Comparing Optimum Non-linear Income Taxation with Optimum Linear Income Taxation: A Numerical Analysis

Seon-Kyou Choi*

As income tax regime changes, the welfare levels of different types of individuals also change. This paper studies the welfare change of individuals with different abilities under different income tax *regimes*: the non-linear income tax system and the linear income tax system. Overall level of social welfare is enhanced in the non-linear system. Though the marginal tax on the high ability people is zero under non-linear income taxation, the simulation results show that high ability type is worse off under reasonable estimates of elasticity of substitution between consumption and leisure. The low and middle ability types are better off under the non-linear system. When the elasticity of substitution is low, non-linear income taxation is Pareto-superior. It makes all types of individuals better off than linear income taxation. (*JEL* Classification: H21)

I. Introduction

The basic theorem of welfare economics suggests that, under standard assumptions, the first-best can be achieved by lump-sum transfers for each individual. However, the calculation of appropriate lump-sum transfers requires information on individuals which they have an incentive not to reveal. This has lead to the theory of optimal income taxation in which all individuals are subject to the same tax schedule, where we assume that only income is observable. The tax schedule is chosen to maximize welfare.

In the linear income tax system, income can be taxed at source. That

*Korea Information Society Development Institute, 966-5 Taechi-dong Kangnam-gu Seoul, Korea. This paper is based on the part of my Ph. D Dissertation. I thank John Wilson, David Wildasin and two anonymous referees for their helpful comments.

[**Seoul Journal of Economics** 1993, Vol. 6, No. 2]

is to say, a dollar of income is taxed at the same rate regardless of the fact who has the dollar of income. The non-linear income tax system, however, taxes a dollar of income depending on the level of individual income. It requires the knowledge of individual income, and individuals have an incentive to be misleading when reporting those incomes to avoid taxes. If one abstracts from such difficulties, an optimal non-linear system can never be worse than an optimal linear system, because the government maximizes welfare with a larger set of available tax instruments in the optimal non-linear system.

As one goes from the linear system to the non-linear system, the level of overall social welfare undoubtedly increases. It is not clear, however, how the welfare of each individual is affected. Does the policy that leads to increased social welfare also improve individual welfare for all types of people? If not, who is better off and who is worse off? This is something we wish to investigate in this paper.

Since the work of Mirrlees (1971), there have been a number of studies on optimal income taxation. In a continuous model where the skill of individuals is continuously distributed, optimal linear taxation was studied by Stern (1976) using constant elasticity of substitution utility function. In non-linear income taxation, Seade (1977) argued that if in the optimum each component of consumption vector is bounded away from zero over the population, each marginal tax must be zero at both ends of the corresponding tax schedule. Tuomala (1984) studied how marginal tax rates varied with income. He found that the zero marginal tax result is very local: marginal tax rates are significantly positive near the top income level. In a discrete model, Stiglitz (1982) studied optimal non-linear taxation when there are two types of individuals. Stern (1982) compared the welfare level of first-best income taxation with errors in classification of types with that of optimal non-linear income taxation and that of optimal linear income taxation. He achieved the conclusion that the level of social welfare for the first-best tax with random classification is similar to that of non-linear taxes.

In the research on optimal income taxation, explicit attention has not been paid to the welfare of each individual. Most works have focused on the structure of marginal tax rates (Seade 1977; Tuomala 1984) and the comparison of overall social welfare (Stern 1982). A comparison of welfare between different age cohorts of individuals has been done in the work of Auerbach and Kotlikoff (1987) when the government switches from the income tax to the consumption tax.

In addition to comparing individual welfare, my research can make a contribution to the study of the tax reforms. One of the proposals in tax reform is that we should move from progressive income taxes (defined in terms of increasing marginal tax rates) to the flat rate income tax with an exemption. Some economists have proposed a flat rate income tax reform for its simplicity of administration because people will reduce their attempts to avoid taxes. Hausman (1981) also claims that it may be possible, by introducing a flat rate tax, to lower the tax rate facing almost all income groups. The great deadweight loss for the high ability workers in the progressive income tax will be significantly reduced in the linear income tax. There is no universal consensus on this matter. However, we can think of another possibility of a reform of the tax structure. One may gain more benefit by reforming any existing tax structure to the optimum non-linear income tax, if the dollar equivalent value of the welfare gains of optimum non-linear income taxation over flat rate income taxation is substantial enough to cover the administrative costs of non-linear income taxation. In this case, one should adopt the non-linear tax rather than the flat rate income tax. We do not attempt to estimate the administrative costs in this paper, but using the estimate of Slemrod and Sorum (1984), we will show that the welfare gain is greater than their estimate of the administrative costs.

II. Model

There are three types of individuals, highly skilled, moderately skilled, and low skilled, indexed by H, M, and L. The advantages of working with three groups over two groups is that we can see how taxes affect the moderately skilled. It also distinguishes my research from the Stern's article (1982). There are two goods, one consumption good and leisure. A consumption good is produced by labor of three types. Everybody is endowed with 1 unit of leisure. Each worker's labor is a perfect substitute for all other labor. Wage rates and labor supplied by individuals are assumed to be unobservable to the government, but income is observable. Each person has the same utility function and an individual of type i maximizes a utility function $U(C_i, L_i)$ subject to the budget constraint. L_i is the amount of labor supplied, C_i his consumption. The indices i take the values H, M, or L. There are α individuals of type H, β individuals of type M, and $3\text{-}\alpha\text{-}\beta$ individuals of type L. (The reason why we use 3 rather than 1 as the total population is that it makes the social welfare maximand simpler form. We can use 1 as the total population without loss of generality, getting the same result.)

Output Y is a function of total labor supplies of each type αL_H, βL_M, and $(3-\alpha-\beta)L_L$. We suppose that the production has the special linear form

$$Y = k_o\,[\gamma\alpha L_H + \delta\beta L_M + (1-\gamma-\delta)(3-\alpha-\beta)L_L],$$

where, γ and δ are competitive shares of the highly skilled and the moderately skilled.

The upshot of a linear production function is that wage rates are constants. It reduces the complexity of the problem, though it ignores the effect of income tax on wage rates. The production function is constant return to scale and the market is perfectly competitive, so that total payments to labor are equal to output.

Suppose that the utility function $U(\cdot)$ has the constant elasticity of substitution form

$$U(C,L)=[\frac{1}{2}C^{-\mu}+\frac{1}{2}(1-L)^{-\mu}]^{-\frac{1}{\mu}},$$

where, the elasticity of substitution ε is equal to $1/(1+\mu)$.

A. The Linear Income Tax System

In the linear income taxation, the government taxes income at the flat rate, and it gives individuals the uniform lump-sum grant. The individual budget constraint is

$$C_i = (1-t)w_i L_i + G,$$

where, w_i is the hourly wage of labor type i, t is the flat marginal tax rate, and G is the uniform lump-sum grant for all individuals.

The indirect utility function corresponding to the CES utility function is given as

$$V(t,G)=(\frac{1}{2})^{\frac{\varepsilon}{\varepsilon-1}}\cdot\{w(1-t)+G\}\cdot[1+\{w(1-t)\}^{1-\varepsilon}]^{\frac{1}{\varepsilon-1}},$$

V is proportional to after-tax 'full' income, $w(1-t)+G$, (after-tax market value of labor endowment of one unit plus lump-sum grant).

The labor supply function has the form of

$$L(t,G)=\frac{1-\dfrac{G}{(1-t)w^{\varepsilon}}}{1+\{w(1-t)\}^{1-\varepsilon}}.$$

The government budget constraint is

$$\alpha G + \beta G + (3\text{-}\alpha\text{-}\beta)G = tY - R,$$

where, R is the revenue requirement.

The government's objective is to maximize W by choosing t and G:

$$\underset{t,G}{\text{Max}}\ W = \frac{1}{\upsilon}[\alpha V_H^{\upsilon} + \beta V_M^{\upsilon} + (3-\alpha-\beta)V_L^{\upsilon}]$$

subject to

$$V_i(t,G) = (\frac{1}{2})^{\frac{\varepsilon}{\varepsilon-1}} \cdot \{w_i(1-t)+G\} \cdot [1+\{w_i(1-t)\}^{1-\varepsilon}]^{\frac{1}{\varepsilon-1}}$$

$$\alpha G + \beta G + (3-\alpha-\beta)G = t[w_H \alpha L_H + w_M \beta L_M + w_L(3-\alpha-\beta)L_L] - R$$

$$L_i(t,G) = \frac{1 - \dfrac{G}{(1-t)w_i^{\varepsilon}}}{1+\{w_i(1-t)\}^{1-\varepsilon}},$$

where, W is a social welfare function; V_i is the indirect utility function corresponding to $U(C_i, L_i)$; L_i, the labor supply function for i individual, and υ is a parameter indicating the government's concern about inequality in utility levels.

B. The Non-linear Income Tax System

In the non-linear income tax system, the marginal tax rates are allowed to change with income levels. Since the tax rates vary with level of income, one have an incentive of disguising himeself as being other type. To overcome this adverse selection problem, the tax system should be incentive compatible. To derive the optimal non-linear income tax structure, the government must use 'the truth-telling mechanism' or 'the revelation principle'. The three groups have to be given choices, such that no one in a group prefers the choice of the other group. That is to say, the government offers three consumption and income packages to individuals such that no individual in a group prefers the consumption and income packages of other groups, and the packages are technologically feasible. The government, then finds the income tax function that supports the consumption and income packages. Formally, the government first solves the following problem.

$$\underset{C_i,Z_i}{\text{Max}}\ W = \frac{1}{\upsilon}[\alpha U^{\upsilon}(C_H, \frac{1}{w_H}Z_H) + \beta U^{\upsilon}(C_M, \frac{1}{w_M}Z_M) + (3-\alpha-\beta)U^{\upsilon}(C_L, \frac{1}{w_L}Z_L)]$$

subject to

$$U(C_H, \frac{1}{w_H} Z_H) \geq U(C_M, \frac{1}{w_H} Z_M) \quad (1)$$

$$U(C_H, \frac{1}{w_H} Z_H) \geq U(C_L, \frac{1}{w_H} Z_L) \quad (2)$$

$$U(C_M, \frac{1}{w_M} Z_M) \geq U(C_L, \frac{1}{w_M} Z_L) \quad (3)$$

$$U(C_M, \frac{1}{w_M} Z_M) \geq U(C_H, \frac{1}{w_M} Z_H) \quad (4)$$

$$U(C_L, \frac{1}{w_L} Z_L) \geq U(C_H, \frac{1}{w_L} Z_H) \quad (5)$$

$$U(C_L, \frac{1}{w_L} Z_L) \geq U(C_M, \frac{1}{w_L} Z_M) \quad (6)$$

$$k_0\{\alpha Z_H + \beta Z_M + (3-\alpha-\beta) Z_L\} = R + \alpha C_H + \beta C_M + (3-\alpha-\beta) C_L \quad (7)$$

where, Z_i is the pre-tax income, $Z_i = w_i L_i$.

Constraints (1) to (6) are the self-selection constraints, and (7) the production constraint. The self-selection constraints are also called the incentive compatibility constraints.

After finding the allocation that is incentive compatible and technologically feasible, we find the tax function which supports the allocation. Formally, we find $T(Z)$ such that $(C_i^*, Z_i^*/w_i)$ is a solution to

$$\underset{C_i, Z_i}{\text{Max}} \; U(C_i, Z_i / w_i)$$

subject to $C_i = Z_i - T(Z_i)$.

$T(\cdot)$ gives the optimum non-linear income tax schedules.

We will discuss the self-selection constraints. The self-selection constraints require that no individual in a given group should have an incentive to choose the income of other group via his choice of work. The constraints (1) to (3) indicates that the more skilled doesn't choose the consumption and income package of the less skilled; The constraints (4) to (6) mean that the less skilled does not envy the packages of the more skilled. With the concave social welfare function, (thus with a lower social marginal utility of consumption for the more skilled), we have

$$W_H \frac{\partial U}{\partial C_H}(C_H, L_H) < W_M \frac{\partial U}{\partial C_M}(C_M, L_M) < W_L \frac{\partial U}{\partial C_L}(C_L, L_L), \quad (8)$$

where, W_t is a partial derivative of W with respect to $U(C_i, L_i)$.

It will be (1) to (3) rather than (4) to (6) that will be relevant (See Stiglitz 1982; Stern 1982). Intuitively speaking, with a decreasing social marginal utility of income, taxation involves redistribution of income from the more skilled to the less skilled. It will be the more skilled that will be worse off because of taxation and considers taking the consumption and income package of the less skilled.

If the government maximizes the maximand W subject only to the production constraint (7), the solution corresponds to the first-best allocation. Mirrlees (1974) has shown that if and only if leisure is a normal good, then the first-best optimum has utility decreasing in skill. In the first-best, the more skilled has the lower utility than the less skilled. The more skilled always can disguise himself as the less skilled only making himself better off. It pays to disguise oneself as the less skilled type in the first-best policy. The first-best allocation is not incentive compatible. The government can not achieve the first-best allocation, unless it has the information about the individual ability.

In the second-best optimal income taxation, at least one of the self-selection constraints, (1) to (3) should bind at the optimum. If all of them hold with strict inequalities, the government can make a lump-sum transfer from the more skilled to the less skilled while preserving the conditions for income taxation. If (8) holds, there is an increase in welfare. Hence, at least one self-selection constraint should bind at the optimum. With specific functional forms on utility and production, one can check, ex post, whether (8) does in fact hold for any solutions of (1) to (7) together with first-order conditions.

To determine which constraint binds at the optimum, we proceed as follows. First, we solve the maximization problem with only one binding constraint. For solutions we check, ex post, whether the solutions satisfy the other two non-binding constraints. If the solution satisfies the other constraints, it is the actual legitimate solution to the optimal income tax, and the process of choosing the right binding constraint stops. But, if the solution violates the self-selection constraints, it is discarded, and we go on to solve the maximization problem with two binding constraints and we repeat the process of checking, ex post, whether the solution satisfy the other non-binding constraint. If no solution satisfies the self-selection constraint, then we move on to the maximization problem with three binding constraints and the solution is selected as the allocation to the optimal income tax system.

With specific functional forms of utility function and production function, the simulation results show that (1) and (3) do bind at the optimum. The constraint, (2) holds with a strict inequality, if (1) and (3) bind. With only one binding constraint, we get a higher maximand but the other two self-selection constraints are violated at those allocations.

C. Equally Distributed Leisurely-equivalent Consumption and Equivalent Variation

We can compare the social welfare between the linear system and the non-linear system, using the notion of the equally-distributed, leisurely-equivalent consumption $°C$, as used in Stern (1982) and defined as follows. Given a certain pattern of utilities resulting from the tax systems, we assign to social welfare W the number $°C$ which is the consumption which, if equally distributed, and when hours of work were zero for everyone, would give social welfare level W.

Levels of the individual's utility between each system are compared using the notion of the equivalent variation. The equivalent variation is the amount of lump-sum income that the government has to give to an individual in the absence of taxation to keep him equally as happy as if he were under taxation. The positive amount of equivalent variation means that the individual is better off after taxation, and vice versa. The larger the equivalent variation, the better off he is after taxation.

The deadweight loss of taxation can be calculated easily using the equivalent variation. The deadweight loss of a distortionary taxation as defined by Kay (1980) is the difference between his tax payments (in the distortionary taxation) and minimal amount of income that the government needs to take away (if the government used non-distortionary taxation) to have an individual reach the same after-tax utility as the distortionary taxation. For the person who is worse off because of a taxation, the equivalent variation for him is negative. The negative number of the equivalent variation is the minimal amount of income that the government needs to take away. The difference between the negative number of the equivalent variation and the tax payment is, thus, the deadweight loss of taxation. For the person who is better off because of a taxation his tax payments (usually negative numbers) will be greater in an absolute term than the equivalent variation (positive numbers for him). It means that the person needs extra amount of income over the equivalent variation because of a distortionary effect of a taxation.

There are a number of parameters to be varied in the simulation; υ, which measures attitudes toward inequality; R, the government revenue requirement; ε, the elasticity of substitution between consumption and leisure; γ and δ, the share of the highly skilled and moderately skilled, respectively; α and β, the proportion of the highly skilled and moderately skilled, respectively. The parameter k_0 is set to unity and not varied.

We define a "base run"; $\upsilon = -1$, $R = 0$, $\varepsilon = 0.5$, $\gamma = 0.6$, $\delta = 0.3$, and $\alpha = \beta = 1$. Parameters are varied one at a time from this base holding the values of the other parameters constant.

The choice of υ is a matter of value judgement. A lower value of υ represents diminishing social marginal utility of full income and increasing aversion to inequality. $\upsilon = 1$ corresponds to Utilitarian principle; $\upsilon = -\infty$ to maxi-min (Rawlsian). The revenue requirement R may be compared with total output Y: for most calculations output, which is endogenous, was between 0.5 and 0.7. Hence a government revenue requirement of 0.1 represents something between 14 and 20% of total output. The elasticity of substitution between consumption and leisure ε, is 0.5 in the base run. Stern (1976) argues that this conforms well with many empirical estimates of labor supply schedules. The case of $\alpha = \beta = 1$ corresponds to equal numbers in each group; with $\alpha = 0.5$, $\beta = 1$, there are 16.6% of the highly skilled, 33.3% of the moderately skilled, and 50% of low skilled in the total population.

As an empirically relevant case, we select the elasticity of substitution, $\varepsilon = 0.5$. For a wage income distribution, we use the data based on the 1986 individual income tax returns published by U.S. Department of Commerce in 1989, estimating the values for γ, δ as 0.6275, 0.2798, respectively (it corresponds to the case where the hourly wage rate is \$30.4, \$13.6, and \$4.5 for the highly skilled, the moderately skilled, and the low skilled, respectively); $\alpha = 0.339$, $\beta = 1.268$, (i.e., 11.3% of the population is the highly skilled, 42.3% the moderately skilled, 46.4% the low skilled); $R = 0.08$ (representing approximate 16% of total output).

There are a number of checks on the reliability of our computations. First, for $\upsilon = -1$, and $\varepsilon = 0.5$, the first-best optimum requires equality of the social marginal utility of consumption and hence in this case, the consumption itself. The optimization routine did indeed give this result. Secondly, we can closely replicate the results the Stern (1982), if we compute the model with only two groups. No problem of multiple local maxima was encountered and computing times were very small.

III. Simulation Results

A. Comparison of Welfare

Since the government is given the additional ability to vary marginal tax rates in the non-linear system, it is expected that it will achieve the higher social welfare in the non-linear system than in the linear system. Mathematically speaking, the government maximizes the maximand with a smaller set of constraints in the non-linear system, hence the value of maximand for the non-linear system should be greater than or equal to that for the linear system. The simulation indeed gives this result. The equally-distributed leisurely-equivalent consumption in the non-linear system is 0.1537 which is larger than 0.1384 in the linear system.

For individual welfare comparison, note the equivalent variations under different tax regimes. For $\upsilon = -1$, (See Table 1-(a)), the government has to give 0.1137 units of a consumption good in the absence of taxation to keep the low skilled equally as happy as if he were under non-linear income tax system, whereas it needs to give him only 0.0724 units if he were under linear tax system. Hence, the low skilled must be better off under non-linear system than under linear system. For the highly skilled, equivalent variation in the non-linear taxation is -0.1333, while it is -0.1145 in the linear taxation. The highly skilled has a higher utility in the linear system. The government's policy of moving from the non-linear system to the linear system makes the low skilled worse off, while making the highly skilled better off.

With constant wage rates, we have simulation results that conform with the standard theorems on optimal income taxation (See Mirrlees 1971 and Seade 1977). One theorem is that the marginal tax on the highest income should be zero, and it is confirmed by $t_H = 0.0000$ in Table 1. Coupled with the positive single marginal tax rate in the linear system, the zero marginal tax rate might lead one to casually think that since the highly skilled faces no marginal tax in the non-linear system and he faces a positive single marginal tax in the linear system, he might be better off in the non-linear system. However, the striking fact is that the highly skilled is actually worse off in the non-linear system.

Here, marginal tax rates have subtle meaning. Although one in the highly skilled faces no marginal tax, he pays positive taxes on the lower

TABLE 1

(a) THE BASE RUN, $\upsilon = -1$

Optimum Non-linear Income Taxation

t_L	G_L	t_M	G_M	t_H	G_H
0.4002	0.1311	0.3839	0.0822	0.0000	−0.1333
Y	$°C_n$	EV_L	EV_M	EV_H	
0.6033	0.1537	0.1137	0.0129	−0.1333	

Optimum Linear Income Taxation

t	G			
0.6038	0.1034 .			
Y	$°C_1$	EV_L	EV_M	EV_H
0.5140	0.1384	0.0724	−0.0055	−0.1145

(b) THE BASE RUN, $\upsilon = -2$

Optimum Non-linear Income Taxation

t_L	G_L	t_M	G_M	t_H	G_H
0.4319	0.1331	0.4180	0.0880	0.0000	−0.1359
Y	$°C_n$	EV_L	EV_M	EV_H	
0.6001	0.1533	0.1145	0.0129	−0.1359	

Optimum Linear Income Taxation

t	G			
0.6407	0.1074 .			
Y	$°C_1$	EV_L	EV_M	EV_H
0.5032	0.1363	0.0756	−0.0078	−0.1255

brackets of his income. The range of income which is exempted from taxation through zero marginal tax rate may be very small, and the simulation results shows that this is indeed the case. Tuomala (1984) studied the locality of the zero marginal tax on the highest income in a continuous type model. He found that the marginal tax rates is far from being zero even for upper 99.9% income brackets, arguing marginal tax rates fall to zero abruptly only in the extremely very high income brackets. The theorem of zero marginal tax rate for the highest income is very local. Our results also suggest that the range of income which is subject to zero marginal tax is very narrow.

The difference of equivalent variations between types is greater in the non-linear system. The differences in welfare between people is smaller in the non-linear system. It suggests that the government achieves more equitable redistribution in the non-linear system.

Marginal tax rates in the non-linear system lie between zero and one.

TABLE 1
(CONTINUED)

(C) THE BASE RUN, $\upsilon = 0.97$

Optimum Non-linear Income Taxation

t_L	G_L	t_M	G_M	t_H	G_H
0.3046	0.1247	0.2760	0.0631	0.0000	−0.1263
Y	$°C_n$	EV_L	EV_M	EV_H	
0.6119	0.1549	0.1107	0.0124	−0.1263	

Optimum Linear Income Taxation

t	G			
0.4045	0.0751 .			
Y	$°C_l$	EV_L	EV_M	EV_H
0.5573	0.1457	0.0508	0.0004	−0.0665

Note: 1. Notations: t_i = marginal tax rate for type i workers
Y = output
υ = parameter measuring attitudes to inequality
$°C_j$ = equally distributed leisurely-equivalent consumption under alternative tax system, j = n(non-linear system), = l (linear system)
EV_i = equivalent variation for the type i worker

2. *The different optima:*
Optimum non-linear income taxation: every individual faces the same income tax schedule although they differ in their wage rates; G_i is the lump-sum grant as given by the tangent to the indifference curve for type i workers.
Optimum linear income taxation: G is the grant common to all individuals.

3. *Other parameters:*
R, ε, γ, δ, α, β are the government revenue requirement, the elasticity of substitution between consumption and leisure, the share for the highly skilled and the moderately skilled, and the proportion of the highly skilled and the moderately skilled, respectively. For these results of Table 1 we have $R = 0$, $\varepsilon = 0.5$, $\gamma = 0.6$, $\delta = 0.3$, $\alpha = \beta = 1$. For variation of parameters see Table 2.

If one looks at the marginal tax rates, the optimal non-linear income tax is regressive in terms of marginal tax rates. The marginal tax rates decreases with income. Table 1 (d) gives the average tax rates for the base run. The average tax rates increases with income. The tax, hence, is progressive in terms of average tax rates. Zero marginal tax tells little about the average tax rates. We note that the absolute value of average tax rates are greater in the non-linear system for everybody than in the linear system. The non-linear system is more progressive than the

TABLE 1
(CONTINUED)

(d) THE AVERAGE TAX RATES

Optimum Non-Linear Income Taxation

υ	the low skilled	the moderately skilled	the highly skilled
-2	-3.301	-0.109	0.342
-1	-3.115	-0.101	0.336
0.97	-2.685	-0.081	0.321

Optimum Linear Income Taxation

υ	the low skilled	the moderately skilled	the highly skilled
-2	-2.312	-0.067	0.299
-1	-1.979	-0.059	0.278
0.97	-0.946	-0.029	0.176

(e) THE DEADWEIGHT LOSS OF TAXATION

Optimum Non-Linear Income Taxation

υ	the low skilled	the moderately skilled	the highly skilled
-2	0.0032	0.0053	0.0000
-1	0.0026	0.0042	0.0000
0.97	0.0013	0.0019	0.0000

Optimum Linear Income Taxation

υ	the low skilled	the moderately skilled	the highly skilled
-2	0.0086	0.0179	0.0312
-1	0.0069	0.0147	0.0261
0.97	0.0019	0.0046	0.0089

linear system.

One interesting point is the value of G_H, the lump-sum grant which is given by the tangent to the indifference curve to the highly skilled. It is equal to the value of equivalent variation for him, EV_H. Since the highly skilled faces no distortionary taxation in the non-linear system, the non-linear income tax has only the income effect on him, and this equality of G_H and EV_H is expected. It is comforting that our optimization routine indeed gives this result. Since the lower skilled and the moderately skilled face positive marginal taxes, thus distortionary taxes, the equality between G_i and EV_i doesn't hold for the low skilled and the moderately skilled.

Table 1-(d) also shows the deadweight loss of taxation under different tax regimes. We immediately see that the deadweight loss for the highly skilled is zero in non-linear taxation. The government can increase the

average tax rate on the highly skilled by increasing the marginal tax rate on the low skilled at the expense of increase in the deadweight loss of the low skilled. At the optimum, the marginal benefit of increasing the marginal tax rate of the low skilled (increase in redistribution through the increase in the average tax rate of the highly skilled) will be exactly equal to the marginal cost of the action (the increase in the deadweight loss of the low skilled). For the marginal tax of the highly skilled, however, there is only an efficiency concern, and the government does not want to impose the distortionary tax for the highly skilled. We can also note that the deadweight loss is smaller in the non-linear system for every type of individual than in the linear system, which suggests that the non-linear taxation is a more efficient tax policy.

B. Variations of Parameters

We begin with the effects of variations in the parameter υ-See Table 1-(a). (The value of $\upsilon = 1$ resulted in a problem of convergence in a numerical optimization method, so we used 0.97 as a proxy.)

The value of υ can be interpreted by looking at the differences of equivalent variations across different values of υ. The difference between equivalent variations of the low skilled and the highly skilled is 0.2504 for $\upsilon = -2$, while it is 0.2470 and 0.2370 for $\upsilon = -1$, and $\upsilon = 0.97$, respectively in the non-linear system. The gap between equivalent variations decreases with υ. Inequality in after-tax utility increases as the gap goes up. It suggests that the government is more concerned with the inequality with a lower value of υ. We can also see that the highly skilled are worse off in the non-linear system for $\upsilon = -2$, and $\upsilon = 0.97$.

We now turn to a discussion of Table 2, which shows, for $\upsilon = -1$, the effects of varying parameters. Table 2-(a) shows the effect of $R = 0.1$ on the economy. An increase in the government revenue requirement imposes greater burdens on the economy. As a result, in both systems, marginal tax rates increase and lump-sum grants decline. Output increases as a result of increases in labor supplies and °C , the welfare level decreases. We have not included any possible benefits from the government expenditure.

Table 2-(b) shows the effects of reduction in the differences in wage rates. A reduction in the share, γ of the highly skilled lowers tax rates and lump-sum grants. The greater the similarity between types of labor, the lower is the desire to redistribute income through taxes.

In Table 2-(c), a reduction in the proportion of the highly skilled in

TABLE 2

(a) VARIATION OF PARAMETERS, $R = 0.1$

Optimum Non-linear Income Taxation

t_L	G_L	t_M	G_M	t_H	G_H
0.4253	0.1106	0.4060	0.0585	0.0000	–0.1769
Y	$°C_n$	EV_L	EV_M	EV_H	
0.6405	0.1356	0.0895	–0.0197	–0.1769	

Optimum Linear Income Taxation

t	G			
0.6504	0.0854 .			
Y	$°C_1$	EV_L	EV_M	EV_H
0.5475	0.1194	0.0480	–0.0400	–0.1628

(b) THE VARIATION OF PARAMETERS, $\gamma = 0.434$, $\delta = 0.333$

Optimum Non-linear Income Taxation

t_L	G_L	t_M	G_M	t_H	G_H
0.2327	0.0657	0.1860	0.0360	0.0000	–0.0316
Y	$°C_n$	EV_L	EV_M	EV_H	
0.6163	0.1527	0.0325	0.0027	–0.0316	

Optimum Linear Income Taxation

t	G			
0.2692	0.0533 .			
Y	$°C_1$	EV_L	EV_M	EV_H
0.5939	0.1497	0.0133	–0.0021	–0.0168

(c) THE VARIATION OF PARAMETERS, $\alpha = 0.5$, $\beta = 1$

Optimum Non-linear Income Taxation

t_L	G_L	t_M	G_M	t_H	G_H
0.3327	0.0969	0.3121	0.0316	0.0000	–0.1846
Y	$°C_n$	EV_L	EV_M	EV_H	
0.4757	0.1296	0.0791	–0.0305	–0.1846	

Optimum Linear Income Taxation

t	G			
0.6247	0.0846 .			
Y	$°C_1$	EV_L	EV_M	EV_H
0.4060	0.1212	–0.0484	–0.0352	–0.1514

Note: 1. $\upsilon = -1$.

2. Parameters not specified at the head of the table are as in Table 1.

the population lowers (raises) marginal tax rates, in the non-linear system (in the linear system) and decreases output and °C, the welfare level in both systems. We can interprete the result as the following way. A relative increase in the proportion of the lower skilled forces the government to concern more about the efficiency loss associated with taxing the lower skilled, hence it lowers the marginal tax for the lower skilled. Also, the desire to redistribute income increases with increasing proportion of the low skilled. In the non-linear system, the government already have achieved more desirable income redistribution than the linear system in the base run and the efficiency concern outweighs the redistribution desire as the population of the low income people grows. However, in the linear system, the need for redistribution is greater than the efficiency loss, hence the government raises the tax rate.

What is commonly true in Table 2 is that the highly skilled is worse off in the non-linear system; the low skilled and the moderately skilled better off.

C. Variations of Parameter: Elasticity of Substitution, ε.

The variation of elasticity of substitution ε deserves a special attention. Table 3-(a) shows the effect of varying ε. The level of welfare, °C of the non-linear system is greater than that of the linear system for all values of ε. The single marginal tax rate decreases with ε in the linear income taxation—the greater the value of ε, the greater the deadweight loss of taxation. The marginal tax rates in the non-linear system don't decrease monotonically with ε. For ε = 0.25 to 0.65, the marginal tax rate for the low skilled actually increases with ε. For ε greater than 0.65, it starts to decrease. For the moderately skilled, the marginal tax rate starts to diminish for the value of ε greater than half.

As the elasticity of substitution changes, the form of utility function also changes. The social marginal utility of consumption depends on the consumption levels as well as the utility function, thus the social marginal utility of consumption also changes with the elasticity of substitution. It implies that the government concerns about not only efficiency but equity with different values of the elasticity of substitution.

Table 3-(b) gives the deadweight loss of taxation under different tax regimes. We see for every value of ε that the deadweight loss of taxation is smaller in the non-linear income taxation for the moderately skilled and for the highly skilled. It is because of the fact that the government can increase the average taxes on them by increasing the marginal tax

TABLE 3

(a) THE VARIATION OF PARAMETER, ε

Optimum Non-Linear Income Taxation

ε	t_L	G_L	t_M	G_M	t_H	G_H	Y	$°C_n$	EV_L	EV_M	EV_H
0.25	0.333	0.164	0.345	0.089	-0.000	-0.164	0.684	0.205	0.143	0.019	-0.164
0.30	0.357	0.157	0.362	0.090	0.000	-0.156	0.668	0.195	0.136	0.017	-0.156
0.35	0.374	0.150	0.373	0.089	-0.000	-0.150	0.652	0.184	0.130	0.016	-0.150
0.40	0.387	0.144	0.380	0.087	0.000	-0.144	0.636	0.173	0.124	0.015	-0.144
0.45	0.395	0.137	0.383	0.085	0.000	-0.138	0.620	0.163	-0.118	0.014	-0.138
0.50	0.400	0.131	0.384	0.082	0.000	-0.133	0.603	0.154	0.114	0.013	-0.133
0.55	0.403	0.125	0.383	0.079	0.000	-0.129	0.587	0.145	0.109	0.012	-0.129
0.60	0.403	0.120	0.381	0.076	-0.000	-0.125	0.570	0.136	0.106	0.011	-0.125
0.65	0.403	0.114	0.377	0.073	0.000	-0.122	0.554	0.129	0.102	0.011	-0.122
0.70	0.401	0.110	0.374	0.070	0.000	-0.119	0.538	0.121	0.099	0.010	-0.119
0.75	0.399	0.105	0.370	0.067	-0.000	-0.117	0.522	0.115	0.097	0.009	-0.117
0.80	0.397	0.101	0.365	0.064	-0.000	-0.114	0.506	0.109	0.094	0.009	-0.114

Optimum Linear Income Taxation

ε	t	G	Y	$°C_l$	EV_L	EV_M	EV_H
0.25	0.984	0.150	0.457	0.152	0.086	-0.062	-0.293
0.30	0.932	0.154	0.496	0.162	0.100	-0.034	-0.240
0.35	0.833	0.145	0.522	0.163	0.099	-0.017	-0.187
0.40	0.739	0.130	0.526	0.157	0.090	-0.010	-0.153
0.45	0.664	0.116	0.522	0.148	0.080	-0.007	-0.131
0.50	0.604	0.103	0.514	0.138	0.072	-0.006	-0.114
0.55	0.553	0.093	0.505	0.129	0.066	-0.004	-0.102
0.60	0.510	0.084	0.495	0.121	0.060	-0.004	-0.092
0.65	0.472	0.076	0.484	0.113	0.055	-0.003	-0.084
0.70	0.439	0.069	0.474	0.106	0.050	-0.002	-0.076
0.75	0.408	0.063	0.464	0.099	0.047	-0.002	-0.070
0.80	0.381	0.058	0.454	0.092	0.043	-0.001	-0.064

Note: 1. Notations: See Table 1.
2. $\upsilon = -1$

on the low skilled. The marginal tax rates for the moderately skilled can be set at the low level and for the highly skilled it is set at zero. It suggests that the government can take care of the equity matter with a lower cost of efficiency loss in non-linear income taxation. For a high value of ε, the government can still take care of equity matter in the non-linear system with a lower efficiency loss. But in the linear system, swamped with the high deadweight loss of the highly skilled, the government cannot pay enough attention to the equity matter. In the linear system, marginal social welfare loss from an efficiency loss by a small increase in ε is always greater than the marginal social welfare

TABLE 3
(CONTINUED)

(b) THE DEADWEIGHT LOSS OF TAXATION

Optimum Non-Linear Income Taxation

ε	t_L	t_M	t_H	DL_L	DL_M	DL_H
0.25	0.3333	0.3453	-0.0000	0.0006	0.0014	-0.0000
0.30	0.3565	0.3620	0.0000	0.0009	0.0020	-0.0000
0.35	0.3742	0.3731	-0.0000	0.0013	0.0026	-0.0000
0.40	0.3868	0.3798	0.0000	0.0017	0.0031	-0.0000
0.45	0.3952	0.3831	0.0000	0.0021	0.0037	-0.0000
0.50	0.4002	0.3839	0.0000	0.0026	0.0042	-0.0000
0.55	0.4028	0.3828	0.0000	0.0030	0.0047	-0.0000
0.60	0.4035	0.3806	-0.0000	0.0034	0.0051	-0.0000
0.65	0.4028	0.3774	0.0000	0.0038	0.0055	-0.0000
0.70	0.4013	0.3737	0.0000	0.0042	0.0058	-0.0000
0.75	0.3993	0.3696	-0.0000	0.0046	0.0061	-0.0000
0.80	0.3970	0.3653	-0.0000	0.0049	0.0064	-0.0000

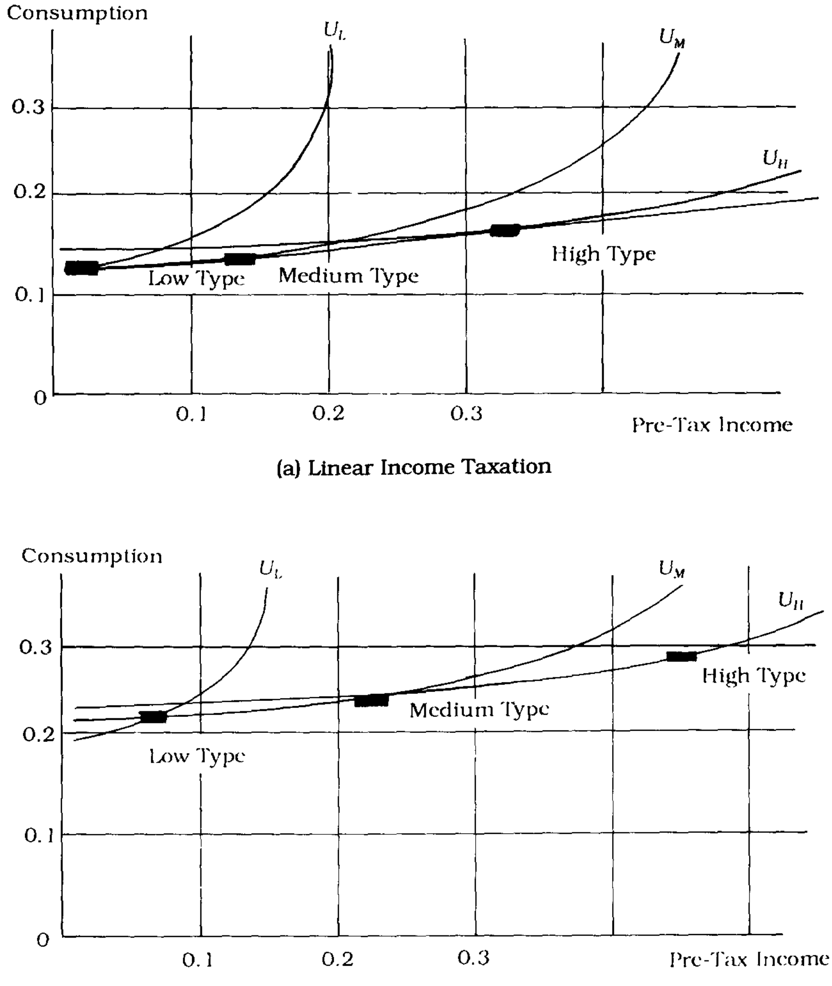
Optimum Linear Income Taxation

ε	t	DL_L	DL_M	DL_H
0.25	0.9837	0.0387	0.0867	0.1439
0.30	0.9320	0.0260	0.0559	0.0923
0.35	0.8332	0.0156	0.0332	0.0564
0.40	0.7388	0.0107	0.0229	0.0398
0.45	0.6641	0.0083	0.0178	0.0313
0.50	0.6038	0.0069	0.0147	0.0261
0.55	0.5535	0.0059	0.0126	0.0224
0.60	0.5102	0.0051	0.0110	0.0196
0.65	0.4723	0.0045	0.0097	0.0173
0.70	0.4386	0.0040	0.0086	0.0154
0.75	0.4083	0.0036	0.0077	0.0138
0.80	0.3810	0.0032	0.0069	0.0124

Note: 1. DL_i is the deadweight loss of taxation for the i type individual,
2. $\upsilon = -1$

benefit of redistribution; vice versa in the non-linear system. The deadweight loss in the non-linear system increases with the elasticity of substitution, but they are smaller than the deadweight loss which decreases with ε in the linear system.

It is helpful to envision this comparison of welfare on the consumption-income plane. Figure 1 shows the indifference curves for each individual under different tax regimes for $\varepsilon = 0.25$; Figure 2 for $\varepsilon = 0.75$. Indifference curve for the high skilled is flatter than that for the low skilled, because the high skilled is an efficient worker; he needs a

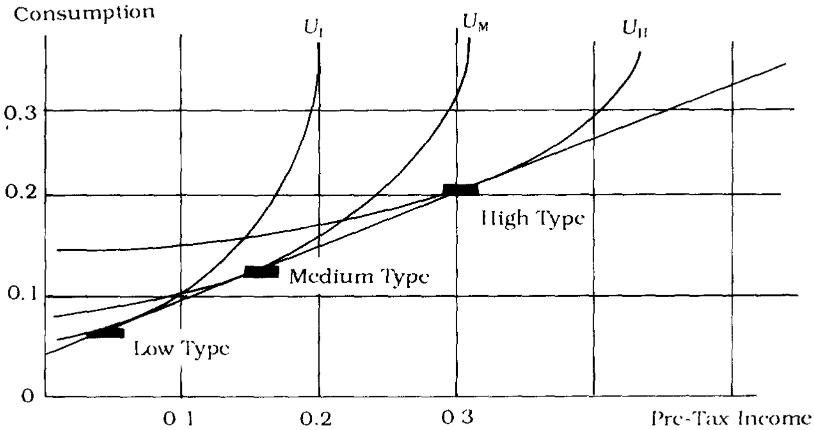


(a) Linear Income Taxation

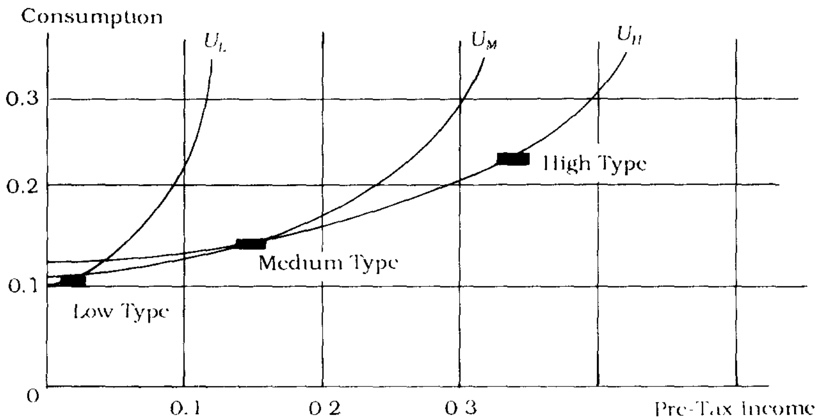
(b) Non-Linear Income Taxation

FIGURE 1
CONSUMPTION AND INCOME OF PEOPLE: ELASTICITY = 0.25

smaller compensation of consumption good to be induced to work additional hour. As the elasticity of substitution gets larger, the indifference curve becomes more of a straight line type indicating a greater deal of substitutability between consumption and leisure. For the low value of ε, the differences between individual consumption levels are smaller, but the amount of individual labor supply is very different.



(a) Linear Income Taxation



(b) Non-Linear Income Taxation

FIGURE 2

CONSUMPTION AND INCOME OF PEOPLE: ELASTICITY = 0.75

Lower values of ε indicate that people think that consumption and leisure are two distinct goods so that the government can redistribute consumption across individuals without much worrying its effect on labor supply. But as ε gets larger, people starts to substitute more leisure when consumption becomes expensive because of taxation. The redistribution of consumption has a higher cost of reducing labor sup-

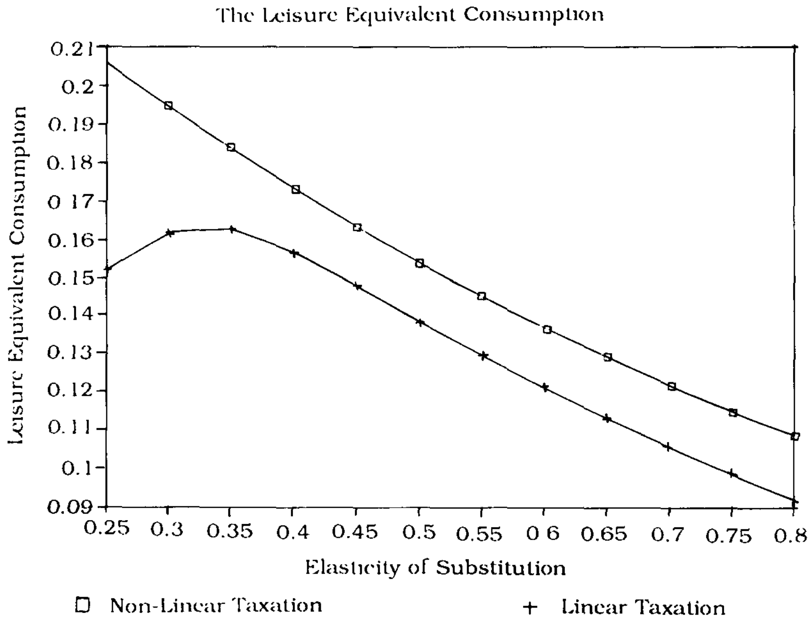


FIGURE 3
THE LEISURE EQUIVALENT CONSUMPTION

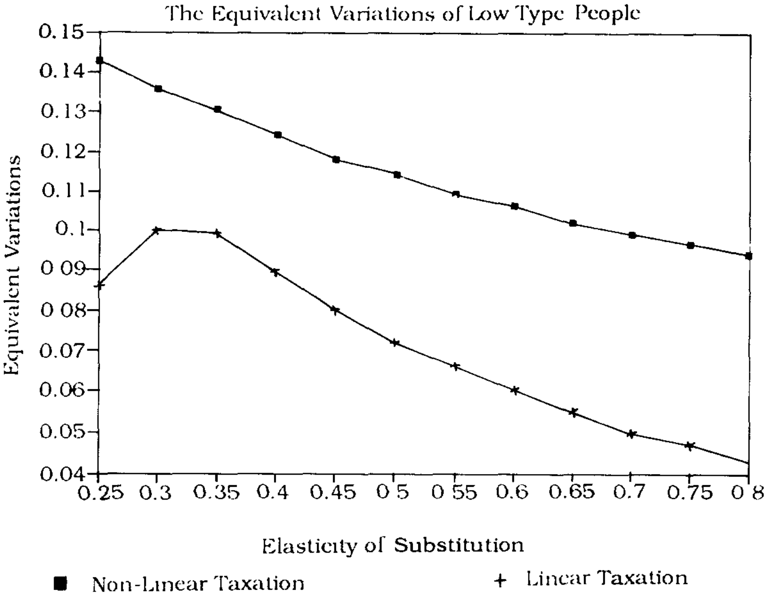
ply for a higher value of ε. The gaps between individual consumption levels is greater with a higher value of ε.

Figure 3 shows the graph of the $°C$ across ε. We did get a similar result as in Stern (1976) for the $°C_1$ graph; in the very lower values of ε, $°C_1$ diminishes.

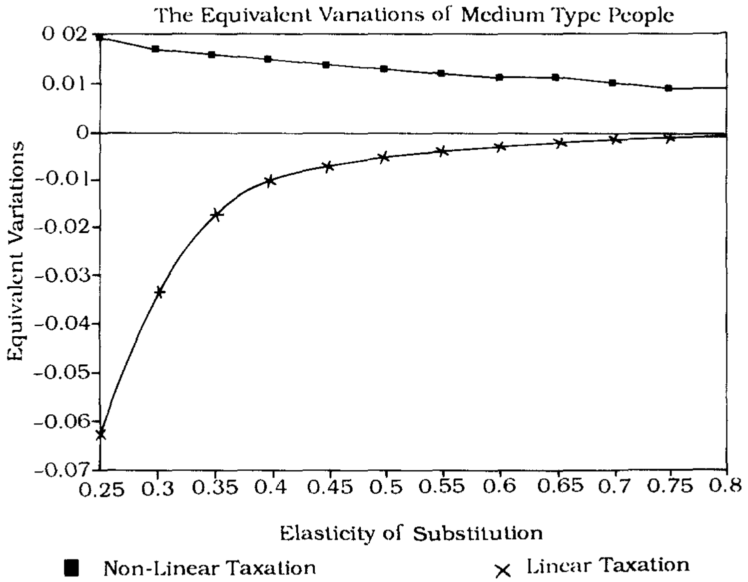
The comparison of individual welfare is depicted in Figure 4-(a) to Figure 4-(c) for each type of individual. The low skilled and the moderately skilled are always better off in non-linear income taxation. For the highly skilled, it is interesting that we have a cut-off point where he is indifferent to alternative income tax system. The cut-off point is the value of between 0.4 and 0.45 for the base run.

The stability of the cut-off point is examined by varying parameters R and υ. Figure 5-(a) shows the equivalent variations for the highly skilled with different values of elasticity of substitution for R = 0.1; Figure 5-(b) for the value of $\upsilon = -2$.

The fact that there exists a cut-off point for the highly skilled suggests that it is crucial to know the elasticity of substitution to evaluate the individual welfare aspects of non-linear income taxation. With a

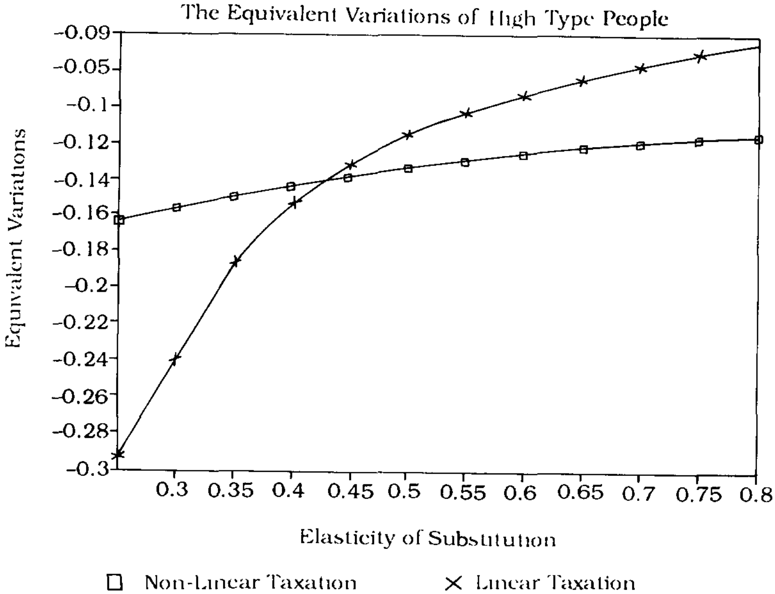


(a) Low Ability Type People



(b) Medium Ability Type People

FIGURE 4
THE EQUIVALENT VARIATIONS

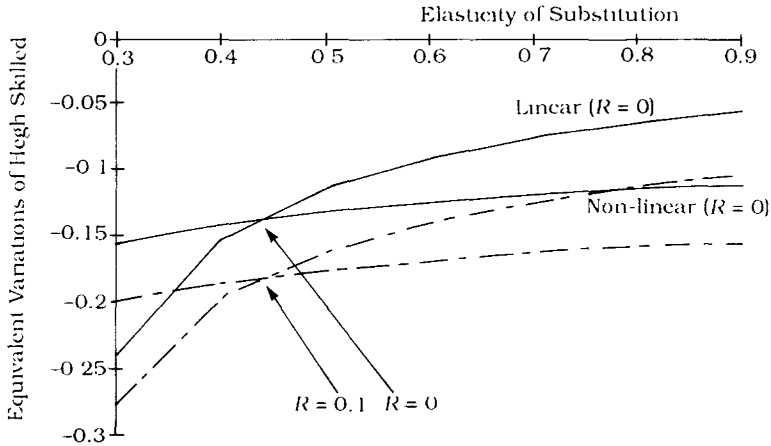


(c) High Ability Type People

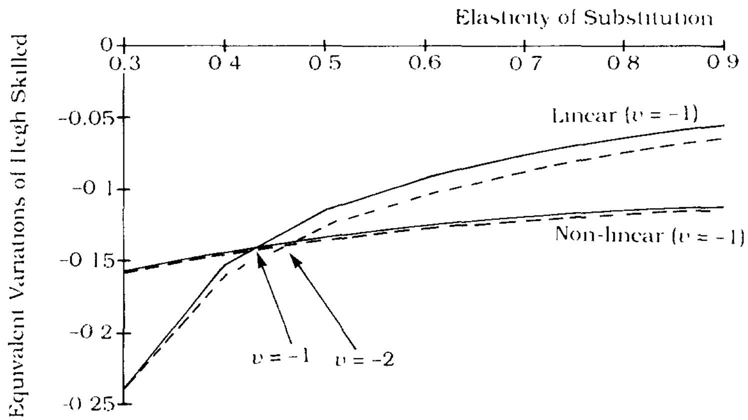
FIGURE 4
(CONTINUED)

low elasticity of substitution between consumption and leisure, even the highly skilled are better off in non-linear income taxation. Non-linear income taxation is Pareto superior. This is because in the linear system, to redistributive income, the government needs to give larger lump-sum grants and thus should tax income at a very high rate (98. 4% of income is taxed for $\varepsilon = 0.25$). When the elasticity of substitution is low, linear income taxation is a very inefficient tool for the redistribution. (See Stiglitz 1988). To provide a reasonable level of lump-sum grants would require imposing a very high tax rate. The lower the elasticity of substitution, the greater is the inefficiency of linear taxation as an income redistribution. The movement toward non-linear income taxation is strongly Pareto superior for the low values of ε.

Stern (1982) argues that $\varepsilon = 0.5$ conforms well with many empirical estimates of labor supply. If this is true, purely redistributive non-linear income taxation makes the highly skilled worse off than in the linear income taxation. The change in the tax policy toward non-linear income taxation from linear income taxation is not Pareto superior. It makes the highly skilled worse off.



(a) The Variations of Government Revenue Requirement, R



(b) The Variations of Attitude toward Inequality, v

FIGURE 5

THE STABILITY OF CUT-OFF POINTS OF EQUIVALENT VARIATIONS OF HIGH SKILLED INDIVIDUALS

D. Evaluating Flat Rate Income Tax Reform

Table 4 gives the simulation result, which is the most empirically rel evant to the U.S case. The parameter values are based on the 1986 U.S

TABLE 4
THE VARIATION OF PARAMETERS (U.S. CASE)

Optimum Non-linear Income Taxation							
t_L	G_L	t_M	G_M	t_H	G_H		
0.6958	0.0912	0.1070	-0.0440	0.0000	-.2611		
Y	$°C_n$	EV_L	EV_M	EV_H	atr_L	atr_m	atr_h
0.4248	0.1067	0.0562	-0.0659	-0.2611	-1.950	0.328	0.569
Optimum Linear Income Taxation							
t	G						
0.7068	0.0678						
Y	$°C_n$	EV_L	EV_M	EV_H	atr_L	atr_m	atr_h
0.4010	0.0986	0.0264	-0.0681	-0.2314	-0.7412	0.2982	0.5233

Notes: 1. Atr_L, atr_m, and atr_h are the average tax rates for the low skilled, the moderately skilled, and the high skilled, respectively.
2. $\upsilon = -1$, $\varepsilon = 0.5$, $\gamma = 0.6275$, $\delta = 0.2798$, $\alpha = 0.339$, $\beta = 1.268$, $R = 0.08$

individual income tax returns. The marginal tax rates for the low skilled is 69.58%. The low skilled are better off, the moderately skilled slightly better off, and the high skilled worse off in non-linear income taxation. The average tax rates are very high; -195% in the non-linear system, -74.12% in the linear system for the low skilled. These magnitudes of average and marginal tax rates crucially depend on one's assumption on the distribution of skills within the population and the sensitivity of labor supply to the changes in the marginal tax rates. (See Atkinson 1973). With different assumptions, we obtain the higher tax rates than Mirrlees (1971).

One interesting application of the simulation is to evaluate the argument for the flat rate income tax reform. There always exists the possibility of reforming the existing tax system to non-linear income taxation rather than to flat rate income taxation. Since one can tax income at source in the flat rate system, it is superior in the aspects of administration. To file for complex non-linear income taxation people also have to spend their time and resources (called the compliance costs) in preparing returns. The compliance costs are borne by taxpayers. The non-linear income system, however, has the welfare gains over the flat rate system. The proponents for the flat rate income tax, hence, should convince others that the administrative costs plus the compliance costs saved by linear income taxation outweigh the welfare gains of non-linear income taxation.

Slemrod and Sorum (1984) have estimated that taxpayers spent from

5 per cent to 7 per cent of the revenue raised by the federal and state income tax systems on compliance costs (the value of their time plus what they paid to tax preparers). The welfare gains (measured by the difference in leisure-equivalent consumption between the linear tax and the non-linear tax) in Table 4 is 10 per cent of the revenue, which suggests that the welfare gains outweigh the compliance costs. If the administrative costs are less than 3 or 5 per cent of the tax revenue raised, it, then, follows that non-linear income taxation is indeed superior to linear income taxation. Flat rate tax reform without considering the possibility of the optimum non-linear income taxation only makes the high income group better off at the expense of low income group. It may be also an inferior form of tax reforms, even if we consider the compliance costs and the administrative costs.

This simulation technique can be used to evaluate the individual welfare when there are two periods. People save in the first period from their endowments and supply labor in the second period. There are three goods in the model; the present consumption good, the future consumption good, and the leisure. They differ in their level of endowments and abilities to supply labor. We can set up two alternative tax systems: the unconstrained non-linear income tax system where the government can tax the wage income and the capital income in any non-linear form; the constrained non-linear income taxation where only the sum of wage income and capital income is taxable in the non-linear form. Constrained non-linear income taxation is often called comprehensive income taxation. We denote unconstrained income taxation as wage-interest income taxation. Choi (1991) compares the welfare properties of the each tax systems, and he finds that even though the preference is separable between consumption and leisure, the government needs to tax interest income for equity purpose.

IV. Conclusion

Non-linear income taxation gives the higher level of social welfare than linear income taxation. Based on the simulation result for a reasonable estimate of parameters, non-linear income taxation makes the low skilled and the moderately skilled better off, the highly skilled worse off compared with linear income taxation with the same revenue requirement. If the elasticity of substitution is small, i.e., the sensitivity of labor supply to the changes in net wage rates is low, non-linear income taxation is Pareto superior. It makes all individuals better off without making no one worse off. Linear income taxation is an inefficient tool for the redistribution compared with non-linear income taxation, in that it has the greater deadweight loss of taxation. Using the estimate of Slemrod and Sorum, it is shown that the welfare gains from switching to non-linear income taxation from linear income taxation outweigh the compliance costs of non-linear taxation.

References

Allen, A. "Optimal Linear Income Taxation with General Equilibrium Effects on Wage." *Journal of Public Economics* 17 (1982): 135-43

Atkinson, A.B. "How Progressive Income Tax Be?" In A. B. Atkinson (eds.), *Social Justice and Public Policy*, Cambridge, Mass: The MIT Press, 1983.

Auerbach, A.J., and Kotlikoff, L.J. *Dynamic Fiscal Policy*, Cambridge: Cambridge University Press, 1987.

Choi, Seon-Kyou. "Comparing Optimum Taxation of Wage and Interest Income with Optimum Taxation of Comprehensive Income and Optimum Taxation of Wage Income: A Numerical Analysis." Ph.D. Dissertation, Bloomington, Indiana: Indiana University, 1991.

Hausman, J.A. "Labor Supply." In H.J. Aaron and J.A. Pechman (eds.), *How Taxes Affect Economic Behavior*, The Brookings Institution, 1981.

Kay, J.A. "Deadweight Loss from a Tax System." *Journal of Public Economics* 13 (1980): 111-9.

Mirrlees, J.A. "An Exploration in the Theory of Income Taxation." *Review of Economic Studies* 38 (1971): 175-208.

____________. "Notes on Welfare Economics, Information and Uncertainty." In M.D. McFadden and S. Wu (eds.), *Essays on Economic Behavior under Uncertainty*, North Holland, 1974.

Seade, J.K. "On the Shape of Optimal Income Tax Schedules." *Journal of Public Economics* 7, (1977): 203-6.

Slemrod, J., and Sorum, N. "The Compliance Cost of the U.S. Individual Income Tax System." University of Minnesota, Mimeo, 1984. (cited from Stiglitz, J.E, 1988, 603)

Stern, N.H. "On the Specification of Models of Optimum Income Taxation." *Journal of Public Economics* 6 (1976): 123-62.

____________. "Optimum Taxation with Errors in Administration." *Journal of Public Economics* 17 (1982): 181-211.

Stiglitz, J.E. "Self-selection and Pareto Efficient Taxation." *Journal of Public Economics* 17 (1982): 213-40.

____________. "Economics of the Public Sector." New York: W.W. Norton & Company, 1988.

Tuomala, M. "On the Optimal Income Taxation; Some Further Numerical Results." *Journal of Public Economics* 23 (1984): 351-66.

U.S. Department of Commerce. "Statistical Abstract of the United States." Washington D.C.: U.S. Government Printing Office, 1989.