University of Wisconsin-Madison

Institute for Research on Poverty

Discussion Papers

Yves Balcer
Efraim Sadka

FAMILY SIZE, PERSONAL INCOME TAX CREDITS AND HORIZONTAL EQUITY

DP #588-79

Family Size, Personal Income Tax Credits

and Horizontal Equity

Yves Balcer
The University of Wisconsin

Efraim Sadka
The University of Wisconsin and Tel-Aviv University

October 1979

This paper employs a utilitarian framework in order to evaluate the common practice of allowing income tax credits (or exemptions) for dependent members of the household. Perhaps the strongest case for this practice can be made when it is assumed that demographic variables (such as family size) manifest themselves in the consumption patterns of various households via demographic translating of the utility function. When this assumption is made, we show that income tax credits for dependent children are an optimal policy if and only if horizontal equity is desirable. However, we show that utilitarianism does not imply the principle of "equal treatment of equals." Therefore, we conclude that income tax credits are not justified in a utilitarian framework.

INTRODUCTION

Many economists recommend income tax credits as a proper way to treat differences in the household size by the income tax system. And indeed, most states and countries have adopted income tax credits or some imperfect variant of them (namely, tax exemptions) for their income tax laws. The purpose of this paper is to examine the soundness of such a policy.

Whether a tax credit is a good or bad policy depends on how differences in family size manifest themselves in the comsumption patterns (or, more generally, the preferences) of households of various sizes. Two of the most popular procedures for incorporating demographic variables into demand systems in general and household size in particular are <u>demographic scaling</u> and <u>demographic translating</u> [see Pollak and Wales (1978a) and (1978b) and their references]. Perhaps the most convincing case for tax credits can be made when differences in family size can be handled by translating. We consider only the latter procedure, although we discuss briefly the method of scaling in the concluding section. For our purposes we can describe the procedure of translating as follows.

Suppose for simplicity's sake that there are only two sizes of households: small households of a size which is normalized to zero and large households of a size which is normalized to one.[1] The procedure of translating

---

*The research reported here was supported by National Science Foundation Grant DAR-7917376 and by funds granted to the Institute for Research on Poverty at the University of Wisconsin-Madison by the Department of Health, Education and Welfare pursuant to the provisions of the Economic Opportunity Act of 1964. The opinions expressed are those of the authors. We wish to thank William Brock and Charles Wilson for some useful suggestions.

in effect makes the assumption that the preferences of a household of size one over bundles (y, x), where y is labor and x is an aggregate consumption good,[2] are the same as the preferences of the smaller household over bundles (y, x-b), where b is a constant. In other words, if we denote the utility of a household of size i, by u^i then the utility functions are related to each other by the identity

$$u^1(y, x) = u^o(y, x - b).$$

In this case, the large family needs, ceteris paribus, an (after-tax) income which is higher than the income of the small family by b dollars in order for the two families to attain the same level of satisfaction.[3] This is why we believe that demographic translating might provide the best support for income tax credits for dependent members of the household.

It turns out, however, that even if one accepts the assumption of demographic translating, one will still come short of showing that allowing income tax credits is an optimal accomodation for family size differences. There are two main reasons for this failure to justify tax credits. First, the distribution of wages (or skills or abilities) within each size class may not be the same. This point is quite obvious and we do not discuss it here except to mention it. The second reason is much more subtle. It has to do with the question of whether an optimal income tax system calls for horizontal equity ("equal treatment of equals") or not. Due to its second-best nature, an optimal income tax does not in general support the principle of horizontal equity[4] and, as we shall see later, it does not therefore support the practice of allowing income tax credits for dependent members of the household.

THE MODEL

As mentioned in the introduction, we consider only two normalized household sizes: zero and one. Consumption is denoted by x, and labor by y. The maximum amount that one can work is $\bar{y}$, so that leisure is $\bar{y} - y$. All the households of size i have the same <u>concave</u> utility function over bundles (y, x), which is denoted by $u^i(y, x)$ where i = 0, 1. The two utility functions, u^o and u^1, are related to each other by

$$u^1(y, x) = u^o(y, x - b), \tag{1}$$

where b is a positive constant (a translating variable). These utility functions are strictly decreasing in y and increasing in x. The cumulative distribution function of real wages (or abilities) of households of size i is denoted by F_i, so that, for instance, $F_o(w)$ is the number of households of size zero whose real wages are less than or equal to w. R is a predetermined level of public expenditures which have to be financed by the income tax system. Although a more general form of social welfare functions can be considered, we restrict our attention to the utilitarian objective which is to maximize the sum of the household utilities. An allocation for the class of households of size i is a (two-dimensional vector) function $[y_i(\cdot), x_i(\cdot)]$, where $y_i(w)$ is the amount of labor supplied by household of size i which earns a wage rate of w, $x_i(w)$ being consumption, i = 0, 1. In order to design an optimal income tax system, the government seeks to find two allocations, $[y_o(\cdot), x_o(\cdot)]$ and $[y_1(\cdot), x_1(\cdot)]$, which maximize the utilitarian objective, which is

$$\int u^o[y_o(w), x_o(w)]\, d\, F_o(w) + \int u^1[y_1(w), x_1(w)]\, d\, F_1(w), \tag{2}$$

subject to some constraints.

First, these allocations must satisfy the government's budget constraint. Noticing that the tax paid by a household of size i which earns a wage w is $wy_i(w) - x_i(w)$, we can write this constraint as

$$\int [wy_o(w) - x_o(w)]\, d\, F_o(w) + \int [wy_1(w) - x_1(w)]\, d\, F_1(w) \geqq R. \quad (3)$$

Secondly, each allocation $[y_i(.), x_i(.)]$ must be attainable by an income tax function. If we denote by A_i the set of such allocations, then the government must also satisfy the constraints

$$[y_i(.), x_i(.)] \in A_i, \quad i = 0, 1. \quad (4)$$

By thus formulating the problem of finding the optimal income tax system, we allow the two classes of households to face two different income tax functions or schedules. We then ask whether an optimal income tax system requires these two tax schedules to differ from each other only by a constraint. If so, then allowing the larger household an income tax credit (which is equal to this constant) over what the smaller household has to pay constitutes an optimal policy. In other words, if we denote the solution to the problem of maximizing (2) subject to (3) and (4) by $[y\overset{*}{i}(.) , x_i^*(.)]$, i = 0, 1, then we can ask whether

$$y_o^*(w) = y_i^*(w) \quad \text{and} \quad x_o^*(w) = x_i^*(w) - b \quad \text{for all w.} \quad (5)$$

If (5) is true, then the two classes of households should be subject to the same tax schedule, except that the larger household is also allowed to claim a tax credit which is equal to b.

THE RELATION BETWEEN HORIZONTAL EQUITY AND TAX CREDITS

The problem of finding the optimal tax system [i.e., maximizing (2) subject to (3) and (4)] will be simplified if we introduce the following transformation:

$$\bar{x}_1 = x_1 - b. \tag{6}$$

Accordingly, we define $\bar{A}_1$ as the set of allocations $[y_1(\cdot), \bar{x}_1(\cdot)]$ such that $[y_1(\cdot), \bar{x}_1(\cdot) + b] \varepsilon A_1$. In view of (1), it follows that

$$A_o = \bar{A}_1 \tag{7}$$

and that

$$\int u^1[y_1(w), x_1(w)]\, dF_1(w) = \int u^o[y_1(w), \bar{x}_1(w)]\, dF_1(w). \tag{8}$$

Let us also define

$$\bar{R} = R + \int b\, dF_1(w). \tag{9}$$

Then the problem of finding the optimal tax system now becomes

$$\max\left\{\int u^o[y_o(w), x_o(w)]\, dF_o(w) + \int u^o[y_1(w), \bar{x}_1(w)]\, dF_1(w)\right\} \tag{10}$$

s.t.:

$$\int [wy_o(w) - x_o(w)]\, dF_o(w) + \int [wy_1(w) - \bar{x}_1(w)]\, dF_1(w) \geqq \bar{R}$$

$$[y_o(\cdot), x_o(\cdot)] \ \varepsilon\ A_o$$

$$[y_1(\cdot), \bar{x}_1(\cdot)] \varepsilon\ A_o.$$

Recall that allowing income tax credits for dependent members of the household contributes an optimal policy whenever (5) holds. In view of the transformation (6), condition (5) becomes equivalent to

$$y_o^*(w) = y_1^*(w) \quad \text{and} \quad x_o^*(w) = \bar{x}_1^*(w) \quad \text{for all } w \tag{11}$$

where $[y_o^*(\cdot), x_o^*(\cdot)]$ and $[y_1^*(\cdot), \bar{x}_1^*(\cdot)]$ are an optimal solution of the optimization problem (10).

The question of the desirability of income tax credits can thus be interpreted as follows. Consider a society consisting of two classes which are identical with respect to preferences (represented by u^o) but not necessarily with respect to wage distributions. (F_o is not necessarily equal to F_1.) Solving (10) yields an optimal income tax system which consists of two tax functions, one for each class. If these two tax functions are the same, then (11) holds and, consequently, tax credits are an optimal way to treat differences in family size.

However, when $F_o \neq F_1$, we do not see any reason why these two tax functions have to be equal to each other. A look at tax formulae available in the literature [e.g. equation (2.17) of Mirrlees (1976)] clearly suggests, as is indeed expected, that the optimal tax schedule depends heavily on the distribution of wages. In other words, when $F_o \neq F_1$ then even if our two classes of households had the same average wage [i.e., $\int w dF_o(w) = \int w dF_1(w)$], it would still not imply that they must be faced with the same tax functions. If, in addition to $F_o \neq F_1$, the average wage itself is not the same for the two classes, then the situation is even more unpredictable.

Let us therefore concentrate on the case where $F_o = F_1$ and see whether in this case (11) has to hold. In order to do this, let us solve problem (10) in two stages. In the first stage we find an optimal tax function for raising a certain amount of revenue (denoted by R_o) from one class and an optimal tax function for raising an amount revenue $\bar{R}_1$ from the other class. Formally, for each R_o and $\bar{R}_1$, we solve:

$$\max \int u^o[y_o(w),\ x_o(w)]\ dF_o(w) \tag{12}$$

$$\text{s.t.:} \int [wy_o(w) - x_o(w)]\ dF_o(w) \geqq R_o$$

$$[y_o(\cdot),\ x_o(\cdot)] \in A_o$$

and

$$\max \int u^o[y_1(w),\ \bar{x}_1(w)]\ dF_o(w) \tag{13}$$

$$\text{s.t.:} \int [wy_1(w) - \bar{x}_1(w)]\ dF_o(w) \geqq \bar{R}_1$$

$$[y_1(\cdot),\ \bar{x}_1(\cdot)] \in A_o.$$

Let us denote by $[y_o(\cdot,\ R_o),\ x_o(\cdot,\ R_o)]$, and $[y_1(\cdot,\ \bar{R}_1),\ \bar{x}_1(\cdot,\ \bar{R}_1)]$ the optimal solutions for (12) and (13), respectively, and by $S_o(R_o)$ and $S_1(R_1)$ the optimal values of the objective functions, i.e.,

$$S_o(R_o) = \int u^o[y_o(w,\ R_o), x_o(w,\ R_o)]\ dF_o(w) \tag{14}$$

and

$$S_1(\bar{R}_1) = \int u^o[y_1(w,\ \bar{R}_1),\ \bar{x}_1(w,\ \bar{R}_1)]\ dF_o(w). \tag{15}$$

When $R_o = \bar{R}_1$, then (12) and (13) are identical, and hence, the two functions S_o and S_1 are the same: $S_o = S_1$.

In the second stage we have to find the optimal amounts of revenue that ought to be raised from each class. Formally, we solve

$$\max\left[S_o(R_o) \;+\; S_o(\bar{R}_1)\right] \tag{16}$$

$$\text{s.t.:}\;\; R_o \;+\; \bar{R}_1 \;\geqq\; \bar{R}.$$

If we denote the optimal solution of (16) by R_o^* and $\bar{R}_1^*$, then the optimal solution of (10), which was denoted by $[y_o^*(.),\ x_o^*(.)]$ and $[y_1^*(.),\ \bar{x}_1^*(.)]$, can be found by

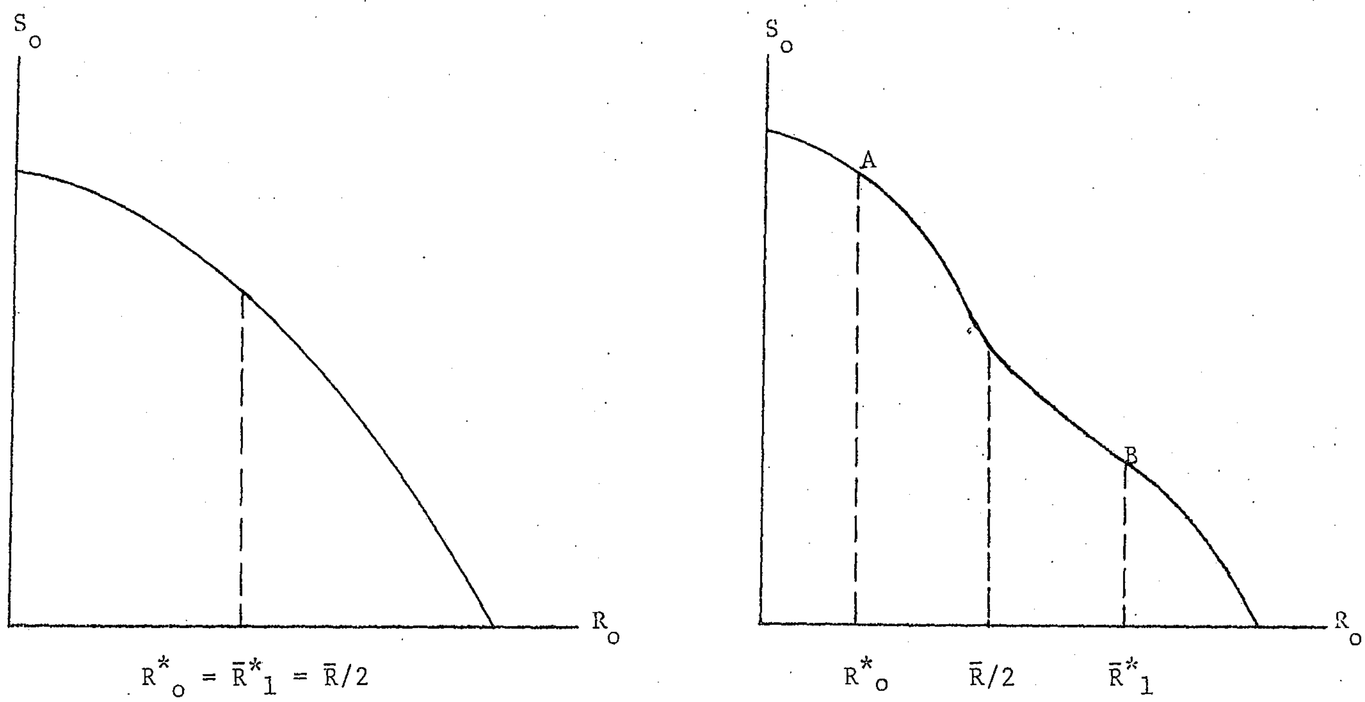
$$y_o^*(w) \;=\; y_o(w,\ R_o^*), \qquad x_o^*(w) \;=\; x_o(w,\ R_o^*)$$

and

$$y_1^*(w) \;=\; y_1(w,\ \bar{R}_1^*), \qquad \bar{x}_1^*(w) \;=\; \bar{x}_1(w,\ \bar{R}_1^*).$$

Since (12) and (13) are identical when $R_o = \bar{R}_1$, it follows that (11) holds if and only if $R_o^* = \bar{R}_1^*$. The latter equality holds if S_o is a concave function. On the other hand, if S_o is not concave, then R_o^* needs not be equal to $\bar{R}_1^*$ (see Figure 1). Recall that our two classes are now identical (same preferences and same wage distribution). Hence if $R_o^* \neq \bar{R}_1^*$ and, consequently, (11) does not hold, we may say that we have a case of horizontal inequity. Thus, the question whether income tax credits for dependent members of the household are desirable becomes a question of whether horizontal equity is desirable under the utilitarian criterion for social welfare.

Let us therefore turn back to the optimization problem (12) [or (13)] and see whether S_o is a concave function or not. Given that the



(a) S_o is concave: horizontal equity.

(b) S_c is not concave: horizontal inequity. (Note: the slope of the curve at A must be equal to the slope at B.)

Figure 1

utility function u^o is concave so that there is a diminishing marginal utility of consumption and increasing marginal disutility of labor, one may expect S^o to be concave. This is indeed true when one considers fully optimal allocations which need lump-sum taxation in order to be sustained. More generally, one can plainly show that S_o is concave if the set A_o is convex.[5] But if A_o is not convex, then S^o needs not be concave, and horizontal inequity might be desirable.

Stiglitz (1976) has also concluded that utilitarianism may lead to horizontal inequity in a very simple context of indirect taxation and one-person economy. In both his case and ours, this inequity stems from the second-best nature of the tax tools which are available. In Stiglitz's case, it is the distortionary nature of indirect taxation (no lump-sum taxes being allowed). In our case, the taxes need not be distortionary. Some sort of lump-sum taxation falls under the category of direct taxation. For instance, a head tax and even some lump-sum taxes which discriminate among individuals are allowed in our case. As a matter of fact, in the example of horizontal inequity which we provide in the next section, the income tax need not be distortionary. (One can move along some portion of the Pareto-frontier with it.) Yet, our income tax fails to achieve a full-optimum; for instance, it fails to equate the marginal utilities of consumption for all households. This is the source of the horizontal inequity result in our case. To see this, consider a situation in which the government treats our two classes of household equally, so that $R_o^* = \bar{R}_1^*$. Now suppose that we consider raising one more dollar from one of the classes and one less dollar from the other. It may be the case that the burden of raising the additional dollar from the first class falls primarily on individuals with low marginal utilities of consumption, while the gain to the other class goes primarily to households with high marginal utilities of consumption. (Recall that the government cannot perfectly control on whom to place the burden of the tax or to whom the gain will go.) Thus, the loss to one of the classes might be lower than the gain to the other: horizontal inequity will be desirable.

We conclude this section by offering some suggestions on how to construct an example of horizontal inequity. Suppose that $y_o(w, R_o)$ and $x_o(w, R_o)$ are linear in R_o. Then:

$$S_o(R_o + \Delta R_o) - S_o(R_o) = \tag{17}$$

$$\int \Big\{ u^o[y_o(w, R_o + \Delta R_o), x_o(w, R_o + \Delta R_o)]$$

$$- u^o[y_o(w, R_o), x_o(w, R_o)] \Big\} dF_o(w)$$

$$\leq \int \Big\{ u_1^o[y_o(w, R_o), x_o(w, R_o)] \quad [y_o(w, R_o + \Delta R_o) - y_o(w, R_o)]$$

$$+ u_2^o[y_o(w, R_o), x_o(w, R_o)] \, [x_o(w, R_o + \Delta R_o) - x_o(w, R_o)] \Big\} dF_o(w)$$

$$= \Delta R_o \int \frac{\partial y_o(w, R_o)}{\partial R_o} \; u_1^o[y_o(w, R_o), x_o(w, R_o)] \, dF_o(w)$$

$$+ \Delta R_o \int \frac{\partial x_o(w, R_o)}{\partial R_o} \; u_2^o[y_o(w, R_o), x_o(w, R_o)] \, dF_o(w),$$

where the inequality sign follows from the concavity of u^o, and the last equality sign follows from the assumption of the linearity of $y_o(w, R_o)$ and $x_o(w, R_o)$ in R_o. Let us now calculate the derivative S_o' of S_o, using (14):

$$S_o'(R_o) = \int u_1^o [y_o(w, R_o), x_o(w, R_o)] \frac{\partial y_o(w, R_o)}{\partial R_o} dF_o(w) \qquad (18)$$

$$+ \int u_2^o [y_o(w, R_o), x_o(w, R_o)] \frac{\partial x_o(w, R_o)}{\partial R_o} dF_o(w).$$

Combining (17) and (18), we conclude that $S_o(R_o + \Delta R_o) - S_o(R_o) \leqq S_o'(R_o)\,\Delta R_o$, which implies that S_o is concave. Hence, when we look for an example of horizontal inequity, we much search for cases where $y_o(w, R_o)$, and $x_o(w, R_o)$ are not linear in R_o. This is exactly what we do in the next section.

AN EXAMPLE OF HORIZONTAL INEQUITY

Consider an economy with two individuals. One of them (the poor man, henceforth) faces a wage rate of w_1; the other (the rich man, henceforth) earns a wage rate of w_2, where $w_2 > w_1$. To find the optimal income tax for this economy we have to solve:

$$\max \left\{ u[y(w_1), x(w_1)] + u[y(w_2), x(w_2)] \right\} \qquad (19)$$

$$\text{s.t.:} \quad w_1 y(w_1) - x(w_1) + w_2 y(w_2) - x(w_2) \geqq R$$

$$[y(\cdot), x(\cdot)] \in A,$$

where A is the set of allocations which can be supported by an income tax function. Denoting the optimal value of the objective function in (19) by S(R), we want to show that S is not everywhere concave.

The income tax in our case can be described as consisting of two linear taxes, one for each individual. Denote by $1-a_i$ and T_i, respectively, the marginal tax rate and the lump-sum components of the linear tax facing individual $i = 1, 2$. Let $v(a_i w_i,\ a_i w_i \bar{y} - T_i)$, $x(a_i w_i,\ a_i w_i \bar{y} - T_i)$, and $y(a_i w_i,\ a_i w_i \bar{y} - T_i)$ be the indirect utility, the consumption demand and the labor supply functions, respectively. The arguments of these functions are the net wage rate $(a_i w_i)$ and the net full-income $(a_i w_i \bar{y} - T_i)$, where $\bar{y}$ is the maximum amount of hours that each individual can work (i.e., the endowment of leisure).

In order for the allocation $[y(\cdot), x(\cdot)]$ induced by these two linear taxes to belong to A, these linear taxes must be restricted in the following way. It must be the case that it does not pay for the rich to decide to be poor (i.e., to decide to earn the gross income of the poor) and vice versa. In order for the rich man to make the gross earning of the poor man, he will have to work only $w_1\ y(a_1 w_1,\ a_1 w_1 \bar{y} - T_1)/w_2$. Therefore, the rich will not decide to be poor if

$$v(a_2 w_2,\ a_2 w_2 \bar{y} - T_2) \geqq u\left(\frac{w_1}{w_2}\ y(a_1 w_1,\ a_1 w_1 \bar{y} - T_1),\ x(a_1 w_1,\ a_1 w_1 \bar{y} - T_1)\right). \quad (20)$$

Similarly, the poor will not decide to be rich if

$$v(a_1 w_1,\ a_1 w_1 \bar{y} - T_1) \geqq u\left(\frac{w_2}{w_1}\ y(a_2 w_2\ ,\ a_2 w_2 \bar{y} - T_2),\ x(a_2 w_2,\ a_2 w_2 \bar{y} - T_2)\right).^6 \quad (21)$$

As is expected, one can show that (21) is not a binding constraint and may be ignored, whereas (20) has to hold as an equality at the optimum (see the Appendix).

The government's revenue constraint is, in this case:

$$\sum_{i=1}^{2} [w_i y\ (a_i w_i,\ a_i w_i \bar{y} - T_i) \ -\ x(a_i w_i,\ a_i w_i \bar{y} - T_i)] \geqq R. \tag{22}$$

Thus, the government chooses a_1, a_2, T_1 and T_2 so as to maximize

$$\sum_{i=1}^{n} v(a_i w_i,\ a_i w_i \bar{y} - T_i) \tag{23}$$

subject to the constraints (20) and (22). For each R, denote the maximum value of the objective function (23) by S(R). Our aim is to find an example where S(R) is not concave.

To do this, we pick some revenue level, say R = R*, and see whether S is convex around R*. However, it turns out to be easier to look at another function S* instead of S. The function S* is defined as follows. Let a_1^* and a_2^* be the optimal values of a_1 and a_2 for R = R*. Now let us maximize (23) subject to (20) and (22), where a_1 and a_2 are constrained to be equal to a_1^* and a_2^*, respectively. Thus, we solve

$$\max_{T_1,\ T_2} \left\{ \sum_{i=1}^{2} v(a_i^* w_i,\ a_i^* w_i \bar{y} - T_i) \right\} \tag{24}$$

subject to:

$$v(a_2^* w_2,\ a_2^* w_2 \bar{y} - T_2) \geqq u\left(\frac{w_1}{w_2}\ y(a_1^* w_1,\ a_1^* w_1 \bar{y} - T_1),\ x(a_1^* w_1,\ a_1^* w_1 \bar{y} - T_1) \right) \tag{25}$$

and

$$\sum_{i=1}^{2} \left[w_i y(a_i^* w_i,\ a_i^* w_i \bar{y} - T_i) \ -\ x(a_i^* w_i,\ a_i^* w_i \bar{y} - T_i) \right] \geqq R \tag{26}$$

The maximum value of (24) is now denoted by $S^*(R)$. Clearly, by the very definition of a_1^* and a_2^*, we must have

$$S^*(R) \leqq S(R) \quad \text{for all } R \tag{27}$$

and

$$S^*(R^*) = S(R^*).$$

In view of (27), if S^* is strictly convex around R^*, then S is strictly convex around R^* (see Figure 2). Thus, it suffices to find an example where S^* is strictly convex.

Consider a linearly homogenous Cobb-Douglas utility function:

$$u(y, x) = (\bar{y} - y)^{1-\alpha} x^{\alpha}, \quad 0 < \alpha < 1. \tag{28}$$

In this case, we have (assuming that R is high enough so that the labor supplies are positive)

$$x(a_i^* w_i, a_i^* w_i \bar{y} - T_i) = \alpha(a_i^* w_i \bar{y} - T_i) \tag{29}$$

$$y(a_i^* w_i, a_i^* w_i \bar{y} - T_i) = \bar{y} - \frac{(1-\alpha)}{a_i^* w_i} (a_i^* w_i \bar{y} - T_i). \tag{30}$$

$$v(a_i^* w_i, a_i^* w_i \bar{y} - T_i) = \alpha^{\alpha}(1 - \alpha)^{1-\alpha} (a_i^* w_i)^{\alpha-1} (a_i^* w_i \bar{y} - T_i). \tag{31}$$

It is simple to show that (26) has to hold as an equality. This can be used in order to solve T_1 in terms of T_2 and R:

$$T_1 = c + eT_2 + kR \tag{32}$$

where

$$c = \frac{-a_1^* \bar{y}[(1 - a_1^*) \alpha w_1 + (1 - a_2^*) \alpha w_2]}{1 - \alpha + \alpha a_1^*}$$

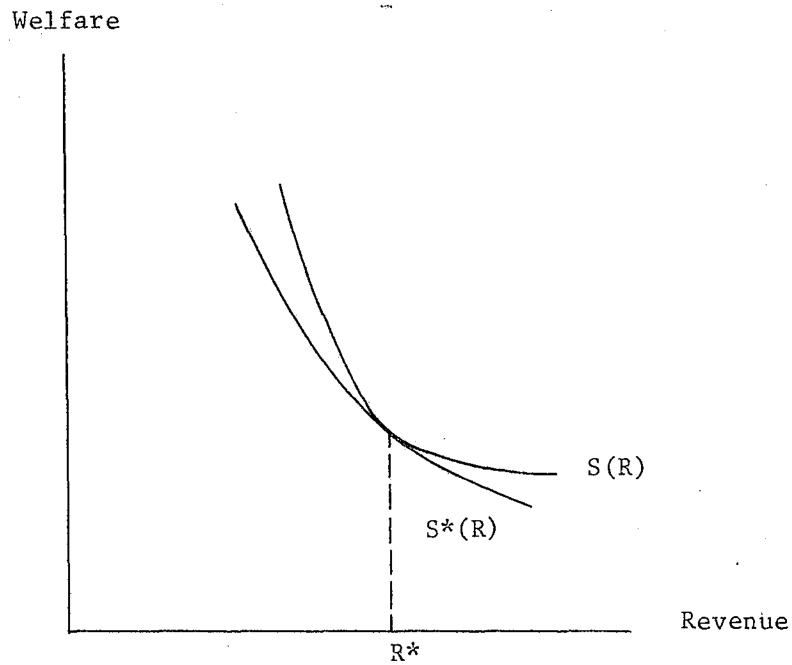


Figure 2

$$e = -\frac{(1 - \alpha + \alpha a_2^*)\, a_1^*}{(1 - \alpha + \alpha a_1^*) a_2^*} < 0$$

and

$$k = \frac{a_1^*}{1 - \alpha + \alpha a_1^*} > 0.$$

The optimization problem of maximizing (24) subject to (25) and (26) is now reduced to

$$\max_{T_2} \left[(a_1^* w_1)^{\alpha-1} (a_1^* w_1 \bar{y} - c - eT_2 - kR) + (a_2^* w_w)^{\alpha-1} (a_2^* w_2\bar{y} - T_2)\right] \alpha^{\alpha}(1 - \alpha)^{1-\alpha}, \tag{33}$$

subject to:

$$(1 - \alpha)^{1-\alpha}(a_2^* w_2)^{\alpha-1} (a_2^* w_2\bar{y} - T_2) \geqq A^{1-\alpha} B^{\alpha}, \tag{34}$$

where

$$A = \bar{y} - \frac{w_1}{w_2}\bar{y} + \frac{(1 - \alpha)}{a_1^* w_2} (a_1^* w_1\bar{y} - c - eT_2 - kR) > 0 \tag{35}$$

and

$$B = a_1^* w_1\bar{y} - c - eT_2 - kR > 0. \tag{36}$$

The first-order necessary condition for this optimization problem is

$$-(1 - \alpha)^{1-\alpha} \left[(a_1^* w_1)^{\alpha-1} e + (a_2^* w_2)^{\alpha-1}\right] - \lambda\left[(1 - \alpha)^{1-\alpha} (a_2^* w_2)^{\alpha-1} - e(1 - \alpha)A^{-\alpha}B^{\alpha} - e\alpha A^{1-\alpha}B^{\alpha-1}\right] = 0, \tag{37}$$

where $\lambda \geqq 0$ is the Lagrange multiplier. Consider what happens when $R = R^*$.

If $\lambda = 0$, then the first-best optimum is achieved at R*, which is impossible. Hence, $\lambda > 0$. Then it follows from (37) that (recalling that $e < \hat{0}$)

$$- (1 - \alpha)^{1-\alpha} \left[(a_1^* w_1)^{\alpha-1} e + (a_2^* w_2)^{\alpha-1} \right] > 0. \tag{38}$$

Denoting the optimal T_2 by $T_2(R)$, it follows that

$$S^*(R) = \alpha^\alpha (1 - \alpha)^{1-\alpha} \left[(a_1^* w_1)^{\alpha-1} (a_1^* w_1 \bar{y} - c - eT_2(R) - kR) \right. \tag{39}$$

$$\left. + (a_2^* w_2)^{\alpha-1} (a_2^* w_2 \bar{y} - T_2(R)) \right],$$

and hence,

$$\frac{d^2 S^*}{dR^2} = - \alpha^\alpha (1 - \alpha)^{1-\alpha} \left[(a_1^* w_1)^{\alpha-1} e + (a_2^* w_2)^{\alpha-1} \right] \frac{d^2 T}{dR^2} . \tag{40}$$

In view of (38), one can conclude that S^* is a strictly convex function of R around $R = R^*$ if and only if T is a strictly convex function of R around $R = R^*$.

Employing (34), which, as we have already mentioned, must hold as an equality at the optimum, we conclude that

$$\frac{dT_2}{dR} = - \frac{k}{e} \frac{F}{E} > 0 \tag{41}$$

and

$$\frac{d^2 T_2}{dR^2} = \frac{k}{e^2} \frac{(1 - \alpha)^{1-\alpha} (a_2^* w_2)^{\alpha-1}}{E^2} \frac{dF}{dR}, \tag{42}$$

where

$$F = (1 - \alpha)^2 A^{-\alpha} B^{\alpha}/(a_1^* w_2) + \alpha A^{1-\alpha} B^{\alpha-1} > 0 \tag{43}$$

and

$$E = F - (1 - \alpha)^{1-\alpha} (a_2^* w_2)^{\alpha-1}/e > 0. \tag{44}$$

Thus, $\frac{d^2T_2}{dR^2} > 0$ if and only if $\frac{dF}{dR} > 0$.

By differentiating (43) with respect to R we obtain

$$\frac{dF}{dR} = \left[e \frac{dT_2}{dR} + k\right] \left[\alpha(1 - \alpha) A^{1-\alpha} B^{\alpha}\right] \left[\frac{1-\alpha}{a_1^* w_2} A^{-1} - B^{-1}\right]^2. \tag{45}$$

Several authors have shown that the richest person must face a zero marginal tax rate in the case of a continuous distribution of wages [see Sadka (1976a), Phelps (1973), Mirrlees (1976), Seade (1977), Cooter (1978) and Brito and Oakland (1977)].[7] It is not difficult to establish that in our case too, $a_2^* = 1$ (see the Appendix). Hence, $- e = k$ and, from (41), (43) and (44), one can conclude that $\frac{dT_2}{dR} < 1$. Thus,

$$e \frac{dT_2}{dR} + k = k \left(1 - \frac{dT_2}{dR}\right) > 0. \tag{46}$$

Then (45) implies that $dF/dR > 0$, and thus, $d^2T_2/dR^2 > 0$ and, consequently, $d^2S^*/dR^2 > 0$.[8]

An intuitive explanation for the convexity of S^* is offered at the end of the Appendix.

CONCLUDING REMARKS

(a) The utility function (28) which we used to demonstrate the optimality of horizontal inequity was linearly homogenous, so that it was concave but not strictly so. However, the utility function u^{ε} becomes strictly concave when $0 < \varepsilon < 1$. Continuity considerations suggest that one can construct an example of horizontal inequity with a strictly concave utility function by taking ε sufficiently close to one.

(b) Another procedure for incorporating household size into demand analysis is demographic scaling, used as the basis for constructing equivalence scales. Demographic scaling in our case amounts to assuming that the utility function of the two classes are related to each other by

$$u^1(y, x) = u^o(y, x/b),$$

where b is a constant greater than one which is called an equivalence scale. In this case, one should certainly not expect income tax credits to constitute an optimal policy.

One can argue that the idea behind allowing joint filing of returns for married couples is to make the optimal allocations $[y_i^*(\cdot), x_i^*(\cdot)]$, $i = 1, 2$, be such that $y_o^*(w) = y_1^*(w)$ and, especially, $x_o^*(w) = x_1^*(w)/b$. However, we do not believe that such a policy is warranted. If we proceed here, as in the case of demographic translating, by making a similar transformation to (6), namely $\bar{x}_1 = x_1/b$, then the situation will be different in the two cases. With demographic translating, the marginal rate of substitution of x_o for $\bar{x}_1$ is one. With demographic scaling, this rate is only $1/b$. In other words, increasing $\bar{x}_1$ by one unit will cost the society b units of x_o. For this reason one should not expect to have an optimal policy where $x_o^*(w) = \bar{x}_1^*(w)$ for all w.[9]

Appendix

We will show here that: (a) the constraint (20) is binding whereas (21) is not; (b) $a_2^* = 1$; and (c) $a_1^* < 1$. We will also offer an intuitive explanation for the convexity of S*.

For this purpose, we will adopt the transformation used by Sadka (1976a), and Mirrlees (1976). By denoting gross income as z = wy, we can describe the preferences of person i over bundles (z, x) by the utility function $U^i(z, x)$ as follows:

$$U^i(z, x) = u(\frac{z}{w_i}, x). \tag{47}$$

One can show that for a large class of utility functions, including the Cobb-Douglas one which we employ, the indifference curves, $U^i(z, x)$ = constant, are steeper for the poor than for the rich:

$$- U_1^1(z, x)/U_2^1(z, x) > -U_1^2(z, x)/U_2^2(z, x) \tag{48}$$

for all (z, x). One can also show that this must imply that $z_1 \leqq z_2$, where z_i is the gross income chosen by person i.

Using this transformation, the constraints that the rich will not choose to be poor and vice versa [namely, (20) and (21)] can now be written as

$$U^2(z_1, x_1) \leqq U^2(z_2, x_2) \tag{49}$$

and

$$U^1(z_2, x_2) \leqq U^1(z_1, x_1). \tag{50}$$

These constraints are illustrated in Figure 3.

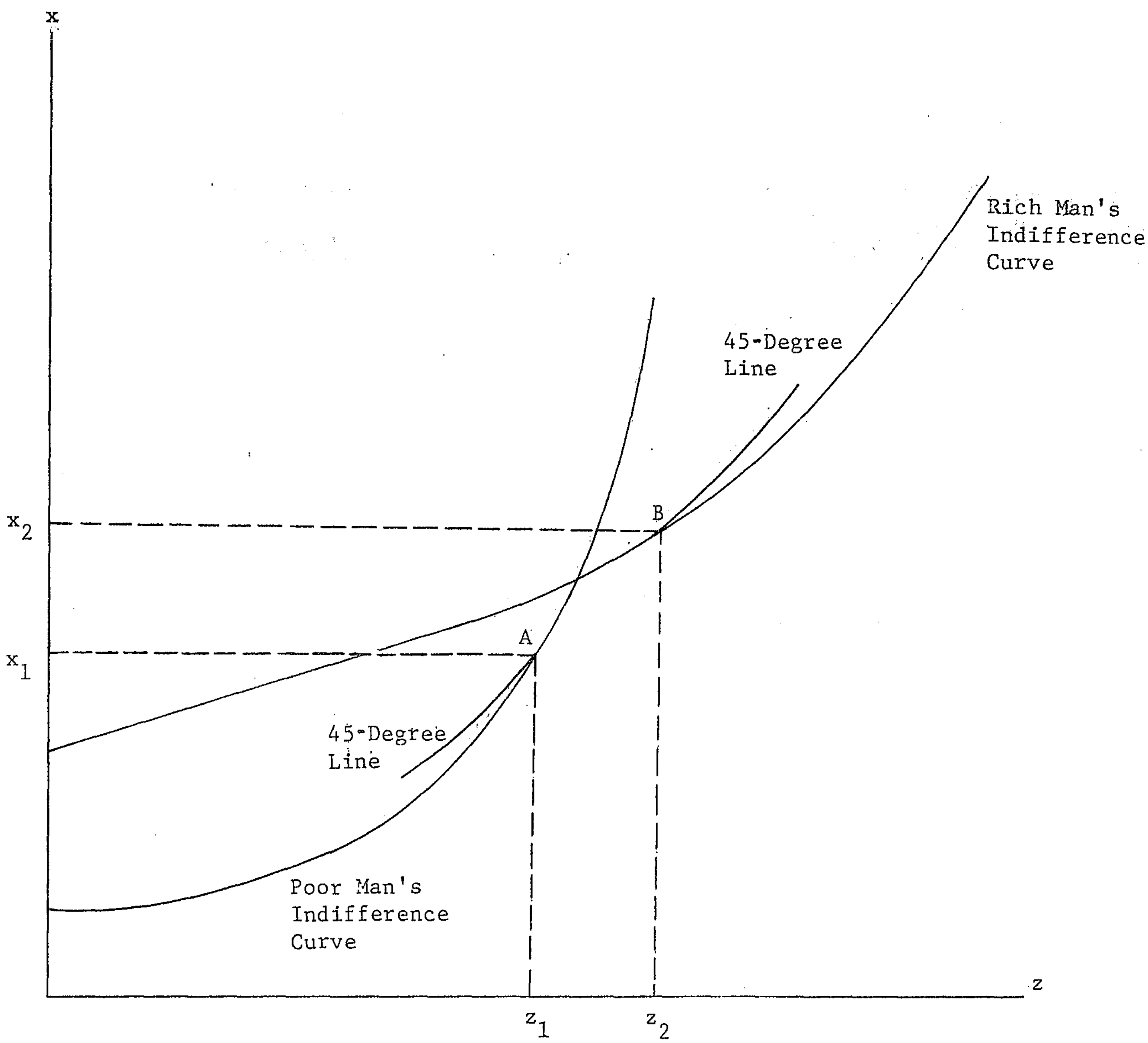

x
Rich Man's
Indifference
Curve
45-Degree
Line
B
x_2
A
x_1
45-Degree
Line
Poor Man's
Indifference
Curve
z
z_1
z_2

Figure 3

The government's revenue constraint (22) becomes

$$z_1 - x_1 + z_2 - x_2 \geqq R. \tag{51}$$

Employing Figure 3 we can show that

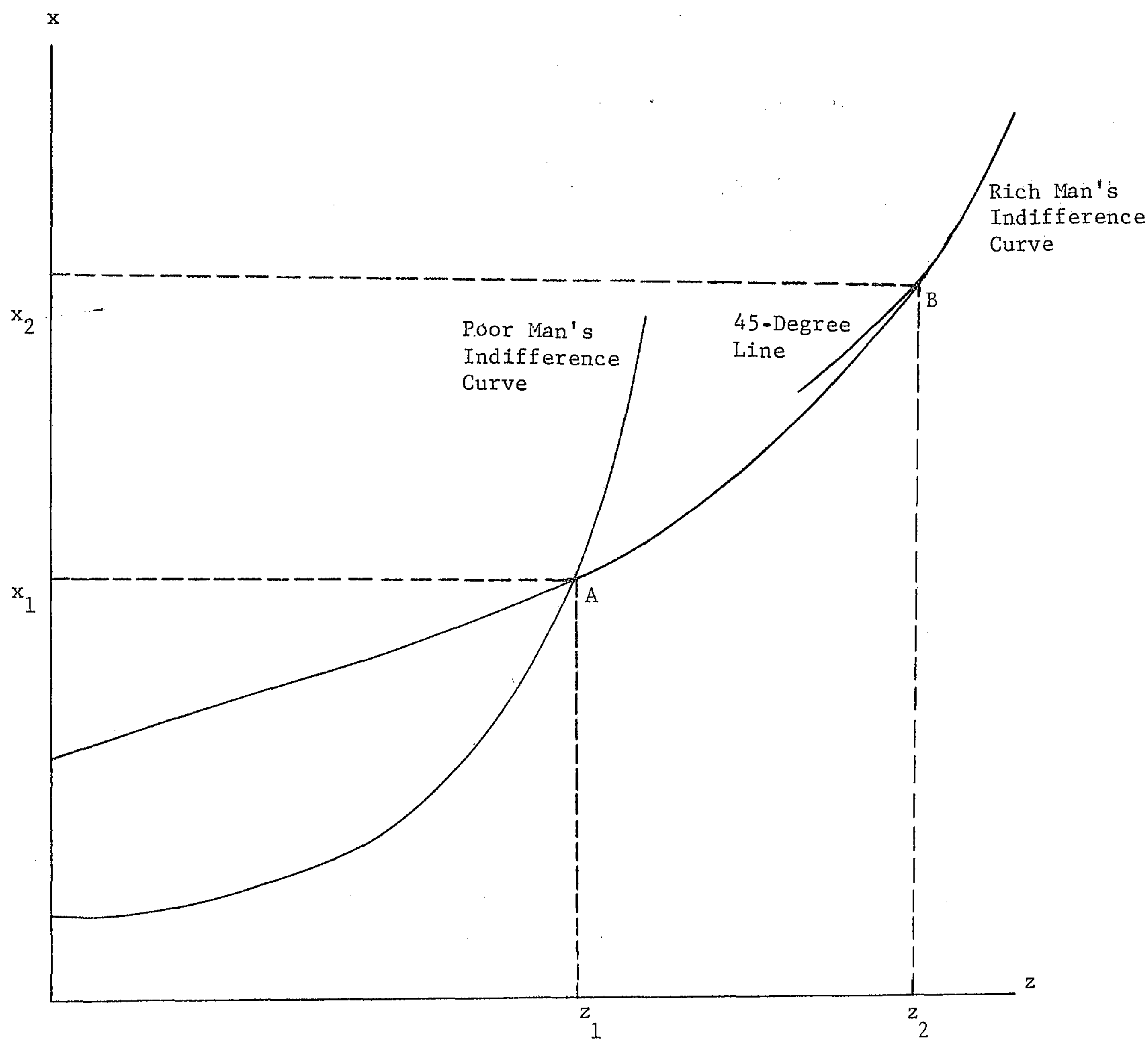
$$a_1^* \leqq 1 \text{ and } a_2^* \geqq 1. \tag{52}$$

If $a_1^* > 1$, then we can let the poor man move slightly along his indifference curve to the left of A. Such a move will increase his tax payment without affecting his utility, which is a contradiction. Similarly, it is shown that $a_2^* \geqq 1$.

It follows from (52) that the net wage rate of the poor $(a_1^* w_1)$ is lower than the net wage rate of the rich $(a_2^* w_2)$. Given the concavity and the linear homogeneity of our Cobb-Douglas utility function, this implies that the ratio $x_1/(\bar{y} - y_1)$ of the poor is lower than that of the rich. But then the poor's marginal utility of x is higher than the rich's. Thus, if the constraint (49) is not binding, welfare will be improved by raising x_1 and lowering x_2 by the same small amount. Hence, (49) must be binding at the optimum. But when (49) is binding, then (50) must not be (see Figure 4). [Note that (52) and (48) imply that $z_1 \neq z_2$ and hence, $z_1 < z_2$.] This proves (a).

We have already shown that $a_2^* \geqq 1$. If $a_2^* > 1$, then the situation is depicted in Figure 4. The rich man's tax payment will increase while his utility will not change if we let him move slightly along his indifference curve to the left of B. Thus, we must have $a_2^* = 1$, which proves (b).

In order to prove (c), it suffices to show that $a_1^* \neq 1$, for, by (52), $a_1^* \leqq 1$. The optimal tax can be found by choosing x_1, z_1, x_2 and z_2, so as



x
x_2
x_1
Rich Man's
Indifference
Curve
Poor Man's
Indifference
Curve
45-Degree
Line
A
B
z
z_1
z_2

Figure 4

to maximize

$$U^1(z_1, x_1) + U^2(z_2, x_2) \tag{53}$$

subject to (49) and (51).

The first-order necessary conditions for this optimization problem are:

$$U^1_1(z_1, x_1) - \theta U^2_1(z_1, x_1) + \lambda = 0 \tag{54}$$

$$U^1_2(z_1, x_1) - \theta U^2_2(z_1, x_1) - \lambda = 0 \tag{55}$$

$$U^2_1(z_2, x_2) + \theta U^2_1(z_2, x_2) + \lambda = 0 \tag{56}$$

$$U^2_2(z_2, x_2) + \theta U^2_2(z_2, x_2) - \lambda = 0 \tag{57}$$

where $\theta \geqq 0$ and $\lambda \geqq 0$ are Lagrange multipliers. If $a^*_1 = 1$, then

$$U^1_1(z_1, x_1)/U^1_2(z_1, x_1) = -1 \tag{58}$$

This, in conjunction with (54) and (55), implies that

$$\theta U^2{}_1(z_1, x_1) = -\theta U^2_2(z_1, x_1). \tag{59}$$

Since θ cannot be zero (for otherwise the first-best optimum is achieved), it follows from (59) that

$$U^2_1(z_1, x_1)/U^2_2(z_1, x_1) = -1. \tag{60}$$

But then (58) and (60) violate (48). This proves (c).

We will now try to help the reader's intuition of why S^* is convex around R*. Since the marginal tax rates are kept constant, then so are the net wage rates ($a_1^*w_1$ and $a_2^*w_2$). Consequently, as R changes, the two persons move along their income-consumption curves (ICC), satisfying the constraint (49) as an equality. Suppose that at R* the rich and the poor are at A_2 and A_1, respectively, in Figure 5. Then, as R is decreased by ΔR, the rich man moves to, say, B_2, and when R is increased by the same ΔR, he moves to, say, C_2. Since the Cobb-Douglas utility function is linear along the rich man's ICC, it follows that the increase in the rich man's utility between B_2 and A_2 is exactly equal to the decrease between A_2 and C_2, if $B_2A_2 = A_2C_2$. A similar result holds with respect to the rich man's tax payment.

From the homotheticity of the Cobb-Douglas utility function, it follows that $B_1'' A_1 = A_1C_1''$ (because $B_2A_2 = A_2C_2$). Hence, $B_1'A_1 = A_1C_1'$. Thus, it is feasible, but not optimal, to move the poor from A_1 to B_1', as R is decreased by ΔR, and from A_1 to C_1', as R is increased by the same ΔR, Since the utility function is linear along the poor man's ICC, it follows that the increase in the poor man's utility from A_1 to B_1' is exactly equal to the decrease from A_1 to C_1'. Thus, the increase in social welfare as R is decreased by ΔR will be exactly equal to the decrease in social welfare as R is increased by the same ΔR, if we indeed let the poor move from A_1 to B_1' and from A_1 to C_1'. But such a policy is not optimal because the constraint (49) will not be satisifed as an equality. Hence, we conclude that the increase in S* as R is decreased by ΔR must be higher than the decrease in S*, as R is decreased by the same ΔR. In other words, S* must be strictly convex around R*.

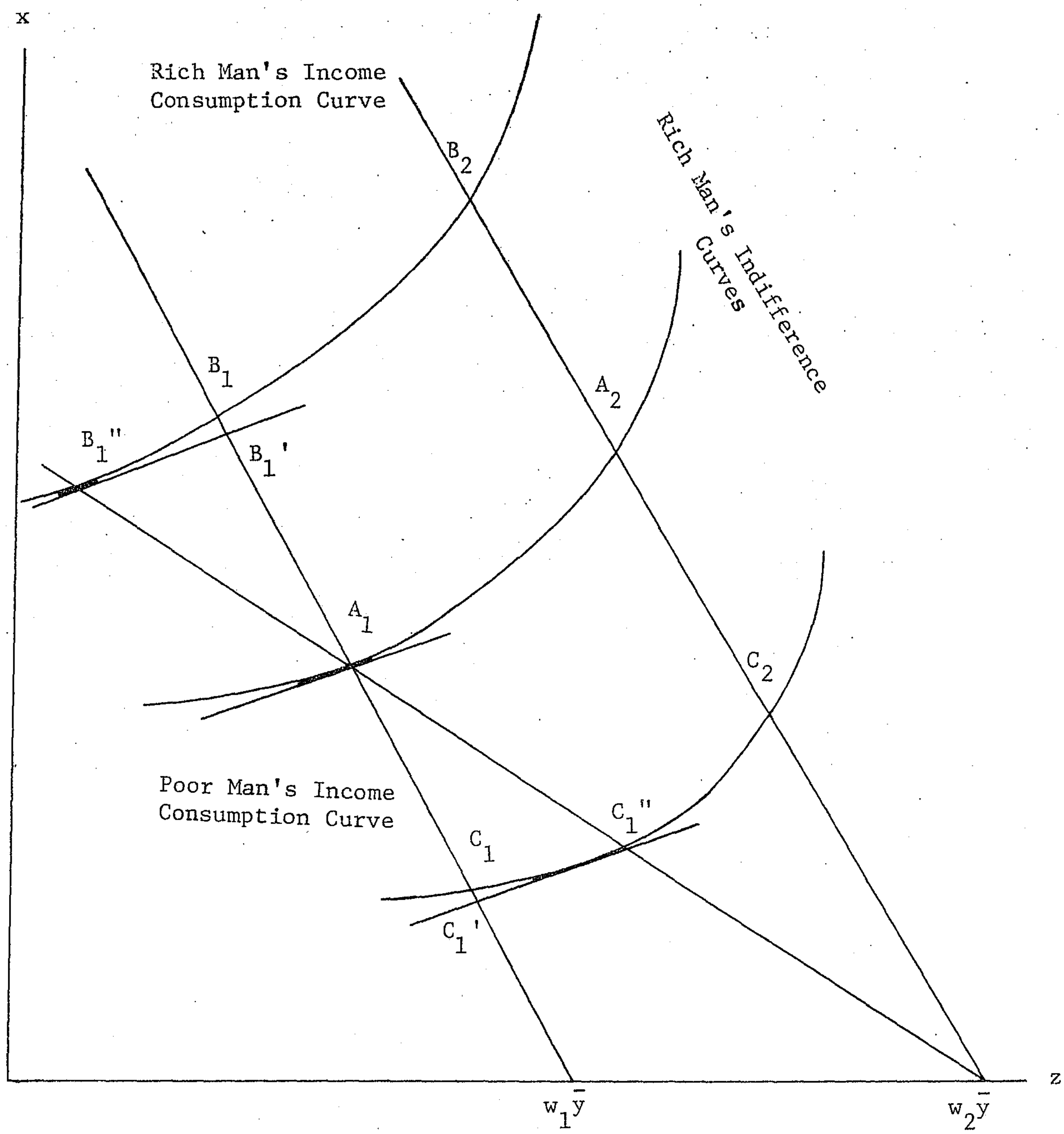

x
Rich Man's Income
Consumption Curve
Rich Man's Indifference
Curves
B2
B1
B1"
B1'
A2
A1
C2
C1"
C1
C1'
Poor Man's Income
Consumption Curve
z
w1ȳ
w2ȳ

Figure 5

NOTES

[1]It is straightforward to extend our analysis to economies with many size classes.

[2]The paper can be easily extended to the case where x is a bundle (vector) of consumption goods and y is a vector of various types of labor services.

[3]For an opposing view, see Pollak and Wales (1978c).

[4]For a similar conclusion in the much simpler context of indirect taxation, see Stiglitz (1976) and Atkinson and Stiglitz (1976).

[5]Note that the set A_o is indeed convex when lump-sum taxation is readily available.

[6]The constraints (20) and (21) are the discrete case analogue of condition (2.3) of Mirrlees (1976).

[7]These authors (except the first two) also show that the marginal tax rate facing the poorest person must, if he works, be zero. In our case of a discrete wage distribution, we can show (see the Appendix) that $a_1^* < 1$, so that the marginal tax rate faced by the poor must be positive, provided that he works.

[8]Interestingly, S cannot be convex in R everywhere. We can show that if R is low enough (and possibly negative) so that it becomes socially optimal for the poor not to work at all, then S must be strictly concave.

[9]This conclusion depends crucially on the fact that the utilitarian criterion (2) does not give higher weights for larger households. While one

is indeed justified not to use weights when the procedure of translating is adopted (as we have done throughout most of the paper), one may argue in favor of using weights in the case of scaling. We do not elaborate on this issue here.

REFERENCES

Atkinson, A.B. and J.E. Stiglitz. 1976. The design of tax structure: Direct versus indirect taxation, Journal of Public Economics, 6, 55-75.

Brito, D.L. and W.H. Oakland. 1977. Some properties of the optimal income Tax, International Economic Review, 18 , 407-32.

Cooter, R. 1978. Optimal tax schedules and rates: Mirrlees and Ramsey, American Economic Review, 68, 756-68.

Mirrlees, J.A. 1976. Optimal tax theory: A synthesis, Journal of Public Economics, 6, 327-58.

Phelps, E. 1973. Taxation of wage income for economic justice, Quarterly Journal of Economics, 87, 331-54.

Pollak, R.A. and T.J. Wales. 1978a. Estimation of complete demand systems from household budget data: The linear and quadratic expenditure systems, American Economic Review, 68, 348-59.

---. 1978b. Comparison of the quadratic expenditure system and translog demand systems with alternative specifications of demographic effects, Discussion Paper No. 78-21, University of British Columbia.

---. 1978c. Welfare comparisons and equivalent scales, (mimeo).

Sadka, E. 1976a. Income distribution, incentive effects and optimal income taxation, Review of Economic Studies, 43, 261-68.

---. 1976b. Social welfare and income distribution, Econometrica, 44, 1239-51.

Seade, J. 1977. On the shape of optimal tax schedules, Journal of Public Economics, 7, 203-20.

Stiglitz, J.E. 1976, Utilitarianism and horizontal equity: The case for random taxation, Technical Report No. 214, Institute for Mathematical Studies in the Social Sciences, Stanford University.