Two Approaches to Determine Public Good Provision under Distortionary Taxation

Abstract - The paper argues that the appropriate approach to determine public good provision financed by distortionary taxes should depend on the available tax regime. If a sufficiently rich tax regime exists, one could rely on the Pareto criterion, which would be less information-demanding than a social welfare approach requiring access to social welfare weights assigned to various groups. The discussion is related to a number of representative tax regimes and cost–benefit approaches in the literature. It is also argued that, whatever the available tax regime, cost–benefit analysis runs into problems unless one can assume that taxes are set optimally.

INTRODUCTION

The overall aim of the present paper is to discuss the role of tax regimes and the optimality of tax policy for the assessment of public projects, and to shed light on previous contributions from this perspective. Despite being a longstanding and central issue in public finance, the issue of what is the most appropriate approach to determine the provision level of public goods financed by taxes appears to be an unresolved one. While for many years the famous Samuelson rule (Samuelson, 1954) was the prevailing guide to public good provision, recognition of inevitable tax distortions has added considerable complexity. The standard argument is that taxes levied to finance public goods will inflict a loss of efficiency on society by inducing behavioral responses that are socially inefficient, even if privately rational. This is a social cost that must be added to the cost in terms of resources needed to produce the public good. A popular statement is that the tax payers will incur a loss of private income that exceeds the revenue raised by the tax collector. There is a marginal cost of public funds exceeding unity.1 This story can be extended and modified along several dimensions. Firstly, it has been noted that the provision of public goods may itself induce behavioral responses that may alleviate, but could also aggravate, the tax distortions, as highlighted for instance in Atkinson and Stern (1974).2 Secondly, it has been pointed out that it is inap-

---

1 Particularly forceful arguments for this case were put forward by Feldstein (1997).
2 Ng (2000) argues strongly that positive incentive effects on the benefit side may offset disincentive effects caused by higher taxes.
appropriate to neglect distributional effects as the reason why distortionary taxes are deployed in the first place is that redistribution is pursued subject to asymmetric information or other constraints that rule out person-specific lump-sum taxes.

The various arguments above have been nicely integrated in the cost–benefit approach of Slemrod and Yitzhaki (2001). This is based on four concepts, two on the cost side and two on the benefit side. On the cost side, a marginal efficiency cost of public funds (MECF) is defined as the aggregate loss of private income per unit of public revenue, and a distributional characteristic \( (DC_c) \) is defined as the weighted average of the social evaluation of the marginal utility of income, weighted by the share of each individual in the burden of raising the tax revenue. On the benefit side, a marginal efficiency benefit of the public project (MEBP) is defined as the monetary value to the individuals per additional dollar spent by the government, and a distributional characteristic \( (DC_b) \) is defined as the welfare-weighted average of the individuals’ shares in the total benefit. The cost–benefit rule requires that \( DC_c \cdot MECF \) be equated to \( DC_b \cdot MEBP \).3

While there is a lot to be said for this elegant social welfare approach, its disadvantage is that it requires access to a detailed social welfare assessment. Beyond the knowledge required by the Samuelson rule and information on behavioral responses, there is a need to know the social welfare weights to be assigned to the respective agents. The latter is information that is not readily available to the cost–benefit analysts. In this paper I want to argue, in the spirit of the Samuelson rule, that the Pareto efficiency criterion in many cases should be considered as an alternative to the Slemrod and Yaitzhaki (2001) approach to determine the public good provision. Which approach is the more appropriate one should be determined by the available tax regime. I will distinguish between “rich” tax regimes, which are capable of preserving the utility levels of all groups when there is an increase in the public provision (permitting any change in tax revenue),4 and “restrictive” tax regimes, which are not.5 The key argument is that when a rich tax regime is available, the cost–benefit analyst can appeal to the Pareto–principle, as does the Samuelson rule, rather than a social welfare objective.

As there may be many Pareto efficient allocations, Pareto efficiency conditions are in general inadequate to determine public good provision, unlike the social welfare approach, which ideally embraces all relevant social welfare concerns. However, the perception of the government as a single entity simultaneously in full command of all policy instruments, aware of all relevant constraints, and taking all aspects of social welfare into account is not entirely appealing. In practice it is hard to dismiss the idea that there should be some division of labor between branches of government being respectively in charge of the “allocation function” and “distribution function” of government in the terminology of Musgrave (see Musgrave and Musgrave (1973)). I will pursue this idea further below.

In the following, and in line with most of the previous literature in this field, I will consider projects that are sufficiently small to be analyzed by first order effects. The paper proceeds as follows. The next section will elaborate on the Pareto

---

3 I have deviated from the notation of the authors by adding the subscripts \( c \) and \( b \) to distinguish between the distributional characteristics on the cost and benefit side, respectively.

4 Alternatively, we may say that the tax system is capable of redistributing income among all groups.

5 As will become clear, a specific tax regime cannot be classified as “rich” or “restrictive” in an absolute sense. The central issue is the scope for policy and, for a given tax regime, that will depend also on the characteristics of the population and the behavior of the agents.
Two Approaches to Determine Public Good Provision under Distortionary Taxation

approach. The subsequent section will review the Slemrod and Yitzhaki (2001) approach to public good provision under the simple regime of a linear income tax. I then address the Pareto approach within the linear income tax model, and contrast the two approaches before extending the discussion to a non-linear income tax. A separate section is devoted to the role of optimum taxation, which has been an obscure issue in the literature on the assessment of public projects. Finally, I establish links to related literature, and offer some concluding remarks.

THE SCOPE FOR PARETO IMPROVEMENTS

Suppose that the tax system is rich enough to allow the policy maker to keep everybody at an unchanged utility level when there is a change in the public good provision. Assuming that everybody derives a non-negative utility from the additional provision, an offsetting extra tax burden would be imposed on every agent. More tax revenue would be collected and the net revenue would increase or not depending on whether the tax proceeds exceed the additional expenditure on public goods. Attaining a surplus is clearly a sufficient and necessary condition for a Pareto improvement as a surplus can be recycled to some or all agents, while a deficit would make it necessary to impose an extra burden (beyond the utility-preserving extent) on at least some agent(s).

If, under a rich tax regime, no Pareto improvement is possible, a welfare improvement might still be feasible but would require a socially favorable redistribution. As, with Pareto improvements precluded, a project would generate a net revenue deficit at constant utilities, we might alternatively assume that the project in the first place is fully funded by tax changes making nobody better off, and some worse off. Then obviously a socially favorable, pure redistribution would be needed to generate a welfare improvement. Such redistribution would not have to rely on any public project, and should not be allowed to interfere with any project appraisal. If, under a rich tax regime, a project cannot yield a Pareto improvement, it should be rejected.

The most prominent example of a rich tax regime is obviously the unrestricted lump-sum taxes underlying the Samuelson rule. Any change in individual utility due to a public project can be offset by adjusting individualized lump-sum taxes, and a revenue surplus is generated unless the Samuelson rule is already valid. With restrictions on individualized lump-sum taxes, say only a uniform lump-sum tax (poll tax) is available, this will no longer hold true. More realistically, we can imagine that a proportional income tax is the only available tax instrument to finance a public good enjoyed by a heterogeneous population. Then there is a unique relationship between the amount of the public good and the tax rate, possibly after ignoring a downward-sloping portion of a humped-shaped Laffer curve that can be dismissed as Pareto inferior. Such a tax regime will seriously restrict, if not entirely eliminate, the scope for Pareto improvements. If the marginal valuation of the public good is either sufficiently steeply or modestly increasing in income, an increase in the tax and public provision level will benefit some individuals at the expense of others. Then a welfare comparison will have to be made to determine whether “large” benefits (losses) for the rich outweigh “small” losses (benefits) for the poor. Even if the Pareto criterion might still rule out certain allocations, say because a tax increase from a very low level might make everybody better off, it provides a very limited and not very helpful guide to public good provision under such restrictive regimes.

The advantage of being able to trace Pareto improvements is that there is no
need to determine social welfare weights in order to accept a project. We can conceive of the allocation branch as an agency in charge of eliminating deviations from Pareto efficiency by carrying out Pareto improvements whenever feasible. It may solely decide whether the public project should be carried out, leaving to the distribution branch to determine the tax changes to fund it. Or, in other words, if we think of the allocation branch as adjusting taxes to maintain the utility of all agents, it will be the task of the distribution branch to recycle any revenue surplus.

A crucial question is, of course, how realistic it is to assume that a sufficiently rich tax regime is available for applying the Pareto principle. It can be noted that some of the tax regimes, considered in the literature on public good provision, are, indeed, rich tax regimes. Before further addressing this question, I will contrast the Pareto approach and the social welfare approach of Slemrod and Yitzhaki (2001) within a simple tax setting.

LINEAR INCOME TAX AND THE SOCIAL WELFARE APPROACH

I will start by considering the frequently analyzed tax regime in which only a linear income tax is available, following Sandmo (1998) and Slemrod and Yitzhaki (2001). This is done partly for illustrative purposes, but a conceivable justification for this simple tax regime might be that other tax regimes are deemed too complicated and too costly in terms of administrative and compliance costs. It is assumed that each agent $i$ has an exogenous wage rate, $w^i$, an exogenous income $a$, which we write as $a = a_o - T$, where $a_o$ is a parameter, possibly zero, and $T$ is a uniform lump-sum tax. The agent is free to choose his labor supply, and his labor income is taxed at a rate $t$. Let $h^i$ denote labor supply, and let $Y^i = wh^i$ be the gross income. An amount of a public good is denoted by $g$. The indirect utility function can then be written as $V(a, (1 - t)w^i, g)$, and the labor supply becomes a function of the same arguments: $h(a, (1 - t)w^i, g)$. Let us assume that the population consists of only two types of agents, and, as a pure simplification, we assume there is an equal number of each type, normalised to unity. Let the social welfare function be

$$W = V(a, (1 - t)w^1, g) + V(a, (1 - t)w^2, g),$$

where the cardinalisation of $V$ is assumed to reflect the distributional preferences of the government. When convenient, we let $V^i$ denote the utility of agent $i$. The public good is assumed to be acquired at a constant marginal cost $k$. We define net tax revenue, $R$, as the revenue collected over and above the expenditure on the public good.

$$R = tw^1h^1 + tw^2h^2 + 2T - kg = tY^1 + tY^2 + 2T - kg.$$  

We find from the welfare function that

$$\frac{\partial W}{\partial g} = \lambda^1 \frac{V^1}{\lambda^1} + \frac{\lambda^2}{\lambda^2} \frac{V^2}{\lambda^2} = \lambda^1 m^1 + \lambda^2 m^2$$

where $\lambda^i = \frac{\partial V^i}{\partial a}$, which for the appropriate cardinalisation of $V$ is interpreted as

---

6 We should note that the latter part of the statement is essential. If we just require the allocation branch to eliminate deviations from Pareto efficiency, the move to a Pareto efficient allocation needs not itself be a Pareto improvement and it is an open question whether the transition is socially welfare enhancing.

7 A further argument might be that the optimal design of a non-linear tax does indeed approach the linear class, but then we can (at least as an approximation) deal with the linear tax as a special case of the general (non-linear) income tax to be addressed below.
the social marginal utility of income, and \( V_i' / \lambda = m_i \) is agent \( i \)'s marginal valuation of the public good. Adopting the results of Slemrod and Yitzhaki (2001), the associated distributional characteristic is then

\[
[4] \quad DC_k = \frac{\lambda_i m_i + \lambda_i^2 m_i^2}{m_i + m_i^2}.
\]

The benefits are more socially beneficial if they accrue to a larger extent to those people to whom we would like to redistribute income, i.e., those with a high marginal social utility of income.\(^8\) The marginal efficiency benefit of the public project is

\[
[5] \quad MEBP = \frac{m_i'^2 + m_i^2}{-\partial R / \partial g},
\]

where

\[
[6] \quad \frac{\partial R}{\partial g} = \left[ tw_i \frac{\partial h_i}{\partial g} + tw_i^2 \frac{\partial h_i^2}{\partial g} \right] - k.
\]

MEBP captures the potential effect that, if the public good stimulates labor supply, will alleviate the pre-existing tax distortion, and the conventional benefit/cost measure should be magnified. Conversely, it will be scaled down if labor supply is discouraged. When \( g \) is financed by adjusting \( t \),

\[
[7] \quad \frac{\partial W}{\partial t} = \lambda_i Y_i + \lambda_i^2 Y_i^2
\]

\[
= \lambda_i Y_i + \lambda_i^2 Y_i^2 \frac{Y_i + Y_i^2}{Y_i + Y_i^2} \quad (Y_i + Y_i^2).
\]

Invoking further results from Slemrod and Yitzhaki (2001), the distributional characteristic is

\[
[8] \quad DC_t = \frac{\lambda_i Y_i + \lambda_i^2 Y_i^2}{Y_i + Y_i^2},
\]

where the subscript refers to the tax instrument. The marginal efficiency cost of public funds is

\[
[9] \quad MECF_t = \frac{Y_i + Y_i^2}{\partial R / \partial t'},
\]

where

\[
[10] \quad \frac{\partial R}{\partial t'} = Y_i + Y_i^2
\]

\[
- \left[ tw_i \frac{\partial h_i}{\partial \omega} \omega_i + tw_i^2 \frac{\partial h_i^2}{\partial \omega} \omega_i^2 \right],
\]

and \( \omega \) denotes the after-tax wage rate \( (1 - t)w \). If the behavioral response to a tax increase is to withdraw labor supply, \( MECF_t \) captures the intuition that there is a marginal cost of public funds exceeding one when a tax increase induces further misallocation of resources. When we rely on the tax parameter \( T \), we find that

\[
[11] \quad -\frac{\partial W}{\partial T} = \lambda_i + \lambda_i^2 = \frac{\lambda_i + \lambda_i^2}{2}.
\]

The distributional characteristic is

\[
[12] \quad DC_T = \frac{\lambda_i + \lambda_i^2}{2}.
\]

The marginal efficiency cost of public funds is

\[
[13] \quad MECF_T = \frac{2}{\partial R / \partial T'},
\]

where

\[
[14] \quad \frac{\partial R}{\partial T'} = 2 + \left[ tw_i \frac{\partial h_i}{\partial T} tw_i + tw_i^2 \frac{\partial h_i^2}{\partial T} tw_i^2 \right].
\]

The social–welfare–based cost–benefit test of a change in \( g \) is then

\[
[15] \quad DC_t MEBP \geq DC_t MECF_t,
\]

when \( t \) is the source of funding. When \( T \) is increased to finance the project, the cost–benefit test is

\[
[16] \quad DC_T MEBP \geq DC_T MECF_T.
\]

\(^8\) The reason why social marginal utilities of income are not equated is that redistribution is costly due to the tax distortions.
Generalization to an arbitrary number of individuals is straightforward. It is easy to realize that we get the same kind of formulae summing over the appropriate number of individuals.

We see that in order to apply the Slemrod–Yitzhaki (2001) criterion, one needs to know the income levels, the labor supply responses to changes in the tax parameters and the public good as well as the welfare weights. We realize that the criterion is quite information demanding. This motivates a discussion of an alternative approach based on the Pareto criterion.

A LINEAR INCOME TAX AND THE PARETO APPROACH

We will now discuss possible Pareto improvements and Pareto efficiency by considering changes in net tax revenue for fixed utility levels for the two types of people. We recall from [2] that

\[ R =tw h^1 + tw h^2 + 2T - kg = tY^1 + tY^2 + 2T - kg. \]

We then consider some fixed utility levels (indicated by a bar) for both types such that \( V(a, (1 - t)w^i, g) = \bar{V}^i \) for \( i = 1, 2 \). We perceive \( t \) and \( T \) as determined by these two constraints and, differentiating w.r.t. \( g \), we find how taxes must respond to increased public good provision in order to restore the initial utility levels as shown in the appendix. We denote the respective derivatives by \( T' \) and \( t' \). Also deriving the effect on the net tax revenue of the simultaneous utility preserving changes in \( g, t \) and \( T \), we get from the appendix

\[ R' = m^1 + m^2 - k \]

\[ -\left[ tw^1 \frac{\partial h^1}{\partial g} w^1 + tw^2 \frac{\partial h^2}{\partial g} w^2 \right] t' \]

\[ + \left[ tw^1 \frac{\partial h^1}{\partial g} + tw^2 \frac{\partial h^2}{\partial g} \right], \]

where \( h^c \) denotes the compensated labor supply (for the fixed utility levels), and

\[ t' = \frac{m^2 - m^1}{Y^2 - Y^1}. \]

The term \( m^1 + m^2 - k \) is the trade-off between the aggregate marginal benefit and the marginal resource cost as captured by the Samuelson rule. The additional terms are due to the second best tax regime and show how the magnitude of the tax distortions are affected by the change in \( g \) (the last term in brackets) and the induced tax changes (the bracketed middle term). \( R' > 0 \) will imply that a Pareto improvement is feasible as the additional tax revenue may be returned to the tax payers to make them better off. The condition for second best Pareto efficiency is \( R' = 0 \).

Contrasting the two approaches, one may ask how the prescription based on [17] differs from that of Slemrod and Yitzhaki (2001) according to [15] and [16]. The answer to that question hinges crucially on the information that is available and the nature of the tax regime. A key point is that, as welfare weights may not be available, or at least may be hard to establish, the Pareto approach, when available, has an informational advantage. Information may be available for applying the Pareto criterion, but not for using the welfare approach. In that case, no further comparison is meaningful.

Suppose on the other hand that full information is available on the effects of the public provision and the corresponding tax changes as well as welfare weights. Then both approaches are applicable. Let us further assume that initial taxes have been set optimally, which implies indifference among the various ways of tax financing the public good at the margin. In that case, the two approaches will yield equivalent results. To see this, first suppose that \( R' > 0 \) and the project is accepted according to [17]. Then tax changes exist, which will fund the project and make everybody better off. As, by the tax opti-
mality, any other tax funding of the project will also yield a welfare improvement when applying the appropriate welfare weights, the welfare criterion will also accept the project. Then assume that \( R' < 0 \) and the Pareto criterion rejects the project. In that case, any feasible tax changes funding the project would make some agent(s) worse off. The only way that the funding the project would make some effect. In that case, any feasible tax changes of individuals 1 and 2 also keeps all other individuals at their initial utility levels. This means that with these special preferences (discussed in more detail below) the Pareto approach can be used even when only a linear income tax can be imposed on a multi–type population.

We should note that \( c \) and \( b \) may be permitted to depend on \( t, T, g \) which are all uniform across individuals. In other respects, stark assumptions are required to obtain a linear relationship of the kind considered above. A class of utility functions that would yield relations with the desired properties would be \( u(B, h, g) = \gamma^{-1} \ln(B - B_\gamma) + \phi(g) + \psi(h) \), where \( B \) is disposable income and \( B_\gamma \) is a parameter. Then \( m = u_y/u_B = \gamma \phi'(g)(B - B_\gamma) \). Inserting \( B = (1 - t)Y - T \), it follows that \( m = \gamma \phi'(g)(1 - t)Y - \gamma \phi'(g)(T + B_\gamma) \), and in terms of the parameter notation above, \( b = \gamma \phi'(1 - t) \) and \( c = -\gamma \phi'(B)(T + B_\gamma) \).

The Pareto approach shares with that of Slemrod and Yitzhaki (2001) the need to estimate behavioral responses. An important information requirement is knowledge of the compensated tax–induced changes in labor supply as opposed to uncompensated effects in the previous case. This hardly makes a big difference when it comes to obtaining empirical results. However, in this respect the two–type case may be a bit deceptive. With several types of tax payers and tax instruments, there is a need to consider the effects of changing many tax parameters in order to restore all utility levels. In the social welfare approach, one may only need to consider a change in one tax instrument in order to finance the public
good if the choice of funding is a matter of indifference. The advantage of the Pareto approach is that cost–benefit analysts do not have to base their assessment on social welfare weights, which are parameters reflecting political preferences and which are not easily available to the analysts.

As we have seen above, if the linear income tax remains the only feasible tax instrument when more groups are considered, the Pareto approach is no longer an available option except in very special circumstances. We then have a clear case of a restrictive tax regime. In practice there will be numerous individuals, and it may be tempting to dismiss the argument for the Pareto approach as with many people a much richer tax system will be required to implement a potential Pareto improvement. However, in practice a policy assessment will have to deal with a fairly limited number of groups.9 (This is an issue I will come back to below.) Extending the tax system from the linear income tax to a mixed tax regime of commodity taxes and possibly a piece–wise linear income tax may result in a sufficiently rich tax system for this purpose.10 In the limit, one may approach a non–linear tax schedule, which is worthwhile addressing as a separate case.

THE NON–LINEAR INCOME TAX

If a non–linear income tax is available in the model considered above, it is always possible to adjust the tax schedule so as to keep everybody as well off as before when there is a change in the public good provision, and once again we can contrast the two cost–benefit approaches. We can illustrate this with the two–type model used by Boadway and Keen (1993) and similar to the one shown above, except that the income tax schedule is different. Before proceeding to the analysis of Pareto efficiency, we should note that, while the two–type model presented here is handy for expositional and pedagogical reasons, the analysis can be generalized to a multiple–agent setting or, indeed, a continuum of individuals, e.g., along the lines of Christiansen (1981).

We assume that there is an unskilled type 1 and a skilled type 2, such that the latter has a higher wage rate: \( w^1 < w^2 \). As above, \( Y^1, B^1, Y^2, B^2 \) denote the gross and net (after–tax or disposable) incomes of the respective persons. The information structure is such that the government only knows the statistical distribution of people, but does not have access to information about individual wage rates. The government is assumed to design a menu of gross and net incomes \((Y, B)\) implicitly defining the tax schedule \((T = Y – B)\) and letting the taxpayers self–select income points by choosing their labor supply and, hence, income (for fixed wage rates). The policy is determined subject to the government budget constraint and the self–selection constraint that no type of person should pick the income point intended for the other type. Normally it is assumed that the binding constraint is that the skilled type should not mimic the unskilled. The utility can now be expressed as a function of disposable income, the amount of the public good, and the gross income, which for a fixed wage rate can be taken as a measure of the labor supply. Thus, we write the utility function as \( V(B, Y, g) \).

The binding self–selection constraint that the skilled person should not be

---

9 It may be argued that more dimensions should be added, for instance by introducing more socio–economic groups, but then there will also be additional instruments, such as child allowances, and other group–targeted benefits and tax provisions. Differences due to heterogeneity of preferences for people on the same income may be assumed to even out (see Ng (2000, p. 258)).

10 Even if a number of taxes are levied primarily in order to internalise external effects, they are still available for fiscal and redistributive purposes.
better off mimicking the unskilled type than choosing "his own" consumption bundle is

\[\tilde{V}^2 = V^2(B^1, Y^1, g) = V^2(B^2, Y^2, g),\]

with the notation "\(\sim\)" referring to the mimicker. The net tax revenue is

\[R = Y^1 - B^1 + Y^2 - B^2 - kg = T^1 + T^2 - kg.\]

The tax regime consists of a tax on type 1, \(T^1\), and a tax on type 2, \(T^2\), and each tax will have distributional and efficiency characteristics, which can be more easily presented after introducing the expenditure function \(e(Y, V)\) that expresses the disposable income that is required to obtain a utility level \(V\) when the gross income is \(Y\). We let subscripts express partial derivatives and indicate the various types of individuals by the same notation as above. As demonstrated in Christiansen (1999), an incremental increase in \(T^1\) will inflict a burden equal to \(1 - (1 - e^1) / (1 - e^2)\) on type 1, but no burden on type 2 (see op. cit. eqs. (11) and (12)).\(^{11}\) We note that \(e^1 = -(V^1 / V^2)\). Omitting superscripts and denoting the marginal tax by \(T\), the standard characterisation of optimal labor supply from the consumer’s private point of view is \(e^1 = -(V^1 / V^2) = 1 - T^1\), and we can take \(1 - e^1\) as a measure of the marginal tax. As shown in Christiansen (1999), a marginal increase in \(T^2\) will inflict a burden equal to \((\lambda^2 / \lambda^2)(1 - e^1) / (1 - e^2)\) on type 1, and a burden equal to 1 on type 2. We can then formulate the following expressions for the various characteristics as defined by Slemrod and Yitzhaki, (2001):

\[MECF_i = 1 - \frac{1 - e^1}{1 - e^2},\]

\[DC_i = \frac{\lambda^1(1 - e^1)}{1 - e^1},\]

\[MECF_2 = \frac{\lambda^2 1 - e^1}{\lambda^2 1 - e^2} + 1,\]

\[DC_2 = \frac{\lambda^1 \lambda^2 (1 - e^1) + \lambda^2}{\lambda^2 1 - e^2 + 1},\]

where subscripts 1 and 2 refer to the respective taxes, \(T^1\) and \(T^2\),

\[MEBP = \frac{m_1 + m_2}{k},\]

and

\[DC_b = \frac{\lambda^1 m_1 + \lambda^2 m_2}{m_1 + m_2}.\]

The key to understanding these results is the recognition that tax changes are constrained by the self-selection constraint. In order to restrain the high–skilled agent from mimicking the low–skilled type, the income bundle of the latter must be distorted in a way that makes it less attractive to the potential mimicker. Hence, the low–skilled agent is facing a tax wedge \(1 - e_i^1\) at the margin.

From the standard single crossing property, \(e^2 < e^1\). It follows that \(MECF^1 < 1\). Imposing a higher tax on type 1 \((dT^1 = 1\), reflected by the former term on the r.h.s. of [21]), the self–selection constraint is being softened as mimicking would now subject the mimicker to a heavier tax burden. The relaxation of the mimicking constraint enables a change in labor supply and material consumption that alleviates the tax distortion and partially relieves

\(^{11}\) The notation deviates slightly from that of Christiansen (1999).
type 1 of the additional tax burden. The relief is captured by the latter term on the r.h.s. of [21]. As is intuitive, the gains from allowing an increase in labor supply are larger for a larger tax wedge \((1 - e_Y^1)\), as confirmed by [21]. However, the scope for allowing type 1 a beneficial joint increase in labor and material consumption is limited by the inclination of type 2 to follow suit. Such mimicking is less attractive to type 2 the larger is the compensation in terms of material consumption, \(\tilde{e}_2\), that he would need in order to want to work more. Accordingly, as seen from [21], a higher value of \(\tilde{e}_2\) would be conducive to relieving the real burden on the low–skilled type.

Increasing the tax liability of type 2, mimicking becomes a more favorable option, and a further distortion must be imposed on type 1 in order to erode the otherwise larger gain from mimicking. This is consistent with the observation from [23] that \(MECF_2\) exceeds unity. Type 1 will be induced to work and consume less. As labor is already under–supplied, the further contraction of labor supply is more harmful the larger is the initial tax wedge, \(1 - e_Y^2\). Type 2 is more inclined to mimic by also choosing a lower labor supply and material consumption the larger is his willingness to exchange material consumption for leisure (\(\tilde{e}_1\)) and the smaller is the decrease in the mimicker’s marginal willingness to forego material consumption when type 2 is demoted to a less favorable indifference curve (as reflected by a larger \(\lambda_2^{2}/\lambda_2^{1}\)). All these effects tend to aggravate the inefficiency as shown in [23].

The optimal tax condition is \(DC_1 \cdot MECF_1 = DC_2 \cdot MECF_2\), implying that at the tax optimum the government is indifferent between collecting a marginal unit of revenue by increasing the tax liability of person 1 or person 2. In detail, we can write the condition as

\[
\lambda_1^1 - \lambda_1^1 \frac{1 - e_Y^1}{1 - \tilde{e}_Y^1} = \lambda_2^2 + \lambda_1^1 \left(1 + \frac{\lambda_2^2}{\lambda_1^1}\right) \frac{1 - e_Y^1}{1 - \tilde{e}_Y^2}.
\]

An inequality averse government would find the loss of one unit of income for the low–skilled type to be socially more costly that the loss of one unit for the high–skilled type (\(\lambda_1^1 > \lambda_2^2\)). The reason why the government will still be indifferent between increasing the two tax liabilities is that the associated distortions will differ in the opposite direction. Increasing the tax liability of type 1 will alleviate the distortion, as there is less inducement to mimic and the effect is to lower the burden on the left–hand side, while increasing the tax liability of type 2 will reinforce the inducement to mimic and add to the burden on the right–hand side. The additional effects on either side will offset the discrepancy between \(\lambda_1^1 \) and \(\lambda_2^2\), and equate the overall social cost of using either tax instrument.

We note that, in order to apply the social welfare approach of Slemrod and Yitzhaki (2001), there is a need to know the marginal benefit and cost parameters \((m_1, m_2, k)\) as well as the marginal trade–offs between net and gross income (the labor–leisure trade–off) of the various people, including the skilled person when mimicking the unskilled, and the social marginal utilities of income.

A Pareto–efficient tax policy is a choice of gross and net incomes \((Y_1^1, B_1^1, Y_2^1, B_2^1)\) that maximizes the tax revenue for fixed utilities and subject to the self–selection constraint. For a change in \(g\), one can adjust the choice to keep utility levels \(V_1^1\), \(V_2^1\), and \(V_2^2\) unchanged. A Pareto improvement is achieved if there is a resulting increase in net tax revenue, which in

\[12 \text{ We note that } \frac{\lambda_2^2}{\lambda_2^1} = \frac{\tilde{e}_2}{\tilde{e}_1} \text{ is an ordinal measure that reflects how the shape of the indifference curve in } Y, B–\text{space changes when moving vertically to a less favorable one, as will happen when type 2 is faced with a larger tax liability.}\]
Two Approaches to Determine Public Good Provision under Distortionary Taxation

Turn makes it possible to cut taxes. Assigning fixed utility levels (indicated by a bar) to the two types of people, enforcing the self-selection constraint, and assuming the tax policy to be Pareto efficient, we have

\[ V'(Y^1 - T^1, Y^1, g) = \bar{V}', \]

\[ V^2(Y^2 - T^2, Y^2, g) = \bar{V}^2, \]

\[ \bar{V}^2(Y^1 - T^1, Y^1, g) = \bar{V}^2, \]

which is a well-known condition for Pareto efficient taxes, usually referred to as zero marginal tax at the top (see, e.g., Stiglitz (1982)). From these conditions, we can derive the effects of a change in \( g \) on \( T^1 \) and \( T^2 \) and, hence, on \( R = T^1 + T^2 - kg \). We find that

\[ \frac{dR}{dg} = \frac{dT^1}{dg} + \frac{dT^2}{dg} - k - m^1 + \frac{m^1 - \bar{m}^2}{1 - e^1_Y - 1} \]

The condition for a Pareto improvement is then that the expression in [32] is positive. The terms \( m^1 + m^2 - k \) are the benefit minus cost effects captured by the conventional cost–benefit criterion (the Samuelson rule). The remaining term is the effect on the self-selection constraint as discussed in further detail in Boadway and Keen (1993). Apart from the resource constraints, it is the self-selection constraint that impedes further Pareto improvements. If \( m^1 > \bar{m}^2 \), the mimicker will benefit less than the unskilled person from an additional unit of \( g \). A tax increase that will keep the unskilled type equally well off will then deprive the mimicker of more than the benefit he derives from the extra unit of \( g \). He is made worse off and mimicking becomes less attractive. By relaxing the self-selection constraint, a Pareto improvement is made achievable.

If there is no effect on the self-selection constraint, we get back to the Samuelson rule. Christiansen (1981) and Boadway and Keen (1993) discussed the conditions under which this will happen. The crucial condition is that labor is weakly separable from private and public goods in the utility function. One may debate how strict this condition is, but in any case it is interesting that the Samuelson rule may be valid even under second best conditions. Kaplow (1996) discussed the same model, brushing aside the deviations from the Samuelson criterion as qualifications that “entail subtle adjustments.” Whether this is a matter of subtleties is, of course, an empirical question. It can be noted that a similar question arises in the discussion of whether commodity taxes should supplement a non-linear income tax where the issue is whether labor is weakly separable from private goods. Browning and Meghir (1991) addressed this issue and found that separability is rejected. In any case, focusing on the case where the last term of [32] vanishes, the emphasis of Kaplow (1996, p. 522) is on the conclusion that the result holds “without regard to whether the initial income tax is optimal.” I will make some further remarks on this generalization below.

The non-linear income tax is clearly a rich tax regime that allows us to take the Pareto approach.13 In these circumstances, it seems awkward to apply the welfare-based criterion along the lines described by Slemrod and Yitzhaki (2001, p. 195): “The introduction of a general non-linear income tax can be interpreted in the case of

---

13 An even richer tax regime obtains if commodity taxes are available in addition to a non-linear income tax, as discussed by Edwards, Keen and Tuomala (1994).
a continuous distribution of endowments as introducing an infinite number of tax parameters, or, in the case of a discrete distribution of endowments, as introducing a number of parameters at least as large as the number of individuals, with each marginal tax rate being a separate tax parameter. The DC of each marginal tax rate is calculated as if it was a lump–sum tax on all incomes higher than the income on which the marginal tax rate is effective. The MECF is determined by the change in tax revenue of all tax payers with income equal to or higher than the one on which the marginal tax rate is imposed."

This approach is awkward to apply in the sense that, for the tax funding that one would like to consider, one would have to take into account and assign welfare weights to income changes for all the agents affected (all incomes higher than the ones for which the marginal tax rates are changed.) As a rich tax regime is available in the form of a non–linear income tax, it appears simpler to avoid the welfare weighting of numerous income changes and exploit the information advantage of the Pareto approach. (For an approach with a continuum of agents, see Christiansen (1981).) This is not to argue that the latter approach is unproblematic or even trivial to apply. In either case, the analyst will need to have empirical knowledge about labor supply responses to changes in public provision and taxes. As is known from Christiansen (1981), the need to know empirical responses is greatly reduced under the assumption that labor is weakly separable from material goods in the utility function. While this may be good news, the bad news is that the existence of separability is in and of itself an empirical question.

The greater is the heterogeneity of the population, the more difficult it will be to keep everybody at an unchanged utility level. In principle there is almost no end to the variety of preferences and socio–economic characteristics that individuals may be endowed with. In the limit, each individual might be considered a unique agent. Available instruments would obviously be inadequate for individual treatment. Accepting these premises, one is easily led to dismiss the Pareto criterion as impracticable. However, the crucial question is what the alternative would be. It is nearly impossible to think of any approach that would be able to allow for all kinds of heterogeneity and individual characteristics.14 If it were the case that a good procedure existed for weighing the interests of all types of individuals, it would not have been of much interest to explore the Pareto approach in the first place. In practice, one will have to limit the number of groups to be considered. This means that, in practice, there is no absolute sense in which the tax system is rich enough to apply the Pareto criterion. The crucial question in practice is whether the number of population groups that is deemed “sufficient” for assessing the project does not exceed the number of groups that can be compensated by appropriate tax adjustments.

Above, we have contrasted the social welfare approach and the Pareto approach in the special discrete case of two types of individuals. By inspecting [32] and [21] – [26] we note that all the information required by the Pareto approach is also needed by the social welfare approach, while the latter also assumes that the cost–benefit analyst has access to information on social marginal utility of income.

---

14 A striking example of the problems posed by heterogeneity is the lack of a theoretical basis for trading off welfare effects on individuals with heterogeneous preferences. A famous statement is the question by Nozick (1974) as to why somebody who prefers looking at the sunset should pay less tax than somebody who has to earn money in order to attain his pleasures. The sparse literature on how to deal with heterogeneous preferences in a tax context is inconclusive already when it comes to whether individuals should be held accountable for their own preferences. For a discussion of the latter issue, see, e.g., Fleurbaey (1995).
The analysis has addressed the provision of a single publicly provided good in various tax regimes. If there are several goods, each satisfying the assumptions I have made, the analysis can be readily applied to each of them. In the case of mutually exclusive projects, a bit of further reflection is required. If only a restrictive tax regime is available, welfare weights are required even in the single good case, and if such weights are, indeed, available, a welfare comparison of several projects can be undertaken by standard welfare analysis. If a rich tax regime is available, we can conceive of compensating tax changes (preserving all utility levels) for each project. As in the analysis above, we can then compute the change in net tax revenue (after offsetting the effect of larger public provision) in each case. Assuming tax optimality, the project for which we obtain the larger increase in net tax revenue is the preferred project. This is confirmed by a simple formal analysis. Let the indirect welfare function be given by $W(t, g_1, g_2)$, where $t$ is the vector of tax rates with typical element $t_i$ and $g_1$ and $g_2$ are the respective amounts of two publicly provided goods. Let the tax revenue be given as a function $R(t, g_1, g_2)$ and normalize $g_1$ and $g_2$ such that the cost of acquiring a unit is one. Tax optimality implies that $W_i = -\lambda R_i$ for all $i$ where subscript $i$ indicates a partial derivative w.r.t. $t_i$ and the positive parameter $\lambda$ is independent of $i$. Consider an incremental change in $g_j$ and denote by $dt_j$ the accompanying change in tax rate $i$ undertaken to finance the project, and let $dW_i$ denote the resulting change in welfare, so that $dW_i = \Sigma W_i dt_j + W_i dg_j$. Then define $\Delta t_j$ as changes in tax rates that restore all utility levels when $g_j$ is increased and accordingly on balance keeps the welfare level unchanged: $\Sigma W_i \Delta t_j + W_i dg_j = 0$. We can then write $dW_i = \Sigma W_i dt_j + \Sigma W_i dg_j - \Sigma W_i dt_j - W_i dg_j = \lambda \Sigma R_i \Delta t_j - \lambda \Sigma \Delta R_i dt_j = \lambda dR_j$, where $dR_j$ is exactly the change in net tax revenue due to tax changes offsetting the effects of larger public provision. As we can see, this is a measure of the welfare effect of the project.

THE ROLE OF OPTIMUM TAXATION

There has been a discussion in the literature about what difference it makes whether the initial tax policy is optimal or not when considering a public project. It is important to recognize that even with a well-defined social objective and well-established decision-making institutions in place, imperfections will exist. Mistakes are being made and corrected. Responses to changes in exogenous circumstances and novel information may be delayed or subject to disturbances. Lack of optimality may be perceived as due to time lag, inertia, or (temporary) lack of initiative. This is the sense in which I will think of failure to achieve optimality. Beyond these rather mild instances of sub-optimality, there may be more fundamental and persistent impediments to tax adjustment, due to institutional shortcomings, administrative problems, or political constraints. If such more fundamental reasons exist for why taxes cannot be freely adjusted, they should count as part of the constraints defining the available tax regime, but a further discussion of these complex issues is beyond the scope of this paper.

Christiansen (1981) and Boadway and Keen (1993) discussed the cost-benefit rule for public goods under the assumption that there is initially an optimal non-linear income tax. Optimality of the initial taxes will imply that to the first order effect it makes no difference how the public good is financed as any incremental change in the taxes is equally good. As mentioned above, Kaplow (1996) generalized these results by showing that the cost-benefit rule of the cited authors is valid even if the initial tax policy is not optimal. It is important to recognize in what sense this is true. The fact is that unless the optimal-
ity condition of Christiansen (1981), and Boadway and Keen (1993) holds, a Pareto sanctioned change of allocation is feasible and, hence, everybody can be made better off. As adherence to the Pareto principle implies that such a Pareto improvement is desirable whatever the initial situation, no initial optimality seems to be required. However, there is potentially a problem with taking this position. Suppose that the (increase in the) public good is provided on the grounds of the Pareto principle in a situation where the tax policy fails to be socially optimal. Then assume that the non-optimality of the tax policy is being redressed. In general, we cannot know that the public good provision is then optimal.\footnote{For instance, we cannot rule out that the effect of a tax reform is to drive the aggregate marginal willingness to pay below the marginal cost under circumstances where the Samuelson rule is valid, which is the case Kaplow (1996) focuses on.} It might have been the case that the provision was driven by a high willingness to pay for the public good among people whose income and willingness to pay are being substantially reduced by the tax reform. Maybe one would rather have shrunk the public good provision as part of the total reform package. Hence, one might regret the more generous provision. It can be noted that this is a problem that is not due to second best constraints, but arises already when applying the Samuelson rule in the first best regime.

We encounter a similar problem for social-welfare-based criteria. Suppose that we are in the two-type world with a linear income tax. One may then finance a public project either by increasing the poll tax component or by increasing the marginal tax rate. If the tax policy is optimized the source of funding makes no difference. If it is not, three outcomes are possible. The project may be socially desirable for both or neither source of funding, or it may be desirable when relying on one tax but not when using the other. In the latter situation, we cannot tell whether the project would, in fact, be worthwhile at the tax optimum. The case for the project may appear stronger if it is found worthwhile for both sources of tax funding. But even then we cannot know for sure when the tax policy is non-optimal. Suppose that the project is considered to be beneficial mainly due to high marginal valuation among the rich. Let us further assume that the tax policy is deemed excessively redistributive, so that the rich are allowed to keep more of their income when the tax policy is redressed. Even if it is plausible that the rich people’s willingness to pay for the public good will then increase, the impact may be small. However, the welfare weight assigned to the well-off will decline, and, if sufficiently strong, this effect may conceivably reverse the outcome of the cost-benefit assessment.

A related problem is that unless the tax policy is optimized, different sources of tax funding will have unequal social costs. Assuming that some tax is strongly under-exploited, its marginal cost may be small, and a project may appear beneficial when relying on this source of funding. The problem is that if the marginal costs of funding were, indeed, equated across taxes, such “cheap” sources of funding might be no longer available, and the cost-benefit assessment might yield a different result.

It appears that whatever criterion the cost-benefit test is based on, we run into problems if the tax policy is not optimized. It is, therefore, important to recognize in what sense results can be generalized. To argue that results obtained under optimum taxation can be generalized to cases in which taxes fail to be optimal does not mean that the tax optimality assumption is redundant or makes no difference. It makes a crucial difference in the sense that even if a public project is part of a reform package yielding a Pareto or social welfare improve-
ment from a non–optimal situation that is assumed to persist, it might not be one if the tax policy were, indeed, redressed.

A further question is whether deviations from tax optimality also include cases of inefficient, or Pareto inferior, taxes. If so, Pareto inefficiency will imply that there is scope for increasing the tax revenue without making anybody worse off. If such additional tax revenue is taken to be the funding of a public project, it will be accepted even if it generates only a small benefit. A public project will always be deemed worthwhile if coupled with a sufficiently beneficial tax reform, which has nothing to do with how good the project is. Whatever the deficiency of the tax policy, I cannot see why any potential public project should be credited with the gain from redressing an inefficient or non–optimal tax policy.

RELATED LITERATURE

It may be of interest to compare the current approach with the use of compensation criteria that has a long tradition in welfare economics due to Kaldor (1939) and Hicks (1940). To recall the criteria, specify a public provision funded in a certain way. This project will satisfy the Kaldor criterion in its original version if everybody could be made better off by using lump–sum taxes and transfers to redistribute the gains from the project. The project will satisfy the original Hicks criterion if no initial lump–sum redistribution in the absence of the project could make everybody better off than they would be with the project (and no other policy reform). In neither case is it a precondition that the lump–sum redistribution actually takes place. This means that according to the Kaldor criterion it is only potentially that everybody ends up as a beneficiary. I will not review the lengthy scrutiny of these criteria by the profession. Today a sufficient reason to dismiss the criteria in their original form is the well–established insight that redistribution in a lump–sum fashion is not feasible. Recasting the criteria in a modern tax/transfer setting, we can define the Kaldor criterion as being satisfied if a feasible redistribution exists in the wake of the project that would make everybody better off than before the project. Similarly we can define the Hicks criterion as being met if, in the absence of the project, no feasible redistribution exists that would make everybody better off than with the project.

Suppose that we specify a public provision funded in a certain way, and let us consider alternative scenarios. If this policy makes everybody better off, it is trivial that the Kaldor criterion is satisfied. However, that the new allocation Pareto dominates the initial one does not necessarily imply that the specified public provision is desirable unless tax optimality prevails (as discussed above). If the new policy makes only a subset of citizens better off while harming others, the Kaldor criterion may or may not be satisfied. Suppose that it is. There is hardly any compelling reason why the analyst should then recommend the specified policy solely on the grounds of the Kaldor criterion. However, it is a crucial implication that, by appropriately choosing a different tax policy than the one originally specified, a Pareto improvement is possible, and, assuming tax optimality, the same case for the public provision can be established as in my analysis above. Suppose the Kaldor criterion is not satisfied. Then no Pareto sanctioned improvement is possible and distributional weights are needed to assess the policy.

Failure to satisfy the Hicks criterion implies that a pure tax reform could make everybody better off than they would be at the allocation generated by the specified policy involving a new provision level, and the latter policy is inefficient. As under tax optimality, no other funding scheme would be less costly than the specified one; no other policy, involving
a new provision level, would be welfare improving.

Coate (2000) argues the case for an efficiency approach, which is explained in the following way. Denote the status quo policy by \( p_0 \) and let \( V_i(.) \) be the utility function of citizen \( i \). A feasible policy change \( \Delta p' \) is then defined as **efficient** if there exists no feasible policy change \( \Delta p \) such that \( V_i(p_0 + \Delta p) \geq V_i(p_0 + \Delta p') \) for every citizen with at least one strict inequality. This is equivalent to the non–existence of any policy that Pareto dominates \( p_0 + \Delta p' \), which by definition implies that the latter policy is Pareto efficient. Coate argues that non–efficient policy changes should be dismissed while the case for an efficient policy change depends on a distributional judgement. This is tantamount to arguing that one should choose a Pareto–efficient policy, whereas which one to choose is a distributional issue.

Coate’s (2000) paper and the present one are in some respects similar in spirit, but take somewhat different routes. Coate advocates an approach where comparisons are made, within a richest possible set, between policy changes that have similar distributional effects with the aim of identifying the Pareto superior one(s). (We might think of this as approaching the Pareto frontier in a particular direction.) I argue that if the tax system is rich enough, an assessment can be founded on the Pareto criterion in the following sense. Provided that the tax–transfer policy is optimal, it makes no difference to the first order welfare assessment how taxes are changed and the distribution is actually affected. If the extra public provision and feasible tax changes would enable a Pareto improvement, the public provision is desirable for any feasible funding scheme. While Coate mainly confines his interest to projects with similar distributional effects, the present paper has highlighted the approach of Slemrod and Yitzhaki (2001) as a framework for dealing with distributional effects.

Coate (2000) argues that in principle one would like to consider a multitude of spending and taxation programs, implying few restrictions on the available set of policy instruments, but goes on to realize that in practice the comparison of alternative changes is going to have to be limited (op. cit. p. 452). Coate also discusses whether political feasibility should be treated as a factor restricting the set of available policies. Again alternative assumptions will imply different sets of feasible policies, “which must be determined on a case by case basis” (p. 453). The present paper highlights the significance of the available set of policy options, contrasting cases in which interesting policy comparisons are confined to alternatives that are all Pareto–efficient due to limited scope for redistribution, and cases in which a much richer set of instruments is available.

One motivation for considering alternative, richer or more restrictive feasibility sets is that the literature does indeed postulate a variety of feasible tax systems (linear or non–linear, one or several alternative or parallel tax bases, etc.), often without recognizing what difference their richness makes for public provision rules. From a more practical perspective, political opportunity sets do vary conditional on circumstances outside the model as evidenced by cross–country differences and changes over time. Whereas import and export duties constituted the major sources of government revenue in many countries about a century ago, a wide variety of direct and indirect taxes are now available in developed countries, while duties and commodity taxes dominate in many less–developed counties. A crucial factor is administrative feasibility in a wide sense, which is likely to depend on development level, the nature of legal institutions, the compliance of tax payers, and the available technology for informing tax payers, establishing tax bases, processing tax returns, and actually collecting various types of taxes, and so forth.
Yet another approach, which is similar in spirit to some of the preceding ones, is the welfare dominance criterion of Slemrod and Yitzhaki (1991), which is based on an analogy with stochastic dominance in the finance literature. This method, too, sparked off from the recognition that no fully specified social welfare function may available to the analyst. The idea is to identify tax reforms that are welfare improving for any social welfare function belonging to a class of permissible functions with presumably widely accepted properties. The key property postulated by Slemrod and Yitzhaki (1991) is that the social marginal utility of income is positive but declining. A similar criterion might conceivably be used for other policy reforms.

CONCLUSION

A central message of this paper is that the appropriate approach to assess public good provision should depend on the available tax regime. I have distinguished between a rich tax regime, which would be sufficiently flexible to keep everybody at an unchanged utility level when a public project is carried out, and a restrictive tax regime, which, in general, is incapable of maintaining all utility levels.

To clarify specific results, it is helpful to start by considering the Mirrlees setting where agents differ only with respect to skill level (reflected by the wage rate), which is purely private information. A tax schedule that discriminates among skill classes can only be designed subject to standard self-selection constraints precluding that one type of agent mimics another. If a non-linear income tax is available and always used optimally, the Samuelson rule applies modified only by how the benefits of the public good affect the self-selection constraint. If the agents’ valuation of the public good is independent of their labor supply, the original Samuelson rule applies. If the tax policy is sub-optimal, the decision rule must be further adjusted to allow for the welfare effects of the particular tax changes implemented to finance the project. But in that case, the rule is a judgement about the benefits of additional public good provision and of the change in the tax system, which should be conceptually disentangled. The decision rule may then reject projects that would be approved (or accept projects that would not be approved) if the sub-optimality of the tax system were redressed. Hence, generalizations of cost–benefit rules to non–optimal tax situations should be considered with great caution.

In general, projects that pass the Pareto test should be undertaken and can be identified without knowing the social welfare weights if a rich tax system is available. If no funding scheme allows the project to pass the test, it should be rejected. As above, initial tax optimality is an important qualification. If only a restrictive tax system is available, the Pareto test is an inadequate criterion for rejecting projects, and weighing the interests of various groups is necessary.

The advantage of the Pareto approach is that the cost–benefit analyst can avoid the need for information on social welfare weights that are rarely easily available. In other respects, the information requirements of the Pareto and welfare approaches are fairly similar or, indeed, identical. Information on income distribution, valuation of the public good, and behavioral responses are needed in either case.

With a population consisting of a large number of distinctive groups, the number of instruments needed for a rich tax system is not likely to be available, and the Pareto approach may appear unrealistic. I have argued that in practice the groups to be addressed must be defined by the analyst and the number must be limited to a tractable order. This is true for any cost–benefit approach. The alternative to
the Pareto approach is not to keep track of extensive heterogeneity where the distinctive features and welfare weights of the various groups are easily identified. In any case, the complexity is progressively increasing in the number of groups that must somehow be constrained. Any manageable analysis must rely on a sufficiently pragmatic attitude. From this perspective, it becomes a matter of judgement whether the tax system is deemed sufficiently rich for the Pareto test to be a valid approach.

Acknowledgments

I am grateful to Agnar Sandmo for comments on the first draft of this paper. I am also indebted to Taran Faehn and other participants in the 2005 Norwegian Tax Research Conference (“Skatteforum”). In particular, my thanks are due to the anonymous referees for their careful reading and insightful comments, and to the editor, Therese McGuire, for her advice.

REFERENCES


Two Approaches to Determine Public Good Provision under Distortionary Taxation


APPENDIX

We consider some fixed utility levels (indicated by a bar) for both types:

[A1] $V'(a,(1-t)\bar{w}^1,g) = \bar{V}^1$;

[A2] $V^2(a,(1-t)\bar{w}^2,g) = \bar{V}^2$.

These equations are assumed to define $t$ and $T$ as functions of $g$. Differentiating w.r.t. $g$, denoting the respective derivatives by $T'$ and $t'$, we get

[A3] $-\lambda^1 T' - \lambda^1 h^1 \bar{w}^1 t' + V^1_g = 0,$

and

[A4] $-\lambda^2 T' - \lambda^2 h^2 \bar{w}^2 t' + V^2_g = 0,$

where $-\lambda = \partial V^i / \partial a^i$ and $m^i = V^i / \lambda^i$.

We can solve [A3] and [A4] for the tax changes to obtain

[A5] $t' = \frac{m^2 - m^1}{Y^2 - Y^1},$

and

[A6] $T' = \frac{Y^2 m^1 - Y^1 m^2}{Y^2 - Y^1}$.

Letting $h^c$ denote the compensated labor supply (for the fixed utility levels), we can derive the effect on the net tax revenue as

[A7] $R' = \frac{dR}{dg} = (Y^1 + Y^2) t' + 2 T' - \left[ tw^1 \frac{\partial h^c}{\partial g} \bar{w}^1 + tw^2 \frac{\partial h^c}{\partial g} \bar{w}^2 \right] t' + tw^1 \frac{\partial h^c}{\partial g} + tw^2 \frac{\partial h^c}{\partial g} - k,$

and, making use of [A5] and [A6],

[A8] $R' = m^1 + m^2 - k - \left[ tw^1 \frac{\partial h^c}{\partial g} \bar{w}^1 + tw^2 \frac{\partial h^c}{\partial g} \bar{w}^2 \right] t' + tw^1 \frac{\partial h^c}{\partial g} + tw^2 \frac{\partial h^c}{\partial g}$.