On the Use of Distributional Weights In Social Cost-Benefit Analysis

Arnold C. Harberger
University of Chicago

In the cost-benefit analysis of commodity taxes or subsidies, distributional weights complicate the standard textbook "triangle analysis." Net benefits are maximized with optimal subsidies in some circumstances, optimal taxes in others. Demand and supply elasticities play important roles in determining these weighted-welfare effects of commodity taxation. I then explore how to use distributional weights (a) to analyze investment projects and (b) to determine an optimum income-tax structure. In all these applications, the use of distributional weights is shown to have very strong and (to many people) intriguing implications. A final section explores alternative "solutions" to these difficulties.

I. Introduction

In this paper I explore the use of distributional weights in applied welfare economics. Section II deals with their application to certain familiar problems of commodity taxation. Some simple results are derived which are closely related to those relating to optimum tariffs and optimal export taxes in the literature of international trade. Section III expands on the critical role which price elasticities of supply and demand play in most applications of distributional weights in cost-benefit work. Three cases are set out which in a sense dramatize this role, and an effort is made to help the reader reach an intuitive understanding of it.

Section IV explores how the use of distributional weights affects the cost-benefit analysis of investment projects, while Section V shows how

This paper was presented at a conference on Research in Taxation sponsored by the National Science Foundation and the National Bureau of Economic Research, Stanford, California, January 1976.

© 1978 by The University of Chicago. 0017-3877/78/0003-0253 $0.50
these weights come into play in some rudimentary problems of income taxation.

Section VI tries to distill the (somewhat disquieting) messages that emerge from the preceding sections into a basic dilemma with which the use of distributional weights almost inevitably confronts us. Finally, in Section VII, three possible alternative "solutions" to this dilemma are presented and explored.

II. Welfare Analysis of Commodity Taxes

In the classical analysis of the efficiency costs of a commodity excise tax, the measurement can be approached in two distinct ways. In figure 1a the problem is viewed from the angle of the quantity axis. The reduction of consumption of \( X^2 \) units entails a loss of output which is valued by consumers as the (vertically shaded) area under the demand curve between \( X^2 \) and \( X_2 \). The resources released have a value, at their marginal supply price to this use, equal to the (horizontally shaded) area under the supply curve. This also measures the value that the owners of these resources place on the alternative use to which the resources will be shifted in the presence of the tax, which in turn is equal to the social benefit of having them in those uses as long as either (a) there are no distortions other than the tax being analyzed, or (b) the distortions which exist in the rest of the economy generate "external" costs and benefits whose net effect is zero (or small in the context of the problem under study). The net effect is the standard triangle of excess burden, ABE.

The second way of looking at the problem is from the angle of the price axis, as shown in figure 1b. Here the loss of consumer surplus (C−) generated by the tax is given by the vertically shaded area \( ABP_1P_2 \), while the loss of producer surplus (P−) is the horizontally shaded area \( AEP_1P_2 \). Since the gain (G+) to the government (other taxpayers, transfer recipients) is given by \( BEP_1P_2 \), there is a net loss of uncompensated consumer surplus equal to \( ABH \) and a net loss of uncompensated producer surplus equal to \( AEH \). Together these give exactly the triangle ABE derived by the alternate method in figure 1a.

The general expression for these "external effects" is

\[
\int_{T_{12}} \int \frac{D_i}{1 + \phi} \frac{dx_i}{dx_2} \, dy_2
\]

where \( D_i \) measures the excess of marginal social benefit over marginal social cost in activity \( i \). If \( D_i \) is constant for the problem in question (i.e., does not vary as a consequence of changes in \( T_{12} \) over the range from zero to \( T_12 \)), this expression reduces to

\[
\int \frac{D_i}{1 + \phi} \, dy_2
\]


Both of these methods make use of the postulate that the monetary value of benefits (+) and costs (−) is summed algebraically across the relevant individuals and groups. The method of figure 1a uses it directly in netting out the area \( S_{AXY} \) and indirectly in failing even to mention the implicit transfers shown as \( (C−, G+) \) and \( (P−, G+) \) which constitute the way in which the postulate is reflected in the method of figure 1b. In considering how the traditional evaluation of welfare costs changes when distributional weights are applied, we adopt the convention that the standard weight (\( \phi \)) is unity and that this weight applies to changes in the
welfare of the average citizen as well as of the government. When a particular group (i) is to be preferred, as against the average citizen, it is assigned a welfare weight \( \phi_i \) greater than one, and when it is to be "dispreferred" (not to use more pejorative terms like "punished" or "discriminated against"), its \( \phi_i \) will be less than one.

A little reflection will show that, when distributional weights are applied, the method of figure 1a cannot be used, while that of figure 1b must be adapted to accommodate the change of assumptions. The reason is that the only "base" to which a distributional weight can reasonably be applied, in an analysis of a group's welfare, is the change in economic surplus experienced by that group.

The way in which distributional weights work, when applied to the classical tax problem, can be illustrated by considering the case in which the demanders of \( X_i \) constitute a "dispreferred" group, while the suppliers are "average citizens." In this case, the analysis surrounding figure 1b would remain the same for the shaded areas lying below \( PT_i A \) but would be altered for the shaded areas above that line. Now the loss to consumers would be given smaller weight than the gain to the government. With distributional weight of \( \phi_i \) for consumers and of unity for the government and for suppliers, the benefit-cost picture now consists of a loss of the full triangle \( ABH \), a loss equal to a fraction \( \phi_i \) of the triangle \( ABH \), and a gain equal to \( (1 - \phi_i) \) since the transfer rectangle \( HBP \) perimeter. Needless to say, under these circumstances the net result can be a gain.

A natural extension of this line of reasoning is to find the optimum tax on \( X_i \), that is, the level of tax, \( T_i \), which produces the largest social gain, assuming the effects of other distortions to be absent or negligible. This case is illustrated in figure 2a. For a small change in the tax \( dT_i \), the following changes in costs and benefits are generated.

1. The lower "flat" rectangle labeled \( (F-, G+). \) For the present problem this generates no change in social benefit because the weights attaching to the owners of the relevant productive factors (\( F \)) and to the government (\( G \)) are equal.

2. The upper "flat" rectangle labeled \( (C-, G+). \) This generates a net social gain equal to \( (1 - \phi_i) X_i dP_i \) because the distributional weight attaching to consumers is less than that applied to changes in revenue to the government.

3. The two "tall" rectangles labeled \( (G-) \). These generate net costs equal to \( T_i dX_i \). The lower of these rectangles is part of the loss to supplying factors, and the upper is part of the loss to demanders, under both \( T_i \) and \( T_i \). Hence these areas do not contribute to the changes in welfare of these private groups, as between the two tax situations. But they do constitute negative components of the change in the command over resources of the government; hence their label \( (G-) \).

The quantitative measure of this contribution to the change in social welfare generated by a small change in the tax is \( T_i dX_i \) (with negative sign because \( dX_i < 0 \)).

Another way to see why these areas are included as cost items is to consider that they pass from being parts of the "transfer" rectangles \( [C-, G+] \) (\( F-, G+ \)) as being parts of the "deadweight loss" triangles \( [C-, F+] \) of a diagram like fig. 1b. In the process, the only group whose net position changes, as far as these particular areas are concerned, is the government.
The net change in social welfare stemming from a small change, $dT$, in the tax on $X_i$ is therefore

$$
\Delta W = (1 - \phi_i) X_i dP_i + T_i dX_i.
$$

Setting this equal to zero and defining $t_i = T_i/P_i$ and $q_i = (P_i/X_i) (dX_i/dP_i)$, we have

$$
\Delta W = (1 - \phi_i) X_i dP_i + T_i dX_i = 0 = [(1 - \phi_i) q_i] + t_i.
$$

The optimum rate of tax is therefore $t_i = -(1 - \phi_i) q_i$, that is, one minus the welfare weight of the consuming group, divided by the absolute value of their elasticity of demand for the product in question.

Figures 1b and 2a can also be used to represent the case where the demanders are ordinary citizens and the suppliers are the "dispreferred" group. Now, in figure 1b, $C$ is the rectangle labeled $(C\rightarrow G\rightarrow F\rightarrow)$ which cancels out, while the rectangle $(F\rightarrow G\rightarrow C\rightarrow)$ contributes a positive benefit equal to the fraction $(1-\phi_i)$ of its area. The two components are the entire triangle labeled $C$ and the fraction $\phi_i$ of that labeled $F\rightarrow$. In finding the optimum tax in this case, the relevant equation is

$$
\Delta W = -(1 - \phi_i) X_i dP_i - t_i dX_i = -[(1 - \phi_i) q_i] + t_i = 0.
$$

Readers with a background in international trade theory may at this point have a sense of déjà vu. For the twin problems of (a) the optimal export tax and (b) the optimal import tariff on a single commodity, as treated in that theory, have closely similar solutions. The optimal export tax, expressed as a fraction of the gross-of-tax price, is $t_i = -1/\eta_i$, where $\eta_i$ is the elasticity of foreign demand for this country's exports of good $i$. Similarly, the optimal import tax, expressed as a fraction of the net-of-tax price, is $t_i = 1/\eta_i$, where $\eta_i$ is the elasticity of the world excess-supply function of good $j$ facing this country's importers. These two taxes exploit, respectively, the monopoly power that this country has in the world market for its exports of good $i$ and the monopoly power that it has vis-à-vis the world-market supply of its imports of good $j$. It is well

\[\text{known, too, that the "optimum tariff" and "optimal export tax" are indeed optimal only from the standpoint of the country in question. From a world-welfare point of view they constitute distortions which taken by themselves would work only to reduce economic efficiency and welfare. Our results coincide exactly with those of international trade theory. When the welfare weight of the consumers of a commodity becomes zero, our optimum tax, } t_i = -(1 - \phi_i) q_i, \text{ on that commodity reduces to the optimum export tax of trade theory. Likewise, when the welfare weight of the suppliers of a commodity becomes zero, our expression for the optimum tax, } t_i = (1 - \phi_i) q_i, \text{ reduces to the formula for the optimal import tariff. It is quite clear, then, that the grand tradition of international trade theory already has dabbled with distributional weights, at least in a rudimentary sort of way, by assigning, for certain problems, a welfare weight of zero to foreigners. It is also gratifying to know that, for similar underlying assumptions, its answers and those of our own analysis are identical.}

The simplicity, and in this sense elegance, of the optimum-tax solution extends to the problem of optimum subsidies, a subject left untouched by the classical literature on trade. In this case we are dealing with situations in which the suppliers and/or demanders of a commodity constitute a "preferred" rather than "dispreferred" group. Figure 2b is an illustration that is relevant to either case. When the consumers of the commodity in question are a "preferred" group ($\phi_i > 1$), while the suppliers are just "ordinary citizens" ($\phi_i = 1$), the areas marked $(G\rightarrow C\rightarrow F\rightarrow)$ generate positive contributions to social welfare as a consequence of a small increase $dS$ in the subsidy on $X_i$. The area marked $(G\rightarrow F\rightarrow C\rightarrow)$, on the other hand, reflect pure transfers between groups of equal weight and hence generate no net benefit or cost. Finally, the areas marked $(G\rightarrow C\rightarrow F\rightarrow)$ represent increments in costs to the government, uncompensated by benefits to either suppliers or demanders. The equation for the change in welfare, which in turn defines the optimal subsidy $\delta_i$, is

\[\Delta W = -(\phi_i - 1) X_i dP_i - 2t_i dX_i = 0 \Rightarrow (\phi_i - 1)/\eta_i + \delta_i.\]

Here $\delta_i$ is the optimal percentage rate of subsidy $[(2t_i)/P_i] = -((\phi_i - 1)/\eta_i)$, and $\eta_i$ is the elasticity of demand for $X_i$.

When the suppliers of factor services are the "preferred" group ($\phi_i > 1$) and demanders have a unitary distributional weight, the corresponding relation is

$$
\Delta W = (\phi_i - 1) X_i dP_i - 2t_i dX_i = 0 = [(\phi_i - 1)/\eta_i] - \delta_i.
$$

Here $\delta_i$ is the elasticity of supply of $X_i$ and the optimal percentage rate of subsidy is $\delta_i = (\phi_i - 1)/\eta_i$.

Not surprisingly, just as the optimum rate of commodity tax turns out to be a partial exploitation of monopoly or monopsony power vis-à-vis
a "dispreferred" group, the optimum rate of subsidy turns out to be a manifestation of a sort of monopolistic or monopolistic preference toward a "preferred" group. That is to say, whereas the exploitation of a monopoly position entails raising the price to demanders by a percentage equal to the reciprocal (1/\(\eta_1\)) of the demand elasticity, a situation in which \(1 > \phi_1 > 0\) calls for raising price to demanders by a fraction of the monopoly percentage, the fraction being \((1 - \phi_1)\). When \(\phi_1 > 1\), however, optimization calls for setting a price that is below marginal production cost—a most unmonopolistic behavior. Yet the traces of the monopoly situation are present, in the sense that the percentage by which price should fall short of marginal cost here once again will vary inversely with the elasticity of demand. When demand is highly elastic, the optimum subsidy will be negligible, even though the consuming group is quite strongly preferred (say \(\phi_1 = 1.2, 1.5, \text{or} 1.4\)), as long as the elasticity of demand is low enough. The same principle applies, mutatis mutandis, in cases where suppliers are the preferred group, the relevant elasticity in this case being that of supply.

I consider the role played by supply and demand elasticities to be a key feature of the use of distributional weights in the analysis of taxes and subsidies as well as in other branches of applied welfare economics. Because of its importance, this role is dealt with in greater detail in the next section.

III. The Interaction of Weights and Elasticities

The bulk of this section (parts A, B, and C) shows how distributional weights should be applied (if they are used at all) in the analysis of three simple and familiar problems of policy economics. In part D, some general lessons are extracted from the earlier exercises.

A. Consider the familiar public-finance problem of designing a progressive excise-tax structure, say for a less developed country whose income-tax administration is not sufficiently strong to warrant its carrying the entire progressive component of the tax system. The common approach to this problem is to array commodities according to their cross-sectional income elasticities, taxing at the highest rates those whose consumption is most highly concentrated in the upper income groups and at progressively lower rates those with lower cross-sectional income elasticities. In terms of the present analysis, this procedure corresponds to attaching distributional weights to the various income brackets (with the weights being less than unity in the upper range of the income distribution and declining as income rises) and taxing commodities at rates which vary inversely with the average distributional weight of those who consume them. 

B. Consider now a hypothetical case in which the government adopts a policy of subsidizing certain agricultural products produced, in the main, by poor farmers. Let us assume, for the sake of the argument, that to be admissible for such a subsidy 75 percent of the production of a commodity would have to be carried out by farmers classified as being below the poverty level, according to certain indicators of socioeconomic status. For our purposes, we can take this to mean that for each commodity meeting the standard (say, barley, corn, and rice) the average distributional weight \((\phi_i)\) of the supplying farmers would be the same—perhaps something like 1.5.

The causal or conventional response in this case would be to set an equal rate of subsidy on the three commodities in question. But that would be as wrong as was a uniform tax in the problem of case A. The correct answer, given that it has been decided to employ distributional weights, is that the subsidies should vary in inverse proportion to the elasticities of supply of the three goods. If \(\phi_i\) is indeed 1.5 for all three, and if the elasticity of supply is 1.0 for barley, 1.5 for corn, and 2.0 for rice, then the optimal subsidy pattern would be a 50 percent subsidy on barley, one of 33\% percent for corn, and one of 25 percent for rice. It is easy, of course, to imagine real-world examples in which the variations among optimal subsidy rates, owing to differential supply elasticities, would be much wider than this.

C. Here we consider a case from the expenditure rather than the tax side of public finance. Assume that in the evaluation of public-sector projects and programs it has been decided to employ distributional weights and that weights in excess of unity are applied to unskilled and low-skilled labor. Figure 3 illustrates a case where a public program employs \(L_1, L_2\) of such workers at the market wage rate. The wage bill of the program is equal to \(L_1L_2W\), the increment of employment is \(L_1L_2\), and the amount earned by this increment is \(L_1L_2BD\); the overall increase in the earnings of this class of workers is \(W_1A L_2B L_1W\). Yet none of these represents the benefit to which the distributional weight premium \((\phi_1 - 1)\) should be applied. The relevant base for that premium is the gain in the surplus of the affected group of workers—the diagonally shaded area in figure 3. Note that only a small part of this benefit (the triangle \(ABD\)
The use of distributional weights

D. In each of the three cases just discussed, the distributional benefits of the policy under review depended on elasticities of supply and/or demand. In particular, such benefits were shown to be zero if the relevant elasticity was infinite. This is an old truth, but one which nonetheless many people find hard to swallow. If eggs were in infinitely elastic supply, a subsidy to their production would do poultry farmers no good; if construction workers were in infinitely elastic supply, a project which increases demand (at the supply price) for their services would do them no good; likewise, demanders will not be hurt by a tax on an item for which very close substitutes are available to which users can costlessly turn as the supply of the taxed item dwindles. All these situations reflect the fact that supply and demand prices reflect indifference, on the part of economic agents, between marginal units of the good or service in question and marginal units of the most immediate and economically interesting alternatives. When the elasticities are finite, this property applies to inframarginal and extra-marginal units as well, over whatever range is covered by the infinite elasticity. When the supply of eggs is infinitely elastic, a subsidy benefits consumers, not producers. When more workers are hired by a construction project, they will be benefited to the extent that the wage they are actually paid is above their supply price, and inframarginal workers will be benefited to the extent that the wage is further bid up above their supply price, but none of these groups will have its welfare improved if the new workers receive just barely their supply price and if the wage paid to the "old" workers does not change. Where a taxed item is in infinitely elastic demand, it is suppliers, not demanders, who are hurt by the tax.

Two examples drawn from the situations of some less developed countries may serve to reinforce the point. The first concerns the capacity of these countries to implement heavy or strongly progressive taxation of the income from capital. In this case the world capital markets have an infinitely elastic demand for the funds of LDC investors. There being no effective way for the LDCs to tax the income generated by such investments made by their nationals, the rate obtainable on these investments provides a floor below which the governments cannot readily drive the net-of-tax return on investments made at home. Thus heavier corporation income taxes in Panama or Guatemala simply mean less capital invested in the corporate sectors there and very likely more capital invested by nationals of those two countries on Wall Street or in the Eurodollar market. The net yield to Panamanian or Guatemalan investors on their investments at home would probably not be affected for long by a rise in the corporation tax rate. Instead, the gross yield would likely rise to reflect the tax, which would thus end up impinging more on the welfare of local consumers than on that of capitalists.

The second case concerns policies designed to deal with the problems of rural underemployment and poverty in certain LDCs. In some places the
opinion is widespread that such underemployment is manifested neither in open unemployment nor simply in a low equilibrium level of real wages. Instead, this line of thought runs, it is reflected in a highly elastic supply of unskilled labor, unwilling to work below the going "conventional" wage but willing to work at it. Accepting these assertions as facts, the welfare-economics conclusion would be that programs to augment the demand for such labor would produce no distributional benefit and could be justified only if they could "bear" a full costing of the labor they employed.

Programs designed to fight poverty by augmenting the demand for labor really come into their own when the facts are quite different—that is, when the conventional market supply and demand curves are quite inelastic. This is a direct application of the analysis surrounding figure 3. The true power of such programs is measured by their capacity, per unit of outlay, to bid up the general level of wages for the given type of labor in the labor market in question. And, as was shown above, the bulk of distributional gains they produce accrue to workers having no direct involvement with the project.

IV. Distributional Weights and the Social Evaluation of Investment Projects

This subject has already been touched upon (Sec. III.C above) in connection with measuring the distributional benefit of a project that increases the demand for labor of a type that is judged to merit favorable treatment (i.e., with a distributional weight \( \phi_i > 1 \)). In that case it was seen that the distributional effect was mainly external to the project itself, hinging upon whether the market wage for workers of the type in question was or was not raised by the presence of the project. Where in fact the market wage was raised, it was seen that the benefits would in most instances accrue mainly to workers employed in jobs having no direct connection with the project. The distributional benefits were in this sense indirect and external to the project in question. In this section we examine situations in which distributional weights impinge more on the direct benefits (Sec. IV.A) or costs (Sec. IV.B) of a project.

A. For the benefit side, consider a class of projects (public roads, parks, etc.) whose benefits accrue mainly "in kind," and for which the beneficiaries do not have to pay a charge or fee. Assume, furthermore, that projects of this class can be replicated at constant cost and that there are no interrelationships in demand among the separate projects (in the sense that the benefits and costs of any one project are independent of whether or not any of the others is undertaken).\(^6\)

---

\(^6\)This set of assumptions is not realistic, but it is useful because it serves the didactic purpose of enabling me to focus directly on the point I am trying to make.
"surplus" according to standard criteria; on this basis project 3 is at the margin, and projects 4, 5, and 6 are not worthwhile.

Once a distributional weight of 2 is applied to the beneficiaries, however, all but project 6 pass the test of social profitability. The critical point here is that now projects pass the test as long as their unweighted benefits amount to more than half their costs. That is to say, the new test accepts a "waste" of up to half the resources devoted to a project as a legitimate "price" to pay for the income transfers that the project carries with it.

It is important that the reader appreciate the sense in which waste is involved. If, for example, the beneficiaries of each project simply were "average citizens," with distributional weights \( \phi_i \) equal to one, then figure 4e would apply, and the acceptance of projects 4 and 5 would genuinely imply the social costs (shaded areas) depicted there for those projects.

Alternatively, if it were known that the projects in question would be financed by additional taxes on people whose distributional weight was the same (2.0) as that of the beneficiaries, then projects 4 and 5 would again be rejected by the social-evaluation criterion. This time their social costs would be twice the areas shown for these projects in figure 4e—twice because in this case both the benefits and the costs of each project would be multiplied by the distributional weight of two.

Thus we see that the only reason why projects 4 and 5 turn out to be acceptable in figure 4b is because the beneficiaries have a higher distributional weight than those who pay the costs of the project. Society ends up, by the criterion of distributional weights, paying a price in terms of efficiency for each incremental distributional benefit it gets.

This seemingly modest point is of extreme importance, for it introduces a whole new element into social-project evaluation, making the evaluation of the project itself depend in a critical way upon the manner in which the project is financed. One of the notable virtues of the traditional methodology—the fact that its measure of the social profitability of a project is largely independent of how the project is financed—is therefore lost.

5. Turning now to cases in which distributional benefits come into play with respect to the direct costs of a project, let us consider once again a situation in which the workers hired on a project have a distributional weight greater than one. But here, as distinct from what was done in Section III.C, let it be true that the projects in question pay a higher-than-

---

THE USE OF DISTRIBUTIONAL WEIGHTS

market wage. Under these conditions, it is likely that the marginal worker employed by the projects will perceive some economic rent, giving rise to a possible distributional benefit (if the relevant \( \phi_i > 1 \)) or cost (\( \phi_i < 1 \)).

To isolate the effect in question from that treated in Section III.C, let us assume that we are dealing with a type of labor that is in infinitely elastic supply (over the relevant range). This rules out a new project's causing the market wage to be higher than it otherwise would be, and at the same time establishes unambiguously the private opportunity cost of any workers hired by the project. For the sake of simplicity, assume that the only costs of the projects considered are for labor of the type in question and that such labor is paid (for working on the projects) 1.5 times the wage it could get elsewhere. The above assumptions are reflected in figure 5a, where the conventional procedures of cost-benefit analysis are applied. Starring complications other than those stated, the social calculus of costs would differ from the strictly monetary one in that labor would be valued at its alternative earnings of \$1.00 rather than the wage actually paid. This is equivalent to assigning to each worker employed a benefit equal to the economic rent (\$0.50) he receives. While only project 7 would pass the strictly monetary cost-benefit test (assuming that project benefits came in the form of cash), both 7 and 8 would pass the conventional social cost-benefit evaluation. Projects 9, 10, 11, and 12 would fail both tests.

Figure 5b depicts the same projects modified to incorporate a distributional weight of 2.0 for workers. This weight must be applied to the amount by which the earnings of labor (taken in all cases to be \$1.50) exceed the opportunity cost as seen by the workers in question. This economic rent is still \$0.50, as in the conventional analysis. But here a social valuation of \$1.00 is placed on the economic rent of workers, in the light of their distributional weight of 2.0. Thus, in the final reckoning, the labor which is being paid \$1.50 has a net social cost of only \$0.50. This shifts projects 9, 10, and 11 into the acceptable category, leaving only project 12 with adjusted costs greater than benefits.

Clearly, the application of distributional weights makes the difference between figures 5a and 5b, but its full meaning may not yet be clear. To explore a bit further, assume now that the wage paid by the projects was \$1.00 rather than \$1.50. Now the laborers would earn no economic rent, hence they would have no benefit to which to apply the distributional weight. Their social opportunity cost would equal their alternative earnings of \$1.00 (regardless of whether a distributional weight were assigned or not, and regardless of the magnitude of whatever weight might be assigned), and only projects 7 and 8 would pass muster.

The argument of the preceding paragraph points up the most troublesome aspect of distributional weights as they impinge on project costs. So long as some costs are for factors whose suppliers have distributional weights greater than one, it is possible to raise the net benefits of a project

---

*The method of financing enters to a degree even in the traditional analysis. For example, the benefits and costs of a given road will differ if a toll is charged or not. Also, a distribution of costs from footnotes which is "social" in a particular project may under certain circumstances give rise to an external benefit attributable to that project. But instances of these types are likely to be infrequent and their impact on the measured social merit of a project small in comparison with what happens under the generalized use of distributional weights.*
The use of Distributional Weights

V. Distributional Weights and the Income Tax

The criterion most often mentioned as the basis for assigning distributional weights is income, the presumption being that, the higher the income of an individual or group, the better for the distributional weight (\( \phi_i \)) to be applied. This naturally brings to mind the income tax, as a possible way of directly—once and for all, as it were—taking the weights into account, rendering it unnecessary thereafter to bring them into play when analysing commodity taxes, individual projects, etc. Unfortunately, the world is not so simple.

A. Consider first the case where the supply of labor is zero elastic with respect to the wage and that of saving with respect to the rate of return thereon. These assumptions rule out marginal adjustments of either leisure or savings and in effect turn the income tax into something like a variable head tax or capital levy, with no incentive effects or efficiency costs whatsoever.

This would seem to be fertile soil for a progressive income tax, and many economists in fact believe that such a tax is what would emerge (analytically) from an application of distributional weights analysis in this case. Also, as many other economists remember from their reading of Pigou and other authors, this is not true. Declining marginal social utility of income (in the absence of incentive effects) not to progressive but to confiscatory taxation, so as to bring about the full and complete equalization of income. If there were no costs associated with the tax-transfer mechanism, the optimal policy under distributional weights would be to tax away all income in excess of \( \bar{P} \) (the average over all the relevant units) and use the proceeds to bring every unit whose income is below \( \bar{P} \) up to that level. If the tax-transfer mechanism carries with it some administrative costs, the optimal policy would presumably be to tax away all income in excess of \( \bar{P} + \delta \) and supplement lower incomes up to \( \bar{P} - \delta \), where \( \delta \) and \( \delta' \) play the same role as transport costs in international trade theory. Basically, the end result of income equalization remains, only slightly modified.

B. Now consider the case where the supply of labor of any given type is upward sloping but where each type of labor is a noncompeting group vis-à-vis the other types. That is to say, the upward slope of the supply curve of each type of labor reflects substitution between labor and leisure of that particular group, not shifts of labor from one group (type, class) to another.

Assume, too, that the labor types are arrayed in a determined order with respect to the incomes they earn, type 2 being richer than type 1, etc., \( \phi_1 \) and type \( \Phi \) being the richest. To each group we can therefore assign a distributional weight, with \( \phi_1 > \phi_2 > \cdots > \phi_{n-1} > \phi_n \).

For the moment let us rule out the idea of taxing (for distributive
persons with distributional weights greater than one; or our analysis will therefore consider (initially only) those groups having distributional weights less than one. Group 1 has a weight \( \phi \), only moderately less than 1; group \( N \) has the lowest weight.

If we set ourselves the problem of determining the "proper" flat-rate tax to apply to each of these noncompeting groups, the analysis of figure 2a holds precisely, assuming that the demanders of each group's services have "average" distributional weights (\( \phi = 1 \)). A rise in the rate of income tax (\( t \)) in group \( i \) produces an uncompensated loss to the government (\( G \)) equal to \( t_i w_i d_i \), while engendering transfers to the government of \( L \cdot d_i \) from demanders, and of \( L \cdot d_i \) from suppliers, of \( L \cdot d_i \), respectively. The loss to demanders nets out against the corresponding gain to the government, but the transfer between suppliers of \( L \cdot d_i \) and the government occasions a social gain of \( (1 - \phi)L \cdot d_i \). The optimum tax on group \( i \) is reached when \( (1 - \phi)L \cdot d_i \) and, or where, defining \( e_i \) as the elasticity of labor supply for group \( i \), \( L \cdot d_i \) is \( w_i e_i d_i \), or where, defining \( e_i \) as the elasticity of labor supply for group \( i \), \( L \cdot d_i \) is \( w_i e_i d_i \).

This fact that \( L \cdot d_i \) (not the tax rate itself, that is equal to \( (1 - \phi)L \cdot d_i \)) is due simply to the fact that the tax rate is here defined as applying to the gross (rather than the net) wage.

The use of distributional weights those with incomes between \$1,000 and \$2,000 pays 20 percent of the excess over \$1,000. The benefits associated with these transfers is divided as follows:

1. A part \( (-1)N \cdot B_i \cdot d_i \) is a transfer between demanders of type \( i \) labor and the government. Since the demanders (direct and indirect) of each type of labor are here assumed to be a representative sample of the whole population, having a distributional weight of 1 (equal to the government's), no distributional benefit or cost is associated with this transfer.

2. A part \( (-1)N \cdot B_i \cdot d_i \) is a transfer from suppliers of type \( i \) labor to the government. The social benefit of this transfer is calculated by applying to it a weight of \( (1 - \phi) \).

3. There remain a series of other transfers, \( N \cdot B_i \cdot d_i \), to the government from taxpayers in the income brackets \( (j > i) \) above the one for which the marginal rate change is being assessed. To each of these transfers, the weight of \( (1 - \phi) \) is applied.

Table 1 carries out the necessary calculations for a simple example, supposed to represent a hypothetical LDC with an income tax that strikes only in the upper part of the income distribution. The lowest bracket covered by the income tax is from \$3,000 to \$5,000. The distributional weight applying to this group is 0.9; therefore, a transfer of 1 from this group to the government generates a "distributional benefit" of 0.1. For higher income brackets, the relevant distributional weights decline, until finally, for the group with incomes between \$18,000 and \$32,000, a transfer of 1 from this group to the government generates a "distributional benefit" of 0.5.

The table seeks to elucidate the calculation of the optimal income tax in this situation. The hypothetical experiment is always a rise of one percentage point in the marginal tax rate, \( \phi \), applying to a specific group. The assumption made regarding the reaction of labor supply to this disturbance is simply that the equilibrium level of \( L \) falls by 0.5 percent. The factor \( 1 \) applied to \( N \cdot B_i \cdot d_i \) makes the approximation that the average taxable income in the \( j \)th bracket is precisely at the midpoint of the bracket. Needless to say, this is only an approximation, and the formula could easily be adjusted to fit the facts of any given case, with any desired degree of precision.

Recall that \( d_i \) is negative for an increase in \( L \).

Elasticities of demand and supply that are approximately equal come close to yielding this result. The precise result is obtained when \( \phi \) to \( (1 + e_i \phi_i)/2 \) (if \( e_i \) is the elasticity of supply, \( \phi_i \) the elasticity of demand \( \phi = 0 \) for labor of type \( i \)).
### TABLE 1

**Determination of Optimal Income Tax for Groups with Distributional Weights Less Than Unity**

**A. "Inframarginal" Benefits Due to Transfers from Higher Income Groups**

<table>
<thead>
<tr>
<th>Income Bracket</th>
<th>Units</th>
<th>Distributional Weight</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>80</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3,000-5,000</td>
<td>80</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5,000-10,000</td>
<td>60</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7,000-15,000</td>
<td>45</td>
<td>4.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10,000-20,000</td>
<td>35</td>
<td>3.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15,000-32,000</td>
<td>25</td>
<td>2.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>270</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**B. Costs and Benefits Associated with Own Group**

<table>
<thead>
<tr>
<th>Income Bracket</th>
<th>Units</th>
<th>Total Income in Group</th>
<th>Social Cost Due to 1/2% Reduction in Labor Supply</th>
<th>Social Benefit Due to Transfer from Own Group</th>
<th>Avg. Transfer of (1)</th>
<th>Own Group Benefit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>80</td>
<td>320,000</td>
<td>1,600</td>
<td>10</td>
<td>1.1</td>
<td>80</td>
</tr>
<tr>
<td>3,000-5,000</td>
<td>80</td>
<td>320,000</td>
<td>1,600</td>
<td>10</td>
<td>1.1</td>
<td>80</td>
</tr>
<tr>
<td>5,000-10,000</td>
<td>60</td>
<td>300,000</td>
<td>900</td>
<td>15</td>
<td>0.2</td>
<td>180</td>
</tr>
<tr>
<td>7,000-15,000</td>
<td>45</td>
<td>250,000</td>
<td>450</td>
<td>25</td>
<td>0.3</td>
<td>240</td>
</tr>
<tr>
<td>10,000-20,000</td>
<td>35</td>
<td>375,000</td>
<td>1,875</td>
<td>30</td>
<td>0.4</td>
<td>500</td>
</tr>
<tr>
<td>15,000-32,000</td>
<td>25</td>
<td>375,000</td>
<td>1,875</td>
<td>30</td>
<td>0.4</td>
<td>525</td>
</tr>
<tr>
<td>Total</td>
<td>270</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**C. Calculation of Optimal Tax**

Incremental Cost = Inframarginal Benefit + Own Group Benefit = Incremental Benefit

<table>
<thead>
<tr>
<th>Incremental Cost</th>
<th>Inframarginal Benefit</th>
<th>Own Group Benefit</th>
<th>Incremental Benefit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,600</td>
<td>80</td>
<td>8</td>
<td>810</td>
</tr>
<tr>
<td>1,600</td>
<td>80</td>
<td>8</td>
<td>810</td>
</tr>
<tr>
<td>1,930</td>
<td>865</td>
<td>8</td>
<td>1065</td>
</tr>
<tr>
<td>2,000</td>
<td>700</td>
<td>7</td>
<td>960</td>
</tr>
<tr>
<td>1,875</td>
<td>450</td>
<td>4</td>
<td>720</td>
</tr>
<tr>
<td>1,875</td>
<td>450</td>
<td>4</td>
<td>720</td>
</tr>
</tbody>
</table>

Optimal Tax Rates:

- $810 / $1,600 = 51%
- $1,065 / $1,930 = 55%
- $940 / $1,700 = 54%
- $750 / $1,875 = 40%
- $525 / $1,875 = 28%

In table 1A, the income brackets, numbers, and distribution weights relevant for the various groups are presented. In the right-hand columns of this section the calculation is made of the distributional benefit associated with the "inframarginal" transfers (from higher-income groups to the government) that are occasioned by a one-percentage-point rise in each given bracket rate. For example, a rise of one percentage point in the marginal rate applying to the first bracket means that each higher income group will have to pay $20 more in tax. For the second group this transfer carries a benefit of $20 × (1 - 0.8) = $4 for each unit, or $240 for all 60 units. The same transfer of $20 carries a per unit benefit of $20 × (1 - 0.3) = $10 for the highest income group, or $150 for all 15 units in that group.

In table 1B are shown the costs and benefits associated with the particular group for which the rise in tax rate is assumed. Recall that the labor supply of any group is supposed to be reduced by 0.5 percent in response to a one-percentage-point rise in that group's marginal tax rate. We measure labor supply in "dollars worth," so that a 0.5 percent reduction in the labor supply of the first group is measured as a reduction of $1,600 (= 0.005 × $320,000), while for the last group a similar 0.5 percent reduction is measured as $1,875 (= 0.005 × $375,000). The efficiency cost of a small rise in the tax is simply this induced reduction in labor supply times the marginal tax rate of the group in question; it is shown in column 3.

In columns 4–6 is measured the social benefit of the transfer from each group to the government as a consequence of a rise of one percentage point in that group's own tax rate. The average transfer given in column 4 assumes that the members of the income class are symmetrically distributed about its midpoint. Thus in the first group a taxing unit with $3,500 of income would pay only $5.00 of extra tax when the rate on income above $3,000 was raised by one point. However, with a symmetrical distribution, to each such family there would correspond another one with $4,500 of income, which would pay $15 of extra tax. In this way we can impute to the group as a whole an average transfer equal to 1 percent of half of the bracket width. The benefit associated with this transfer (shown in col. 6) is equal to the transfer itself (col. 4) times the benefit per dollar transferred (1 - $\beta_1$, given in col. 5), times the number of taxing units in the group (col. 1).

Table 1C presents the calculation of the optimal marginal tax rate for each group. This is done by setting the incremental cost of a one-point rise in the marginal rate against the sum of the two types of incremental benefit—the inframarginal benefit associated with the transfers from higher-income groups (table 1A, totals) plus the social benefit accruing as a result of transfers from the same group for which the tax is being set (table 1B, col. 6).

The results of the calculation show optimal tax rates ranging down from 57 percent for the lowest income group shown in the table to 28 percent for the highest. That is to say, in spite of distributional weights that decline from 0.9 to 0.5 as income rates, the optimal schedule of marginal tax rates for the income groups shown is uninterruptedly regressive! The
reason for this result is not hard to find—it lies in the "triangular" pattern of columns 4-7 of table 1A. When the marginal tax rate is raised for a given income group, all higher groups also pay the extra tax on that particular bracket. When the given group is low on the income scale, there are many higher groups that pay this extra tax. When a high-income group is concerned, however, the number of groups still higher on the income scale will certainly be fewer, and may be (in the case of a tax rise for the highest bracket) none at all.

D. The example of Section V.C above considered the setting of an optimal tax schedule for groups with distributional weights less than unity. Though the restriction of income taxation to such groups may seem at first glance so obvious as to go without saying, further reflection reveals that no income-tax system based on the strict application of this criterion would meet this criterion. Every income-tax system that is fully consonant with distributional weights that decline with income must assess positive marginal tax rates against at least some income ranges for which the distributional weight is greater than or equal to unity.

This proposition follows from precisely the same characteristic that tends to make the optimal tax system regressive among groups with distributional weights that are less than 1. Consider the income group whose distributional weight is precisely equal to 1. There is (by definition) no "own group" benefit stemming from the transfer of tax funds between this group and the government. There is, however, an "inframarginal" benefit from raising the marginal tax rate above zero for this group—for such a rise will cause the higher-income groups to pay more tax, without at the same time distorting their marginal choices between effort and leisure. Since this inframarginal benefit, from a small rise in the group's tax rate above zero, is finite (or of the first order of smalls), while the social cost due to the reduction in the group's labor supply is infinitesimal (or of the second order of smalls), it will always pay society to set the tax on this group above zero.

With respect to groups with weights that are strictly greater than 1, similar considerations apply, modified by the fact that the own group benefit is now necessarily negative and that the inframarginal benefit will now normally have some negative components. Nonetheless, barring quite unusual discontinuities either in the pattern of distributional weights or in the income distribution itself, there will typically be a substantial range of below-average incomes which "qualify" for significant and positive marginal tax rates.

Such a case is illustrated in table 2, which can be regarded as a continuation of table 1 to cover groups with distributional weights greater than 1. (For convenience, the subscripts used to distinguish the different income groups are keyed to their distributional weights, so that the group with a distributional weight of 1 is given a subscript of zero; that with a distributional weight of 1.1 has a subscript of −1; that with a distributional weight of 0.9 has a subscript of 1; and so on.)

---

**TABLE 2**

**DETERMINATION OF OPTIMAL INCOME TAX FOR GROUPS WITH DISTRIBUTIONAL WEIGHTS ≥1**

<table>
<thead>
<tr>
<th>Income Bracket (f)</th>
<th>Distributive Weight (φ)</th>
<th>Distributive Benefit ($) from Transfer of</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1-1.000</td>
<td>120</td>
<td>1.1</td>
</tr>
<tr>
<td>1.000-1.100</td>
<td>60</td>
<td>1.2</td>
</tr>
<tr>
<td>1.100-1.200</td>
<td>80</td>
<td>1.3</td>
</tr>
<tr>
<td>1.200-1.300</td>
<td>100</td>
<td>1.4</td>
</tr>
<tr>
<td>1.300-1.400</td>
<td>80</td>
<td>1.5</td>
</tr>
<tr>
<td>1.400-1.500</td>
<td>60</td>
<td>1.6</td>
</tr>
<tr>
<td>1.500-1.600</td>
<td>120</td>
<td>2.1</td>
</tr>
<tr>
<td>1.600-1.700</td>
<td>120</td>
<td>3.1</td>
</tr>
<tr>
<td>1.700-1.800</td>
<td>180</td>
<td>4.1</td>
</tr>
<tr>
<td>1.800-2.000</td>
<td>240</td>
<td>5.1</td>
</tr>
<tr>
<td>2.000-2.200</td>
<td>320</td>
<td>6.1</td>
</tr>
<tr>
<td>2.200-2.500</td>
<td>400</td>
<td>7.1</td>
</tr>
</tbody>
</table>

**C. CALCULATION OF OPTIMAL TAX**

- Incremental Cost = Inframarginal Benefit + Own Group Benefit = Incremental Benefit
- 1,750 / 40 = 44.5
d | 600 / 40 = 15
- 560 / 24 = 23
- 500 / 24 = 21
- Optimal Tax Rates | \( T_f = 493 / 2,500 = 0.19 \)
- \( T_f = 275 / 680 = 0.40 \)
- \( T_f = 143 / 360 = 0.40 \)
- \( T_f = 55 / 520 = 0.10 \)
Table 2A shows the inframarginal benefits generated by a one-percentage-point increase in the tax on the lower-income groups. Column 4 shows the effects of a transfer of $10 per taxpaying unit, a raising of a one-percentage-point increase in the rate in the $0-$1,000 income group. In column 5 are displayed the effects of a transfer of $6.00 per taxpaying unit, generated by a one-percentage-point rise in the marginal tax rate of the $1,000-$1,400 income group. Similar interpretations should be given for the figures in column 6 and 7, here referring to rises in the tax rates on the $1,400-$2,000 and the $2,000-$3,000 income groups, respectively.

Table 2B follows directly the format of the corresponding part of table 1, using (for col. 3) some assumption concerning labor supply—namely, that for each group a one-percentage-point rise in its tax rate causes a fall of 0.5 percent in its labor supply. The own group benefit shown in column 6 is naturally zero for the group whose distributional weight is unity and negative for those with distributional weights of less than 1.

Table 2C displays the calculation of the optimal marginal tax rate for each group. In spite of the fact that the direct distributional benefit of transfers from groups 0, -1, and -2 to the government is zero or negative, it is surprising to find that the optimal marginal tax rate for these groups is 40 percent—equal to that (q) corresponding to the second highest income bracket ($12,000-$16,000) and higher than that (q) corresponding to the highest income group (see table 1C).

Thus we can see quite clearly that the use of distributional weights that vary inversely with income need not give rise to a generally progressive tax structure and can very well conduce to relatively high income taxes even for the poor.*

It should be noted, too, that the weights employed in tables 1 and 2 had a relatively modest range of variation and that the general attributes of the hypothetical economy there treated were not unrealistic, at least for the poorer countries of the world.*

VI. What is Wrong with Distributional Weights?

There can be no doubt that distributional weights (particularly weights which vary inversely with income) have a strong appeal to those nurtured in the grand tradition of economics. The notion of a diminishing marginal utility of income pervades the classical and neoclassical economic literature; and although most economists have come to shun interpersonal comparisons of utility whenever they feel they can avoid or circumvent them, they usually make them when they have to. And when they do, they most commonly rely on the assumption that the loss of a dollar means less to the rich man than to his poorer brother.

Now, it is difficult indeed to tell the difference between a system of distributional weights that decline with income and an alternative system of measurable utility made interpersonally comparable on some basis and characterized by diminishing marginal utility of wealth or income. In short, I think it is fair to say that the notion of distributional weights (that decline with income) has been lurking for a long time in our professional unconscious.

Yet the implications for policy of a thorough and consistent use of distributional weights turn out to be quite disturbing—even to those who (perhaps as today's representatives of our utilitarian heritage) are attracted to the idea. These implications include

1. With respect to commodity taxation (see Secs. II, IIIA, III-B), it is only by accident that an optimal excise tax (or subsidy) would be zero. For most commodities, the distributional effects of a marginal increment of tax would be different from zero, starting from an untaxed, unsubsidized situation. An optimum would be found only when the point was reached where the distributional gains from a small further increase in the tax or subsidy in question were just matched by the additional efficiency losses it entailed.

2. With respect to investment projects (Sec. IV), the implications of distributional weights are quite similar; where the estimated net effect of a project is a transfer to people with distributional weights higher than 1, the distributional benefit assigned on this ground represents the price that society "should" (by the logic of distributional weights) be willing to pay in terms of economic efficiency and still approve the project. That is to say, when a project carries distributional benefits, the straightforward application of distributional weighting schemes would lead us to use as a criterion the balancing (at the margin) of distributional benefits against efficiency costs.

3. Finally, with respect to income taxation (Sec. V), we find that such taxation can readily turn out to be regressive over significant ranges of income, that it will always apply positive marginal tax rates to groups from which the direct distributional consequence of a transfer to the government is
The use of distributional weights is the "waste" takes place in the process of effectuating the transfer itself, with the beneficiaries actually receiving only a fraction of what the transferors give up. This comes closer to what actually happens with inefficient investment projects or other economic policies (such as taxes) or programs (such as the U.S. sugar program).

An example I have found useful for conveying the essential elements of this sort of transfer is a project to send ice cream on camelback across the desert from a richer oasis to a poorer one. If \( \lambda \) is equal to 1 and the richer oasis has twice the per capita income or consumption of the country as a whole (this makes \( \phi_1 = 1/2 \)), while the poorer oasis has a per capita income equal to half the national average (for a \( \phi_2 \) of 2), it would be possible for up to three-fourths of the ice cream to melt en route without causing the project to fail: this is the distributionally weighted cost-benefit test. (The resource costs of the camel transport are neglected in this example.)

If \( \lambda \) were equal to 3, the \( \phi_1 \) of the richer oasis would be 1/3 and that of the poorer one would be 1. In this case (again neglecting the resource costs of transport), up to 63/64 of the ice cream could melt away without causing the project to fail the test.

The lesson from these examples is clear: When distributional weights are used together with weighting functions of the type most commonly employed in writings on the subject, the result is to open the door to projects and programs whose degree of inefficiency by more traditional (unweighted) cost-benefit measures would (I feel confident) be unacceptable to the vast majority of economists and of the informed public.

The dilemma, then, is that, when the differences in weights are small, distributional considerations are reflected only to a minor degree in the evaluation process. When, on the other hand, the differences in weights get to be large, it is all too easy for considerations of distribution to swamp those of efficiency altogether, and for grossly inefficient policies, programs, and projects to be deemed acceptable.

VII. In Search of a Solution

There are a variety of ways in which one can approach the dilemma presented in the preceding section (and, indeed, throughout this paper). I shall here present three distinct alternatives, the main differences among them being their degree of fidelity to the basic idea that the right way to take distributional considerations into account in social cost-benefit analysis is by way of distributional weights.

A. Improving Both Tests Simultaneously

One sure way to avoid paying an exaggerated price in terms of lost efficiency for the redistributive benefits that a weighting scheme would
THE USE OF DISTRIBUTIONAL WEIGHTS

components should continue to pass, as in the traditional approach, the unweighted cost-benefit test, while requiring a nonnegative weighted net benefit only for the project as a whole. In this case individual components could have weighted costs greater than weighted benefits, though they would not be permitted to have unweighted costs greater than unweighted benefits. This type of hybrid criterion implicitly gives priority to the efficiency criterion; the opposite hybrid—insisting that each separable component pass the weighted test while only the global program or project must pass the unweighted test—gives relative priority to the distribution criterion.

B. A Single Premium Magnifying the Net Benefits (Costs) of Those Below a Poverty Line

The path that leads to this "solution" is devious but (I think) interesting. If one wants to justify the traditional cost-benefit procedures (which neglect distributional considerations), the best approach is surely to adopt the assumption that lump-sum transfers can be effectuated at will and without any cost in terms of economic efficiency. This is, at least in one dimension, the key assumption underlying the Hicks-Kaldor principle under which programs or projects are to be approved as long as the gainers could in principle compensate the losers and still be no worse off than before.

Of course, such compensations and transfers are rarely made—a fact that is subject to at least two interpretations: (a) "society" could costlessly effectuate the transfers but chooses not to do so because it is satisfied with the status quo, and (b) the compensations and transfers are not costless after all. Clearly, the second of these reasons carries more weight (and with more people) than the first.

Yet the fact that it costs more than zero to effectuate a transfer—once a decision in that direction has been reached—does not mean that we need to accept resource waste of one-half or three-fourths or seven-eighths or 85/64 of the amounts involved in order to bring the transfer about. While the most eminent minds in the economic profession would be hard put to find ways of effectuating quantitatively important transfers at efficiency costs of less than 1 percent of the amounts involved, most beginning graduate students could in one afternoon invent a hundred ways of bringing about major transfers at an efficiency cost of less than 20 percent of the amounts involved.

On this reasoning, it seems quite sensible to think of there being something like a critical level (say, 10 percent) of efficiency cost per dollar transferred. If a way is found of effectuating desirable transfers more cheaply than this—either through a conscious program or as a side effect of other programs or projects whose main aims may be quite different—it should (on this line of reasoning) be welcomed. But no merit
should accrue to a program or project for its distributive achievements if their implicit cost is above the critical level (here 10 percent of the amount transferred).

This view of the world fits neatly into modern economics. The 10 percent premium may be thought of as reflecting a shadow price of public funds that is different from unity. Alternatively, one can reach the same conclusion by following one of the cardinal principles of cost-benefit analysis: Never attribute to an action a benefit that exceeds the alternative cost of achieving the same result.

So far, this line of argument would lead us to try to make explicit the various transfers that might be implicit in a program or project of other policy decision and to array them in descending order of the net benefits that would be assigned to them using distributional weights. Where the net benefit was less than 10 percent of the amount transferred, it would stand unaltered in the cost-benefit calculation. But where the net benefit from an implicit transfer was greater than 10 percent of the amount transferred, it would be stricken from the balance and replaced by the alternative cost figure of, in the present example, 10 percent of the amount transferred.

Up to this point, we are still being quite faithful to the use of distributional weights. All we have done is to recognize the alternative cost principle and bring it to bear upon our problem. But here the argument starts to follow a different and less strictly theoretical track. Relative magnitudes become important here. How rich are the payers relative to the recipients of the implicit transfers in question? How rapidly do the distributional weights fall as income (or consumption) rises over this range? How large is the critical premium (10 percent in our example) relative to the range within which the distributional weights vary? And finally, does “society” really “care” about transfers taking place among the wealthy and the reasonably well-to-do?

The specific “facts” that we presume to exist, in making the next leap in the argument, are the following:

1. The distribution of income has a wide span—say, a factor of 10 or more.
2. Distributional weights are not all tightly clustered about 1. For a 10-fold range of income distribution, the range of weights might be 0.50-2 or 0.33-3 or 0.25-5—but not 0.95±1.05.
3. The critical premium is not much more than 10 percent (say, 20 percent at the outside).
4. “Society” is substantially indifferent concerning transfers among those with above-average incomes.

The use of distributional weights

Presumption (4) enables us to neglect transfers within the well-to-do group. Presumptions (1) and (2) indicate that transfers from those significantly above a “poverty line” to those below it will carry substantial benefits when distributional weights are used. Presumption (3) says, however, that these benefits will be supplanted, in the cost-benefit calculus, by the standard premium (10 percent) based on the alternative cost of effectuating transfers.

This critical premium dominates the likely calculus of distributional costs and benefits to such an extent that the precise pattern of decline of the distributional weights as income, consumption, or wealth rise turns out to be quite unimportant. In other words, it does not matter much whether $I$, as defined in Section VI, has a value of 0.5 or 1 or 3 or 5—a 10 percent premium based on the alternative cost of effectuating transfers would likely dominate the cost-benefit calculus in any event.

All of this leads to a robust and (to me at least) appealing simplification—to work with standard (unweighted) cost-benefit criteria in general, but to calculate also a net change (plus or minus) in surplus for the group below the poverty line. The alternative cost premium times this net change would then be treated as the net distributional benefit, or cost, of the program under study.

This simplification has the advantage of reducing the estimation of distributional effects to something much less than (and therefore much more feasible of estimation than) a full catalog of how every group’s surplus would be affected. It also has the advantage of being quite adaptable to different ideologies or visions of the good society. A conservative with a strong belief in the work ethic might set the poverty line in the United States at $2,000 and the alternative-cost premium at 3 percent. A soft-hearted liberal, on the other hand, might set the poverty line at $10,000 and the alternative-cost premium at 20 percent. Between these two there is ample range to capture the implicit value systems of, say, Goldwater or Reagan on the one hand as well as Humphrey or McGovern on the other. From my own discussions with professional colleagues on these matters, I find that the more seriously they contemplate the full range of issues and problems involved, the more they incline toward this solution as against that of Subsection A.

C. New Light on the Traditional Solution

Having come this far, we may better appreciate some of the pitfalls that await efforts to take systematic account, within the formal procedures of social cost-benefit analysis, of distributional and other nonefficiency considerations. In this concluding subsection, I shall only sketch a line of thought that can quite readily lead to a reaffirmation of the traditional efficiency-oriented calculus.
We begin with a hypothetical experiment, or, alternatively, an appeal to our own experience and observation, concerning transfers within family groups—or, even better, within groups of grown siblings. What sort of circumstances elicit major transfers? How are these transfers effectuated? Could they be rationalized or justified in terms of a set of distributional weights applied by and within the family?

Taking the last question first, let me emphatically deny that the behavior of siblings can be plausibly rationalized or explained in terms of distributional weighting schemes of the kind we have been analyzing. Such schemes would in principle call for all windfalls to be shared among siblings—rises in pay, stock market successes, lottery winnings, etc. Yet we do not observe systematic transfers taking place in such cases.

The transfers that we do observe among siblings have much more to do with major negative windfalls than with positive ones. The family coalesces to help when a brother loses his job or a sister has a serious accident or a house burns down, etc.—but not when an expected raise fails to come through or when a sibling’s automobile is smashed up.

Looking deeper into transfers between siblings, they seem to be oriented more toward financing specific expenditures or “needs” rather than toward providing a generalized sharing of a common budgetary fund. Basic food, clothing, and shelter in a time of unemployment or distress, emergency medical care at almost any time, education of the family’s children up to some accepted level—these appear to be the needs that call forth intrafamily transfers with the greatest reliability.

Such transfers, then, are typically not generalized and systematic “sharing of the wealth” transfers in cash, such as would be implied by commonly suggested distributional weighting functions. Instead, they tend to be transfers that are, for all practical purposes, “in kind” and oriented toward meeting a certain minimum (which obviously varies from case to case) standard of food, clothing, shelter, medical care, and education.

Having drawn this conclusion from a hypothetical look at family behavior, I for one was quite surprised at how well it fit the observed pattern of societal transfers. Free public education, minimum medical care for those who cannot afford it, food and minimal shelter for the destitute—all have roots going back long before distributional weights were invented.

In contrast, the economist’s favorite gimmick—the cash transfer—has had rough going in virtually every modern society. Free education, subsidized medical care, food and housing for the poor—all these are standard fixtures in today’s world—developed and developing countries alike. Efforts to convert these subsidies to a cash equivalent independent of the purpose for which it might be spent have everywhere faltered and failed. Might it not be that there is a message here for economists trying to

THE USE OF DISTRIBUTIONAL WEIGHTS

I believe so, and the message really tells us to be more modest and humble, less ready to arrogate to our profession and discipline the solution of all of society’s problems.

There are plenty of grounds for modesty and humility as we approach these basic issues. We may all agree that the poor should be helped, but who are the poor? Is income a test? Most graduate students are poor by that standard. We want to help alleviate the human suffering connected with unemployment without inducing people to go casually and habitually “on the dole”—as many do. We want to see it so that the needy have an adequate standard of medical care, but who is to say that the long waiting lists observed in the public clinics of many countries do not perform a quite useful function in restricting the beneficiaries of the free service to people who are really poor and really ill?

All that need be said here is that distributional weights, as commonly understood in our professional literature, simply do not enter into these dimensions. And the more important or relevant are these dimensions of the problem, the less useful are distributional weights as a way of grappling with it.

In the end, then, we cannot condemn as crass or unfeeling the idea our profession’s possibly moving toward a “consensus” based on the traditional criterion of efficiency. On the contrary, such a result might well reflect a greater and more sensitive understanding of the value systems of our citizens and our societies, as well as a more modest and realistic appreciation of our own professional role.

Bibliography

Limitations of space as well as considerations of the scope and purpose of the present paper have precluded an adequate review of the growing literature on distributional weights. In recognition of this fact, the following partial bibliography is provided.


———. Lectures on Public Economics. Forthcoming.


———. “Optimal Taxation and Public Production: Tax Rules II.” AER 61