272-75

INSTITUTE FOR RESEARCH ON POVERTY DISCUSSION PAPERS

A DIAGRAMMATIC EXPOSITION OF THE THEORY
OF OPTIMAL INCOME TAXATION

Efraim Sadka

UNIVERSITY OF WISCONSIN - MADISON

A DIAGRAMMATIC EXPOSITION OF THE THEORY OF OPTIMAL INCOME TAXATION

Efraim Sadka

June 1975

The research reported in this paper was supported 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 Equal Opportunity Act of 1964. The views expressed are those of the author.

ABSTRACT

This paper is a survey of the recent literature on the optimal taxation of earned income. It focuses on the issues of wage-subsidy, a negative income tax, and progressive taxation. Emphasis is placed on the impact of alternative criteria of distributional equity on these issues.

I. Introduction

In modern economies, income taxation serves as the most important tool for raising revenues and redistributing income. Yet, a theoretical study of it has begun only in the last few years (Mirrlees [4], Sheshiuski [9], Phelps [6], and others). At present, we still know very little about the optimal tax. Furthermore, due to the mathematical complexity of these studies, even their results are not readily accessible to standard economists. The purpose of this paper is to present as simple an exposition as possible of the theory of optimal income taxation.

The adjective "optimal" is, of course, meaningless without a reference to the objectives of society. Here we will discuss both the classical sum-of-utilities criterion, that is, an additive social welfare function, and Rawls's [7] max-min criterion.[1]

Two different models are employed in the literature on this subject. In the first model, which was introduced by Mirrlees [4] and which will be referred to here as the labor model, the income of an individual depends on the number of hours worked. In the second model, which was introduced by Sheshinski [9] and which will be referred to here as the education model, the level of education that determines one's income.

In both models an income tax has two effects: (a) on the size of the national income through either the incentive to work or to obtain further education (depending on the model) and (b) on the distribution of aftertax income. Both social ordering criteria favor, in a way to be explained later, equality over inequality in the distribution of income. Both also favor higher national income. On the other hand, an income tax usually works in opposite directions with respect to these objectives. The derivation of the optimal tax therefore involves a determination of the optimal trade-off between the two objectives.

In general, an optimal tax possesses very few properties. Trivially, we may always assume that the marginal rate of any tax (not necessarily the optimal one) is not higher than 100 percent. It is also known that the optimal tax should have a nonnegative marginal rate everywhere (i.e., no wage subsidy).[2] This is true in either model and under either social ordering. For the max-min criterion it is further proven that the marginal rate has to be positive (in both models). Some weak results are established on the question of the progressivity of the optimal tax; it is shown that progressivity has nothing to do with equality and, in fact, no tax can be optimal if it is progressive everywhere, in either of the models and under either of the social ordering criteria. When we confine ourselves to linear (but not necessarily proportional) taxes, we are able to show that the marginal rate should be higher under the max-min criterion than under the additive criterion.

We will be concerned primarily with Mirrlees's labor model. The education model has very similar mathematical properties, but it also has a peculiar attribute that makes it less interesting. Only at the very end will we briefly describe the latter model.

II. The Labor Model

A. A Description of the Model

1. Tastes

There are only 2 commodities: a composite consumption good, x, and labor services, y. All have the same tastes over (x,y), represented by the common utility function u(x,y). The marginal utility of x is positive ($u_x > 0$) and falling ($u_{xx} < 0$). Work generates marginal disutility ($u_y < 0$) which is increasing in y (that is, $u_{yy} < 0$).[3] We also assume that the marginal utility of x increases as y increases (that is, $u_{xy} > 0$).[4] The indifference curves are, as usual, smooth and convex. (A map of such indifference curves is depicted in Figure 1.) All have the same endowment of leisure--A. It is assumed that the slope of any indifference curve $\left(\frac{-u_y}{u_x}\right)$ tends to $+\infty$ as y approaches A. The indifference curves are therefore asymptotic to the dashed line along which y = A (Figure 1). Both consumption and leisure are normal goods, which means that an upward shift of the budget line will increase x and decrease y.

2. Skills

Individuals differ in their skills or abilities to produce income. Skill is denoted by n. The range of values of n is an interval between two points, N_1 and N_2, inclusive, where $N_1 < N_2$, $N_1 \geqq 0$ and $N_2 < \infty$. It is assumed that there is a continuum of individuals,[5] so that the distribution of skills is continuous. We denote by F(n) the number of persons with skill n or less. F has all the properties of a cumulative

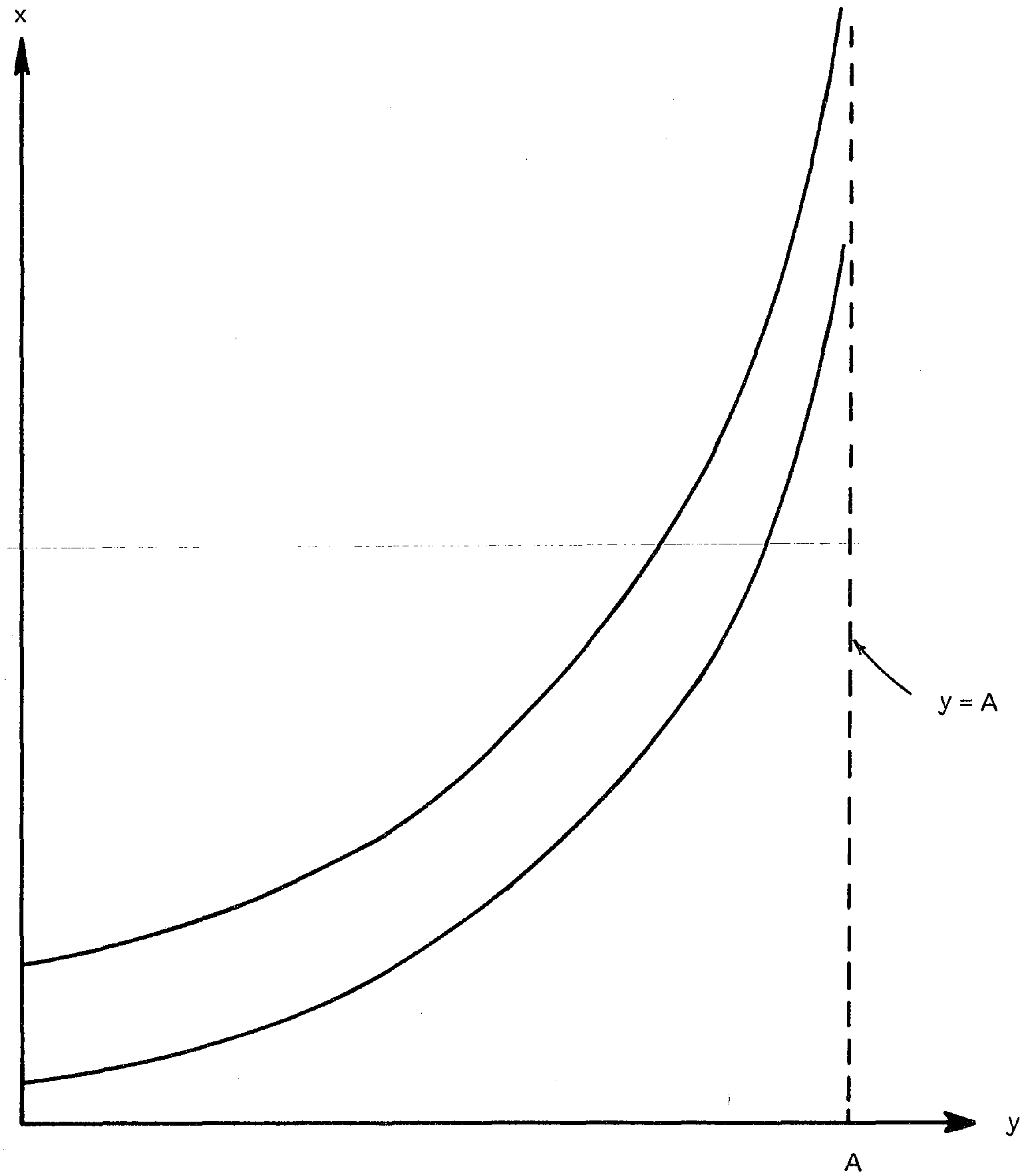


Figure 1

distribution function (of skills), except that $F(N_2)$ is equal to the number of persons in the society rather than to 1. We further assume that F is strictly increasing[6] and differentiable, and we denote F' by f. f is therefore like a frequency function; for instance, the number of people who have skills between any two points, n_1 and n_2, is $\int_{n_1}^{n_2} f(n)dn$.

The skill of a person is what determines his ability to produce income. Specifically, it is assumed that the labor services supplied by different individuals differ in their productivities according to the skills of these individuals. A common "efficiency" unit is therefore used and it is assumed that an individual with skill n who works y hours supplies labor services at the amount ny, as measured in efficiency units. His gross income is then $z = wny$, where w is the real wage rate per an efficiency unit (in terms of the consumption good).

3. Technology

In general, income taxation may have an impact on the real wage rate, w. In order to avoid this undue complication and concentrate on the issues at hand, we will assume a linear technology, in which case w is fixed. To simplify the notation we further assume an appropriate choice of units so that $w = 1$.

4. Utility Maximization

Let us henceforth use the term "person n" for "a person with a skill n." When person n works y hours his gross income is $z = ny$. His consumption is then equal to his net income, that is,

$x = z - T(z)$, where T is the income tax function. As usual, each individual is assumed to maximize utility. Thus, person n chooses a bundle (x,y) so as to maximize $u(x,y)$ subject to the constraints: $x = z-T(z)$ and $z = ny$ (or, equivalently, $x = ny-T(ny)$).

Because of skill differences, individuals who work the same number of hours earn different incomes. Thus, they face different budget lines, when plotted in the x-y plane (see Figure 2). As n increases, the budget line rotates upward, because $z = ny$ increases (except at $y = 0$).[7] Therefore individuals will, in general, choose different bundles of (x,y). We denote the utility-maximizing bundle of person n, given an income tax T, by $[x_T(n),\ y_T(n)]$. Such choice gives rise to a gross income of $z_T(n) = ny_T(n)$. The functions, $x_T(n)$, $y_T(n)$ and $z_T(n)$ are called, respectively, the consumption, the labor supply, and the gross income functions (under T). By our assumption that the indifference curves are asymptotic to the vertical line $y = A$, it follows that corner solutions with $y = A$ are excluded:

$$y_T(n) < A \text{ for all } n \text{ and } T. \tag{1}$$

We denote by $u_T(n)$ the maximum utility enjoyed by n under T, that is, $u_T(n) = u[x_T(n),\ y_T(n)]$.

5. <u>The Government</u>

The only tax available to the government here is an income tax. The government has to raise revenues to finance some predetermined level of public consumption, B. When the government imposes a tax T, it collects an amount $T[z_T(n)]$ of revenue from person n. Total tax revenues are therefore $\int_{N_1}^{N_2} T[z_T(n)]f(n)dn$ and they have to be at least as large as B:

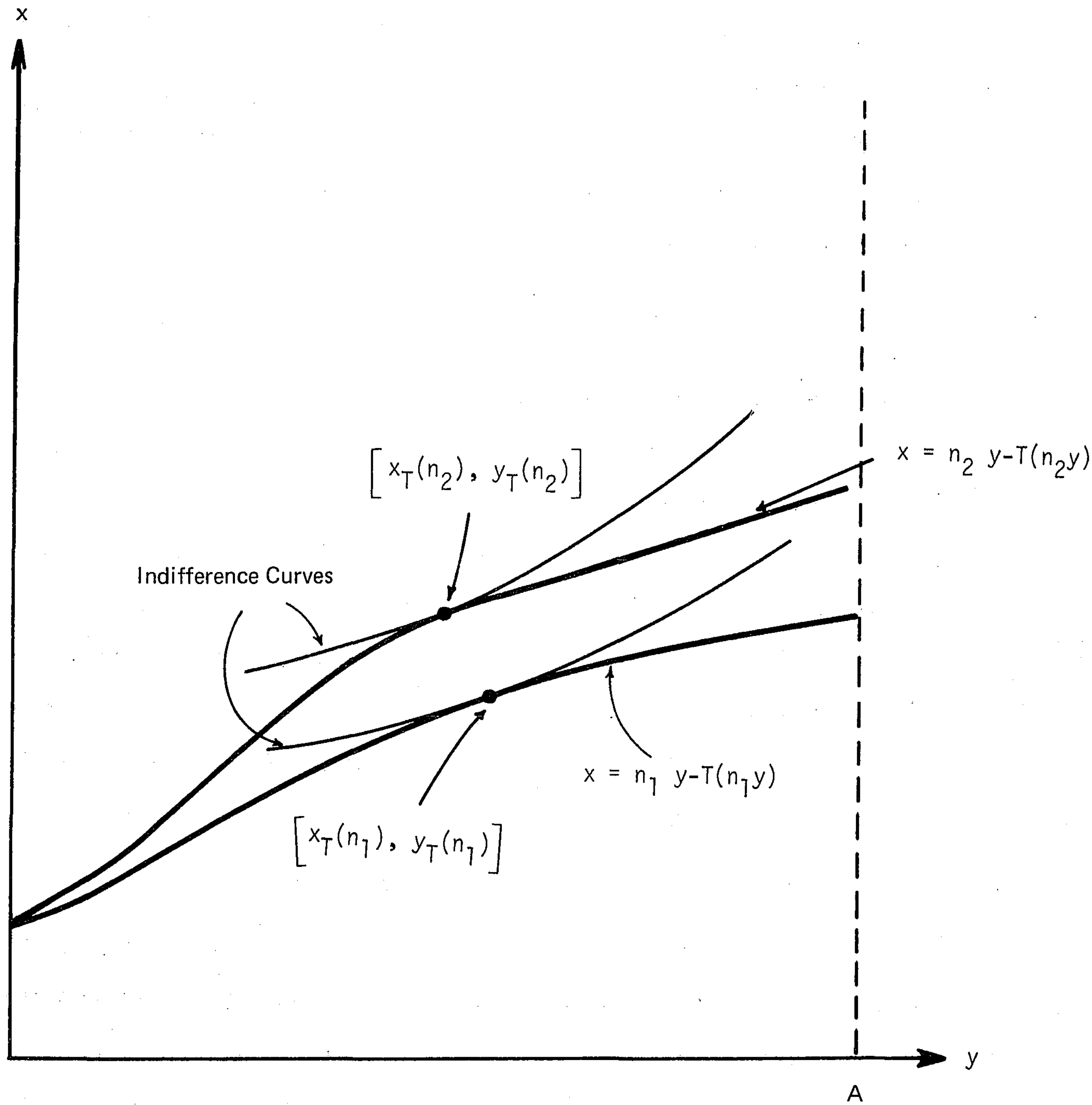


Figure 2: $n_1 < n_2$

$$\int_{N_1}^{N_2} T[z_T(n)]f(n)dn \geqq B \tag{2}$$

(Henceforth all integrals are assumed to be taken from N_1 to N_2, unless otherwise stated.)

The government is supposed to choose the best tax function, according to the criteria defined in the next subsection, which satisfies its own budget constraint (2).[8]

6. The Social Ordering Criteria

As mentioned above, two alternative criteria of social ordering are considered here: (a) an additive social welfare function and (b) the max-min criterion.

~~(a) An Additive Social Welfare~~ Function. We refer to the pair of functions, $[x(\cdot), y(n)]$. For example, the pair $[x_T(\cdot), y_T(\cdot)]$ is the allocation that actually prevails when there is a tax T. The utility level that individual n enjoys under the allocation $[x(\cdot), y(\cdot)]$ is $u[x(n), y(n)]$. The social welfare associated with this allocation, denoted by $W[x(\cdot), y(\cdot)]$, is the sum of the individual utilities:

$$W[x(\cdot), y(\cdot)] = \int u[x(n), y(n)]f(n)dn. \tag{3}$$

One may, as did Harsanyi [3], construct an axiomatic theory of social choice which leads to (3), but that is beyond the scope of this paper. Instead, we will only draw the implications of (3) for the distribution of income (or rather consumption, since it is equal to income, after tax).

Consider first the case of independent marginal utilities, that is, $u_{xy} = 0$. In this case u is of the form $u(x,y) = S(x) + R(y)$, where S and R are some concave functions ($s' > 0$, R', S'', $R'' < 0$). The social welfare function (3) then becomes:

$$W[x(\cdot),\ y(\cdot)] = \int S[x(n)]f(n)dn + \int R[y(n)]f(n)dn. \qquad (4)$$

Since the marginal utility of x declines ($s'' < 0$), it follows that the marginal utility of x is lower for the rich than for the poor. Therefore, a small transfer of income from the rich to the poor will increase the sum of their utilities. Such transfer may be called a mean-preserving concentration (because it does not affect the mean of the distribution) and it is socially desirable according to (4). For more details the reader is referred to Atkinson [1].[9] An important thing to remember here is that the transfer must not be too large, for otherwise the rich may become poorer than the former poor and the poor richer than the former rich; such a transfer spreads rather than concentrates the distribution of x. For instance, if A and B have incomes of, respectively, \$100 and \$200, then a transfer of \$10 from B to A will bring their income closer together and improve social welfare. But a transfer of \$110 from B to A will make A's income equal to \$210 and B's income to \$90. Such a transfer spreads the distribution of income and worsens social welfare.

When $u_{xy} > 0$, things become more complicated. If A has more x than B then this tends to make A's marginal utility of x lower than that of B. But, if A also has a higher y, then this tends to make his marginal utility of x higher than that of B (because $u_{xy} > 0$). Therefore, we cannot a priori say that the marginal utility of x is higher for the rich than for the poor. However, when considering an allocation which is generated by an income tax function (i.e., an allocation of the form $[x_T(\cdot),\ y_T(\cdot)]$ for some T), it was shown by the author [7] that some transfer of income from the rich to the poor is still socially desirable according to (3).

(b). The Max-Min Criterion. When dealing with a discrete number of individuals this rule is very simple (Rawls[6]): when comparing two social states, choose that state in which the minimum utility enjoyed by anyone is higher; in case of a tie, go to the next to the minimum and so forth. . . . For example, if there are three individuals, as shown in Table 1, then state I is preferred to state II because the minimum utility in state I (namely, 4) is higher than the minimum utility in state II (namely, 3). States III and IV have the same minimum utility (namely, 7) but the next-to-min utility in state IV (namely, 11) is higher than the next-to-min in state III (namely, 10); therefore state IV is preferred to III.

TABLE 1

	Utility			
Individual	State I	State II	State III	State IV
A	5	3	7	11
B	10	20	10	7
C	4	300	400	15

A generalization of this definition to the case of a continuum of individuals is not difficult, in principle. We will not do this here because there is a technical difficulty that arises in the continuous case, due to the fact that the concept of the "next-to-min" is not always well-defined in this case. A rigorous definition is given in [8].

An important feature of both of these social ordering criteria is that they are individualistic, or in other words, comply with the Pareto Principle; namely, if some people prefer state A to B and all others are indifferent between the two states, then A is preferred to B according to either one of these criteria.

B. Preliminary Results

One cannot a priori determine whether the number of hours that an individual chooses to work will increase or decrease as his skill level increases. To see this, let us write down again the utility-maximization problem facing individual n: max $u(x,y)$, s.t.: $x = ny - T(ny)$. Compare this problem to the one discussed in length by Musgrave [5, ch. 11] of a single individual facing a wage rate of r: max $u(x,y)$, s.t.: $x = ry - T(ry)$. One can see at once that an increase in n in the first problem is equivalent to an increase in r in the second one. For the same reason that we cannot a priori determine what happens to the supply of labor of an individual when his skill level increases. In other words, we do not know whether $y_T(n)$ will increase or decrease as n increases.

Nevertheless, we can still say that gross income, namely $z_T(n) = ny_T(n)$, will increase (or, at least, will not fall) as n increases. Thus, even if $y_T(n)$ actually falls as n increases, this fall cannot be "very large" for the product $ny_T(n)$ must not fall. To prove this result, as well as the properties of the optimal tax, it is convenient to switch the analysis to the x-z plane from the x-y plane with which we have so far dealt.

In the x-y plane all have the same indifference curve map but face different budget lines (as in Figure 2). In the x-z plane all face the same budget line because all face the same tax schedule. Thus, if both individuals n_1 and n_2 earn an income of z, then they will both consume the same x (namely, $x = z - t(z)$). What will be different is their utility levels: individual n, has to work $y_1 = \frac{z}{n_1}$ hours to earn an income of z while individual n_2 has to work $y_2 = \frac{z}{n_2}$ hours to earn this same income; therefore n_1 enjoys a utility level of $u(x,y_1)$ while n_2 enjoys $u(x,y_2)$. In general, for person n_1 the point (x,z) in the x-z plane corresponds to the point $(x,\frac{z}{n})$ in the x-y plan. Given an indifference map in the x-y plane, therefore, the utility person n assigns to each point in the x-z plane can be determined (according to the relation $y = \frac{z}{n}$) and hence the map of his indifference curves between x and z. It is important to realize that this map of indifference curves in the x-z plane is merely a translation of the indifference curve map which holds in the x-y plane; we look upon z in the x-z plane as merely representing a certain number of hours that an individual has to work in order to earn z (namely, $y = \frac{z}{n}$ hours if his skill is n).

In the new plane, however, the indifference curve map differs for different individuals, since, as we have already seen, the same point in the x-z plane represents different utilities for different persons. The first thing to demonstrate about the relationship between the indifference curves of two different persons is that they become flatter as the skill increases:

<u>Lemma 1</u>: Let (x_o, z_o) be a point in the (x,z) plane and let $n_1 < n_2$ (see Figure 3(b)). The slope of the indifference curve of n_2

at the point (x_o, z_o) is then smaller than the slope of the indifference curve of n_1 at the same point.

Proof: The point (x_o, z_o) in the x-z plane (point A in Figure 3(b)) corresponds to point C in the x-y plane (Figure 3(a)) for person n_1 and to point B for person n_2. B is to the left of C since $y_2 = \frac{z_o}{n_2} < \frac{z_o}{n_1} = y_1$ (because $n_1 < n_2$). If we now add Δx to each person's consumption, then person n_1 and n_2 have to work Δy_1 and Δy_2 more hours, respectively, if their utilities are to remain constant. Thus to keep his utility constant, person n_1 has to earn (by working) an additional income of $\Delta z_1 = n_1 \Delta y_1$; similarly, $\Delta z_2 = n_2 \Delta y_2$ for person n_2. It follows from the normality-of-leisure assumption that the slope of the indifference curves increases as we move from left to right along any horizontal line in Figure 3(a). Therefore $\Delta y_2 > \Delta y_1$ and since $n_2 > n_1$, it follows that $\Delta z_2 = n_2 \Delta y_2 > n_1 \Delta y_1 = \Delta z_1$. Hence:

$$\frac{\Delta x}{\Delta z_2} < \frac{\Delta x}{\Delta z_1} . \qquad \text{Q.E.D.}$$

Two corollaries are immediate consequences of this lemma. Consider, again, two individuals, n_1 and n_2, with $n_1 < n_2$. Suppose that n_1 regards the point (x_1, z_1) to be at least as good as the point (x_o, z_o) and that $z_o < z_1$ (see Figure 4). Then, since person n_2's indifference curve which passes through (x_o, z_o) is flatter at that point than person n_1's indifference curve, it follows that n_2 actually prefers (x_1, z_1) to (x_o, z_o). This proves corollary 1 below.

Corollary 1: Suppose that individual n_1 regards point (x_1, z_1) to be at least as good as point (x_o, z_o) and that $z_o < z_1$. If $n_2 > n_1$, the individual n_2 prefers (x_1, z_1) to (x_o, z_o).

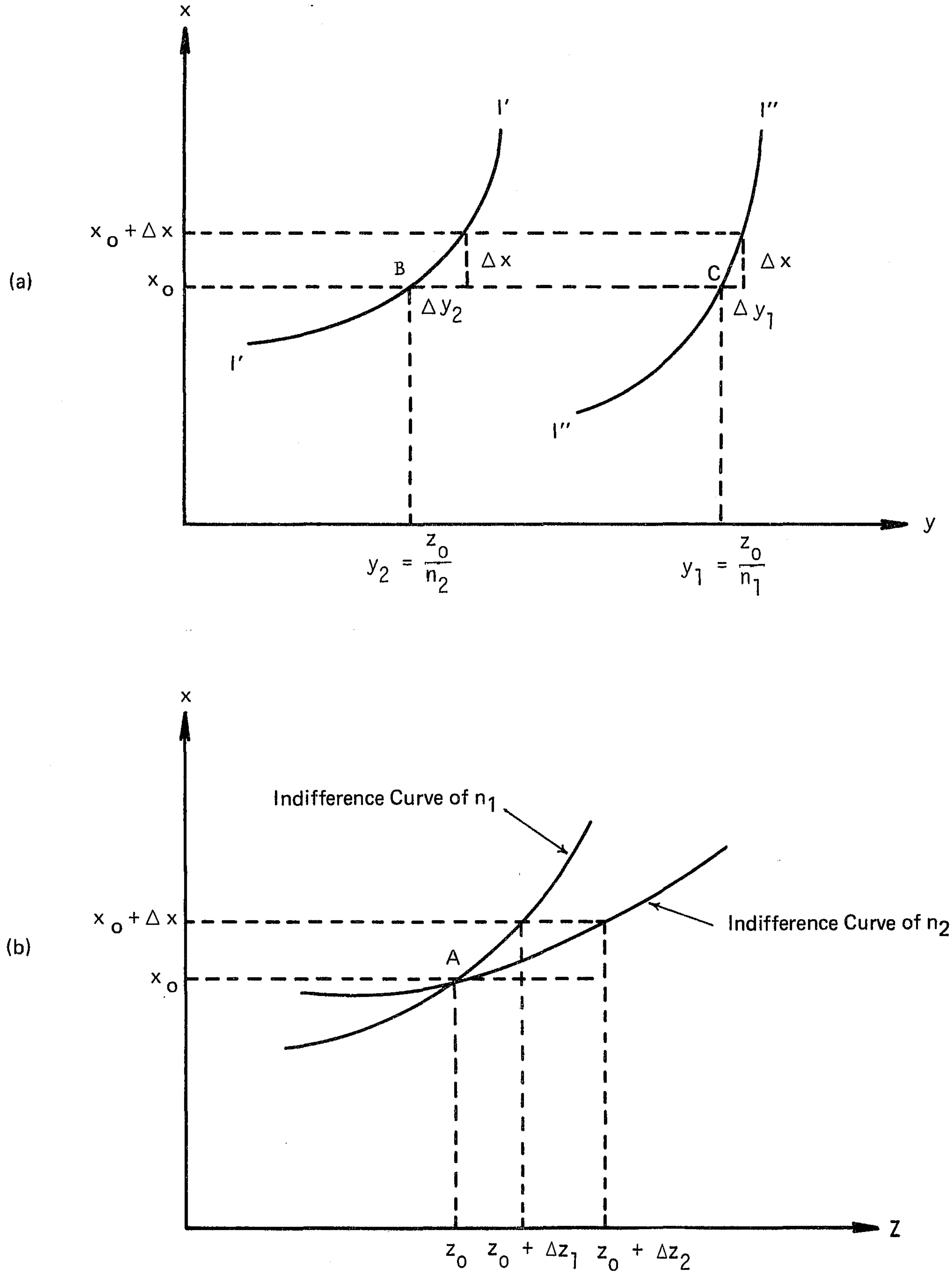
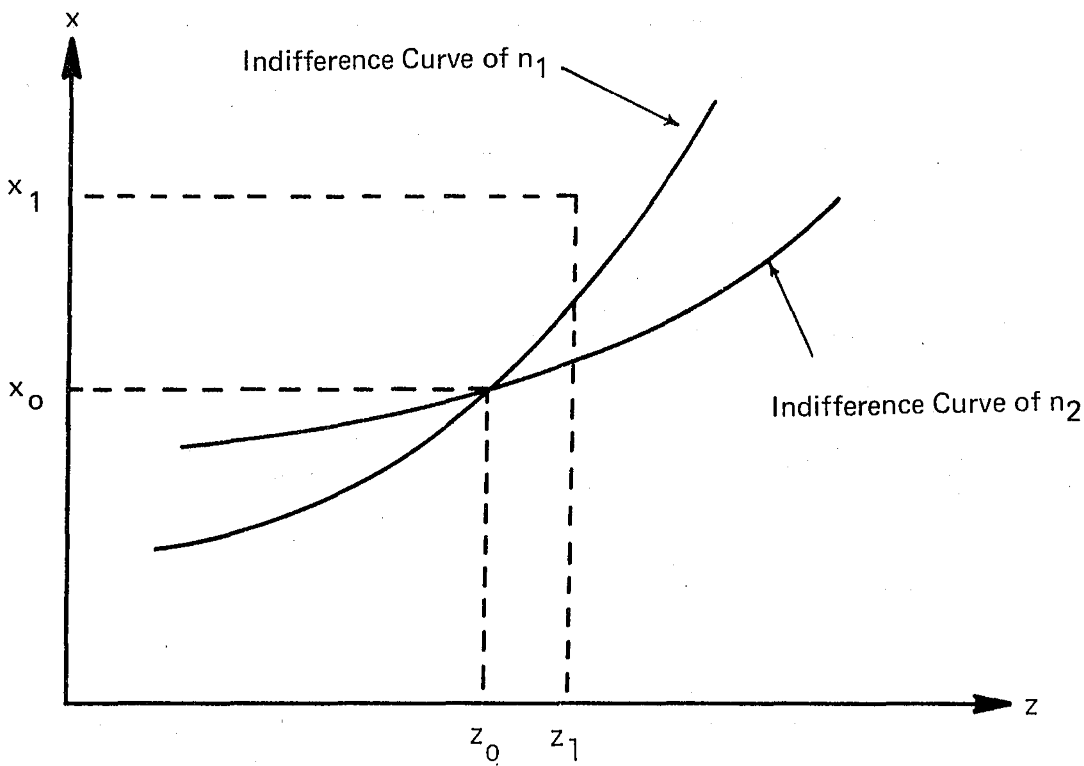

x
I'
I''
$x_o + \Delta x$
(a)
x_o
B
Δx
C
Δx
Δy_2
Δy_1
I'
I''
y
$y_2 = \frac{z_o}{n_2}$
$y_1 = \frac{z_o}{n_1}$
x
Indifference Curve of n_1
$x_o + \Delta x$
Indifference Curve of n_2
(b)
A
x_o
Z
z_o
$z_o + \Delta z_1$
$z_o + \Delta z_2$

Figure 3



x
Indifference Curve of n_1
x_1
x_o
Indifference Curve of n_2
z
z_0
z_1

Figure 4

To maximize his utility each person chooses that point in the x-z plane which lies on the highest indifference curve among all the points on the budget line. Thus, if we earlier defined $[x_T(n), y_T(n)]$ as person n's consumption-labor choice under a tax T, then $[x_T(n), z_T(n)]$, where $z_T(n) = ny_T(n)$, will be his consumption-income choice under the same tax. Now, since person n_1 chooses the bundle $[x_T(n_1), z_T(n_1)]$, it follows, by definition, that he considers this point to be at least as good as any other point on the budget line. In particular, $[x_T(n_1), z_T(n_1)]$ is at least as good, from his point of view, as any point on the budget line with a lower z. By corollary 1, person n_2 $(n_2 > n_1)$ actually prefers $[x_T(n_1), z_T(n_1)]$ over all other points on the budget line which involve a lower z. Therefore his consumption-income choice, $[x_T(n_2), z_T(n_2)]$, must not have a lower z: $z_T(n_2) \geqq z_T(n_1)$. This proves what we have stated at the beginning of this section:

Corollary 2: (Mirrlees [4]): For each T, $z_T(n)$ is nondecreasing in n (see Figure 5).

It is obvious from the skill-earning pattern of this model that a higher skill entitles the individual owning it to a higher utility level (under any tax). This is stated in the next lemma:

Lemma 2: For each T, $u_T(n)$ is increasing in n.

In this model working is unpleasant. Therefore no one will ever increase his work effort to the point where such an increase leaves him with a lower net income. In other words, no one will ever choose to be in an income bracket where the marginal tax rate exceeds 100 percent. Thus, if the curve ABCDE in Figure 6 is the graph of $z - T(z)$, then no one will choose to be in the income bracket between z_1 and z_2, where

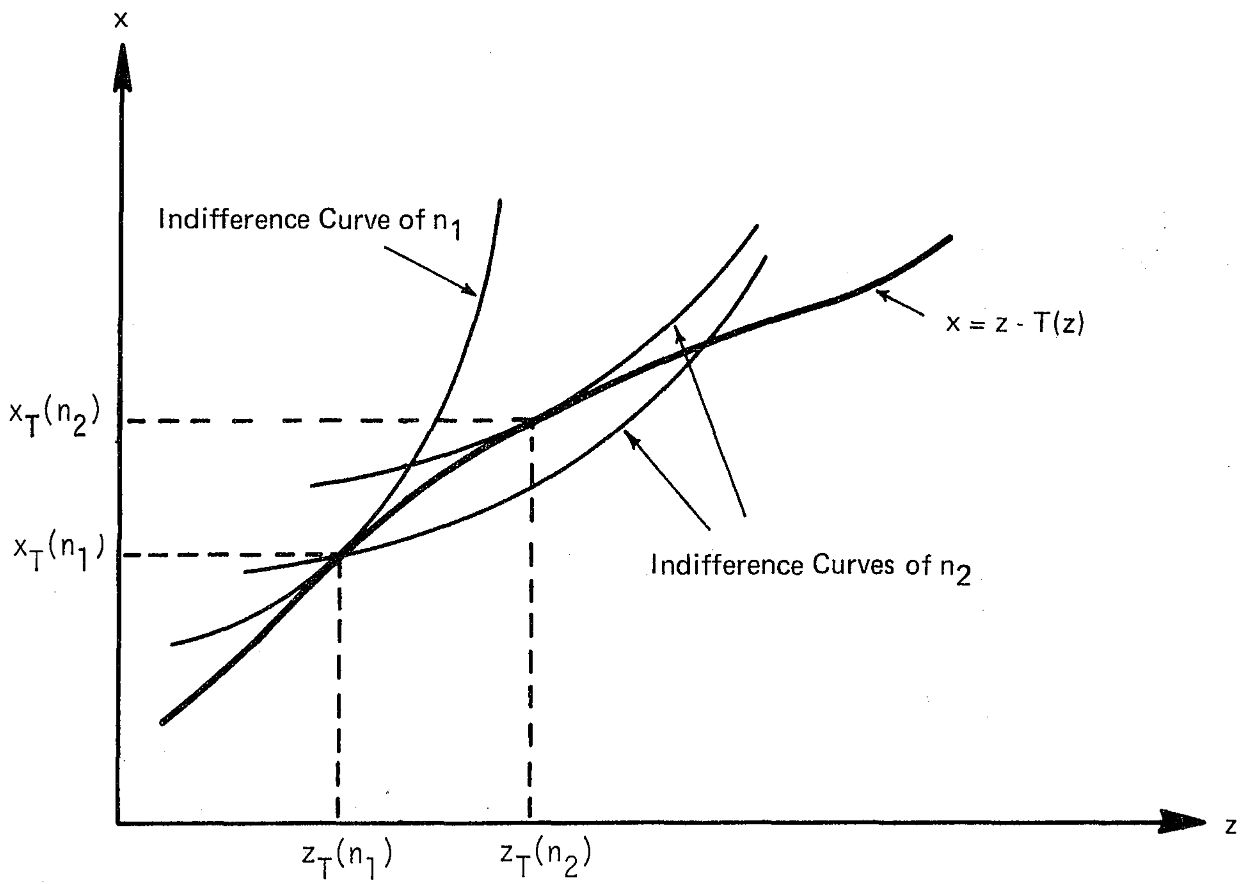


Figure 5: $n_1 < n_2$

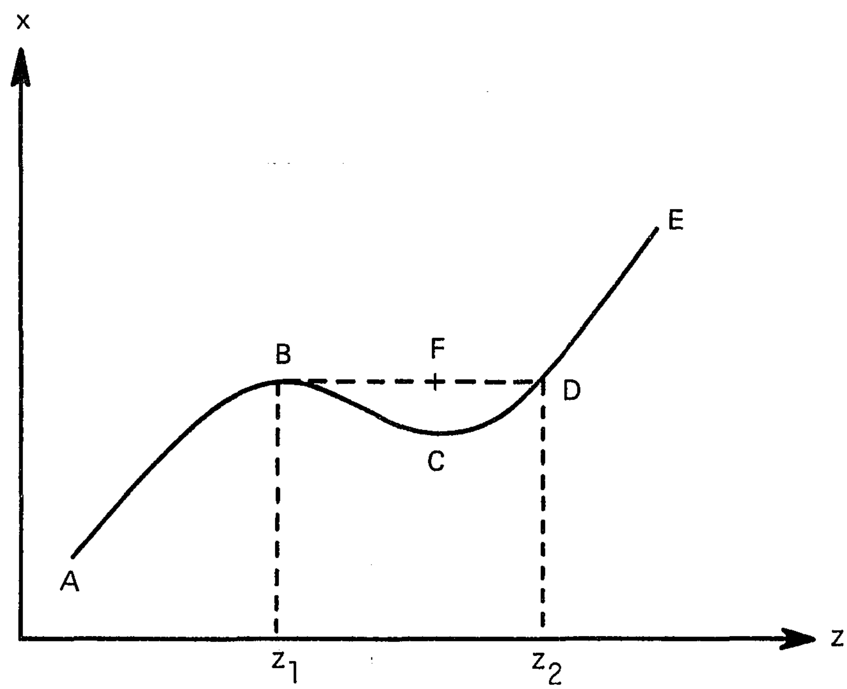


Figure 6

the marginal tax rate exceeds 100 percent. With no loss of generality, we may therefore assume that the graph of $z - T(z)$ is ABFDE (the segment BFD replacing BCD). This means that, with no loss of generality, we may assume that the marginal tax rate does not exceed 100 percent anywhere, that is, $z - T(z)$ is nondecreasing in z (see also Mirrlees [4]). Thus, that the marginal tax rate cannot exceed 100 percent is not a property of an optimal tax but is rather an immediate consequence of consumer sovereignty.

With these results at hand, we are now ready to investigate the properties of the optimal tax.

C. The Monotonicity of the Optimal Tax

The central property of the optimal tax under either social ordering criterion is its monotonicity. Specifically, the optimal tax is nondecreasing with income which means that the marginal rate is nonnegative. We shall emphasize here again that it is only a negative marginal tax rate that is excluded but not a negative tax. As we shall see in the next section, a negative income tax such as the tax T_1 which is drawn in Figure 7 might well be optimal. Under this tax, all people with gross incomes below some level (namely, $\bar{z}$) pay a negative tax (that is, receive payment from the government). What is excluded is a tax such as T_2 in Figure 7 which is falling with income, that is, the rich pay less than the poor.

The nonnegativity of the marginal rate of the optimal tax was first proved by Mirrlees [4] for the case of an additive social welfare function. Here we shall sketch a simple proof by showing that a tax which does not possess this property can be improved.

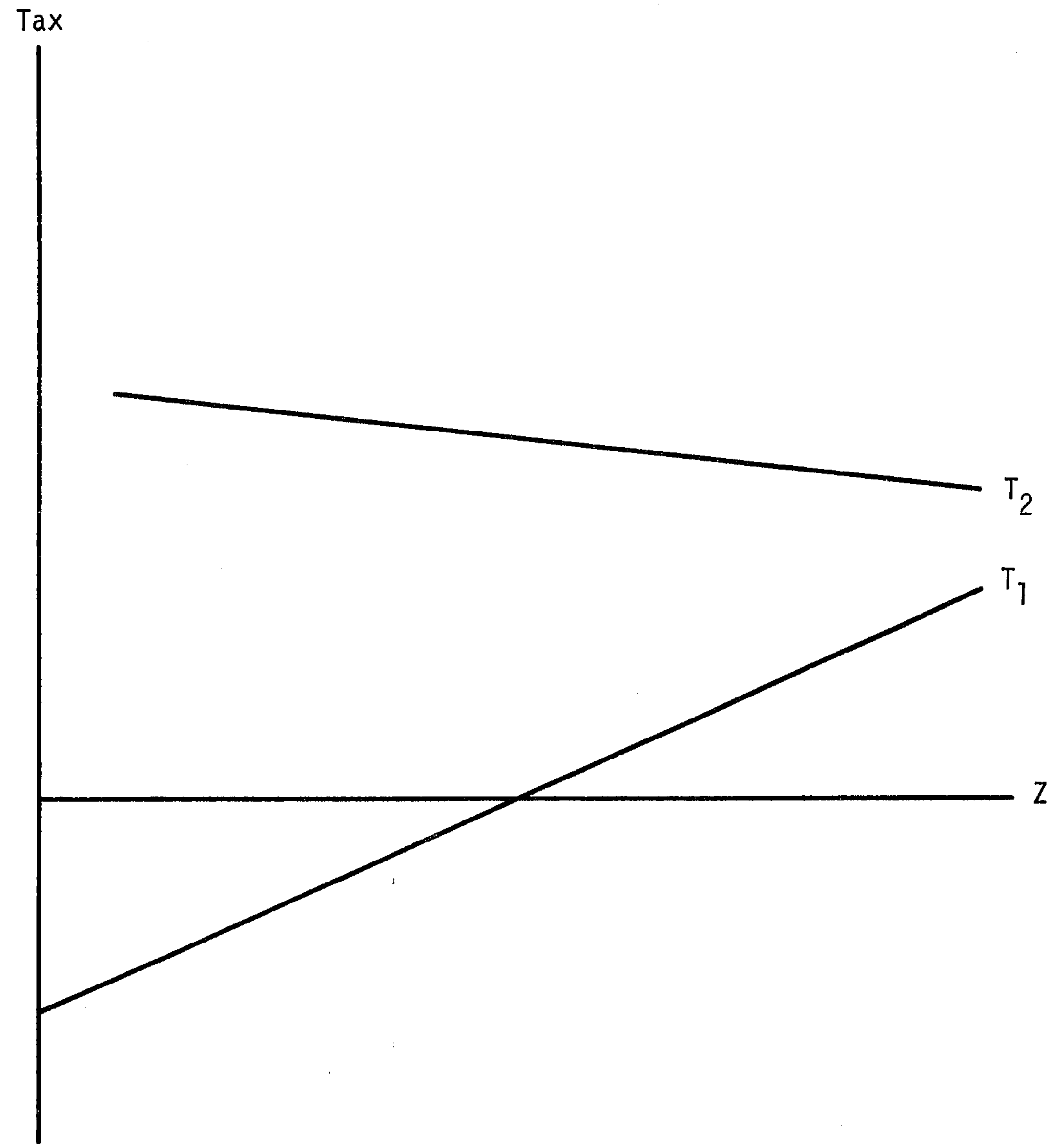


Figure 7

Theorem 1: If $T(z)$ is optimal under an additive social welfare function, then T is nondecreasing in z, that is, the marginal tax rate is nonnegative.

A sketch of a proof. The form of the proof will be to show that for every T for which the marginal tax rate is negative somewhere there exists a T_1 which is nondecreasing in z, is feasible in the sense that it yields the required revenue (see Equation (2)), and is socially preferred to T. To facilitate presentation of the argument it will be assumed that T is linear,[10] but this assumption is not necessary to the proof. In Figure 8, let ASBPC be the graph of $z - T(z)$. The slope of ASBPC is greater than one, corresponding to a negative marginal tax rate.[11] We will now construct another tax, T_1, which is feasible (that is, satisfies (2)) and which is socially preferred to T.

The line ERBMD is the graph of a net income function corresponding to some tax function, T_1 (namely, ERBMD is the graph of $z - T_1(z)$). The essential characteristic of T_1 is that it has a zero marginal rate everywhere. Hence the slope of ERBMD is unity. If we make the assumption for the moment that if T_1 were to replace T, individuals would not alter their labor supplies and hence their incomes, an individual who was at S would be moved to R and an individual at P would be moved to M. That is, the new tax increases the burden on the rich and reduces the burden on the poor. Since, by assumption, gross income is unchanged, the intercept of ERBMD with the vertical axis can be chosen in a way that insures that T_1 collects as much revenue as T: what is lost by the rich is gained by the poor. Since, as explained earlier, a transfer of consumption from the rich to the poor, total income unchanged, improves social welfare, T_1 is socially preferred to T.

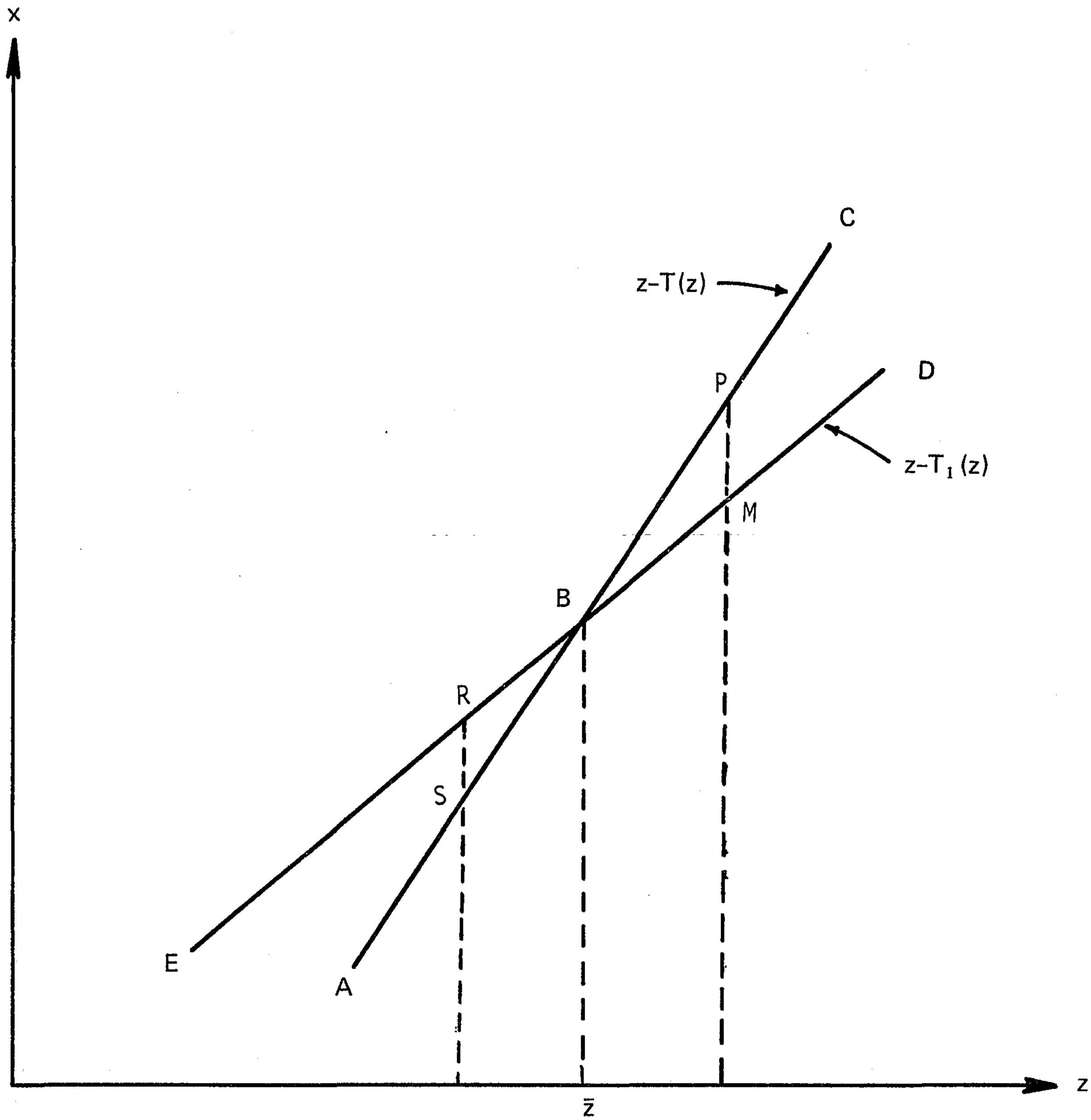


Figure 8

We now relax the assumption that persons do not adjust their incomes in the new tax regimes. Each person responds to the new tax by moving to the point on ERBMD which maximizes his utility. Social welfare is thus increased yet further, but the question now arises as to whether or not tax revenues remain unchanged. This is indeed the case since the marginal tax rate along ERBMD is zero, and hence as individuals readjust their labor supplies and hence incomes, taxes paid are unaffected. The theorem is thus proved.

The structure of the proof of theorem 1 can be used to establish a similar theorem for the max-min case. If T has a negative marginal rate somewhere, we can again construct the tax T_1 of theorem 1 which is feasible; and, by its construction (see Figure 8), all people who had gross incomes below $\bar{z}$ are better off under the new tax. This proves theorem 2 below.

Theorem 2: If T(z) is optimal under the max-min criterion, then it is nondecreasing in z (that is, nonnegative marginal tax rates).[12]

Theorem 2, which was obtained as a by-product of theorem 1, states that the marginal rate of the optimal tax under the max-min criterion should not be negative. But it is possible to prove a still stronger result, namely, that the marginal rate should be positive. To prove this result we will need lemma 3 below which we now explain. Lemma 2 states that individual N_1 is always (that is, under any tax) going to be the least well-off in the society. Under the max-min criterion, we should therefore enhance his well-being as much as possible. For this purpose we should maximize the tax payments of the rest of the society, so that he could pay as little tax as possible (or receive as much welfare as possible). Therefore, if T is an optimal tax under the max-min criterion and T_1 is another tax which is equal to T at the income level of person N_1 (that is, $T_1[z_T(N_1)] = T[z_T(N_1)]$), then T_1 cannot collect more revenues than T; for otherwise, these extra revenues can be used to lower the tax burden borne by N_1 (or increase the welfare payments received by him). This proves lemma 3.

Lemma 3: Let T be optimal under the max-min criterion and let T_1 be another tax such that $T_1[z_T(N_1)] = T(z_T(N_1)]$. Then T_1 cannot collect more revenues than T.[13]

Theorem 3: Let T(z) be optimal under the max-min criterion. Then T is strictly increasing in z, that is, the marginal tax rate is positive everywhere.[14]

Proof: We will present here a proof which shows how, under the max-min criterion, one can improve any tax which is not strictly increasing. Suppose that T is not strictly increasing: there then exist z_1 and z_2 such that $z_1 < z_2$ and $T(z_1) \geqq T(z_2)$. Since $T(z_1) > T(z_2)$ is excluded by theorem 2, it must be the case that $T(z_1) = T(z_2)$. Then if z is between z_1 and z_2, we must have $T(z_1) = T(z) = T(z_2)$, for all other possibilities are, again, excluded by theorem 2. Hence T is constant and, consequently, $T' = 0$ on the interval between z_1 and z_2. Thus $c' = 1$ on this interval, where $c(z) = z - T(z)$. Let ABCDEHK in Figure 9 be the graph of c; the slope of BCD being 1. Choose some z_3 between z_1 and z_2 and let n_1, n_3 and n_2 be, respectively, the skills of persons for whom the points B, C and D are utility-maximizing bundles under T (in other words, $z_1 = z_T(n_1)$, $z_3 = z_3 = z_T(n_3)$ and $z_2 = z_T(n_2)$). Corollary 2 then implies that $n_1 \leqq n_3 \leqq n_2$, but the strictly convex indifference curves imply that $n_1 < n_3 < n_2$. Let QBREJ be the indifference curve of person n_3 which passes through B. It follows from the

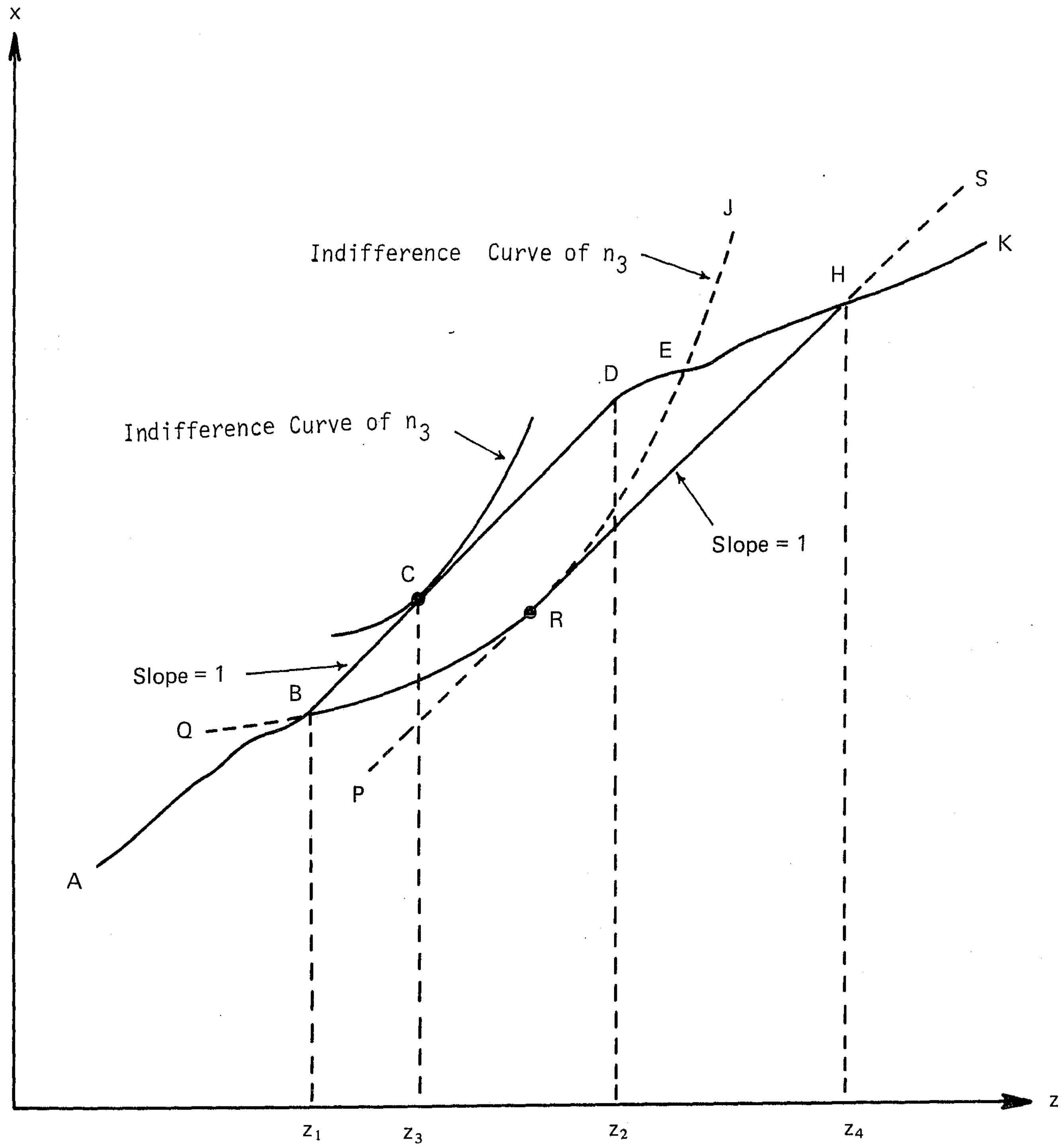

x
S
J
K
Indifference Curve of n_3
H
E
D
Indifference Curve of n_3
Slope = 1
C
R
Slope = 1
B
Q
P
A
z
z_1
z_3
z_2
z_4

Figure 9

normality assumption that the slope of this indifference curve is equal to 1 at some point, say R, to the southeast of C. At point R draw a line PRHS which is tangent to the indifference curve QBREJ and has therefore a unity slope. Let n_4 be the skill of a person for whom point H is a utility-maximizing bundle under T (i.e., $z_4 = z_T(n_4)$). Again, corollary 2 implies that $n_2 \leqq n_4$.

Now define a new tax T_1 such that the graph of its corresponding net income function, $e_1(z) = z - T_1(z)$, is ABRHK. By revealed preferences, we can certainly conclude that no one with skill below n_1 or above n_4 will change his position under the new tax. In particular, points B and H remain utility-maximizing bundles under T_1 for n_2 and n_4, respectively. Therefore, all persons with skills between n_2 and n_4, inclusive, will find their equilibrium positions under T_1 along the curve BRH (by corollary 2). Since, by construction, person n_3 considers point B to be at least as good as any other point along BRH, it follows from corollary 1 that all persons with skills between n_1 and n_3 actually prefer B over any other point along BRH. Therefore, they will choose B under T_1 (that is, $z_1 = z_{T_1}(n)$ for all n between n_1 and n_3). Since the slope of BC is 1, it follows that their tax payments do not change (any point along BC results in the same tax payment). Since n_3 is indifferent between R and B, it follows that all persons with skills between n_3 and n_4 prefer R over B (by corollary 1). They therefore choose their equilibrium positions under T_1 along the line segment RH, increasing their tax payments beyond what they paid under T. Thus, no one decreases his tax payment under T_1 while some increase their tax payments. This means that T_1 collects more revenues than T, contradicting lemma 3. Q.E.D.

D. The Case for a Negative Income Tax

It is an interesting question whether a negative income tax can be optimal. Suppose first that public consumption, B, is zero, so that the government needs to raise no revenue and will therefore use income taxation for a redistribution of income only. In this case the optimal tax (under either social ordering criterion) cannot be positive everywhere, for then the government collects a positive amount of revenues which is more than it needs (namely, zero). A tax which is everywhere negative is not feasible because it yields a negative amount of revenues. Sheshinski [10] showed that some taxation is preferred to no taxation at all in the case of an additive social welfare function. Theorem 3 implies that some taxation is preferred to no taxation at all in the max-min case too (because no taxation at all implies a zero marginal tax rate everywhere). Thus, neither $T > 0$ everywhere, nor $T = 0$ everywhere is optimal, nor $T < 0$ everywhere is feasible. Thus, the optimal tax (under either criterion of social welfare) has to be positive somewhere and negative somewhere else. Since the optimal tax must not be decreasing, it therefore has to be negative at the lowest income brackets and positive at the highest. Under the additive social welfare function, it is possible for the optimal tax to be just zero at a middle income bracket (containing more than one point), but this is excluded by theorem 3 in the max-min case (since the marginal tax rate will be zero in that bracket). We have thus proven the following theorem:

Theorem 4: Suppose that public consumption, B, is zero and let T be optimal under the additive social welfare function. There then exist $\bar{z}$ and z^* such that $z_T(N_1) < \bar{z} \leqq z^* < z_T(N_2)$ and such that

$$T(z) \text{ is } \begin{cases} < 0 & \text{for } z < \bar{z} \\ = 0 & \text{for } \bar{z} \leqq z \leqq z^* \\ > 0 & \text{for } z > z^* \end{cases}$$

With the addition that $\bar{z} = z^*$ (that is, the tax is zero at only one point), this characterization of the optimal tax is also valid under the max-min criterion.

When public consumption is not zero, the question whether a negative income tax is optimal or not depends on the magnitude of public consumption[15] and skills. With little public consumption it will presumably be still optimal to have a negative income tax. When $N_1 = 0$, which means that N_1 can afford to buy some x only if he receives welfare payment, an optimal tax under the max-min criterion will be negative at the lowest income bracket. With some further assumptions (for instance, that $u \to -\infty$ as $x \to 0$), the optimal tax will be negative at the lowest income bracket under the additive social welfare function too. When both B and N_1 are positive and large in magnitude, we cannot a priori say whether a negative income tax is optimal or not.[16]

E. The Progressivity of the Optimal Tax

Two alternative, nonequivalent, definitions of progressivity or regressivity of an income tax coexist in the literature.

Definition I: An income tax is said to be progressive over some interval of gross incomes if its marginal rate is increasing in that interval. Similarly, an income tax is said to be regressive over some interval if its marginal rate is decreasing. Thus, according to this definition a progressive (respectively, regressive) tax function is convex (respectively, concave) while the corresponding net income function is concave (respectively, convex). Examples of progressive taxes, according to this definition, are given in Figures 10(a) and 10(b); regression taxes are shown in Figures 10(c) and 10(d).

Definition II: An income tax is said to be progressive (respectively, regressive) over some interval of gross income if its average rate is increasing (respectively, decreasing) over that interval. Thus T_1 of Figure 11 is progressive according to the latter definition, while T_2 is regressive.

Consider first definition I. Since both of our social ordering relations in some sense favor equality over inequality in the distribution of (net) income, we may, at first thought, expect the optimal tax to be progressive under both social ordering criteria. A second look, however, will show that there is no clear relationship between the desire for an equal distribution of income and the progressivity of the income tax. A society which prefers equality would like to see after-tax income (that is, net income or consumption) distributed as equally as possible. In other words, such a society favors a narrowing of the gap between $x_T(N_2) = z_T(N_2) - T[z_T(N_2)]$ and $x_T(N_1) = z_T(N_1) - T[z_T(N_1)]$. Roughly speaking, this means that the net income schedule should be as flat as possible. A flat net income schedule would result from high marginal tax rates. The term "progressivity," however, does not refer to whether the marginal rates are high or low, but rather to how they change with changes in gross income (namely, whether they increase or decrease). Therefore, there is no reason to expect that the social desire for equality will necessarily be reflected in a progressive optimal tax.

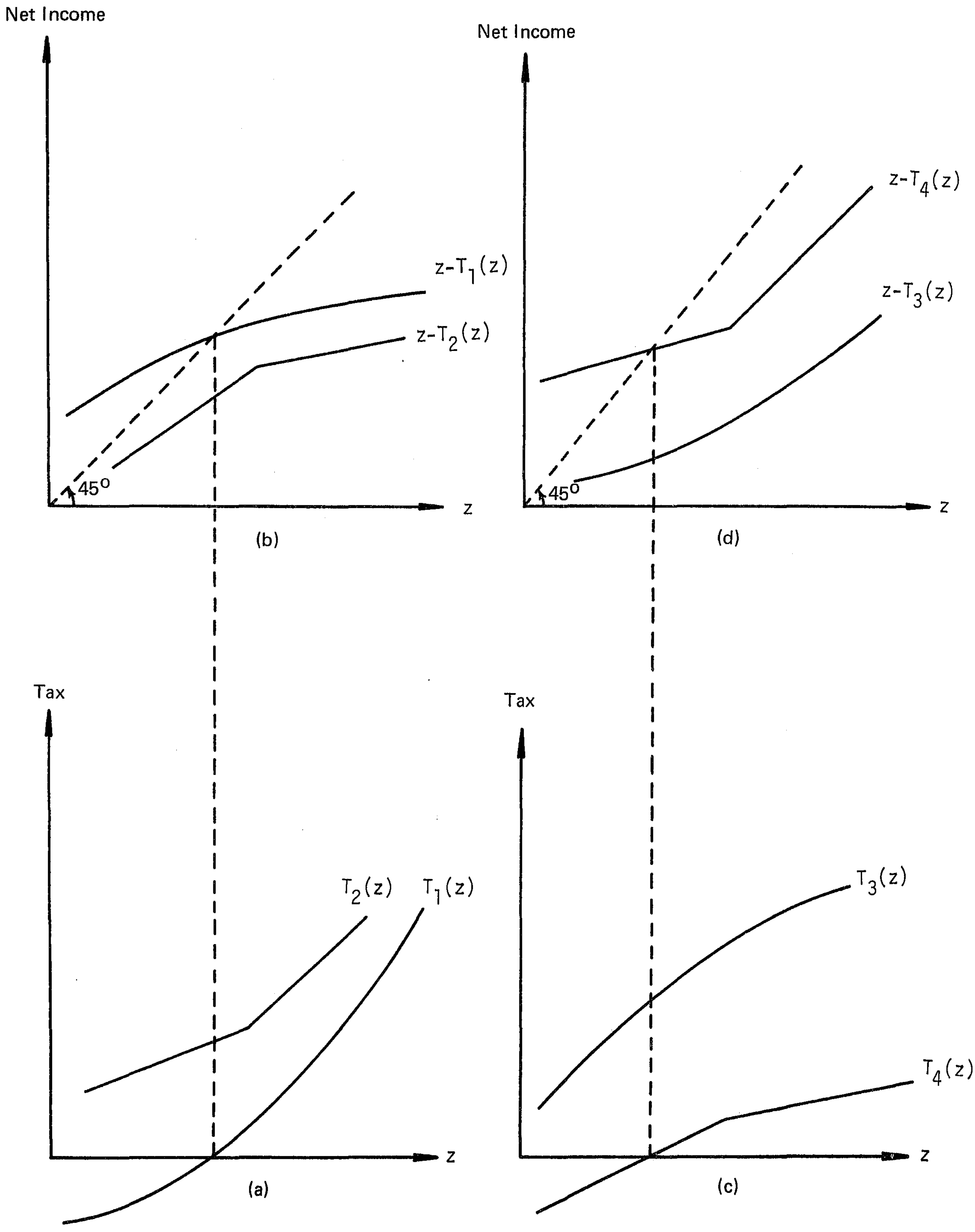

Net Income
z-T1(z)
z-T2(z)
45°
z
(b)
Net Income
z-T4(z)
z-T3(z)
45°
z
(d)
Tax
T2(z)
T1(z)
z
(a)
Tax
T3(z)
T4(z)
z
(c)

Figure 10

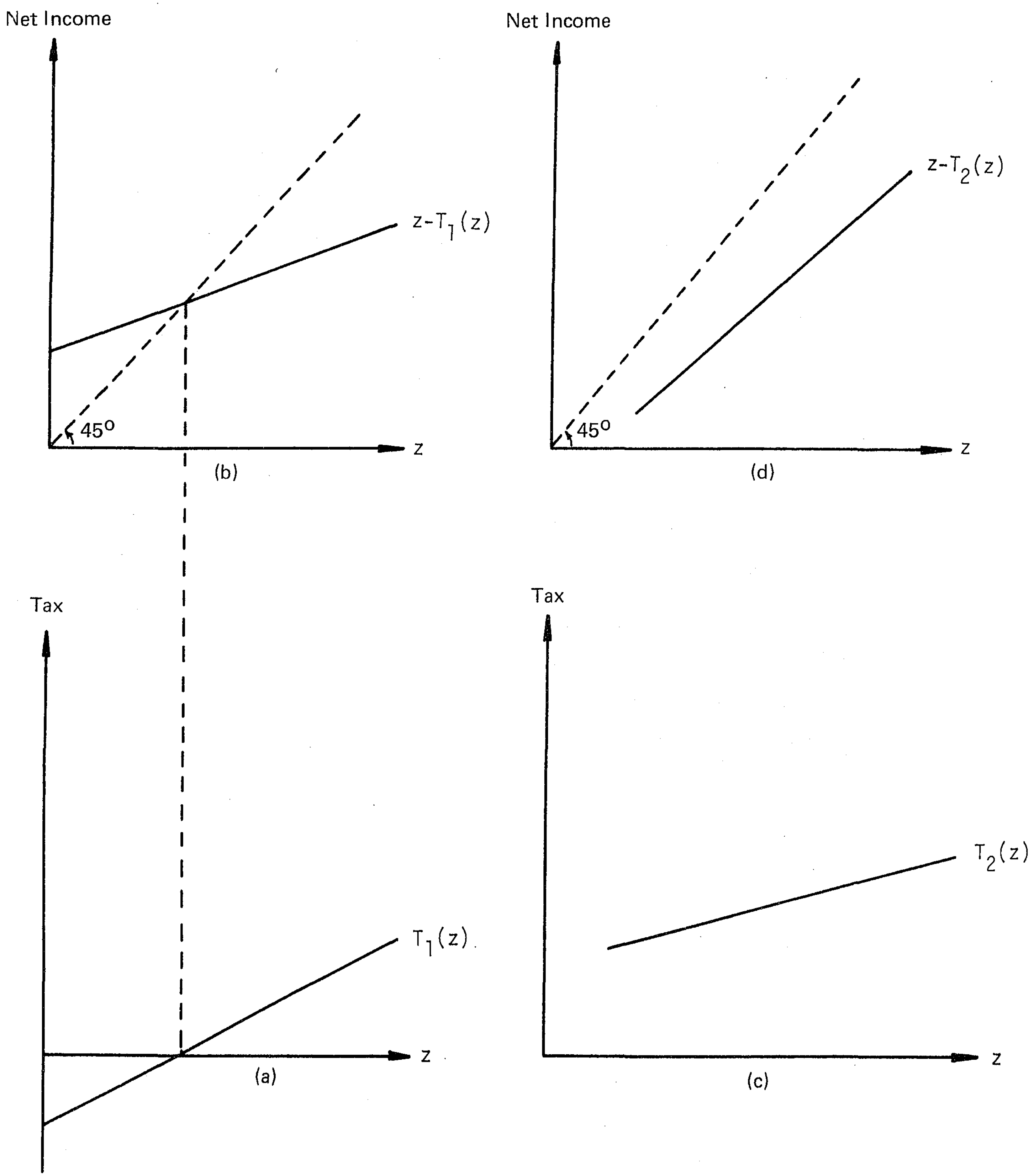

Net Income
z-T1(z)
45o
z
(b)
Net Income
z-T2(z)
45o
z
(d)
Tax
T1(z)
z
(a)
Tax
T2(z)
z
(c)

Figure 11

A diagram (Figure 12) will clarify the argument. The convex curve SKR is the graph of a net income function corresponding to some regressive tax T_1 (that is, SKR in the graph of $z - T_1(z)$). Suppose, for the sake of the argument, that individual labor supplies are independent of the tax function. Now draw a concave curve EKM which may be thought of as the graph of a net income function corresponding to a progressive tax T_2. We can certainly draw EKM in such a way that T_2 will generate the same revenues to the government as T_1. However, the graph of $z - T_1(z)$ (namely, SKR) is "flatter" than the graph of $z - T_2(z)$ (namely, EKM). In fact, as compared to the regressive tax T_1, the progressive tax T_2 amounts to transferring consumption from the poor to the rich!

Theorem 5 and corollary 3 to be developed below actually exclude the possibility of having an income tax which is progressive everywhere.

Theorem 5: Let T be optimal under either one of our social ordering relations. Then $T' = 0$ at $z_T(N_2)$. In other words, the marginal rate applicable to the richest person in the society must be zero.

Proof: Theorems 1 and 2 imply that T' must not be negative anywhere. Hence, $T'[z_T(N_2)] \geqq 0$. Therefore, all we have to show is that $T'[z_T(N_2)]$ is not positive. Suppose, to the contrary, that $T'[z_T(N_2)] > 0$. In this case it is possible to construct another tax T_1 which is socially preferred over T, a contradiction. Let T_1 be as follows:

$$T_1(z) = \begin{cases} T(z) \text{ for } z \leqq z_T(N_2) \\ T[z_T(N_2)] \text{ for } z \geqq z_T(N_2) \end{cases}$$

In other words, T_1 coincides with T up to an income level of $z_T(N_2)$;

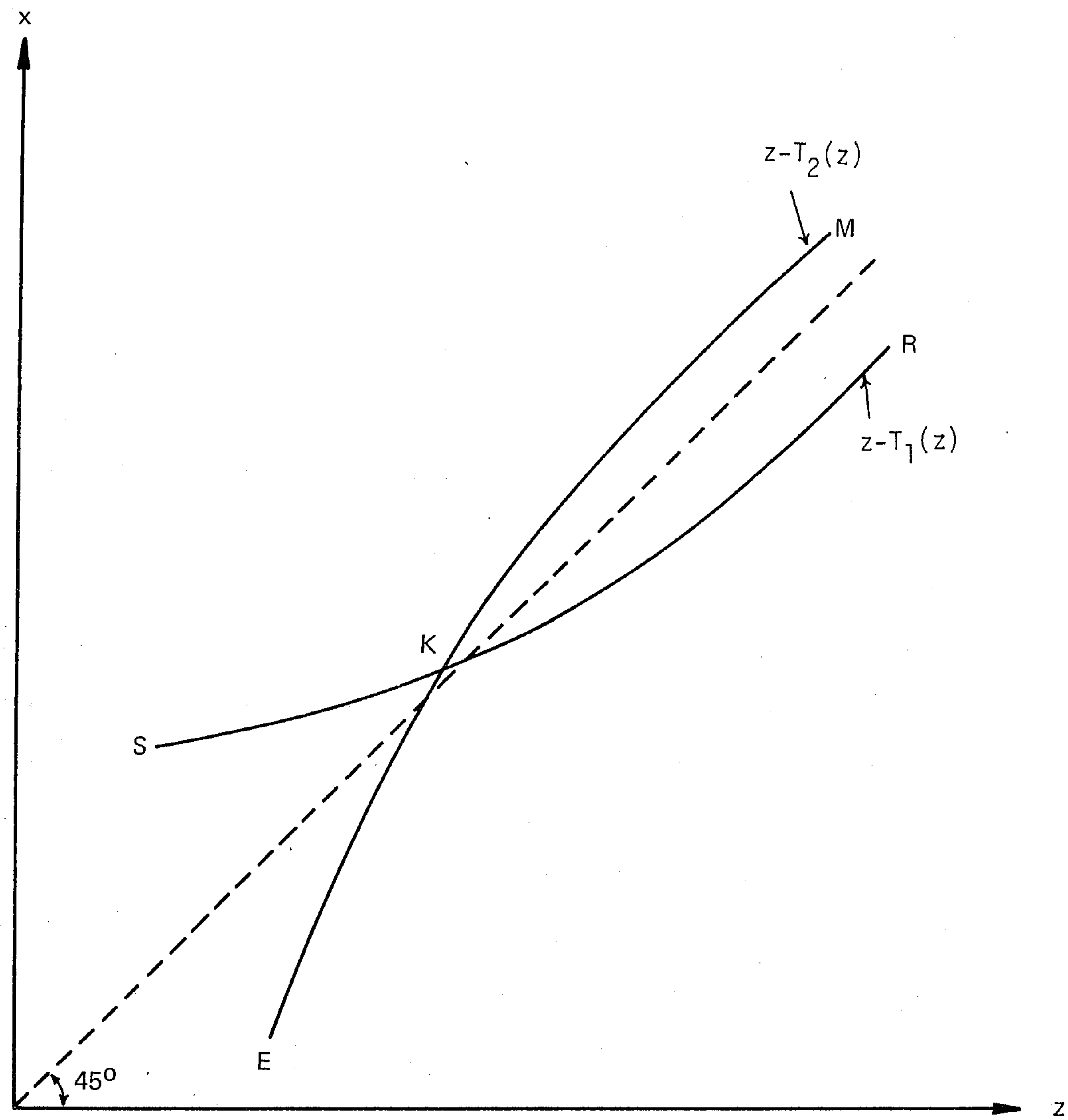


Figure 12

beyond $z_T(N_2)$, the marginal rate of T_1 becomes 0 (T_1 becomes constant at the amount of tax paid under T by individual N_2). Let ABC in Figure 13 be the graph of $z - T(z)$; the slope at B is less than 1 (because $T' > 0$ at that point). ABKD is then the graph of $z - T_1(z)$, the slope of BKD being 1.

It is clear from Figure 13 that individual N_2 will not stay at B under T_1 but rather will move to K. (He can do so because, by (1), $y_T(N_2) < A$ which means that an increase in his work effort is possible.) Clearly, N_2 is better-off under T_1. Clearly, no one is worse-off under T_1. Thus, some individual (namely, N_2) prefers T_1 and no one prefers T.[17] Therefore, T_1 must be socially preferred to T according to any individualistic social ordering criterion (including the two criteria considered here).

It remains to be shown that T_1 is feasible. Anyone who changes his equilibrium position under T_1 will do that by moving to some point along BKD; for all other points were available under T, too. By theorems 1 and 2, point B involves at least as much tax payment as any other point along AB, while, by the construction of T_1, all points along BKD result in the same tax payment (because the slope of BKD is 1). Therefore, any point along BKD results in at least as much tax payment as any point along AB. This proves that T_1 collects at least as much tax revenues as T and is therefore feasible. Q.E.D.

The last theorem states that the marginal rate of the optimal tax must be zero at $z_T(N_2)$. Therefore, the marginal rate will have to be negative at income levels below $z_T(N_2)$, if it is to be increasing in the vicinity of $z_T(N_2)$. But this is impossible by theorems 1 and 2, which establishes the next corollary:

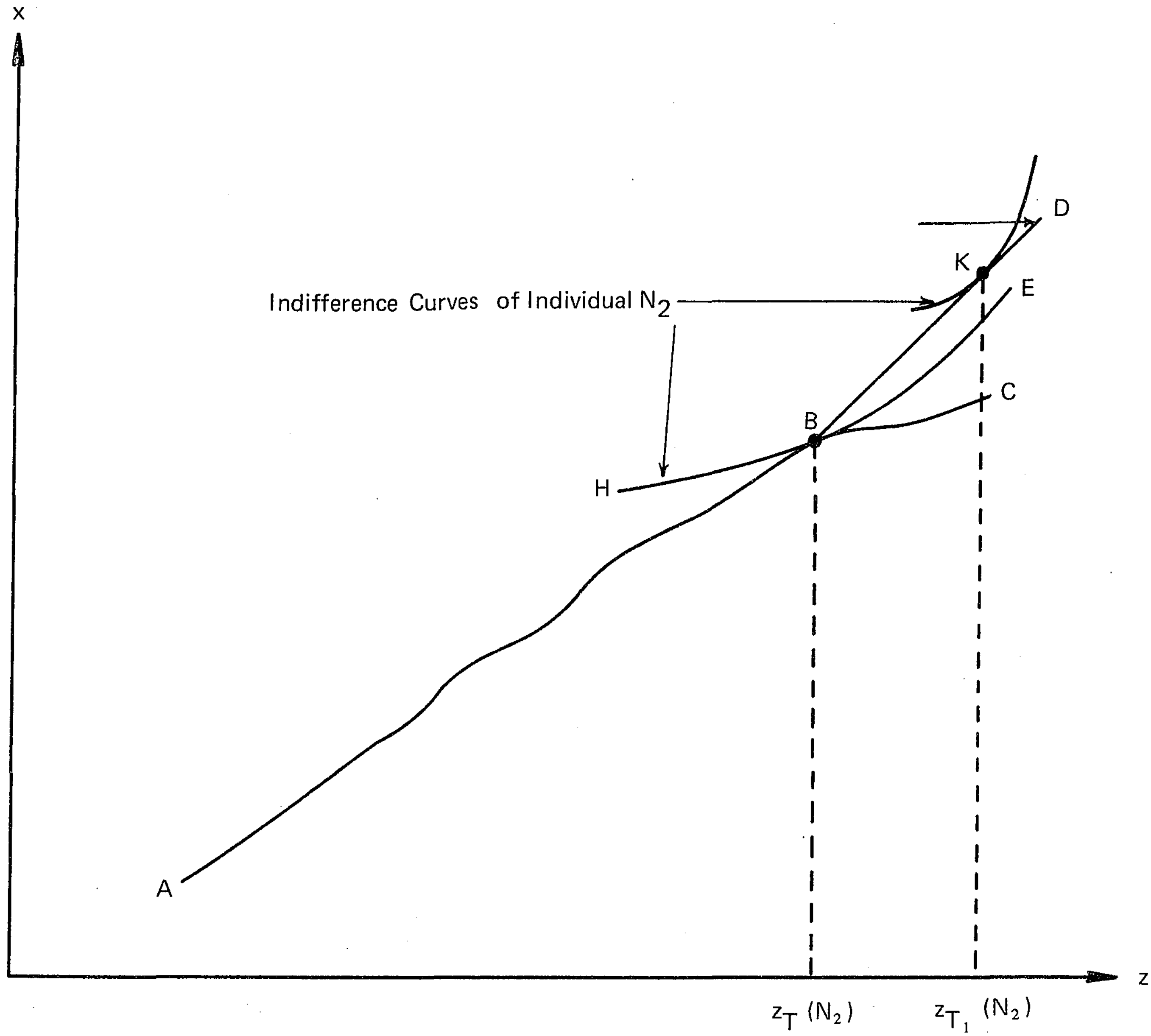

x
D
K
E
Indifference Curves of Individual N_2
C
B
H
A
z
z_T (N_2)
z_{T_1} (N_2)

Figure 13

Corollary 3: Under either one of our social ordering relations, the optimal tax cannot be progressive everywhere according to definition I of progressivity.

Let us now turn to definition II of progressivity. The relationship between any marginal and average values is known to be such that the average value rises whenever the marginal value is greater than the average value.[18] Therefore, a tax is progressive, according to definition II, if $T' > T/z$. We know from theorem 5 that the optimal tax, T, has a zero marginal rate at $z_T(N_2)$ while the tax itself is positive at that point (see footnote 16). Therefore the marginal tax rate (which is zero) is lower than the average tax rate (which is positive) at the income of $z_T(N_2)$. This implies that the optimal tax (under either one of the two ordering criteria) is regressive at the highest income bracket according to definition II of progressivity.

Corollary 4: Suppose that T is optimal under any one of our social ordering relations. Then T is regressive at the vicinity of $z_T(N_2)$ according to definition II of progressivity. If, in addition, $B = 0$ (that is, zero public consumptions), then T is progressive on the interval from $z_T(N_2)$ to $\bar{z}$, where $\bar{z}$ is defined in theorem 4.

Proof: The first part of the corollary has been already established. When $B = 0$, we know from theorem 4 that T must be negative on the interval from $z_T(N_2)$ to $\bar{z}$, while $T' \geqq 0$ by theorems 1 and 2. Thus, the marginal tax rate (which is nonnegative) exceeds the average tax rate (which is negative) on the interval from $z_T(N_1)$ to $\bar{z}$, implying progressivity. Q.E.D.

F. A Comparison Between the Additive and the Max-Min Criterion

Throughout this section we will denote by T_1 and T_2 the optimal taxes under, respectively, the additive social welfare function and the max-min criterion. The first thing to observe about the relationship between T_1 and T_2 is that individual N_1 must be at least as well-off under T_2 as under T_1; for if he is better-off under T_1, then the latter tax would be the optimal tax under the max-min criterion and not T_2. This proves part (a) of theorem 6. The second part states that if T_1 and T_2 are both linear, then T_2 must have a lower intersect ($z - T_2(z)$ must have a higher intersect) and a higher marginal rate.

Theorem 6: Let T_1 and T_2 be optimal under, respectively, the additive social welfare function and the max-min criterion. Then: (a) $u_{T_1}(N_1) = u_{T_2}(N_1)$; (b) if both T_1 and T_2 are linear and we write them as $T_1(z) = -a_1 + b_1 z$ and $T_2(z) = -a_2 + b_2 z$, then $a_1 \leqq a_2$ and $b_1 \leqq b_2$.

Proof: (a) was already proved. To prove part (b) let us first suppose that $N_1 = 0$. In this case individual N_1 will never go to work, because he is incapable of producing any positive income. Thus, $y_{T_1}(N_1) = y_{T_2}(N_1) = z_{T_1}(N_1) = z_{T_2}(N_1) = 0$. Consumption by N_1 is therefore equal to his net income: $x_{T_1}(N_1) = z_{T_1}(N_1) - T[z_{T_1}(N_1) = z_{T_1}(N_1) - [-a_1 + b_1 z_{T_1}(N_1)] = a_1$ and likewise $x_{T_2}(N_1) = a_2$. Now, the only way in which person N_1 is going to be at least as well-off under T_2 as under T_1 is when his consumption is at least as high under T_2 as under T_1 (under both taxes N_1 does not work). This proves that $a_1 \leqq a_2$. The graph of $z - T_2(z) = a_2 + (1-b_2)z$ intersects the vertical axis at a

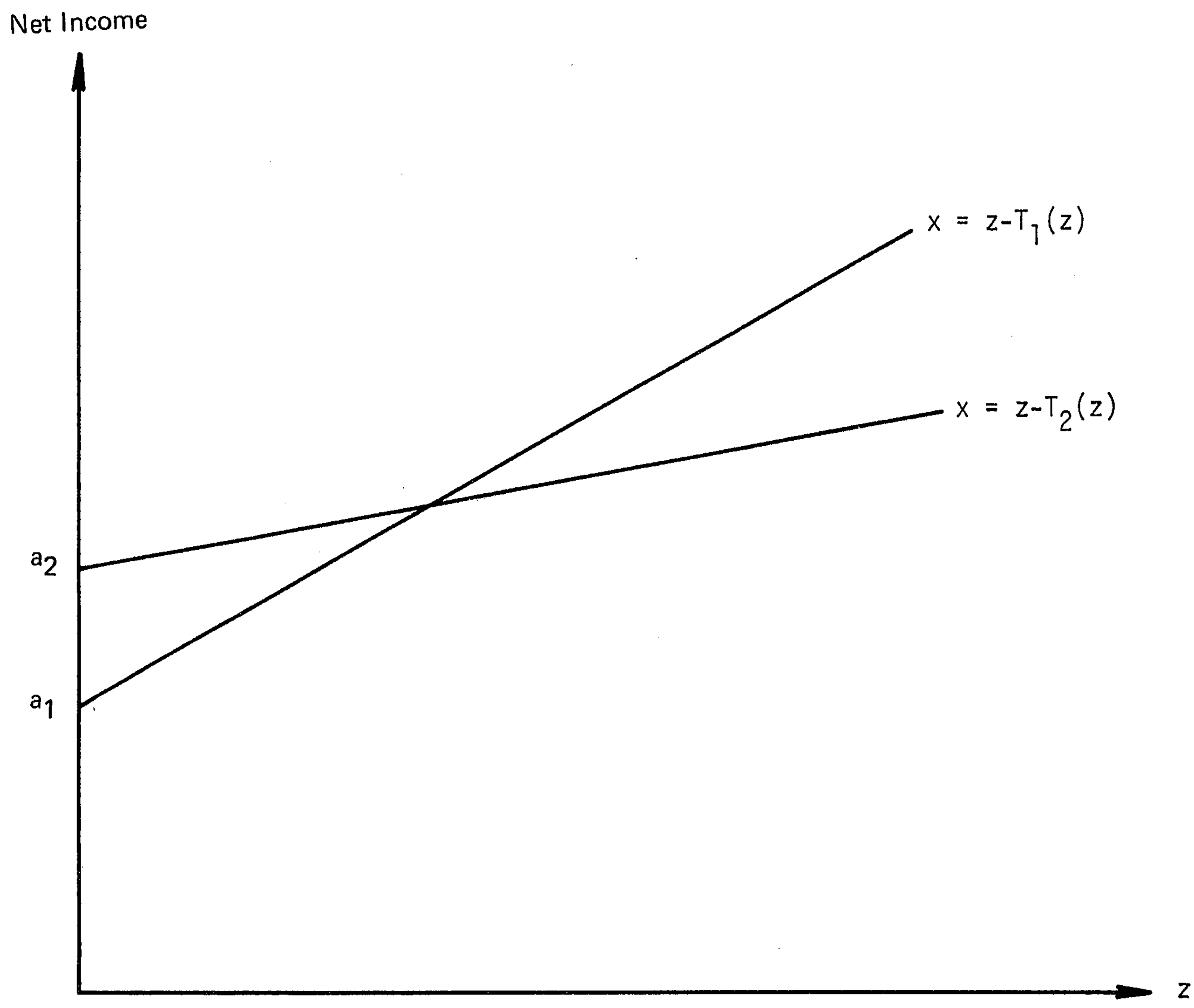


Figure 14

point which is at least as high as the point of intersection of the graph of $z - T_1(z) = a_1 + (1-b_1)z$ with the same axis (see Figure 14). Now, the graph of $z - T_2(z)$ cannot be everywhere above the graph of $z - T_1(z)$ for then everyone would prefer T_2 over T_1, and T_1 would not therefore be optimal under any individualistic criterion of social ordering. This means that the graph of $z - T_2(z)$ must be at least as flat as the graph of $z - T_1(z)$, implying that the marginal rate of T_2, which is b_2, must be at least as high as the marginal rate of T_1, which is b_1. The generalization of this proof to the case $N_1 \geqq 0$ is straightforward and left to the reader. Q.E.D.

G. Numerical Results

For the additive social welfare function, Mirrlees carried out several numerical examples under alternative assumptions about the distribution of skills and of individual preferences. On the basis of these examples, he concluded that "an approximately linear income-tax schedule, with all the administrative advantages it would bring, is desirable" and that "the income-tax is a much less effective tool for reducing inequalities than has often been thought." [4, p. 208]. Another feature of Mirrlees' results is that the marginal tax rates are rather low (15-25 percent in most cases) and tend to fall with income (a regressive tax according to definition I). Atkinson [2] presented similar examples for the max-min case. His results show marginal tax rates which are considerably higher than Mirrlees' for most of the income range (and this conforms to our theorem 6, part (b)). Also, the optimal tax is no longer approximately linear.

III. The Education Model

In this model (Sheshinski [9]) there is only one final good--consumption; denote it by x. People differ again in their skills (denoted by n), and their incomes depend on their skills and on the levels of education they acquire. An individual with skill n who acquires education at level y, incurs an education cost of g(y), where g is a strictly increasing and convex function (namely, the marginal cost of education is positive and increasing); his gross income is $z = ny$ and his consumption is then $ny - T(ny) - g(y)$. Thus, taxable income is ny and not $ny - g(y)$, which means that education costs are not tax deductible.

This model is very similar in its mathematical characteristics to the former model. Thus, all we have done in the labor model can be carried out in most cases, except for some changes in interpretation, to the education model. For such a treatment of the education model, the reader is referred to [8]. Here we will mention only one serious disadvantage of this model: a crucial feature in it is the assumption that educational costs are not deductible from taxable income. If educational costs were deductible, the optimal tax would have a 100 percent marginal rate everywhere for then the tax would have no effect on how much education individuals acquire.

IV. Conclusion

Two criteria of distributional equity--the additive social welfare function and the max-min criterion--and two sorts of incentive effects that an income tax may have--through the consumption-leisure choice and through the profitability of acquiring education--were examined in this paper. We have seen that, in general, the marginal tax rate must be bounded between 0 and 100 percent. This result rules out wage-subsidy but leaves open the possibility of a subsidy based on other variables, such as, for instance, innate ability. We have also examined the desirability of a negative income tax and concluded that it depends on the size of public spending, which was taken as an exogenous variable throughout this paper. A better understanding of this problem would be obtained in a framework where the level of public consumption and the structure of taxes needed to finance it are both examined simultaneously. Another interesting question concerning an income tax is whether it is progressive or regressive. We have seen that although both our criteria of distributional justice favor equality, the optimal tax need not be progressive everywhere and, furthermore, must not be so at the highest income brackets.

FOOTNOTES

[1]No other criterion is discussed in the literature.

[2]We should emphasize that it is only a negative marginal rate which is excluded but not a negative tax (see section D below).

[3]Notice that $u_{yy} < 0$ means that as y increases, u_y becomes "more negative." Therefore, $u_{yy} < 0$ indeed implies increasing marginal disutility of y.

[4]It was elsewhere shown by the author (see [8]) that the additive social welfare function makes no sense unless it is assumed that $u_{xy} > 0$.

[5]The purpose of this assumption is to guarantee that the tax schedule can not be tailor-made for each individual. With a discrete number of individuals, for instance, it might be possible to design a tax schedule which is essentially a lump-sum tax.

[6]The assumption that F is strictly increasing is a restatement of the assumption of a continuum of individuals.

[7]It is assumed here that the marginal tax does not exceed 100 percent (see the next section).

[8]This tax function is further confined to the class of continuous functions.

[9]Atkinson [1] employed the term "mean-preserving spread" to denote a transfer from the poor to the rich.

[10]In fact, Mirrlees has found that the optimal tax is approximately linear in a wide variety of examples.

[11]The slope of $z - T(z)$ is $1 - T'$ and when $T' < 0$, this slope is greater than 1.

[12]This theorem was first proved by Phelps [6] under the strong assumption that $N_1 = 0$, which means that the lowest-skilled individual in society is incapable of producing any income, no matter how much he works.

[13]A special case of this lemma in which $N_1 = 0$ was first observed by Phelps [6].

[14]Footnote 12 is applicable here too.

[15]Notice that the kind of optimality we deal with in this paper is in some sense a partial one. We fix the level of public consumption and ask what is the optimal tax. A full optimum requires a simultaneous determination of the tax and the level of public consumption.

[16]All we can say in this case is that if an optimal tax is negative somewhere, this must be so at the lowest income brackets.

[17]Actually, not only N_2 prefers T_1. Continuity considerations show that all persons with skills in a neighborhood of N_2 are better-off under T_1.

[18]Such is, for instance, the relationship between the marginal cost cruve and the average cost curve in the theory of the producer. For a direct proof in our case let us differentiate the average tax, T/z, with respect to income, z:

$$\frac{d}{dz}\left(\frac{T}{z}\right) = \frac{T'z - T}{z^2} = \frac{1}{z}\left(T' - \frac{T}{z}\right)$$

This derivative is positive when the marginal tax, T', exceeds the average tax, T/z.

REFERENCES

[1] Atkinson, A. B., "On the Measurement of Inequality," Journal of Economic Theory, 2 (September 1970):244-263.

[2] Atkinson, A. B., "'Max-Min' and Optimal Income Taxation," 1972 (mimeo).

[3] Harsanyi, "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility," Journal of Political Economy, August 1955.

[4] Mirrlees, J. A., "An Exploration in the Theory of Optimum Income Taxation," Review of Economic Studies, 1971, 175-208.

[5] Musgrave, R. A., The Theory of Public Finance, (McGraw-Hill, 1959).

[6] Phelps, E., "Taxation of Wage Income for Economic Justice," Quarterly Journal of Economics, 87 (August 1973):331-354.

[7] Rawls, J., A Theory of Justice (Cambridge, Mass.: Harvard Univ. Press, 1971).

[8] Sadka, E., "Income Distribution, Incentive Effects and Optimal Income Taxation," unpublished Ph.D. dissertation, M.I.T., May, 1974.

[9] Sheshinski, E., "On the Theory of Optimal Income Taxation," Journal of Public Economics (forthcoming).

[10] Sheshinski, E., "The Optimal Linear Income-Tax," Review of Economic Studies, 1972, 297-302.