

**Simple Formulae for Optimal Income Taxation and the Measurement of Inequality:  
An Essay in Honor of Amartya Sen**

Joseph E. Stiglitz<sup>1</sup>

I. Introduction

It is a pleasure to contribute this essay in honor of Amartya Sen, whose intellectual work has been an inspiration, and whose friendship I have valued for more than four decades. Amartya's interests, especially in inequality and development, have overlapped with mine. At a time when so many developmental economists were advocating Washington Consensus policies—almost completely ignoring the consequences for poverty and inequality—Amartya stood out as a lonely voice, at least within the West. Today, many, if not most, of those who pushed these policies have recognized that the policies often have not promoted growth, and even when they have led to growth, not all have benefited. They recognize the seriousness of the failure to pay attention to the distributive consequences of economic policies.<sup>2</sup>

For this festschrift, however, I want to focus on another area in which Sen has made an important contribution—the measurement of inequality.<sup>3 4</sup> Some thirty five years ago, Michael Rothschild and I asked when one can say that one distribution is “riskier” than another—for any risk-averse individual.<sup>5</sup> The same logic could be applied to income distributions: when can one say that one income distribution is more unequal than another—for any inequality averse social welfare function? Atkinson noticed that our

earlier results had a natural interpretation in terms of the standard Lorenz curve: an income distribution is more unequal than another (such that any inequality averse Benthamite social welfare function preferred one distribution to the other) if and only if one Lorenz curve lies inside the other.<sup>6</sup> Dasgupta, Sen, and Starrett and Rothchild and I then extended the analysis to more general inequality averse social welfare functions.<sup>7</sup>

Sen made two even more important contributions to the policy and conceptual debate. The first was to emphasize that, in assessing the performance of the economy (or more broadly, society), one should not look only at outcomes (the incomes or consumptions that individuals enjoy), but also at the freedom they give for human action, which in turn depends on the capabilities of individuals and the scope they have for participating in the decisions that affect their lives.

The emphasis on capabilities is congruent with the emphasis on “opportunity” on which modern political discourse has centered.<sup>8</sup> If all individuals had the same opportunities, including education and access to basic necessities of life (e.g. health care or food, so that, for instance, they do not suffer the lifelong consequences of lack of medical care or malnutrition), inequalities in income or consumption would not be of great concern; in such circumstances, inequalities would simply reflect the fact that some individuals work harder and others do not.<sup>9</sup> The inequalities that are so marked in our societies are, however, only partly the result of differences in preferences; they reflect differences in opportunity sets and in “luck,” that is, even when individuals *on average* have the same opportunity set, even when individuals work just as hard, some win one of life’s many

lotteries, and wind up wealthy, and others lose. Parsing out the relative importance of these different sources of inequality is not easy; but there are obviously fundamental differences in the appropriate policy responses.<sup>1</sup>

The difficulties of ranking opportunity sets (when one opportunity set is “more equal” than another) are even greater than those we discussed earlier in ranking probability distributions of income. Focusing on just two commodities, it is easy to assess whether one individual’s opportunity set is better than another’s (see figure 1): if it lies outside the other, it clearly is better. But even if one individual’s opportunity set lies outside the other, the extent to which it does so depends on the axis (vertical or horizontal) on which we make our measurements.

However, the situation changes if budget constraints cross. Then some individuals might prefer one budget constraint, others the other. The point becomes particularly relevant when we think about the two goods as “consumption” and “leisure,” for all individuals have the same endowment of leisure (though not necessarily with the ability to use it well.) (See Figure 2).

One aspect, the importance of which can shift greatly over time, is the provision of public goods. To the extent that public goods are truly public goods and are equally enjoyed by all individuals, societies where a greater fraction of GDP is spent on public

---

<sup>1</sup> One of my own earliest papers was concerned with trying to understand the dynamics of inequality, the forces which give rise to its persistence. See J. E. Stiglitz, “Distribution of Income and Wealth Among Individuals,” *Econometrica*, 37(3), July 1969, pp. 382-397. Interestingly, it appears that many conservatives think that most inequality arises from differences in choices; “liberals” emphasize the importance of luck. See Matthew Miller, *The Two-Percent Solution: Fixing America’s Problems in a Way that Conservatives and Liberals can Love*, Public Affairs, 2003. Many successful individuals obviously attribute their success to their hard work and their insight (itself a result of investments in say education, rather than the “luck” of being born with good brain power).

goods will, in a fundamental sense, be more egalitarian; focusing on the inequality of *private consumption* may, accordingly, misrepresent the degree of inequality in society.<sup>10</sup>

This highlights the fact that there are many dimensions to an individual's well-being or assessing that of society. One of Sen's most important practical contributions was the role he played in the construction of the Human Development Indicator (HDI), a center piece of the annual *Human Development Report* published by the UN Development Programme (UNDP). The HDI provides a metric of societal well-being that goes beyond GDP per capita and includes, for instance, measures of health and education. It is a measure which has made an enormous positive contribution to the policy debate by drawing attention to aspects of society which would receive inadequate attention were GDP the main metric of success. What you measure is what you strive for. By noting that some societies that may rank towards the top on GDP per capita (like the U.S.) score much more poorly in the HDI (more poorly, in particular, than all of the Scandinavian countries), it at least raises questions about the appropriate direction of policy.<sup>2</sup>

One of the dimensions of societal well-being that is not captured at all in measures of *average* GDP per capita is how access to goods is distributed among those in society. Atkinson went on from his analysis of the question of when one society has a "better" distribution of income than another to construct a simple and powerful measure of

---

<sup>2</sup> There are broader objections to the use of GDP; it does not, for instance, take into account resource depletion or environmental degradation; it looks at the output produced within a country, not the well-being of the citizens of the country. See, for instance, J. E. Stiglitz, *Making Globalization Work*, New York: W.W. Norton, 2006.

inequality: how much society would be willing to give up in order to eliminate inequality.<sup>11</sup> The amount was simply related to society's inequality aversion (a measure of the concavity of the social welfare function, analogous to the Arrow-Pratt measure of risk aversion) and the magnitude of the inequality itself.

Just as Rothschild and Stiglitz's work highlighted that often, one could not rank two distributions (one might be preferred by some risk averse individual, and another by another risk averse individual), so too Atkinson's work highlighted that whenever two Lorenz curves crosses, some inequality averse individual would prefer one distribution, another the other. And just as Rothschild-Stiglitz highlighted that standard measures of risk (like variance) may be misleading, Atkinson's work highlighted that standard measures of inequality (like the Gini coefficient) may be misleading. These limitations in ranking were not the result of an incompleteness in the theory; they called attention to the fact that one could "improve" the distribution of income at one place (such as reducing poverty), and at the same time, make it worse (transferring income from the middle class to the rich.)

This measure (often referred to as the Atkinson-Dalton measure, because of its earlier use by Hugh Dalton<sup>12</sup>) may, however, not be the best measure of inequality, for several reasons.<sup>13</sup> There are two that we do not take up in this paper. The first is that the Atkinson-Dalton measure is a static measure, a glimpse of inequality at one point in time. If there is individual variability of income over time, and individuals can smooth their consumption, then this overstates the degree of inequality of "well-being" (or

consumption). If there is an increase in year to year income variability, it may seem as if there is an increase in inequality, but this may in reality have little consequence. On the other hand, if capital markets are imperfect (as they are), so that individuals cannot effectively smooth consumption over time, and if there are imperfect insurance markets (as there are), so that individuals cannot insure against this variability, then the increased year-to-year variability will have adverse welfare consequences, though it is still the case that looking at income inequality (or even consumption inequality) at a certain point in time exaggerates the degree of life-time inequality. Year-to-year variability gives rise to insecurity, which can be very costly to risk averse individuals. Indeed, much of the political debate in recent years has focused on how to reduce various forms of insecurity.<sup>14</sup>

The second issue we do not discuss here is related, but in some sense has even more profound social consequences: the Atkinson-Dalton measure is not a measure of social mobility and therefore does not capture the dynamics of equality of opportunity, that is, to what extent the life chances of someone born at the bottom of the income distribution differ from someone born at the top. A society's social mobility is characterized by a mobility matrix, which describes, for instance, the probability distribution (measured by the decile of the population) of a person born in a particular decile. If the probability distributions of each decile were the same, there would be true equality of opportunity. One concern is that there is some evidence that opportunities for upward mobility in the U.S. may be decreasing.<sup>15</sup> Just as we can ask when one income distribution is more egalitarian than another, we can ask when one mobility matrix is more egalitarian than

another (i.e., preferred by a society with an inequality averse social welfare function.)

The difficulties noted by Rothschild, Stiglitz, and Atkinson in ranking income distributions (or probability distributions) are compounded in ranking mobility matrices.<sup>16</sup>

However, there are several limitations to the Atkinson-Dalton measure which this paper does address. The Atkinson-Dalton measure illustrates the percentage of national income which society would be willing to sacrifice if all inequality were eliminated. It is, in other words, a measure of the total cost of inequality. We wish to develop a *marginal* measure of inequality, i.e., a measure of how much society is willing to sacrifice to reduce inequality a given amount.

As in most aspects of economic analysis, it is marginal valuations, not total valuations, which are crucial to resource allocation. The question which is often of relevance is how society should trade-off distribution and efficiency. That there is a trade-off between “distribution” and “efficiency” has, for instance, long been recognized in the literature on income taxation. More progressive tax structures have greater disincentive effects but, in principle at least, result in more egalitarian income distributions.

It has sometimes been suggested that some countries have gone too far in their attempts to obtain more egalitarian income distribution, that the cost in loss of efficiency and disincentives is too great relative to the benefits in attaining a “better” income distribution, at least at the margin. Such a statement involves judgments of two sorts:

(a) Empirical judgments concerning the order of magnitude of the “costs,” the disincentive effects, and benefits, i.e., the change in the income distribution; and

(b) Value judgments concerning one’s attitudes towards inequality.

The Atkinson-Dalton measure addresses the second question. But no one proposes eliminating all inequality—the Atkinson-Dalton measure tells us how much we would be willing to give up to achieve this, but such a number is not relevant for policy. Rather, the question is posed *at the margin*: what is the marginal cost of more progressivity, and what is the marginal benefit? There are standard ways of calculating the marginal costs, the increments in the dead weight loss. But it would be useful to have a measure of the marginal benefit from reducing inequality; we need, in short, a *marginal measure* of inequality. This is what we provide here.

Since the marginal value of a reduction in inequality may, in some sense, exceed the total value, previous measures have understated the social cost of inequality, or perhaps more accurately, the social gain from reducing it. That is, intuitively, if, in the process of transferring money from the rich to the poor, one loses a fraction of the resources being transferred, and if society has a great deal of inequality, one might be willing to pay a great deal to reduce inequality a little bit, but the fraction of resources that could be lost that would make the transfer socially acceptable would get smaller as the degree of inequality is reduced.

There is a second problem with these measures of inequality: they assume implicitly either that the supply of labor is inelastic or that the only source of inequality is in

inherited income. In other words, differences in income arising out of differences in ability when labor is elastically supplied—and the consequent differences in leisure consumed—are not appropriately taken into account.

The first objective of this essay then is to construct a new measure of inequality, a measure which focuses, on the one hand, on the *value of marginal reductions in equality* (rather than its total elimination), and on the other hand, on differences in abilities to earn. While focusing on the evaluation of marginal changes in inequality suggests that the Atkinson-Dalton measure understates the costs of inequality, our analysis of the consequences of wage inequality with an elastic labor supply suggests that the conventional measures overstate the cost of inequality.

There is a second objective of this essay: To relate the optimal tax structure to some simple observable parameters and to our newly developed measure of inequality. Our results provide a simple formula for determining the optimal tax rate.

The analysis proceeds within the utilitarian framework for analyzing the desirability of redistribution and the design of tax structures, a tradition dating at least back to Edgeworth and Bentham.<sup>17</sup> Whether this is an appropriate framework for analyzing these questions is an issue we discuss briefly in the concluding section of the paper, which will bring us back to issues on which Sen has made fundamental contributions.

## 2. On the Cost of Inequality and the Benefits of Redistribution: A New Measure of Inequality

Social attitudes about the importance of inequality have varied greatly. At times, there has been a general consensus that economic growth, while not eliminating poverty itself, will increase the standard of living of the poor far more than would be possible under any redistributive scheme; this led to the belief that attention should be directed at growth rather than inequality. More recently, however, there has been a widespread feeling that some of the worst manifestations of poverty are a result of the *relative* position of the poor, and, if that is the case, growth which does not eliminate the degree of inequality will not alleviate the problems of poverty.<sup>18</sup>

What do statements such as “inequality is a significant problem in the US” mean? We can measure the degree of inequality, say, by statistical measures such as the coefficient of variation. But how can we decide what is a large number and what is a small number? One way of attacking the problem is a suggestion, originally put forward by Dalton, and developed by Kolm, Atkinson, and others. Assume you did not know where in the income distribution you were going to be. One could calculate the expected utility of income  $Eu(Y)$ . If individuals are risk averse, the expected utility is lower than that which individuals would have received if the same total income were distributed equally, i.e.,

$$(2.1) \quad Eu(Y) < u(E(Y)) \text{ if } u'' < 0.$$

Thus, to eliminate inequality, one would be willing to give up a fraction of total income,  $\hat{\mu}$ , where

$$(2.2) \quad Eu(Y) = u((1 - \hat{\mu})EY).$$

Thus, the measure of inequality is given by

$$(2.3) \quad \hat{\mu} = 1 - \frac{u^{-1}[Eu(Y)]}{EY}.$$

As a first order approximation, it is easy to show that<sup>19 20</sup>

$$(2.4) \quad \hat{\mu} \approx \frac{Rs_Y^2}{2}$$

where

$$(2.5) \quad R = -\frac{u''Y}{u'}$$

is the elasticity of marginal utility (sometimes referred to as the measure of inequality aversion), and

$$(2.6) \quad s_Y = \frac{\sqrt{E(Y - \bar{Y})^2}}{\bar{Y}}$$

is the coefficient of variation.

Thus, if  $R=1$ , and the coefficient of variation of income is .5, then one is willing to sacrifice approximately 12.5% of national income to eliminate inequality, a seemingly large figure (although still only equal to a few years' growth).

This provides a measure of the total gains to eliminating all inequality and is thus a measure of the total degree of inequality. But for many problems, a marginal notion is more useful. Assume we could take one percent of all incomes and redistribute a fraction of the amount collected equally to all individuals. The marginal measure of inequality is defined as the fraction which just leaves social welfare unaffected:

$$(2.7) \quad Eu'(Y)(\bar{Y} - Y) = \hat{m}\bar{Y}Eu'$$

or

$$(2.8) \quad \hat{m} = \frac{Eu'(\bar{Y} - Y)}{\bar{Y}Eu'}$$

The LHS of (2.7) is the gain from the redistribution.  $Eu'(Y)$  is the loss in expected utility in subtracting a dollar away from each individual.  $\hat{m}$  can be rewritten in a slightly different way:

$$\hat{m} = \frac{-E[u' - \bar{u}'] [Y - \bar{Y}]}{\bar{Y}u'}$$

$\hat{m}$  is just the normalized covariance between the marginal utility of income and income.

$\hat{m}$  is generally greater than  $\hat{\mu}$ , reflecting diminishing marginal returns to redistribution.

Although  $\hat{m}$  and  $\hat{\mu}$  are defined differently, for small variance  $\hat{m} \approx 2\hat{\mu}$ :

$$(2.8) \quad m \approx \frac{u'(\bar{Y})E(\bar{Y} - Y) - u''(\bar{Y})E(Y - \bar{Y})^2}{\bar{Y}u'(\bar{Y})} \approx Rs_Y^2 = 2\hat{\mu}.$$

This means, for instance, that if  $R = 1$ , and  $s_Y = .5$ , a redistribution scheme which took 1% of each person's income away and redistributed just over three quarters of that amount back again, but this time equally to everyone, would increase social welfare.<sup>21</sup>

Of course, while all complete reductions of inequality are the same, every reduction of inequality is marginal in its own way. The amount society would be willing to pay to move a little bit of income from the very bottom to the very top is obviously much larger than it would be willing to pay to move a little bit of income from someone just above the median to someone just below the median. The particular marginal measure we have used is tailored made for analyzing the optimal *linear* income tax; for greater progressivity in that context entails taking away a fraction of each individual's income, and returning the proceeds as a uniform lump sum (demi-grant) to each individual. We know that there is a distortion as a result of the higher marginal tax rate—a dead weight loss. Our marginal measure is designed to answer the question, how large can this deadweight loss be (as a fraction of the revenue raised), for it to be still worthwhile (at the margin) to raise the tax rate still further?

### 2.1. Measures of inequality with wage differences

The previous section developed a marginal measure of inequality; but like the earlier Atkinson-Dalton measure, we ignore leisure. But individuals enjoy leisure, just as they enjoy goods; and individuals differ in their consumption of leisure, just as they differ in their consumption of goods.

The measurement of inequality when individuals “consume” goods and leisure poses a classical index number problem. The natural question is, can we develop measures of inequality analogous to Atkinson’s total measure and our marginal measure which take into account leisure? What relationship would such a new measure have to the older measures?<sup>22</sup>

Intuitively, it would seem that ignoring leisure leads the Atkinson measure to overstate the significance of inequality. This may be seen in several ways. Clearly, if all individuals had the same ability (received the same wage) but differed with respect to their preferences for goods and leisure, the conventional measures of inequality, focusing just on difference in consumption of goods, would provide a gross overestimate of the significance of inequality in “welfare.” If hours of leisure have increased generally, and particularly if they have increased more for low wage workers than high wage workers, then the conventional measure may understate the reduction in inequality over time.

When the source of inequality is a difference in wages, the endowments of leisure are still the same. Hence, if we formulated a simple index of welfare consisting of a weighted average of leisure and consumption

$$W = \delta (1 - \ell) + (1 - \delta) C$$

where the time available is normalized at unity,  $\ell$  is the percentage of time spent working, and  $C$  is consumption, then it is clear that if  $\ell$  is, say, constant (zero elasticity of labor supply), then the coefficient of variation of  $W$ ,  $s_W$  is less than that of  $C$ ,  $s_C$ :

$$s_W = \frac{(1 - \delta)\bar{C}}{\delta(1 - \bar{\ell}) + (1 - \delta)\bar{C}} s_C$$

where  $\bar{C}$  and  $\bar{\ell}$  (bar) are the mean values of labor and consumption.

But even this exaggerates the magnitude of inequality. Inequality in wages affects the ability to transform leisure into consumption commodities. The more willing individuals are to substitute leisure for consumption goods (i.e., the greater the elasticity of substitution) the less the “cost” of inequality. Moreover, if the elasticity of labor supply is positive, the inequality in incomes will be greater than the inequality in wages.<sup>23</sup>

The problem is highlighted in the case of a high elasticity of substitution, for which it is immediately apparent that *a mean preserving increase in the inequality of abilities*<sup>24</sup> (which preserves the mean level of ability) actually raises expected utility (social welfare).<sup>25</sup> Consider what happens with an infinite elasticity of substitution. Originally, assume that all individuals have the same productivity (= “unity”) and that the marginal rate of substitution between leisure and consumption goods is unity. Then assume that one-half the population has a productivity of  $1 + \Delta w$  and one-half the population has a productivity of  $1 - \Delta w$ . Obviously those who have a lowered productivity simply consume “leisure” and are no worse off than before, but those who have a higher productivity are better off. Thus the cost of “productivity” inequality would appear to be negative.<sup>26</sup> As figure 3 makes clear, the result is more general: the gain in utility with

wages (productivity) is increased more than offset the losses when it is decreased symmetrically.

For the remainder of this paper we focus on inequalities arising from differences in ability. This is a very important assumption that has become conventional within the optimal income tax literature. Yet ironically, its implications for the measurement-of-inequality literature do not seem to have been explored. All workers are assumed to be perfect substitutes for one another; i.e., they differ only in the number of “efficiency units of labor” which they embody. We choose our units so that an individual of “unit” efficiency receives a wage of unity. Hence, the before-tax income of somebody of efficiency “ $w$ ” is

$$(2.9) \quad Y = wL$$

where  $L$  is the amount of labor supplied. We again define the total measure of social loss to be that amount of consumption goods one would be willing to give up to obtain complete equality. To do this we let  $U(C,L)$  be the level of utility as a function of consumption and work.

$$U_1 > 0, U_2 < 0$$

We then define the indirect utility function

$$(2.10) \quad V(w,I) = \max U(C,L)$$

$$s.t. C = wL + I$$

giving the maximum level of utility attainable as a function of the wage rate and “exogenous” income,  $I$ . Then our new “total” measure of inequality,  $\mu$ , is given by

$$EV(w,0) = V(I, -\mu Y)$$

where we have normalized  $Ew = 1$ . We show (in Appendix A) that

$$(2.11) \quad \mu \approx \frac{V_{11}}{2V_2} s_w^2 = \frac{s_w^2}{2(1+n_L)^2} \left( R - n_L - \frac{\partial L}{\partial I} \right)$$

where

$$n_L = \frac{d \ln L}{d \ln w}, \text{ the supply elasticity of labor,}$$

$$R = \frac{-V_{22}C}{V_2} \quad \left( = \frac{-u'C}{u'} \text{ if the utility function is additive} \right),$$

the elasticity of marginal utility of income,  $s_w$  is the coefficient of variation of abilities.

Note that if  $n_L = 0$  and  $\partial L/\partial I = 0$ , then (2.11) is identical to (2.4) as we would have expected; otherwise, the measures differ. And as we noted intuitively earlier, if  $R$  is small (low measure of income inequality aversion) but the labor supply is highly elastic,  $\mu$  may be negative; wage variability increases expected utility.

More generally, the percentage of income that one would be willing to forego to eliminate wage inequality is a function not only of  $R$  and the coefficient of variation of income,  $s_Y^2$ , but also of the wage elasticity of labor and the income elasticity of labor. The greater the wage elasticity, the greater the magnitude of the overestimate of the degree of inequality yielded by the Atkinson measure.

In the appendix, two special cases are presented:

a) Additive utility function

$$(2.12'') \quad \mu \approx \frac{s_Y^2}{2(1+n_L)^2} \left[ R - n_L \left( \frac{1-2R}{1-R} \right) \right], \quad R \neq 1$$

$$\mu \approx \frac{s_Y^2}{2} \left( 1 - \frac{\partial L}{\partial dI} \right), \quad R = 1.$$

(b) Homothetic indifference curves

$$(2.12''') \quad \mu \approx \frac{s_Y^2}{2(1+n_L)^2} [R + (1-L) - n_L]$$

where  $L$  is to be interpreted here as the fraction of the time available that is worked.

If, in addition, we assume a constant elasticity utility function, we obtain

$$(2.12''''') \quad \mu \approx \frac{s_w^2}{2} [R + (1-L)(2-\sigma)] = \frac{s_Y^2 [R + (1-L)(2-\sigma)]}{2[1 + (\sigma-1)(1-L)]^2}$$

where  $\sigma$  is the elasticity of substitution.

In Table 1 we provide some estimates of  $\hat{\mu}$  for different values of the parameters. Notice that the measure may differ considerably from the Atkinson measure, for even small values of  $n_L$ . Indeed, as we suggested earlier, if  $n_L$  is large and  $R$  small,  $\hat{\mu}$  may be negative.

More generally, in a command economy (where the government assigns everyone a level of consumption and work) expected utility is always higher if abilities are unequally distributed than if all individuals have the same productivity, equal to the mean productivity in the situation with inequality. That is, each individual could work the same amount in the situation with inequality as without it, and then net output would be the same, so each individual could receive the same consumption goods. However, this will not in general be optimal for the economy with inequality, so that the attainable level of expected utility must be higher. Indeed, *everyone* can be made better off. Whether they are or not depends on what kinds of redistribution taxes are introduced. If we allow lump sum taxes but insist on everyone remaining at the same level of utility, the social gain from inequality in abilities is approximately equal to

$$\frac{1}{2} \left( \frac{\partial L}{\partial w} \right)_u s_w^2$$

(see Appendix C).

With the linear income tax, there may be either a social gain or social loss, although for reasonable values of the relevant parameters the gain or loss appears to be small.

To see this, we must first analyze the optimal linear income tax, to which we turn in section 3.

First, however, we must define the analog to  $m$ , our marginal measure of inequality, for the case where there is an elastic labor supply. Two such measures will prove useful in the subsequent discussion.

$$(2.13) \quad m_c \equiv \frac{EU_c(\bar{C} - C)}{ECEU_c} = \frac{E(U_c - \bar{U}_c)(\bar{C} - C)}{\bar{C}\bar{U}_c}$$

the normalized covariance of consumption with marginal utility. If we had an additive (Benthamite) social welfare function and took one percent of consumption away from everyone and redistributed it equally to everyone,  $m_c$  tells us the fraction of consumption we could lose in the process and still be better off.

If there is an additive utility function

$$(2.14) \quad m_c \approx Rs_c^2 \quad .$$

A natural extension of this measure is to the case where the social welfare function is of the form (where  $f(w)$  is the density distribution of abilities)

$$\int W(U(C, L))f(w)dw$$

in which case we obtain, instead of (2.13)

$$(2.13') \quad m_C = \frac{EW'U_C(\bar{C} - C)}{ECEW'U_C}$$

We define analogously

$$(2.15) \quad m_Y = \frac{EW'U_C(\bar{Y} - Y)}{EYEW'U_C}$$

For some purposes, however, these are not good measures: when we take income away from individuals or give it to them, individuals adjust their effort supply, and this affects tax revenue and the amount we can redistribute. Thus, the net social gain from giving an individual a dollar is not  $W'U_C$ , the social marginal utility of income, but  $W'U_C + \lambda \tau w$  ( $dL/dI$ ), where  $\lambda$  is the shadow price on government revenue and  $\tau$  is the (marginal) tax rate. If we define

$$\beta \equiv \frac{W'U_C}{\lambda} + \tau w \frac{\partial L}{\partial I}$$

as the marginal social utility of income to an individual of wage  $w$ , then we can define

$$(2.16) \quad m_\beta = \frac{-E(\beta - \bar{\beta})(Y - \bar{Y})}{\bar{\beta}Y} .$$

### 3. Optimal Linear Income Tax

In this section, we shall derive several simple alternative expressions for the optimal linear income tax rate, in terms of our measure of inequality and certain characteristics of individuals' behavior (in particular, the elasticity of supply of labor). It should be observed that although we shall write down expressions in which the tax rate appears explicitly on the left hand side of an equation, it often appears implicitly on the right hand side. For instance, the elasticity of supply of labor will in general not be constant. The alternative expressions may be useful in deriving the optimal tax rate, depending on what assumptions one makes about what parameters in the system are constant (e.g., whether the compensated or the uncompensated demand curves have constant elasticity).<sup>27</sup>

A linear income tax rate is characterized by an equation of the form

$$(3.1) \quad T = \tau Y - I$$

When  $T$  is total tax payment,  $\tau$  is the marginal tax rate, assumed to be constant,  $Y$  is before tax income, and  $I$  is the value of the tax credit. The essential characteristic of the linear income tax is said to be progressive if  $I > 0$ , i.e.,

$$(3.2) \quad \frac{T}{Y} = \tau - \frac{I}{Y}$$

is an increasing function of income. It is regressive if  $I < 0$ .

Mirrlees has shown some examples in which the optimal tax structure is approximately linear.<sup>28</sup> If this conclusion is correct more generally, then focusing on linear income taxes may not be too restrictive.

The national income identity takes on the form

$$(3.3) \quad \int wL_w f(w)dw = (1 - \tau) \int wL_w f(w)dw + I + G$$

where  $f(w)$  is the density function of the population by ability,  $G$  is per capita expenditure on government services, and  $L_w$  is the labor supplied by an individual of wage  $w$ .<sup>29</sup>

Alternatively this can be written as

$$(3.4) \quad \tau = \frac{I + G}{\int wL_w f(w)dw} .$$

For given government revenue ( $I + G$ ) there are in general at least two values of  $\tau$  satisfying (3.4). That is, the government revenue is not a monotonic function of  $\tau$ ; at  $\tau =$

0, government revenue equals zero, and at  $\tau = 1$ , revenue equals zero.<sup>3</sup> With multiple solutions, it is clear that the relevant solution is the one with the lowest value of  $\tau$ .

For simplicity, we assume as we did in the previous section that all individuals have the same utility function and that the utility from public goods is separable from that of private goods, i.e.  $U^* = U(C, L) + H(G)$ . For this part of the analysis, we assume  $G$  is fixed.

Individuals maximize utility, (3.5),

$$(3.5) \quad U = U(C, L), \quad U_1 > 0, \quad U_2 < 0$$

where  $C$  = consumption, subject to the budget constraint:

$$(3.6) \quad \begin{aligned} C &= I + (1 - \tau)wL \\ &= \tau \bar{Y} - G + (1 - \tau)Y \end{aligned}$$

where  $Y = wL$ , so

$$\bar{Y} = \int wL_w f(w) dw.$$

The maximized value of this, expressed as a function of after tax wage and  $I$  is, as we mentioned before, the indirect utility function,  $V(w(1 - \tau), I)$ .

The government has a social welfare function of the form<sup>30</sup>

$$(3.7) \quad \int W(U_w f(w) dw$$

with  $W' > 0$ ,  $W'' \leq 0$ .

---

<sup>3</sup> This obvious result, which I noted in the early 1970s, became popularized under President Reagan, in the 1980s, by Arthur Laffer, and came to be called the Laffer curve. It provided the basis of supply side economics, arguing that lowering taxes could lead to more revenues and increased welfare.

The problem of the optimum income tax may be formulated as

$$\max \int W(V(w(1-\tau), I))f(w)dw$$

subject to the constraint (3.3) Thus, letting  $\lambda$  be the Lagrange multiplier associated with the constraint (3.3), we obtain

$$(3.8a) \quad EWV_I = \lambda E(1 - \tau w \frac{\partial L}{\partial I})$$

or

$$E\beta = 1$$

and

$$(3.8b) \quad EWV_w = \lambda E[wL - \tau w^2 \frac{\partial L}{\partial w}]$$

or

$$(3.9) \quad \tau = \frac{-E(\beta - \bar{\beta})(Y - \bar{Y})}{\bar{Y}\bar{\beta}} \times \frac{1}{{}_u \bar{n}_L} = \frac{m_\beta}{{}_u \bar{n}_L}$$

where

$${}_u \bar{n}_L = \frac{EY_u n_L}{\bar{Y}} = \text{weighted average compensated labor supply elasticity.}$$

(3.9) is the key result of the paper, providing a remarkably simple formula for the optimal linear income tax; it relates the optimal tax rate to the normalized covariance between income and the marginal social utility of income—what we have identified earlier as our marginal measure of inequality—and the weighted mean value of the compensated

supply elasticity of labor. The greater the marginal measure of inequality, the greater the tax rate, and the smaller the elasticity of supply—the less the loss in welfare from the income tax—the greater the tax rate. This is just as one might expect, although the simplicity of the expression may be seen as somewhat surprising. *The optimal tax rate is equal to the marginal measure of inequality divided by the (average) compensated supply elasticity of labor.*

In some special cases, (3.9) might be a useful expression for obtaining some idea of the magnitude of the optimal tax rate. Assume, for instance, that there were constant marginal disutility of labor, and  $u'$  is of constant elasticity,  $R$ . If  $I = 0$ ,

$$n_L = \frac{1 - R}{R} .$$

Assume, moreover, that the distribution of wages is lognormal. Then, straightforward manipulations yield

$$\tau = R \left( 1 - e^{-\frac{\sigma^2}{R}} \right) \approx \sigma^2 - \frac{\sigma^4}{2R}$$

If  $\sigma = .4$ , the optimal tax rate is 16%, essentially independent of  $R$ , provided  $R$  is not too big. (If  $R=1$ ,  $\sigma^4 / 2R = .0128$ ) If  $I$  is small, we can rewrite this as

$$\tau \approx \frac{\left( 1 - e^{-\sigma^2(1+n_L)} \right)}{1 + n_L} .$$

Note too that in the case of an additive (Benthamite) social welfare function (and with government expenditures having no effect on labor supply or the marginal social utility of income), the tax rate does not depend at all on government expenditures; but this is obviously a limiting and special case. <sup>31</sup>

Although (3.9) provides a simple and intuitively appealing formula, it suffers from having the tax rate explicitly involved in the definition of  $m_\beta$ . An alternative formula is derived directly from (3.8)

$$\begin{aligned} -E\tau w^2 \frac{\partial L}{\partial w} &= \frac{EW'V_I wL}{\lambda} - \bar{Y} \\ &= \frac{EW'V_I wL}{EW'V_I} \left(1 - \tau Ew \frac{\partial L}{\partial I}\right) - \bar{Y} \end{aligned}$$

so

$$\begin{aligned} (3.10) \quad \tau &= -\frac{EW'V_I(Y - \bar{Y})}{EW'V_I \bar{Y}} \left[ \frac{1}{\bar{n}_L - \frac{EW'V_I Y}{EW'V_I EY} E \frac{dY}{dI}} \right] \\ &= m_Y \left[ \frac{1}{\bar{n}_L - (1 - m_Y) \frac{d\bar{Y}}{dI}} \right] \end{aligned}$$

(3.10) stresses on further aspect of the determination of the optimal tax rate; if total income is reduced as a result of the lump sum payment (because of income effects), the

optimal tax will be lower. The importance of this term depends on the degree of inequality; with a very high degree of inequality, this term is insignificant.

Notice that in (3.10) it is the uncompensated elasticity which is used. However, if the aggregate labor supply equation approximately satisfies the Slutsky equation, so we can write

$${}_{u}\bar{n}_L \approx \bar{n}_L - \frac{d\bar{Y}}{dI} ;$$

then (3.10) becomes

$$(3.10') \quad \tau \approx m_Y \left[ \frac{1}{{}_{u}\bar{n}_L + m_Y \frac{d\bar{Y}}{dI}} \right].$$

A third expression for the optimal tax formula is in terms of the marginal measure of inequality of  $C$ . Using (2.13), we obtain directly from (3.10)

$$(3.11) \quad \tau(1-\tau)\left[\bar{n}_L - \frac{d\bar{Y}}{dI}\right] + \tau m_C \frac{dY}{dI}(1-g) = m_C(1-g)$$

where  $g$  is the proportion of optimal income spent on government expenditures. This is a quadratic equation in  $\tau$  which can be solved in a straightforward manner.

If  $\tau$  is small, the solution can be approximated by

$$(3.12) \quad \tau \approx \frac{m_c(1-g)}{\bar{n}_L - \frac{d\bar{Y}}{dI} + m_c(1-g)}$$

Still, a fourth expression makes use of the *total* elasticity of government revenue with respect to the tax rate:

$$(3.13) \quad \gamma \equiv \frac{d \ln \tau \bar{Y}}{d \ln \tau} .$$

This elasticity involves two effects: both because of the lump sum payment and because of the substitution effect of the higher tax rate, labor supply is reduced. It is possible to show that<sup>32</sup>

$$(3.14) \quad 1 - \tau = \frac{m_c(1-g)}{1-\gamma} .$$

The numerator is just the product of the marginal measure of consumption inequality and the share of consumption in national income. The denominator is just one minus the revenue elasticity of the tax.

Upon further manipulation, the revenue elasticity can be related to the mean income and wage elasticity's of labor<sup>33</sup> supply:

$$1 - \gamma = \frac{\frac{\tau}{1-\tau} \frac{\partial \ln \bar{Y} - \tau}{\partial \ln (1-\tau)\tau - g} \frac{\partial \ln \bar{Y}}{\partial \ln I}}{1 - \frac{\tau}{\tau - g} \frac{\partial \ln Y}{\partial \ln I}} .$$

Thus, if  $g = 0$ , we obtain, upon substituting (3.14)

$$(3.16) \quad \tau = \frac{m_C + (1 - m_C) \frac{\partial \ln \bar{Y}}{\partial \ln I}}{\left( \frac{\partial \ln \bar{Y}}{\partial \ln (1 - \tau)} \right)_u} .$$

The expressions derived thus far have been general expressions, valid for any utility function and<sup>34</sup> valid regardless of the magnitude of the degree of inequality. There are two approaches that one might take at this point: one could specify a particular utility function and distribution of abilities and then compute precisely the optimal tax rate (depending on the social welfare function posited, and the level of government expenditures). Alternatively, we can use the approximations derived earlier to obtain approximate numerical magnitudes for the optimal tax rate and to ascertain what parameters are likely to be important in determining the optimal tax rate. This is the approach taken here.<sup>35</sup>

For deriving our approximations, the most convenient expression to use is (3.10). We then obtain, for small variance, with an additive utility function and assuming no government expenditure

$$(3.17) \quad \frac{\tau}{1 - \tau} \approx \frac{R s_Y^2}{u \bar{n}_L} .$$

With government expenditures we obtain

$$(3.18) \quad \frac{\tau}{1-\tau} = \frac{R s_Y^2}{\bar{u} n_L (1-g)} .$$

(3.17) displays several of the properties which we would expect of an optimal tax rate: the tax rate is higher the greater the degree of equality or equality aversion and smaller the magnitude of the compensated price elasticity. What this expression also makes clear is that the optimal tax rate is going to be extremely sensitive to estimates of the compensated price elasticity and the magnitude of inequality aversion ( $R$ ).

Estimates of the compensated supply elasticity vary greatly, although the estimates are almost all small. Ashenfelter and Heckman, for instance, estimate the compensated supply elasticity as .12, so that if  $R = 1$ ,  $s_Y = .4$ , and  $g = 0$ , we obtain as our optimal tax rate  $4/7$ . On the other hand, if we assume the compensated price elasticity is approximately one-half, then the optimal tax rate is .4, while if the compensated price elasticity is unity, the optimal tax rate is  $2/7$ . Higher degrees of inequality aversion will clearly result in much higher tax rates, making our approximations less valid.<sup>36</sup>

#### 4. How Much Good Does the Optimal Linear Income Tax Do?

In the first section of this paper, we derived an estimate of the social loss from ability inequality. How much is this reduced by the optimal linear income tax? Obviously, there is a gain in inequality at a cost of deadweight loss from the tax. As a first approximation, the deadweight loss is just equal to

$$\frac{\bar{Y} \tau^2}{2 \bar{u} n_L}$$

where  $\bar{n}_L$  is the compensated price elasticity. The gain relative to national income is equal to

$$\tau \bar{Y} m_Y \approx R s_Y^2 (1 - \tau) \tau \bar{Y} .$$

Thus the net gain is just (if  $g = 0$ )

$$\approx \tau^2 \bar{Y} \left\{ \frac{1 - \tau}{\tau} R s_Y^2 - \frac{\bar{n}_L}{2} \right\}$$

$$\approx \frac{\tau^2 \bar{Y}}{2} \bar{n}_L$$

$$\approx \frac{R^2 s_Y^4}{2 \bar{n}_L} ,$$

i.e., if  $s_Y = .4$ ,  $\bar{n}_L = .5$ ,  $R = 1$ , then the net social gain from imposing the optimal linear income tax (at a rate of 32%) is only 2.5% of net national product. Obviously, higher values of inequality aversion increase the net gains.

## 5. Concluding Comments

It is always difficult to think rationally about a subject which has all the political and moral overtones that inequality does. Is there “too much inequality,” as many economists and the popular press often suggest? What can one mean by such a statement? This paper has re-examined these issues, within a framework that has long been put forward as providing a justification for progressivity. Two results of the analysis may, at one level, seem particularly disturbing: (a) The “cost” of inequality may be much smaller when the sources of inequality are differences in ability rather than when it is some exogenously

provided wealth endowment; and (b) The net social gain from an optimal linear income tax may be relatively small.<sup>37 38</sup> These propositions will undoubtedly be the subject of debate; using other values of some of the parameters may yield different results.

But more than that, this analysis puts into question the usefulness of the standard optimal income tax framework for thinking about these essential questions. That framework has emphasized differences in wages (abilities) as giving rise to inequalities. But to the extent that inequality is the result of other factors—“luck,” the good fortune of a farmer having a good crop, of a real estate investor buying a plot of land next to where a road is constructed years later, of an entrepreneur coming up with the right product at the right time, or of an investor betting on the right stock—then the trade-offs on which we have focused are not the key determinants of the optimal income tax. Analyses based on these determinants of inequality would suggest far higher optimal income tax rates within a utilitarian framework.

Sen’s work, however, leads us to the conclusion that the conventional presumptions about the importance of inequality and the desirability of progressivity in taxes probably should be justified on grounds other than the individualistic utilitarian framework used by Edgeworth, Lerner, Samuelson, and others. We should focus our attention not so much on the distribution of income itself, but on the processes by which it is generated; that our concern should be more with equality of opportunity rather than with equality of incomes; that we should look at opportunities to participate in political processes, as well, and that the distribution of goods (economic power) has consequences for the distribution

of political power, with further consequences for the outcomes of these political processes; and, finally, that what is at stake is both the nature of society itself and the ability of individuals to live meaningful lives, living up fully to their potentials.

## Appendices

### A. Calculation of Social Loss From Inequality of Wages: No Redistribution

The measure of social loss is defined by

$$(A.1) \quad EV(w,0) = V(1, -\mu\bar{Y}).$$

Taking a Taylor series expansion, we obtain

$$(A.2) \quad V(w,0) \approx V(1,0) + V_1(w-1) + \frac{V_{11}}{2}(w-1)^2,$$

$$(A.3) \quad EV(w,0) = V(1,0) + \frac{s_w^2}{2}V_{11},$$

and

$$(A.4) \quad V(1, -\mu\bar{Y}) = V(1,0) - \mu\bar{Y}_2(1,0).$$

Thus,

$$(A.5) \quad \mu \approx -\frac{V_{11}}{2V_2} \frac{s_w^2}{\bar{Y}}$$

Our problem is to interpret  $V_{11}$ . Recall that

$$(A.6) \quad V_1 = LV_2$$

$$(A.7) \quad V_{12} = V_2 \frac{\partial L}{\partial I} + LV_{22}$$

$$V_{11} = LV_{21} + V_2 \frac{\partial L}{\partial w} = V_2 \left( \frac{\partial L}{\partial w} + L \frac{\partial L}{\partial I} \right) + L^2 V_{22}$$

Defining

$$R = - \frac{V_{22}^C}{V_2}, \quad \eta_L = \frac{\partial \ln L}{\partial \ln w}$$

we obtain

$$(A.8) \quad \mu \approx \frac{1}{2} \left( R - \frac{dL}{dI} - \eta_L \right) s_w^2$$

Thus, since the coefficient of variation

$$s_Y^2 = \frac{E(wL - EwL)^2}{(EwL)^2} \approx s_w^2 (1 + \eta_L)^2,$$

$$(A.9) \quad \mu \approx \frac{s_Y^2}{2} \frac{\left( R - \eta_L - \frac{dL}{dI} \right)}{(1 + \eta_L)^2}$$

If the individual has homothetic indifference curves between leisure goods and consumption and if we normalize the length of the day at unity, then (if  $w = 1$ )

$$-\left( \frac{dL}{dI} \right)_{I=0} = 1 - L$$

If there is a constant elasticity of substitution utility function of the form

$$[\delta C^{-\rho} + (1 - \delta)(1 - L)^{-\rho}]^{-\frac{1}{\rho}},$$

with  $1/(1 + \rho) = \sigma$ , the elasticity of substitution, then

$$\left[\frac{1 - L}{I + wL}\right]^{1+\rho} = \frac{1 - \delta}{\delta w}.$$

Straightforward differentiation leads to, when  $I = 0$ ,

$$\eta_L = -\frac{\rho}{1 + \rho}(1 - L) = (\sigma - 1)(1 - L).$$

Substituting into (A.8) we obtain (2.12''').

If there is an additive utility function,

$$U = u(C) - Z(L)$$

$$u'w = Z'$$

$$\left(\frac{\partial L}{\partial w}\right)_{I=0} = - \frac{u''C + u'}{u''w^2 - Z''}$$

$$\frac{dL}{dI} = - \frac{u''}{u''w^2 - Z''}$$

Hence

$$\left(\frac{\partial L}{\partial w} + L \frac{dL}{dI}\right)_{\substack{w=1 \\ I=0}} = - \frac{u'}{u''w^2 - Z''} (1 - 2R)$$

$$= \left(\frac{\partial L}{\partial w}\right) \left(\frac{1-2R}{1-R}\right) \text{ if } R \neq 1$$

## B. Calculation of Social Loss From Inequality: Linear Income Tax, Additive Utility

### Function

The measure of social loss is

$$(B.1) \quad EV((1 - \tau)w, I) = V(1, -\hat{\mu}\bar{Y})$$

or

$$(B.2) \quad \max Eu((1 - \tau)wL + \tau\bar{Y}) + Z(L) = \max U(L(1 - \hat{\mu})) + Z(L)$$

where  $Z(L)$  is the disutility of labor. The left-hand side can be approximated by

$$EV(w,0) + \tau \left( \frac{dEV}{d\tau} \right)_{\tau=0} + \frac{\tau^2}{2} \left( \frac{d^2EV}{d\tau^2} \right)_{\tau=0} .$$

At  $\tau = 0$ ,

$$\frac{dEV}{d\tau} = EU'(\bar{Y} - wL) + EU' \tau \frac{d\bar{Y}}{d\tau}$$

$$\approx -U'' \left( \frac{dY}{dw} (1 - \tau) \right) \frac{dY}{dw} s_w^2$$

$$= [R(1 + \eta_L)^2 s_w^2] U' \bar{Y} (1 - \tau)$$

$$\tau^2 \left( \frac{d^2EV}{d\tau^2} \right)_{\tau=0} \approx \tau^2 \frac{d\bar{Y}}{d\tau} U' = -U' \tau^2 u_L \eta_L \bar{Y} .$$

Thus

$$\hat{\mu} = \hat{\mu} + (R(1 + \eta_L)^2 s_w^2 \tau - \frac{\tau^2 u_L \eta_L}{2}) .$$

### C. Calculation of Social Loss (Gain): Equal Utility

We let the lump sum subsidy (tax)  $I(w)$  be such that

$$(C.1) \quad \frac{dV(w, I(w))}{dw} = V_1 + V_2 I' = 0$$

and

$$(C.2) \quad EI(w) = 0 \quad .$$

Thus

$$(C.3) \quad I' = -L \quad .$$

We define the social loss (gain) as  $I(1)$  since

$$EV(w, I(w)) = V(1, I(1)) \quad .$$

From (C.2)

$$I(1) + \frac{I''}{2} s_w^2 \approx 0$$

so

$$I(1) \approx -\frac{I''}{2} s_w^2$$

$$I'' = -\left(\frac{\partial L}{\partial w}\right)_{\bar{u}} \quad .$$

Hence, the social “gain” is just equal to

$$\frac{1}{2} \left( \frac{\partial L}{\partial w} \right)_{\bar{u}} s_w^2 .$$

D. Quadratic Utility Function, With Constant Marginal Disutility of Labor

$$U = \frac{b(\hat{c} - c)^2}{2} - L$$

Then

$$b(1 - \tau)w(\hat{c} - c) = 1$$

Hence

$$c = \hat{c} - \frac{1}{b(1 - \tau)w} = I + (1 - \tau)wL$$

$$L = \frac{\hat{c} - I}{(1 - \tau)w} - \frac{1}{b((1 - \tau)w)^2}$$

$$Y = wL = \frac{\hat{c} - I}{1 - \tau} - \frac{1}{b(1 - \tau)^2 w}$$

$$\bar{Y} = \frac{\hat{c} - I}{1 - \tau} - \frac{1}{b(1 - \tau)^2} \int \frac{1}{w} f(w) dw$$

assuming that the distribution is sufficiently concentrated that everyone works. (The modifications in the general case are straightforward but tedious.) Hence the national income identity can be written:

$$\tau \bar{Y} = \frac{\tau(\hat{c} - I)}{1 - \tau} - \frac{\tau A}{(1 - \tau)^2} = I + G$$

Or

$$I = \tau(\hat{c} - \frac{A}{1 - \tau}) - G(1 - \tau)$$

where

$$A = \int \frac{1}{bw} f(w) dw$$

The maximization problem thus becomes

$$\max \frac{1}{2b(1-\tau)^2} E \frac{1}{w^2} - EL$$

Since

$$L = \frac{\hat{c}}{(1-\tau)w} - \frac{\tau}{(1-\tau)w} \left( \hat{c} - \frac{A}{1-\tau} \right) - \frac{1}{b((1-\tau)w)^2} + \frac{G}{w}$$

$$\bar{L} = bA(\hat{c} + G) + \frac{A^2\tau b}{(1-\tau)^2} - \frac{bB}{(1-\tau)^2}$$

where

$$B = \frac{1}{b^2} \int \left( \frac{1}{w} \right)^2 f(w) dw$$

Thus optimality requires

$$\max -A(G + \hat{c}) - \frac{\tau A^2}{(1-\tau)^2} + \frac{3}{2} \frac{B}{(1-\tau)^2}$$

i.e.,

$$-\tau A^2 + \frac{3}{2} B - \frac{(1 - \tau)A^2}{2} = 0$$

or

$$\tau = \frac{3B}{A^2} - 1 .$$

Observe that

$$\frac{d\tau}{dG} = 0 .$$

so

$$\frac{dI}{dG} = -1 .$$

An increase in government expenditure reduces the progressivity of the tax.

### E. Cobb-Douglas Utility Function

$$U = \frac{\ln C}{b} + \ln(\hat{L} - L) .$$

Then

$$\frac{(1 - \tau)w}{bC} = \frac{1}{\hat{L} - L}$$

or

$$\begin{aligned} L &= \hat{L} - \frac{bI}{(1 - \tau)w} - bL \\ &= \frac{\hat{L}}{1 + b} - \frac{Ib}{(1 + b)(1 - \tau)w} , \end{aligned}$$

$$\bar{Y} = \frac{\hat{L}}{1 + b} - \frac{Ib}{(1 + b)(1 - \tau)}$$

(using the normalization,  $Ew = 1$ ). Thus

$$\tau\bar{Y} = \frac{\tau\hat{L}}{1 + b} - \frac{\tau Ib}{(1 + b)(1 - \tau)} = I + G$$

or

$$\frac{\tau\hat{L}}{1 + b} = G + I\left(1 + \frac{\tau b}{(1 + b)(1 - \tau)}\right)$$

or

$$I = \frac{(\tau\hat{L} - G(1 + b))}{(1 + b - \tau)}(1 - \tau) .$$

Also

$$C = \frac{1}{1+b} (\hat{w}\hat{L}(1-\tau) + I) = \frac{(1-\tau)}{(1+b)(1+b-\tau)} \cdot [\hat{L}(w(1+b-\tau) + \tau) - G(1+b)]$$

Our problem is thus to maximize

$$\phi(\tau, G) \equiv E \ln [\hat{L}(w(1+b-\tau) + \tau) - G(1+b)] - \ln(1+b-\tau) + \frac{\ln(1-\tau)}{1+b}$$

$$\frac{d\tau}{dG} = - \frac{\phi_{\tau G}}{\phi_{GG}} > 0,$$

since

$$\phi_G = -E \frac{1+b}{\hat{L}(w(1+b-\tau) + \tau) - G(1+b)}$$

$$\phi_{G\tau} = E \frac{(1+b)\hat{L}(1-w)}{[\hat{L}(w(1+b-\tau) + \tau) - G(1+b)]^2} > 0$$

When  $G = \hat{L}/(1+b)$ ,  $I = 0$ . We can solve for the value of  $\tau$  for which  $\phi_{\tau} = 0$ , when  $G = 0$ .

$$\phi_{\tau} = E \frac{1-w}{w(1+b-\tau)} + \frac{1}{1+b-\tau} - \frac{1}{(1+b)(1-\tau)} = 0$$

or

$$\frac{\tau}{1-\tau} = \frac{1+b}{b} (E \frac{1}{w} - 1)$$

If there is a lognormal distribution, and if individuals in the absence of taxation would have spent approximately half of the available time working (so  $b = 1$ ), then

$$\frac{\tau}{1 - \tau} = 2s_w^2 .$$

If  $s_w = .4$ ,  $\tau = .24\%$ .

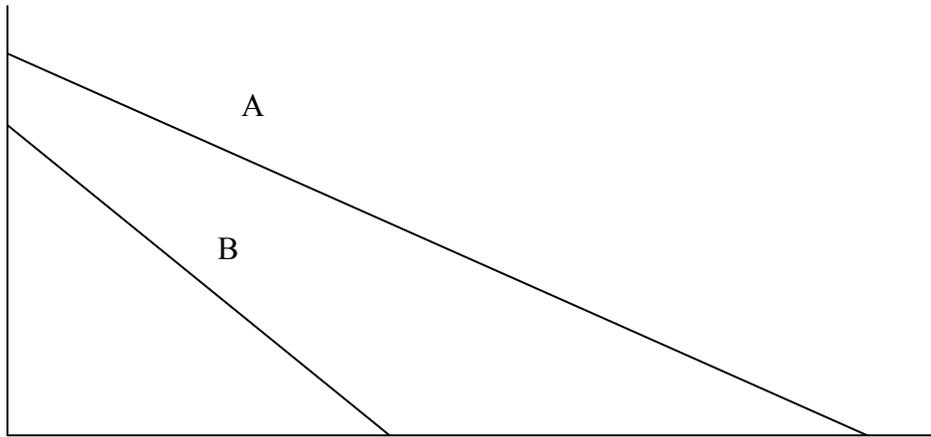
Table 1

Social Losses from Inequality,  $s_y = .4$

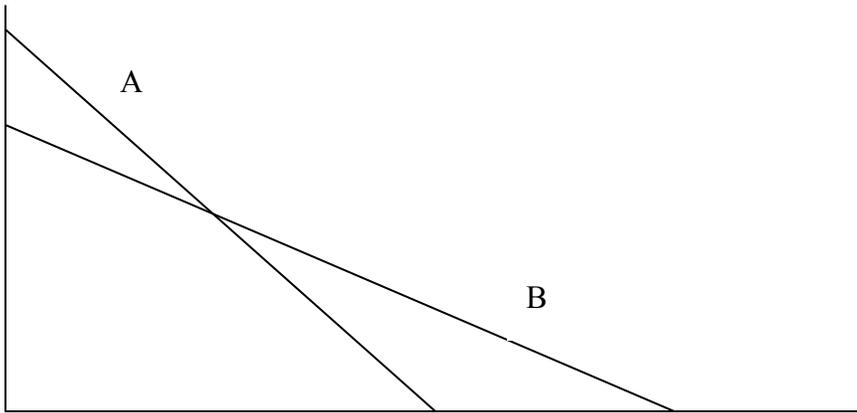
$$R = 1, \eta_L = 0, \bar{u}_L = \frac{1}{2} \qquad R = \frac{3}{2}, \eta_L = \frac{1}{4}, \bar{u}_L = \frac{1}{2}$$

Atkinson measure	8%	12%
No tax: relative to zero inequality	12%	7.7%
Optimal linear income tax rate	32%	48%
Gains from optimal linear tax	2.6%	5.8%
Egalitarian lump sum tax: relative to zero inequality	-4%	-2.6%

Figure 1



Panel A. Opportunity set “A” is better than Opportunity Set “B”: all individuals (regardless of their preferences) would prefer A over B.



Panel B: When the budget constraints cross, one cannot rank the opportunity sets: some individuals would prefer “A” to “B”; others would prefer “B” to “A”.

Figure 2

Measuring Inequality with Different Wages: All Individuals have the same endowment of leisure

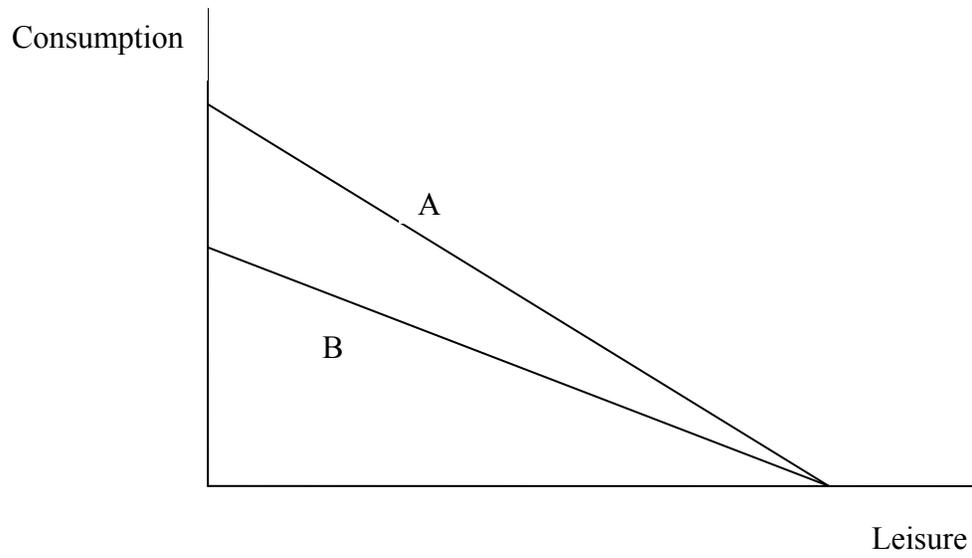
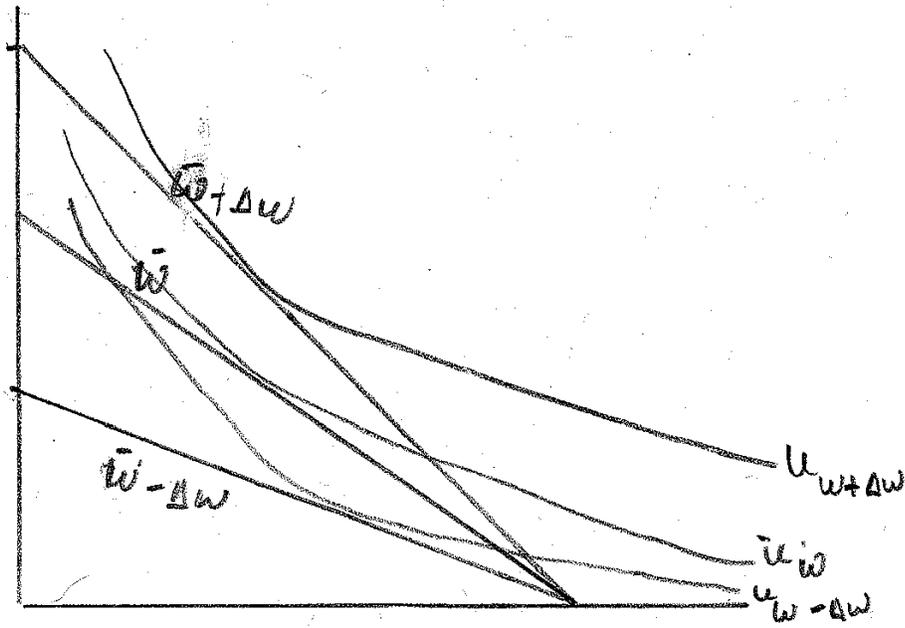


Figure 3

Increasing variability in wages may lead to an increase in welfare. The loss in welfare from a reduction in the wage is much smaller than the increase in welfare from a symmetric increase in the wage.





---

<sup>1</sup> The author is indebted to P. Diamond, J. Flemming, J. Mirrlees, Arjun Jayadev and R. Kanbur for helpful discussions. This work was supported in part by the National Science Foundation, the MacArthur Foundation, and the Ford Foundation. The author is University Professor, Columbia University.

<sup>2</sup> For a broader discussion, see J. E. Stiglitz, “The Post Washington Consensus Consensus,” in *The Barcelona Consensus*, edited by N. Sarra and J. E. Stiglitz, Oxford University Press, 2008.

<sup>3</sup> See, for instance, P. Dasgupta, A. Sen, and D. Starrett, “Notes on the Measurement of Inequality,” *Journal of Economic Theory*, 6(2), April 1973, pp. 180-187, and M. Rothschild and J. E. Stiglitz, “Some Further Results on the Measurement of Inequality,” *Journal of Economic Theory*, 6(2), April 1973, pp. 188-204. Rothschild and I showed that Atkinson’s results held for a much more general class of social welfare functions than he had assumed—they hold for any symmetric, monotone, locally equality preferring social welfare function (i.e. separability is not required, nor even quasi-concavity.) Our analysis also highlighted the problems in measuring inequality which arise when there are more than one commodity and relative prices are not fixed—one of the subjects on which this essay focuses.

<sup>4</sup> This paper is a revised and extended version of an earlier working paper of the same title, IMSSS #215., Stanford University, August, 1976. The fuller version of the paper provides alternative formulae.

<sup>5</sup> M. Rothschild and J.E. Stiglitz, “Increasing Risk: I. A Definition,” *Journal of Economic Theory*, 2(3), September 1970, pp. 225-243, and M. Rothschild and J.E. Stiglitz, “Increasing Risk: II. Its Economic Consequences,” *Journal of Economic Theory*, 5(1), March 1971, pp. 66-84.

<sup>6</sup> A.B. Atkinson, “On the Measurement of Inequality,” *Journal of Economic Theory*, 2(3), September 1970, pp. 244-63.

<sup>7</sup> S. Kolm, “The Optimal Production of Social Justice,” J. Margolis and H. Guitton (eds.), *Public Economics*, London: Macmillan, 1968.

<sup>8</sup> It should be clear, however, that the focus on economic opportunity sets touches on only one aspect of Sen’s broader conception that includes participation rights and opportunities, which are seen as more than just instrumental in ensuring equality (or fairness) in outcomes.

<sup>9</sup> There are, of course, deeper concerns: did those individuals who “chose” not to work really have the freedom to make those choices, given their upbringing, their inherited psychological make-up, etc.?

<sup>10</sup> Of course, many goods provided by the public sector are really publicly provided private goods (in the terminology of Atkinson and Stiglitz), and there needs to be an independent assessment of how these are distributed within a society. But even pure public goods can be goods that are more valued by the poor or by the rich, and therefore, in a fundamental sense, either reduce or increase, the overall degree of inequality in society.

<sup>11</sup> Corneo and Fong find, on the basis of survey data, that on average, individuals would be willing to sacrifice 20% of income to eliminate inequality. Giacomo Corneo and Christina Fong “What’s the Monetary Value of Justice?” *Journal of Public Economics*, forthcoming.

<sup>12</sup> H. Dalton, “The Measurement of the Inequality of Incomes,” *Economic Journal*, 30(119), 1920, pp. 348-361.

<sup>13</sup> Sen himself provided a critique of the Atkinson measure in *Inequality Reexamined*, Cambridge. Harvard University Press, 1992, pp. 97-99.

<sup>14</sup> Without adverse incentive effects. The UN has promoted a “Human Security” agenda, in which Amartya Sen has played an important role. See the UN Commission on Human Security, chaired by Sadako Ogato and Amartya Sen and its final report, *Human Security Now: Protecting and Empowering People*, United Nations: New York, 2003. The World Bank’s *World Development Report, 2000* emphasized the importance of insecurity among the poor.

<sup>15</sup> See, in particular: C.I. Lee and G. Solon, “Trends in Intergenerational Income Mobility,” unpublished manuscript, 2004, University of Michigan; D.I. Levine and B. Mazumder, “Choosing the Right Parents: Changes in the Intergenerational Transmission of Inequality Between 1980 and the Early 1990s,” Federal Reserve Bank of Chicago Working Paper 2002-08, June 2002; B. Mazumder, “Earnings Mobility in the US: A New Look at Intergenerational Inequality” Federal Reserve Bank of Chicago Working Paper, 2001-18, March 2001; S. Bowles and H. Gintis, “The Inheritance of Inequality,” *Journal of Economic Perspectives*, 16(3), Summer 2002, -pp. 3-30; L. Chadwick and G. Solon, “Intergenerational Income Mobility Among Daughters,” *American Economic Review*, 92(1), March 2002, pp. 335-344; A. Fertig, “Trends in Intergenerational Mobility.” Center for Research on Child Wellbeing. Working Paper #01-23, Princeton University, 2001; T. Hertz, “Understanding Mobility in America,” Center for American Progress and American University, April 26, 2006; Daniel P. McMurrer and Isabel V. Sawhill, *Getting Ahead: Economic and Social Mobility in America*. Urban Institute Press, 2007. ck

<sup>16</sup> See, for instance, R. Kanbur and J. E. Stiglitz, “Mobility and Inequality: A Utilitarian Analysis,” Economic Theory Discussion Paper 57, University of Cambridge, May 1982 and R. Kanbur and J.E. Stiglitz, “Intergenerational Mobility and Dynastic Inequality,” Princeton University Economic Research Program Research memorandum No. 324, April 1986. .

<sup>17</sup> See F.Y. Edgeworth, “The Pure Theory of Taxation,” *Economic Journal*, 7(28), December 1897, pp. 550-571, and resuscitated in the “New Public Finance Literature,” of Mirrlees, Diamond-Mirrlees, Stiglitz-Dasgupta, Atkinson-Stiglitz, Feldstein, and others.

<sup>18</sup> See in particular, Martin Ravallion “Inequality Is Bad for the Poor,” in *Inequality and Poverty Re-Examined*, edited by J. Micklewright and S. Jenkins Oxford: Oxford University Press, 2007; Martin Ravallion, “Economic Growth and Poverty Reduction: Do Poor Countries Need to Worry about Inequality?” 2020 Focus Brief on the World’s Poor and Hungry People. Washington, DC: IFPRI, October 2007; and C. Citro and R. Michael, eds., *Measuring Poverty: A New Approach*, National Academies Press: Washington, DC, 1995.

<sup>19</sup> For the lognormal distribution with a logarithmic utility function, we can calculate  $\hat{\mu}$  precisely:

$$\mu = 1 - \frac{1}{(1 + s_Y^2)^{\frac{1}{2}}}$$

$$\hat{\mu} = 1 - \frac{1}{(1 + s_Y^2)^{\frac{1}{2}}}$$

If, for instance,  $s_Y = .5$  and  $\hat{\mu} = .11$

<sup>20</sup> If we wish to take the Taylor series expansion one more term, we obtain

$$(2.4') \quad \hat{\mu} \approx \frac{-1 \pm \sqrt{1 + 2R^2 s_Y^2 - 2\left(\frac{RU'''}{3U'Y}\right)E(Y - \bar{Y})^3}}{2R}$$

If  $R = 1$  and we have a lognormal distribution with  $s_Y = .5$ , the approximation (2.4) gives

$$\hat{\mu} \approx .125$$

while the approximation (2.4') gives

$$\hat{\mu} \approx -1 + \sqrt{1.12}$$

The first approximation is slightly in excess, the second slightly less than true value.

<sup>21</sup> For the lognormal distribution with a logarithmic utility function

$$m = \frac{\bar{Y}E\frac{1}{Y} - 1}{\bar{Y}E\frac{1}{Y}} = \frac{e^{\sigma^2} - 1}{e^{\sigma^2}} = \frac{s_Y^2}{1 - s_Y^2}$$

Thus, if  $s_Y = .5$ ,  $m = 1/3$ .

We can calculate a higher order approximation:

$$m \approx R s_Y^2 + \frac{U'''}{6U'Y} E(Y - \bar{Y})^3$$

<sup>22</sup> We could also ask, when will one distribution of goods-cum-leisure be more equal than another? This is a question which Kolm has addressed. It is clear that looking at the distribution of each good (leisure) separately is not appropriate. If there were two goods ( $x_1, x_2$ ), then in A, the distribution of  $x_1$  could be more unequal (in the sense of Rothschild-Stiglitz in "Some Further Results on the Measurement of Inequality," *op cit.*) than in B, and the distribution of  $x_2$  could be more unequal than in B. But if utility,  $u = x_1 + x_2$ , it is clear that  $u$  may be more equally distributed. If everyone has the same tastes, the utility function provides the natural basis for aggregation.

<sup>23</sup> This sounds like a proposition which could easily be verified. Unfortunately, this is not the case, for what we mean by the "wage" rate is the payment for a standard unit of effort for one hour. Although differences in wages per hour are easy to observe, differences in effort are difficult, if not impossible, to observe and quantify.

<sup>24</sup> For a discussion of the concept of a mean preserving increase in inequality, see Rothschild and Stiglitz, "Some Further Results on the Measurement of Inequality," *op cit.*

<sup>25</sup> There have been numerous extensions in understanding mean preserving changes in distributions, notably in the literature on income polarization. These (like the earlier cited references on the role of relative deprivation) highlight that there may be mean preserving changes in distributions which raise inequality but which also increase, or decrease, polarization, or increase, or decrease, relative deprivation. These important dimensions of changes in income distribution are not captured in the standard utilitarian model, and in measures that are derived from that model. See, for instance, J.Y. Duclos, J. Esteban, and D. Ray, "Polarization: Concepts, Measurement, Estimation," *Econometrica*, 72(6), November 2004, pp. 1737-

1772; and J. Esteban and D. Ray, "On the Measurement of Polarization," *Econometrica* 62(4), July 1994, pp. 819-851.

<sup>26</sup> It should be clear by now that the problem of ability inequality is much like the problem of price uncertainty, just as the problem of income inequality is much like the problem of income uncertainty. See, for instance, J.E. Stiglitz, "Portfolio Allocation with Many Risky Assets," in Szego-Shell (eds.) *Mathematical Methods in investment and Finance*, North Holland Publishing Co, 1972, and Rothschild and Stiglitz, "Further Results in the Measurement of Inequality," *op cit.*

<sup>27</sup> Saez calculates optimal tax formulas based on compensated and uncompensated elasticities of earnings with respect to taxes. E. Saez, "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies*, 68(1), January 2001, pp. 205-229. Other recent theoretical explorations of the issue of optimal taxation include P. Diamond, "Optimal Income Taxation: An Example with a U-shaped Pattern for the Optimal Marginal Rates," *American Economic Review*, 88(1), March 1998, pp. 83-95. R. Kanbur and M. Tuomala, "Inherent Inequality and the Optimal Graduation of Marginal Tax Rates," *Scandinavian Journal of Economics*, 96(2), 1994, pp. 275-282; and R. Kanbur, M. Keen, and M. Tuomala, "Optimal Non-Linear Income Taxation for the Alleviation of Income Poverty," *European Economic Review*, 38(8), October 1994, pp. 1613-1632.

<sup>28</sup> J.A. Mirrlees, "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 38(2), April 1971, pp. 175-208.

<sup>29</sup> We are thus assuming all individuals of a given ability are identical. Alternatively, we could interpret  $L_w$  as the *average* labor supplied by individuals of wage level  $w$ .

<sup>30</sup> The government's social welfare maximization problem is actually slightly different: It is to maximize  $E\{W(U^*)\}$ , which differs from (3.7) only because of the valuation of the public good. So long as  $G$  is fixed, the problems are identical. However, when  $G$  changes, one must be careful in interpreting the results. If  $W'' = 0$ , then the analysis proceeds unchanged; but if  $W'' < 0$ , an increase in  $G$  affects the covariance (in the notation below,  $m_\beta$ ).

<sup>31</sup> In a more general analysis, there are several effects going on. Assume that, at one extreme, public goods were a perfect substitute for private goods, i.e. the utility function was of the form  $U(C + G, L)$ . Then an increase in  $G$  is equivalent to increasing the endowment of "I" by the same amount. More generally, a change in the supply of a public good can affect the marginal rate of substitution between private goods and leisure. (In the case of the separable utility function on which we focus, this effect is absent.)

<sup>32</sup> (3.14) can be derived directly from our maximization problem: substituting (3.6) into (3.7) we obtain

$$\max_{\tau} \int W(U(1 - \tau)wL + \tau(\bar{Y} - G), L) f(w) dw$$

or

$$-\int W'_1 U_1 (Y - \bar{Y}) f(w) dw + \tau \frac{d\bar{Y}}{d\tau} \int W'_1 U_1 f(w) dw = \bar{Y}_g \int W'_1 U_1 f(w) dw$$

$$\frac{d \ln \bar{Y}}{d \ln \tau} = \frac{\partial \ln \bar{Y}}{\partial \ln I} \frac{d \ln I}{d \ln \tau} + \frac{\partial \ln \bar{Y}}{\partial \ln \tau} ;$$

from the national income identity

$$\frac{dI}{d\tau} = \tau \frac{d\bar{Y}}{d\tau} + \bar{Y}$$

so

$$\frac{d \ln \bar{Y}}{d \ln \tau} (1 - \tau \frac{d\bar{Y}}{dI}) = \tau \frac{\partial \bar{Y}}{\partial I} + \frac{\tau}{\bar{Y}} \frac{\partial \bar{Y}}{\partial \tau} .$$

<sup>34</sup> With the exception of expression (3.12')

<sup>35</sup> Further expressions, in terms of the consumption elasticities, may be derived in a straightforward manner. Since, however, numerical estimates tend to be expressed in labor supply elasticities, we have chosen to express our formulae in those terms.

<sup>36</sup> There are several reasons for believing that conventional cross section estimates of the supply elasticity of labor may not provide an accurate estimate the magnitude of the elasticity. First, if real wages consist of a pecuniary and nonpecuniary component

$$w^* = w + \varepsilon$$

with  $\varepsilon$  uncorrelated with  $w$ , then very high wages will on the average represent a lower real wage, and low wages will represent a higher real wage. Thus, for reasons analogous to those analyzed by Friedman in his study of the consumption function, the estimated elasticity of the supply of labor will be biased downwards.

One of the reasons that real wages and measured wages will differ, besides non-pecuniary characteristics of the job, is that some jobs have more training associated with them. Evidence of the potential importance of this effect is provided by the differences in supply elasticities calculated for older workers.

Secondly, if there are some jobs which systematically have long work weeks (doctors) and others which have short work weeks (differences in training costs would explain why such differences are to be expected) then, if the jobs were otherwise identical, on the average those individuals with a smaller aversion to work would choose the longer work week jobs. Thus, the differences in the wages between the two jobs are less than the difference in wages that would be required to induce an individual with a low work week job to work a long week.

<sup>37</sup> A third set of results, on the relationship between the degree of progressivity and the level of government expenditures, is highly dependent on the assumptions of separability between government expenditures on the one hand, and consumption and leisure on the other. How the result is altered under more general assumptions is a subject of ongoing enquiry.

<sup>38</sup> Matters may even be worse. Stiglitz showed that the optimal income tax, within a utilitarian framework with two classes (skilled and unskilled workers), entails negative marginal tax rates at the top (J. E. Stiglitz, "Pareto Efficient Taxation and Expenditure Policies, With Applications to the Taxation of Capital, Public Investment, and Externalities," presented at conference in honor of Agnar Sandmo, January 1998.)