ON UTILITARIANISM AND HORIZONTAL EQUITY AS WELFARE-SYMMETRY IN INCOMES

Jesús Seade

NUMBER 197

WARWICK ECONOMIC RESEARCH PAPERS

DEPARTMENT OF ECONOMICS

UNIVERSITY OF WARWICK
COVENTRY

ON UTILITARIANISM AND HORIZONTAL EQUITY
AS WELFARE-SYMMETRY IN INCOMES

Jesús Seade

NUMBER 197

July 1981

1. Introduction

The principle of horizontal equity, that people with equal (full) incomes should be treated alike by the taxman, is widely regarded in the traditional public-finance literature as one of the central guiding lights of good tax design, "called for by the principle of equal justice under the law" (Musgrave (1976; p.4)). Yet, very little attention has been given to the study of this principle, in particular its necessity or even admissibility within a more general welfare-theoretic framework.

The first question that arises is why horizontal equity? Some might answer that there is something to be said for the distribution of welfare in the absence of tax, whose ranking is to be preserved. This line of thought would go along historic notions of justice and deserts, held by some economists and philosophers (e.g. Nozick) but I hope not by too many; such a stand seems, to me, gratuitous, unjustified. Welfare, or at any rate the bulk of formal welfare economics, has to do with actual allocations, with (end-) results. Perhaps changes in these results should in some cases be made to matter, or comparisons with other groups or countries. But I see no reason why the comparison with the "primitive state" should be given a central role in choosing policy. All potential distributions of income and of tax burdens should be treated symmetrically, just as possibilities in the menu that they all are. Accordingly, if horizontal equity is to arise as desirable, it must be as a result of other underlying criteria which we may agree upon as constituting the social good.

This takes us to the second justification that has been suggested for the principle of horizontal equity, namely as a consequence of the more widely accepted principle of (concave) utilitarianism (Feldstein (1976, p.82)), rather than as a primitive moral value in its own right. This implication is well-known to be incorrect if not properly qualified, and indeed for applications. It implicitly assumes identical tastes (people differ in income alone) and more specifically it requires convexity of the set of possibilities, which does follow from technological convexity only in a world of first-best. The latter point has been recognized and illustrated by Stiglitz (1976), and more recently by Balcer and Sadka (1980), who provide examples of income-tax situations with non-convexities (in tax-parameter space), where all too easily it may be optimal to "convexify" the set of (expected) possibilities through randomization of individual outcomes or other similarly inequitable policies. This rejection of the principle, interesting as it is, does not seem to me very persuasive, however. Firstly, the horizontal-equity discussion does not really refer to the tax-treatment of identical people -a natural constraint on tax schemes would be that such people be treated alike, if only as a condition for good decentralizability- but to that of different people with equal incomes. The real question is whether their different preferences for different goods should in themselves be grounds for differential tax treatment. Secondly, the validity of the horizontal-equity principle as such, or its basic relationship to other welfare-theoretic concepts, say to utilitarianism, should not be made subordinate to the tools available: it should be settled in a first-best context, as a discussion of social views rather than possibilities.

The purpose of this paper is to study this relation, i.e. to search for conditions that make utilitarian optimization yield horizontal equity. Equivalently, the inquiry is for conditions under which (utilitarian) welfare is symmetric in individual incomes, that is, conditions under which income-equality amongst a (sub-) group of consumer is not to be disturbed just in response to their heterogeneity in terms of preferences.[1/2/]

Of course it will come as a surprise to no-one that when differences in tastes are <u>not</u> assumed away, horizontal equity and utilitarianism are generally incompatible. But the exercise is of interest, we hope, if only to confirm this surmise, that is, to illustrate how very special requirements must be met for it to be optimal (to a utilitarian) not to treat (say tax) different consumers differently, whether their ex-ante budget sets are the same or not. Furthermore, it is not a-priori clear whether horizontal equity will only "usually" fail to obtain under utilitarianism, or whether this will in fact be a <u>generic</u> occurence: we shall see that the latter is not the case when "generic" is defined in the space of budgets for given preferences, whereas in the wider space where preferences are allowed to vary alongside incomes, horizontal equity does become the non-generic feature of optima one might have expected from the outset.

On the other hand, looking for practical rather than formal motivations for the exercise, it may be useful to bring out the kinds of conditions under which redistribution among income-equals is to be dispensed with in general, to help our intuition on these questions. Similarly, the no-taxation possibility will normally be a central one in what concerns the optimal tax-treatment across income-equals, possibility around which the desired redistribution will lie. The present form of analysis can then be used to deduce directly the qualitative looks of optimal price or tax schemes for given cases, which is useful, for analytical results are otherwise hard to find by direct consideration of the optimum[3/].

The remainder of the paper is organized as follows. After a brief introduction of the model, section 3 derives conditions, for the general case, under which horizontal equity follows from utilitarianism. These conditions are, however, hard to interpret and not very useful in themselves, but rather an input for the rest of the paper. In sections 4 and 5 we accordingly specialize considerably, in two different ways that yield more meaningful conditions. In the former the conditions obtained refer to the way consumer preferences differ, without imposing any restriction on the nature of these preferences (nor, essentially, on the cardinalization chosen by the government); these conditions do not seem like naturally emerging in applications, but are intuitively telling. On the other hand the results in section 5 refer essentially to the chosen cardinalization: assuming isoelastic utilities across commodities, with consumers attaching different weights to different goods (different time-discounts, say), social weights ϕ are found for individual utilities in a welfare maximand $\Sigma_h \phi_h u_h$, under which horizontal equity prevails. Except in the logarithmic case in which the weights that emerge seem rather natural, the weights found are price-related (they cannot be price-<u>dependent</u> for a reason to be noted later) and so is the eventual appropriateness of horizontal equity. Conversely it is shown that, within the class of additively separable utility functions with consumers differing as indicated above, it is essentially for the logarithmic case alone (in fact the linear expenditure system) that horizontal equity holds and not exceptionally in terms of budgets: price-independent weights can be found which do the trick. Lastly in section 6 these results are used to find the optimal pattern of wealth taxes for a simple model of wealth-taxation, when weights adopted by the government are arbitrary rather than being those that would call for the horizontally-equitable absense of such taxation.

2. The model

We consider an economy whose individual members are, for simplicity, described by a vector $\underline{h}$ (for 'household'), which is meant to capture whichever central differences among consumers a model is to concentrate on —in the present case it is tastes one has in mind for interpretation, but the restriction is formally unnecessary and $\underline{h}$ could just as well include a 'vertical' trait such as ability. We assume $\underline{h}$ to be a continuous variable with convex support S.

Since $\underline{h}$ incorporates all forms in which consumers differ, they must otherwise have identical preferences on the vector of consumption goods $\underline{x}$, parameterized by $\underline{h}$: $u(\underline{x}; \underline{h})$. I take this $u(\, ; \,)$ to be smooth in all its arguments, strictly concave and monotonic in $\underline{x}$, and to directly represent the cardinalizations of $\underline{h}$-utilities chosen by the government (utilitarian). On the other hand I impose the convention that, unlike their preferences, consumers' budget sets are all identical: we want to see whether such equalization of incomes is optimal or not.

Budget sets being the same $\forall \, \underline{h}$, utility maximization by consumers requires that each $\underline{h}$ derives no less utility from his own bundle $\underline{x}(\underline{h})$ than from the bundle chosen by other people:

$$u(\underline{x}(\underline{h})\ ;\ \underline{h}) - u(\underline{x}(\underline{h}')\ ;\ \underline{h}) \geq 0 \tag{1}$$

$(\forall \underline{h}, \underline{h}' \in S)$. This expression, as a function of $\underline{h}'$, must accordingly attain a minimum at $\underline{h}$, which under differentiability of allocations $\underline{x}(\underline{h})$ (easy to establish under linearity of the budget constraint) requires, writing $\underline{x}_{\underline{h}}$ for the gradient matrix $(\partial x_i/\partial h_j)$,

$$u_{\underline{x}} \cdot \underline{x}_{\underline{h}} = 0. \tag{2}$$

Since $\underline{x}_{\underline{h}}$ must lie flat on the frontier of the budget set, (2) is clearly no more than a tangency requirement. This is surely in principle not sufficient for individual maximization but I shall ignore this point: sufficiency is easy again under linearity, as one has in the full optimum.

As far as the production side of the analysis is concerned, all we need to consider directly is the vector of prices $\underline{p}$, somehow normalized, for the vector of goods in the economy. We can treat $\underline{p}$ parametrically since only second-order changes of allocations need to be considered in checking for optimality: that is, small changes over small $\underline{h}$-neighbourhoods.

3. Conditions for full equity in the general case

For horizontal equity/first-best optimality, we require[4/]

$$u_{\underline{x}^T} = \underline{p}^T \qquad \forall\ \underline{h}, \tag{3}$$

while from individual optimization (2) obtains. That is, the question is when does independence of the gradient from $\underline{h}$, in (3), imply independence from $\underline{h}$ of total expenditure in consumption $\underline{p}.\underline{x}$, as required by (2) under the substitution indicated by (3). Differentiating $u_{\underline{x}} = \underline{p}$, [5/]

$$u_{\underline{xx}} \cdot \underline{x}_{\underline{h}} + u^T_{\underline{xh}} = 0. \tag{4}$$

From here, and by strict concavity of $u(.; h)$, we can solve for $\underline{x}_{\underline{h}}$:

$$\underline{x}_{\underline{h}} = - u^{-1}_{\underline{xx}} \cdot u^T_{\underline{xh}} \tag{5}$$

so that the decentralization-requirement (2) becomes

$$u_{\underline{x}} \cdot u^{-1}_{\underline{xx}} \cdot u^T_{\underline{xh}} = 0 . \tag{6}$$

To simplify this expression we can proceed in either of two ways, namely grouping the first and second or the second and third terms together for manipulation. Doing the latter first, consider

$$u^T_{\underline{x}} = u^T_{\underline{x}}(\underline{x}; \underline{h}) \tag{7}$$

as an implicit function for $\underline{x}$ (in terms of the left-side gradient), which exists by concavity of $u(.; \underline{h})$. Differentiating it at constant gradient, yields

$$0 = u_{\underline{xx}} \cdot \frac{\partial \underline{x}}{\partial \underline{h}}\bigg|_{u_{\underline{x}}} + u^T_{\underline{xh}} , \tag{8}$$

where the arguments kept constant are as indicated. Hence (6) becomes

$$\underline{p} \cdot \frac{\partial \underline{x}}{\partial \underline{h}}\Big|_{u_{\underline{x}}} = 0 \,, \tag{9}$$

substituting $\underline{p}$ for $u_{\underline{x}}$.

This can be integrated as a function of $\underline{h}$ alone, at constant gradient which therefore appears in the constant of integration:

$$\underline{p} \cdot \underline{x} = k(u_{\underline{x}}) \,, \tag{10}$$

re-expressing the condition in a natural form: demands by different people at the points where their gradients are the same must have the same total cost.

Alternatively, we first group together the first two terms of (6) and notice that their product is proportional to income derivatives of demands for the different elements of $\underline{x}$, namely:

$$\frac{\partial \underline{x}^T}{\partial b}\Big|_{\underline{p}} = \eta \, u_{\underline{x}} \cdot u_{\underline{xx}}^{-1} \,, \tag{11}$$

where b is income and $\eta \equiv \lambda / u_{\underline{x}} . u_{\underline{xx}}^{-1} \cdot u_{\underline{x}}^{T}$ ($\lambda \equiv$ marginal utility of income), which of course depends on $\underline{x}$ (or more precisely on prices and income) but is equal for all goods.[6/] From this expression and (6), we get

$$u_{\underline{xh}} \cdot \frac{\partial \underline{x}}{\partial b}\Big|_{\text{expansion path for } \underline{h}} = 0. \tag{12}$$

Just as $u_{\underline{x}}$ is orthogonal to the constant-u surface through a point, $u_{\underline{xh}}$ is orthogonal to the constant-$u_{\underline{h}}$ surface, in x-space. Hence (12) says that the expansion path for each $\underline{h}$, at his demands in the undistorted equilibrium, lie on the constant-$u_{\underline{h}}$ surface through that point. That is, with two goods x and y ,

$$\left.\frac{dy}{dx}\right|_{u_{\underline{h}}} = \left.\frac{dy}{dx}\right|_{u_y/u_x} \quad \text{(exp. path)}. \tag{13}$$

The requirement takes a simple form- for the general case, and it seems to be amenable to interpretation, but I have not found it. In order to transform these expressions into directly interpretable assertions on utilities it is necessary to specialize. It may be noticed, however, that the right side of this expression relates to preferences alone while the left side depends critically on the u-cardinalization adopted, so that any small change in the latter will in all probability upset the condition and make some redistribution away from equality desirable. This will hopefully come out more clearly below.

I shall now explore, in the next two sections, the implications of these conditions through two special cases, the first one mainly illustrative and the second more relevant for applications, emphasizing respectively the role of consumers' preferences and of the government's evaluation of these preferences.

4. Cost-neutral differences amongst consumers

Let us consider for convenience, in this and the following sections, the two-commodity case, with goods x and y and prices p and q, and $\underline{h}$ for simplicity a scalar, h. Equation (9) becomes

$$p\ \partial x/\partial h\Big|_{u_x,u_y} + q\ \partial y/\partial h\Big|_{u_x,u_y} = 0. \tag{14}$$

Now suppose that the two terms in this expression are not only equal to the negative of one another, but are functions of h alone. That is, momentarily setting $p = q = 1$ to simplify notation,

$$\partial x/\partial h\Big|_{u_x,u_y} = -\partial y/\partial h\Big|_{u_x,u_y} = \psi(h). \tag{15}$$

In this case we can integrate these two expressions separately, as before at constant gradient:

$$\begin{aligned} x &= \Psi(h) + \alpha(u_x,u_y); \\ y &= -\Psi(h) + \beta(u_x,u_y), \end{aligned} \tag{16}$$

where $\Psi(h) \equiv \int \psi(h)\ dh$. For each fixed h, this expression is invertible for the gradient if the corresponding Jacobian is non-vanishing. But this is again ensured by concavity of $u(\cdot;h)$: writing (16) in vector form

$$\underline{x} = \underline{\Psi}(h) + \underline{\alpha}(u_{\underline{x}})$$

and differentiating with respect to $\underline{x}$, yields

$$I = \underline{\alpha}_{\underline{u}_{\underline{x}}} \cdot u_{\underline{\underline{xx}}} ,$$

so that $\underline{\alpha}_{\underline{u}_{\underline{x}}}$ is precisely $u_{\underline{\underline{xx}}}^{-1}$, and invertible. Hence, from (16),

$$\left.\begin{aligned} u_x &= A(x - \Psi(h) , y + \Psi(h)); \\ u_y &= B(x - \Psi(h), y + \Psi(h)), \end{aligned}\right\} \qquad (17)$$

Lastly, integrability requires these expressions to be derivatives of the same function. Hence,

$$u = u(x - \Psi(h), y + \Psi(h)) \qquad (18)$$

(plus a constant of integration $c(h)$ which does not matter in any way).

Under these preferences, no distortionary taxation should be imposed. Notice that no restriction is placed on the nature of the preference map for a given consumer. The restriction primarily refers to the way these preferences differ across consumers. A particular aspect of the cardinalization of utilities is not restricted either: any transform of the usual type $\tilde{u}(u)$ is allowed (e.g. Cobb-Douglas vis-à-vis addilog utilities), the same for all h, but not one of the type $\tilde{u}(u, h)$, such as will be considered in the following section.

It is not difficult to see why this example behaves they way it does. Indifference maps for different h are, under (18), identical up to a shift of the origin along a -45° line. Hence when consumers face producers' prices, which is again a -45° line (by $p = q$), their different points of tangency correspond to the same value of u, the very same shifted indifference curve. Furthermore their gradients are also the same at these points which is what matters, for distances among indifference curves are not affected by moves of the origin. The full optimum is achieved at equal cost; no need for horizontally inequitable redistribution arises.

Our use of prices $p = q = 1$ was for convenience only, but the form of the no-taxation requirement clearly depends on this: intuitively, the shifts of indifference maps discussed above would more generally have to be in the precise direction of equal costs as determined by the given prices. This is easily checked revising the argument above: $\Psi(h)$ must enter (18) in value terms, i.e. as

$$u = u(x - \Psi/\bar{p},\ y + \Psi/\bar{q}), \tag{18'}$$

where the constants $\bar{p}, \bar{q}$ are the values of the prices at the given equilibrium.

5. Additive Separability: the allocation of h-weights to consumers

There is no special reason to suppose consumers' choices to differ, in any given context, in the form just described. Let us now change our approach and adopt from the outset a structure of preferences that is commonly used, namely additive separability, and moreover assume that consumers only differ in the relative weights they give to the subutility functions to be added. That is, consumers' preferences (whose cardinalization is yet to be finalized) can be captured by

$$u(x, y; h) = U(x) + hV(y) , \tag{19}$$

where h is by definition b/a from the apparently more general form $aU + bV$.

A particular example where horizontal-equity considerations are frequently raised is the taxation of wealth or consumers' savings: whether intertemporal choices should be basis for differential tax payments or not. On the other hand intertemporal utilities are for some reason often taken to be additively separable. It thus seems natural to give the wealth-formation and wealth-taxation interpretation to the above specific structure, bearing in mind that only the 'choice' component of differences in savings is being considered, to the neglect of differences in individual incomes and other important aspects of the problem.

The government may wish to use (19) directly as its cardinalization of preferences and according to it proceed to find the optimal tax or transfer policies. The same private preferences, however, could alternatively be put as

$$\tilde{u}(x,y;h) = U(x)/h + V(y), \tag{19'}$$

which is just the previous index divided by h. It is clear that, with the first index, the optimum effects transfers of purchasing power up the scale of y-consumption, at all levels, thus favouring relatively light consumers of x: the thrifty, as it is given by $U' = \text{const.}$, $hV' = \text{const.}$ In sharp contrast, were the government to use the second index, (19'), the optimal distribution of real income and of benefit would be entirely the opposite, penalizing the thrifty. In both cases redistribution flows monotonically up or down the scale, in opposite directions.

This wide difference in the nature of the optimum for the two indices considered is not surprising, for 'productivity' of social utility (level and margin) increases in h in the first case and decreases in the second. But there is no a-priori reason why either of the two forms should be used: a fundamental ambiguity immediately arises which is not present when the difference among consumers has a physical (or 'real') meaning, as the wage or 'ability' does in models of income taxation. Differences in tastes are trickier. The question we now ask, in the context of the present exercise, is which is the h-factor $\phi(h)$ by which we should have multiplied the first expression introduced above, instead of multiplying it by $1/h$ as we did, in order to render optimal divergences from full equality indentically zero.[7/] That is, we wish to find the allocation of weights that underlies full respect for the 'natural' state of affairs among income-equals.

Let us further assume, for the first exercise that follows, (equal) iso-elastic sub-utility functions for the two periods. So, (19) is replaced by

$$u(x,y;h) = \phi(h)x^{1-\gamma}/(1-\gamma) + h\phi(h)y^{1-\gamma}/(1-\gamma) \; , \tag{20}$$

where $\gamma > 0$ is the inverse of the (constant) elasticity of substitution between x and y .

For this utility function, the two sides of (13) directly become, respectively;

$$-\frac{\phi' \, x^{-\gamma}}{(h\,\phi' + \phi)y^{-\gamma}} = \frac{y}{x} \, . \tag{21}$$

Using (3), namely $u_{\underline{x}} = \underline{p}$, or here

$$\phi \, x^{-\gamma} = p \; , \; h\phi y^{-\gamma} = q \; , \tag{3'}$$

and after some manipulations, (21) becomes

$$-\phi'/\phi = 1/h\left[1 + R^{(1-\gamma)/\gamma} \, h^{-1/\gamma}\right] , \tag{22}$$

where R is the relative price q/p, in the intertemporal interpretation the discount factor in production. Integrating (by change of variable: $h^{1/\gamma} = \zeta$), this finally yields

$$\phi(h) = K/(h^{1/\gamma} + R^{(1-\gamma)/\gamma})^{\gamma} \, . \tag{23}$$

The constant of integration K is totally irrelevant and can be set equal to 1 , but of course not so the constant R . That is, for horizontal equity to prevail, a different welfare function must be used for each set of relative prices.

But at the same time the above weighting function is not allowed: price-dependent transforms of utility are non-Paretian,[8/] hence incompatible with our assumption throughout that we wish to be utilitarian. The reason for this can best be seen looking at indirect utilities for consumers: $v(p)$ and $F(v(p),p)$ will generally not correspond to the same preferences. Intuitively, the same allocation of physical quantities but looked-at twice, each time associated to a different price vector, should be considered equally good by a Paretian, but this will not be so under price-dependent transforms of utilities. By continuity, an allocation which is preferred by all consumers to another allocation can, if each is paired with suitably chosen prices, be deemed to be the socially inferior of the two. To recover our claim to be utilitarians in the paper, we simply must reinterpret R in (23) as being a constant which at the given equilibrium happens to equal the price ratio q/p, but which stays put at its "current" value if and when prices themselves change. I thus dub these price-related as opposed to price-dependent weights.

The above assignment of weights is surely special, as any particular set of weights would be, but it is also peculiar, in that it asks us to evaluate the distribution of consumption taking data from the production side of the economy into account, instead of having the latter determining the constraint only. Or equivalently, a great (non-generic) coincidence between certain social-preference-parameter and production-data is required.

It is immediately apparent, however, that there is a special case where this direct dependence on prices disappears: this is under Cobb-Douglas preferences in (20), $\gamma = 1$, for which (23) (setting $K = 1$) reduces to

$$\phi(h) = 1/(h + 1). \tag{24}$$

Hence the commonly used addilog utility function does yield 'horizontal equity' under utilitarianism,when consumers differ and their utilities are weighed in the form indicated. This is, moreover, the only member of the isoelastic family (20) that does so independently of prices. But we do have an interesting if special case of preferences for which horizontal equity is a generic property of optima in the space of prices.

It is of interest to note that the last result above generalizes considerably: the logarithmic case is (essentially) the only member of the much wider additively separable family (19) (weighed by some $\phi(h)$) that behaves in this form.

To show this, consider

$$u(x, y; h) = \phi(h)\ U(x) + h\phi(h)\ V(y)\ . \tag{25}$$

We require that first-best gradient

$$\phi(h)\ U'(x) = p\ ,\ h\phi(h)\ V'(y) = q \tag{26}$$

imply equal cost $(px + qy)$ for all h, i.e.,

$$\xi(h, p, q) \equiv p\, U'^{-1}\left(\frac{p}{\phi}\right) + q\, V'^{-1}\left(\frac{q}{h\phi}\right) = \text{constant}, \tag{27}$$

but now we insist that this hold <u>for all</u> p, q. That is, in addition to $\xi_h = 0$, we impose $\xi_{hp} = \xi_{hq} = 0$. Hence,

$$\frac{\partial^2}{\partial p \partial h} [px(p, h)] = 0, \tag{28}$$

where $x(p,h)$ is defined by $\phi(h)\, U'(x(p, h)) = p$. This transforms into

$$0 = \frac{\partial}{\partial p} [p\, x_n(p, h)]$$

$$= -\frac{\partial}{\partial p} \left[\frac{p\phi'\, U'}{\phi\, U''}\right] \quad \text{(differentiating (26a) w.r.t. h)}$$

$$= -\frac{\partial}{\partial p} \left[\frac{\phi'(U')^2}{U''}\right] \quad \text{(by 26a)}$$

$$= -\phi' \frac{d}{dx} \left[\frac{(U')^2}{U''}\right] \frac{\partial x(p, h)}{\partial p} \quad \text{(indicating chain rule)}$$

$$= -\frac{\phi'}{\phi U''} \frac{d}{dx} \left[\frac{(U')^2}{U''}\right] \quad \text{(differentiating (26a) w.r.t. p)}.$$

Hence, $U(\cdot)$ must be such that 9/

$$d[(U')^2/U'']/dx = 0, \tag{29}$$

which upon triple integration yields

$$U(x) = a \log(x - \bar{x}) \tag{30}$$

(for some constants a; $\bar{x}$), plus a third, irrelevant constant. A similar analysis for y yields

$$V(y) = b \log(y - \bar{y}). \tag{30'}$$

Hence the most general additively separable utility function of the form (25) for which price-independent weights can be found such that the horizontal status quo is optimal, is the linear expenditure system,

$$u(x, y; h) = \phi(h) \left[a \log(x - \bar{x}) + bh \log(y - \bar{Y})\right] , \tag{31}$$

which is rather restrictive.

The weights that do the trick can be found proceeding as we did for the isoelastic case; these weights are[10/]

$$\phi(h) = 1/(a + bh) . \tag{32}$$

6. Optimal wealth-taxes for 'normalized' weighting functions

We noticed in p.13 above that rather different patterns of optimal redistribution follow from using the index $U + hV$ as opposed to weighting this expression by $1/h$ to use $U/h + V$. We later found the particular weights that would yield a non-redistributive optimum, namely the price-dependent function (23) (for isoelastic underlying preferences). But there is no compelling reason why any of these, or indeed any other particular weighting system should be used: it all depends on the government's views on whose consumption contributes most to social utility at the margin.

Now let us suppose, merely for the sake of argument, that the weights used are the simple ones given by (24) ,

$$\phi(h) = 1/(h + 1) ,$$

in which case some form of redistribution is generally in order. We wish to find the qualitative pattern this redistribution will take when effected, in a second-best manner, through some combination of savings- and poll-taxes and subsidies.

To motivate the exercise we offer two arguments. Firstly, that the qualitative looks of optimal nonlinear schedules usually are very hard to establish other than through numerical computation of solutions for particular examples, so that alternative more direct arguments to deduce such features are worth looking for.

Second, now on the assumption that it is precisely the weights (24) that the government adopts, we notice that this weighting-function has the property that it makes total utilities be weighted averages of spot-utilities from the two periods: the weights given to these add up to unity. This seems sensible in the absence of reasons to the contrary. We may think of this $\phi(h)$ as a normalization factor, which somehow offsets the unpalatable feature of $U + hV$ (or $U/h + V$) of making high (low) h's produce more utility out of any given pair (x, y) . In contrast, under

$$u = U(x)/(1 + h) + hV(y)/(1 + h) , \tag{33}$$

utility from a given pair (x, y) increases with h (increasing preference for the future) if and only if this pair embodies more (discounted) future than present utility: $hV > U$. But of course we do not need to push this too far: we use (24) as an example but the arguments below can be adapted to other cardinalization and weighting systems the government may wish to use.

Let us suppose again that preferences are isoelastic, as in (20), so that no distortion would be imposed <u>if</u> weights (23) were those the government wished to use. But it uses (24) . Denote by $\omega^o(h)$ the sum of utility-weights for his two periods an h-man receives under (23) , namely

$$\omega^o(h) = (1 + h)/(h^{1/\gamma} + R^{(1-\gamma)/\gamma})^\gamma . \tag{34}$$

We now proceed as follows. Start from the no-redistribution, symmetric-welfare allocation optimal under (34), as if these were the 'total weights' we wished to give to people. All gradients of income-equals are in that case identical. Now change these total weights to their true values under the weighting-system actually adopted: under the 'normalization' (24) this is a constant across h (the value 1 in particular, but that is irrelevant). Clearly, before adjusting the budget set, and hence demands accordingly, unweighted marginal utilities $U'(x)$, $V'(y)$ have not changed for any h , and it follows that the gradient rises (in both components) for those whose total weight under (34) was low, for these weights all become the same now. Conversely, those who previously 'required' a relatively high weight to be as deserving as the others, now become less deserving gradient-wise. By continuity, higher gradients remain higher after some (sufficiently small) adjustment towards the optimum has been performed. On the other hand it can be shown that, as one would expect, optimal redistribution through marginal taxes follows the gradients closely: deservingness gradient-wise is deservingness tax-wise, and it follows[11] that the value of consumption is expanded (contracted) for consumers whose gradients are scaled up (down) by the withdrawal of relatively low (high) weighting factors.

Accordingly, what we need to do in order to find out what the distribution of taxes is at the new optimal equilibrium, is to determine the shape of $\omega^{o}(h)$. It is easy to see that

$$\text{sign} \quad (d\omega^{o}/dh) \quad = \quad \text{sign} \quad (R^{\beta} - h^{\beta}) \ , \tag{35}$$

where $\beta \equiv (1 - \gamma)/\gamma$, and from here, that

$$\frac{d\omega^{o}}{dh} \lesseqgtr 0 \quad \text{as} \begin{cases} R \lesseqgtr h & \text{if} \quad \gamma < 1 \\ \\ R \gtreqless h & \text{if} \quad \gamma > 1 \end{cases} \tag{36}$$

Therefore the allocation of total weights is as indicated in figure 1, perhaps with more bends than those shown but always strictly monotonic on either side of R, i.e. with a single stationary point at $h = R$, the value of h of the person whose utility-discount is the same as producers'.

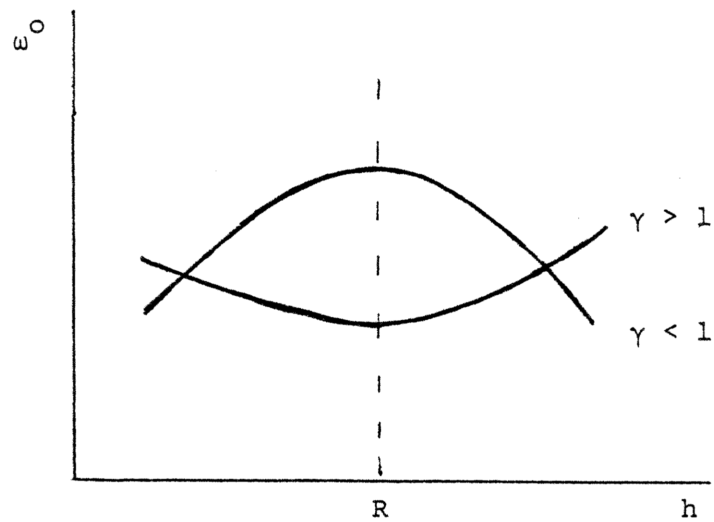

Figure 1: distribution of 'total weights' for horizontally equitable optima.

Hence, transferring resources from high to low values of ω^o , we finally arrive at the budget sets of figure 2, where again the slope of isocost lines is the inverse of the shadow discount factor and x and y present and future consumption (net compounded savings).

Case (ii) in the figure, first, corresponds to $\gamma > 1$: elasticity of substitution less than unity. Borrowing from studies in other areas of consumer choice amongst similarly 'essential' commodities or aggregates (e.g. leisure/consumption), this may be

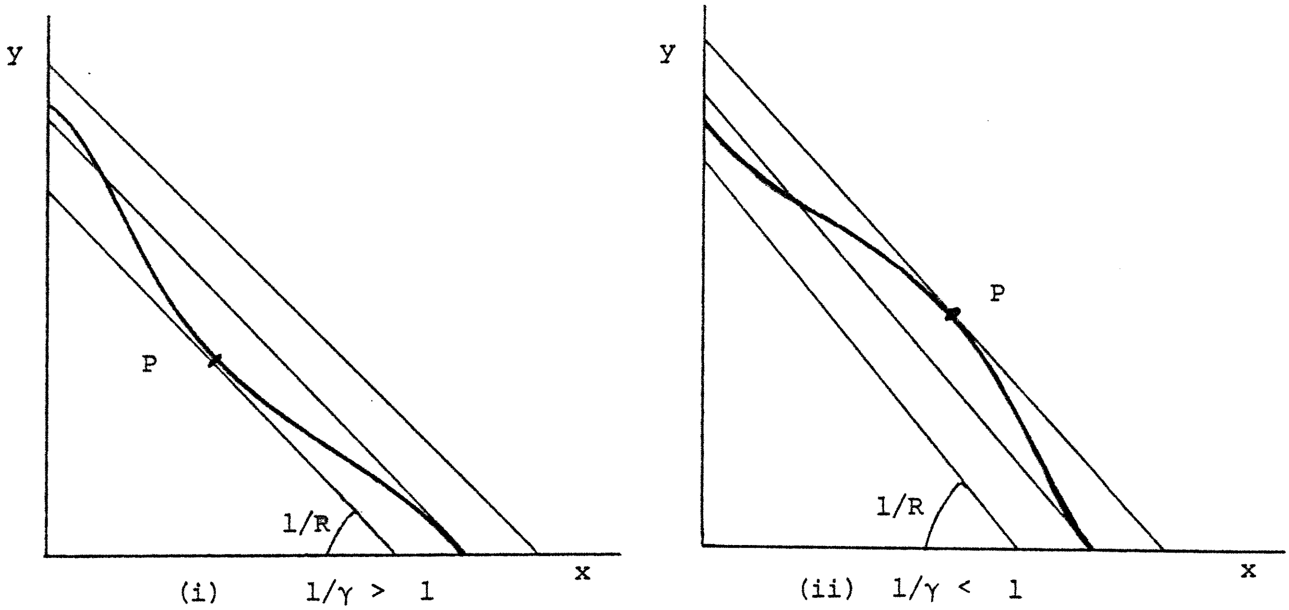

Figure 2: 'typical' wealth taxes under the weight-normalization (24) , for elasticity of substitution (= $1/\gamma$) (i) greater than and (ii) less than one.

deemed as the empirically more interesting of the two cases. The budget set subsidizes savings at the margin from the bottom up to point P , say through contributions to the retirement fund, after which taxes on savings (net of the said contributions) are applied.

The consumer receiving the largest transfer (or paying the least tax) is at P , where all the benefits from lower-ranges subsidies have been received and no (non-poll) tax is yet being paid. This point clearly corresponds, as a first approximation, to $h = R$. The 'balanced' pattern of life-cycle consumption of this person is rewarded and, through that reward, encouraged: more people will consume on the 'hump' around p under this budget set than would under its linearization through that point.[12/]

In contrast, part (i) of the figure penalizes point P most heavily, inducing consumers in some range around $h = R$ to move up or down the scale, where more subsidy is received or less tax is paid, respectively. This case corresponds to 'high' substitutability between x and y . Deservingness does not then have a single maximum, as would seem to be a sound feature to expect from (or impose on) policies, but two: at the extreme values of h in the population. With high substitutability consumers seem to be doing well, socially, by specializing their consumption.

Footnotes

Prepared for the second Latin American Meeting of the Econometric Society (Rio de Janeiro, July 1981). I am grateful to Peter Hammond, Andreu Mas-Colell, Marcus Miller and particularly Jim Mirrlees for helpful comments or suggestions on previous versions of the paper. Errors and limitations are, indeed, my own.

1/ Somewhat related in spirit is Peter Hammond's (1977) penetrating study of the relation between utilitarianism and (vertical) egalitarianism. The same author has also investigated, in unpublished work, the implications of welfare symmetry.

2/ The above is my conception of the horizontal-equity problem, which I feel I share with many though precise statements in this area are not very common. In particular, I am not adopting an alternative common definition of horizontal equity, put in terms of preserving the ranking of utilities rather than that of real expenditure power at producer prices.

3/ An exercise along these lines is presented in sec. 6 below.

4/ Or more precisely $u_{\underline{x}} = \lambda^{\underline{h}}\underline{p}$, but a condition for optimality is $\lambda^{\underline{h}} = \lambda \; \forall \; \underline{h}$, so that (3) obtains by simple choice of numeraire units.

All this, of course, for a utilitarian SWF.

5/ Or, rather, $u^T_{\underline{x}} = \underline{p}^T$, to permit differentiation with respect to $\underline{x}$ (a column), adopting the convention that differentiation of vectors (w.r. to scalars) preseves the arrangement and (of scalars) with respect to vectors transposes it.

6/ This is equation (12) in Brown and Deaton (1972).

7/ We are thus concentrating on a particular dimension of the choice of cardinalization by the government: the additive structure of (19) is preserved, to the exclusion of non-affine transforms of this expression. That is, the otherwise more general transformation $\tilde{u}(u, h)$ is restricted to the form $\phi(h)u$.

8/ I am grateful to Peter Hammond for drawing this point to my attention.

9/ A second solution of the previous equation is $\phi' = 0$, but this is ruled out by the analysis to determine ϕ, below, where it implies the spurious $\phi \equiv 0$.

10/ Without loss of generality we may relabel bh/a as h alone, and write (31) as

$$u = \phi(h) \ \{\log(x - \bar{x}) + h \ (\log(y - \bar{y})\} \tag{31'}$$

with $\phi(h)$ now given by (24): $\phi(h) = 1/(h + 1)$.

11/ For optima close enough to the undistorted state of affairs; further away there may well be tax-reversals in some savings-ranges from what these arguments predict.

12/ We have not said anything about densities, however. Perhaps all h lie entirely to the left of R , say (generalized impatience over producers'). In that case all weight-adjustments would fall on a monotonic branch of the corresponding curve in fig. 1, and the budget set would accordingly look like the portion to the right of P in figure 2-i: savings are encouraged at all levels.

References

Balcer, Y. and E. Sadka (1980), "Horizontal equity in models of self-selection with applications to income taxation and signalling", Foerder Institute for Economic Research (Tel-Aviv University) working paper no. 44-80; November.

Brown, A. and A. Deaton (1972), "Models of consumer behaviour: a survey", Economic Journal 82, 1145-1236.

Feldstein, M. (1976), "On the theory of tax reform", Journal of Public Economics 6, 77-104.

Hammond, P.J. (1977), "Dual interpersonal comparisons of utility and the welfare economics of income distribution", Journal of Public Economics 7, 51-71.

Musgrave, R.A. (1976), "ET, OT and SBT", Journal of Public Economics 6, 3-16.

Stiglitz, J.E. (1976), "Utilitarianism and horizontal equity: the case for random taxation", Stanford University IMSS technical report no. 214.