgment, treat individuals as "selfish" and so ignore any distributional externalities, if they exist. Second, assume a Social Darwinist view of the social contract. The social and economic system represents an environment (or jungle) and people "should" be permitted to survive as best they can on their own. As a value judgment, one should not consider any equity criteria in social decision making since that would be at odds with the principle of "survival of the socially fittest." In this case, $\frac{\partial W}{\partial D} = 0$ and
efficiency criteria are all that is relevant. One presumably flips a coin to choose among efficient points, or perhaps attests that what the political and economic systems choose is best since the outcome (somewhat tautologically) reflects the power of the "socially fittest."

3. CONCLUSION

One theme of this article has, like many criticisms of modern welfare economics, been destructive in tone. One must give up the traditional equity-efficiency dichotomy except under very restrictive conditions. Traditional benefit-cost analysis is premised on invalid assumptions and the problem is not just one of value judgments. The existence of distributional externalities alone is enough, unless one makes very strong countervailing assumptions about what determines social welfare. However, some solutions are presented, with costs.

First, one can analytically specify (a) how distributional externalities affect individual welfare and (b) the form of the social welfare function. Second, one can, if possible, restrict the analysis to distributionally neutral changes. Third, one can include distribution statistics directly into the social welfare function. The first solution is empirically difficult but potentially feasible if one restricted the analysis to a few groups and assumed that the intra-group distributions were acceptable. The second is feasible under certain circumstances, if one could persuade social decision makers to go along. The third is quite practical, but involves abandoning the Pareto value judgment -- that society is always better off if one person is made better off without making anyone else worse off -- which has been perhaps the most fundamental assumption underlying modern welfare economics. The three approaches are,
of course, not mutually exclusive.

The advantages of the third approach seem to me to be worth the cost, at least for some purposes. The approach is practical and could be incorporated, for example, in optimizing models and benefit-cost analysis. It also provides a framework within which broader views of equity can be incorporated than is possible with a utilitarian, Paretoian social welfare function.

In any case, it is clear that it is not possible to ignore equity and deal only with efficiency. Such an approach is not more "scientific," it is simply based on invalid assumptions. It seems far more productive to try to provide a framework in which one can explicitly specify different equity judgments than to try to pretend that such judgments can be separated from the rest of the analysis. Concentrating on efficiency implicitly, and in actual practice, involves a value judgment either that only efficiency matters or that the workings of the market and the political system will always achieve the "best" distribution regardless of what practitioners of benefit-cost analysis recommend. Both of these judgments would seem, at best, not to be "widely acceptable."

Sherman Robinson
Princeton University
FOOTNOTES

1. See, for example, Sen [1970].

2. It should be noted that even the definition of "better off" is not obvious. For an argument that the definition must involve normative value judgments, see Gintis [1969].

3. See Bergson [1938].

4. The notation $\frac{\partial U}{\partial Y_i}$ denotes the derivative of $U$ with respect to $Y_i$. The partial derivative symbol $\partial$ is used even when, as in this case, it is applied to a single variable function. The symbol "d" (e.g., $dU$) will always denote a total derivative.

5. This is precisely what Atkinson [1970] has done in his derivation of a measure of inequality from a social welfare function. Atkinson uses incomes, not utilities, as the arguments in the social welfare function. He thus neatly avoids any questions of externalities by simply implicitly substituting out utility. For an elegant use of the Benthamite function, see Arrow [1971]. Arrow assumes everyone has the same utility function with declining marginal utility of income. Fair [1971] uses two welfare functions; (1) the sum of utilities and (2) their product which is the sum of log utilities.


7. See Harberger [1971], Little and Mirlees [1974],

8. For a survey of this literature, see Mishan [1960]. See also Math [1969] and Mishan [1973].

9. If all the $\partial$ statistics are homogeneous of degree zero in all incomes, this represents a functional relationship among the columns of $J$ and so results in a vanishing Jacobian determinant. See Chiang [1967], p. 192.
10. See, for example, Hirschman and Rothschild [1973].
11. See Hochman and Rogers [1969].
12. See Musgrave [1970].
14. For a similar analysis see Winch [1971], pp. 34-35.
15. See Rawles [1967] and [1971]. Rawles ends up with an implicit social welfare function that is non-Paretian. For a discussion, see Phelps [1973].
BIBLIOGRAPHY


