NBER WORKING PAPERS SERIES

MENUS OF LINEAR INCOME TAX SCHEDULES

Alberto Alesina

Philippe Weil

Working Paper No. 3968

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 1992

We are grateful to Patrick Bolton, Daniel Cohen, and Eytan Sheshinski for helpful discussions. We thank Phillip Swagel for excellent research assistance. The original version of this paper was written while Alesina was an Olin Fellow at the NBER. Alesina gratefully acknowledges financial support from the Olin and Sloan Foundations, and Weil from the National Science Foundation (SES-8823040). This paper is part of NBER's research program in Taxation. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper #3968
January 1992

MENUS OF LINEAR INCOME TAX SCHEDULES

ABSTRACT

Relative to traditional piecewise linear income taxation schemes, it is possible to increase government revenues by offering to consumers a menu of linear income tax schedules. In the resulting Pareto-superior equilibrium, consumers sort themselves out according to their (unobservable) productivity level, with high productivity agents choosing the tax schedules with low marginal tax rate and high intercept. This scheme extracts from the economy an unexploited source of revenue which, in contrast with standard supply-side proposals, does not depend on the economy being on the downward-sloping side of the Laffer curve.

Alberto Alesina
Department of Economics
Harvard University
Cambridge, MA 02138
and NBER

Philippe Weil
Department of Economics
Harvard University
Cambridge, MA 02138
and NBER

Lump sum income taxes, while non distortionary, are considered undesirable because of their regressivity. Thus, in designing income tax schedules, benevolent governments face a trade-off between the efficiency cost of distorting labor supply decisions and the distributional benefit of progressivity (or proportionality) of the tax burden.[1]

In most developed countries, including the United States, income tax schedules are progressive and take, undoubtedly for simplicity, the form of continuous piecewise linear functions.[2] In this paper, we demonstrate that any fiscal system with a continuous linear (or piecewise linear) tax schedule can be Pareto improved by the introduction of a second tax schedule, and by letting the taxpayers select their preferred tax function on the menu of linear schedule presented to them. The additional tax schedule should have a lower marginal rate than the first one and a higher intercept. In more colorful terms, by introducing the second tax schedule, the government offers to "sell" a reduction of the marginal tax rate for the price of a lower lump sum transfer (the higher intercept of the new tax schedule). The tax payers who select to "buy" the reduction in the marginal rate, i.e., who choose the new tax schedule, will be the most productive workers: under the new tax schedule they will work and consume more, and some (or all) of them will pay more taxes. In fact, we derive simple conditions which insure that the Pareto improving introduction of the second tax schedule does not reduce total tax revenues. More generally, additional revenue-neutral or revenue-increasing Pareto improvements can be achieved by offering to the tax payers a menu of N linear tax schedules (with N greater than two). The only limit to the number of schedules is the complexity of the tax system. It should be stressed that, in contrast with standard supply-side arguments, this Pareto improving increase in tax revenues can be achieved even if the economy is on the upward sloping side of the Laffer curve and the additional fiscal revenue is wasted. These results hold under assumptions on utility functions which are quite general and standard in the optimal taxation literature.

Our results can be viewed as a practical proposal to implement the "no distortions at the top" principle.[3] We show that a version of this principle may often be applied "globally" so as to offer lower marginal income tax rates to a substantial

1. On the theory of income taxation, see Mirrlees [1977], Sadka [1977], Seade [1977] and Atkinson and Stiglitz [1980].
2. See Sheshinski [1972] and Sheshinski [1989] on the theory of optimal linear tax schedules.
3. Sadka [1977], Seade [1977].

fraction of the population—without affecting government revenues or relying on supply-side effects. We provide simple conditions under which this ensuing Pareto-improvement can be implemented for linear tax schedules.

Our results on the welfare benefits of a menu of linear income taxes are reminiscent of some recent theoretical developments in the principal-agent literature, which analyzes a situation where a principal employs agents without being able to observe their effort. Laffont and Tirole [1986] show that under certain conditions and with specific functional forms, the optimal contract is one in which the principal offers a menu of wage functions which are linear in output and the agents are free to choose any of these functions as their contract. The analogy between our results and those of Laffont and Tirole is that, in both cases, the menu of opportunities is such that high productivity workers choose to work more. In the principal-agent situation this occurs because by doing so high productivity workers take advantage of a more profitable wage function; in the case of income taxes they take advantage of lower marginal tax rates, even though they may pay more taxes in the end.

It should be pointed out that results similar to ours have been obtained independently by Slemrod et al. [1991].

This paper is organized as follows. In Section 1 we introduce the model and review the standard case with a single linear income tax schedule. Section 2 shows how a menu of linear income tax schedules may improve welfare without reducing tax revenues. In Section 3 we develop a numerical example and discuss the empirical relevance of our proposal. Section 4 discusses several extensions. The last section concludes.

1 Unique linear income tax schedule

In this section we analyze, to fix the notation and for future reference, the equilibrium in the presence of a unique linear income tax schedule. We describe first the behavior of consumers and of the government, and then characterize the equilibrium.

1.1 Consumers

The economy consists of one-period lived consumers-workers who differ only in their labor productivity, w.[4] Let $F(w)$ denote the cumulative distribution function of productivities. Clearly, $F(0) = 0$, $F(\infty) = 1$ and $\int_0^\infty w \, dF(w) = 1$.

We assume that all agents are productive ($w > 0$). Productivity is truly exogenous (a consumer cannot choose to be less productive), and cannot be unobserved by fiscal authorities.

All consumers have an equal unit endowment of leisure time, and rank bundles of consumption (c) and labor (ℓ) according to the same utility function $u(c, \ell)$, with $u_1 > 0$ and $u_2 < 0$. In addition to strict quasi-concavity, differentiability, monotonicity and local non-satiation of $u(.,.)$ in c and $1 - \ell$, we impose:

Assumption 1

$$\lim_{c \to 0} u_1(c, \ell) = +\infty \;\; \forall \, \ell \in (0, 1)$$

$$\lim_{\ell \to 1} u_2(c, \ell) = +\infty \;\; \forall \, c > 0$$

Assumption 2

$$\lim_{\ell \to 0} \frac{u_2(c, \ell)}{u_1(c, \ell)} = 0 \;\; \forall \, c \geq 0.$$

Assumption 3

$$u_1 u_{22} - u_2 u_{12} < 0 \;\text{ and }\; u_2 u_{11} - u_1 u_{12} > 0.$$

Assumption 4

$$u_1^2 + (u_1 u_{12} - u_2 u_{11})\ell > 0.$$

4. We could as well study consumers who only differ in their tastes. What matters is that we restrict ourselves, for simplicity, to one-dimensional heterogeneity. Our notation closely follows Sheshinski [1989].

The first assumption rules out corner solutions at zero for consumption or leisure. The second assumption guarantees that, at a consumption optimum, all consumers work a strictly positive number of hours.[5] The third assumption ensures that consumption and leisure are normal goods, and the fourth implies that optimum labor supply increases with the wage rate and decreases with the marginal tax rate on labor income.

One can think of consumers agents as yeoman farmers with different production functions; the production function of a consumer-worker of productivity w is $q = w\ell$. Alternatively, one can imagine that the consumption good is produced by competitive firms which employ different types of labor, with one unit of type w labor input yielding w units of output; in that interpretation, assuming that firms can perfectly monitor their workers' marginal product, w is the unit wage of consumer w. There is no capital.

1.2 Government

The government is modelled as an entity (possibly a social planner) which must raise enough tax revenue to finance an exogenously determined supply of a public good which is not an argument of the consumers' utility function. As we shall be exclusively concerned with revenue-increasing and Pareto-improving changes in the tax structure, we avoid the arbitrariness of explicitly specifying, in this economy with heterogeneous agents and without voting, the social planner's objective function. We however limit ourselves (and therefore limit the fisc) to the class linear income tax schemes.

Suppose that there is a unique linear income tax schedule τ

$$t = -\alpha + (1 - \beta)y. \tag{1.1}$$

Taxes paid by a consumer with income y have accordingly two components: a lump-sum transfer $\alpha > 0$ (a lump-sum tax if $\alpha < 0$), and a distortionary labor income tax at the marginal rate $1 - \beta \in [0, 1]$. Note that the tax system is progressive when $\alpha > 0$.

---

5. If one feels uncomfortable with this restriction on the utility function, one can alternatively stipulate, instead of assumption 1, that productivity levels are bounded from below by $\lim_{\ell \to 0}[\beta u_2(c, \ell)]/[u_1(c, \ell)]$. Either assumption will rule out non-active workers.

The social planner sets the tax structure taking into account the reaction of the consumers. Taxes cannot be changed after labor supply and consumption decisions have been taken.

Tax revenues are constrained to finance the exogenous government expenditure $G > 0$:[6]

$$R = \int_0^\infty t(y^*)\, dF(w) \geq G, \tag{1.2}$$

where $y^* = y(\alpha, \beta w)$ denotes the equilibrium labor income of a consumer of productivity w who faces tax schedule τ.

1.3 Equilibrium and comparative statics

A consumer of type w maximizes $u(c, \ell)$ subject to the constraint that consumption cannot exceed after-tax labor income:

$$c \leq \alpha + \beta w \ell, \tag{1.3}$$

and that

$$c \geq 0 \ , \ \ell \in [0, 1]. \tag{1.4}$$

Because of assumption 2, and assuming that the following second-order condition holds

$$(\beta w)^2 u_{11} + 2\beta w u_{12} + u_{22} \leq 0, \tag{1.5}$$

the solution to this problem is unique and interior for every w, and satisfies the first-order condition:

$$\beta w u_1 + u_2 = 0. \tag{1.6}$$

Let $c^* = c(\alpha, \beta w)$ and $\ell^* = \ell(\alpha, \beta w)$ denote that solution. Because of assumption 3, optimal consumption and leisure increase when the lump-sum transfer α increases:

$$c_1^* > 0 \quad \text{and} \quad \ell_1^* < 0, \tag{1.7}$$

where, for instance, c_1^* denotes $\partial c^* / \partial \alpha$.

Because of assumption 4, optimal labor supply increases as the marginal tax rate $1 - \beta$ decreases or as productivity w rises:

$$\ell_2^* > 0. \tag{1.8}$$

6. As we choose to abstract ourselves from intertemporal considerations, we assume that all expenditures are tax-financed.

This obviously implies that pre-tax labor income rises with productivity.

The maximum utility attained by a consumer of type w is therefore $u^* = u(c^*, \ell^*)$. Using the envelope theorem,

$$\frac{\partial u^*}{\partial \alpha} = u_1^* \text{ and } \frac{\partial u^*}{\partial \beta} = w\ell^* u_1^*. \tag{1.9}$$

Taxes paid, at the optimum, by a consumer of type w are

$$t^* = -\alpha + (1 - \beta)\ell^*. \tag{1.10}$$

As a consequence,

$$\frac{\partial t^*}{\partial \alpha} = -1 + (1 - \beta)w\ell_1^* < 0 \tag{1.11}$$

$$\frac{\partial t^*}{\partial \beta} = w[(1 - \beta)w\ell_2^* - \ell^*]. \tag{1.12}$$

2 A menu of linear income tax schedules

Now suppose that the fisc introduces a second linear income tax schedule, and offers to taxpayers a *choice* between two schedules. The first schedule, $\tau = (\alpha, \beta)$, is the one described in the previous section. The second schedule, $\tau' = (\alpha', \beta')$, proposes to consumers a lower marginal tax rate ($\beta' > \beta$) in exchange for a lower lump-sum transfer ($\alpha' < \alpha$). In others terms, tax authorities sell at a "price" of $p = \alpha - \alpha' > 0$ the right to a lower marginal tax rate.

The introduction of this second tax schedule cannot but be Pareto-improving.[7] The welfare of consumers who do not wish to take advantage of the opportunity to buy a reduction in their marginal tax rate is unaffected, while the utility level reached by taxpayers who choose to use the new tax schedule τ' is necessarily higher (they would not otherwise choose the new schedule).

A decrease in the marginal tax rate from $1 - \beta$ to $1 - \beta'$ does not provide the same utility benefit to all consumers: highly productive consumers (those with a

7. Strictly speaking, introducing this second schedule cannot lead to a Pareto-inferior equilibrium. As we shall see below, one can however always construct τ' in such a way that it makes at least one consumer strictly better off.

high w) value it more than low productivity workers. Because the cost for this reduction in the marginal rate is the same (p) for all consumers,[8] only consumers with a high enough productivity, for whom benefit exceeds cost, will choose the new tax schedule.

The central point of this paper is that, for any initial tax schedule, for any income distribution and for any desired reduction in the marginal tax rate, one can always find a reduction p in the lump-sum transfer component of the tax which will result in increased tax revenues—provided taxpayers are presented with both the τ and τ' schedule.

The formal proof, which we now present, follows the logic of the preceding verbal argument.

2.1 Welfare

Let $\bar{y} > 0$ be the income level at which both tax schedules produce the same revenue $\bar{t}$ [i.e., τ and τ' intersect at $(\bar{y}, \bar{t})$ in (y, t) space]:

$$-\alpha + (1 - \beta)\bar{y} = -\alpha' + (1 - \beta')\bar{y} = \bar{t}. \tag{2.1}$$

For a given initial tax schedule τ and a given $\bar{y}$, the second tax schedule τ' is therefore fully characterized by the choice of β'.

Let $\bar{w}$ denote the productivity of a consumer who would have pre-tax labor income $\bar{y}$ under the high marginal rate tax schedule τ, and $\bar{w}'$ denote the productivity of a consumer who would have pre-tax labor income $\bar{y}$ under the low marginal rate tax schedule τ':[9]

$$\bar{w}\ell(\alpha, \beta\bar{w}) = \bar{y} \tag{2.2}$$

$$\bar{w}'\ell(\alpha', \beta'\bar{w}') = \bar{y}. \tag{2.3}$$

The following lemma establishes that a lower productivity level is required under the τ' schedule to achieve the pre-tax income level $\bar{y}$:

Lemma 1 $\bar{w} > \bar{w}' > 0$.

8. A productivity-specific price is not feasible, as productivities are assumed to be unobservable. An income-specific price, equivalent to introducing non-linear income taxation, might be Pareto-superior to our scheme but would be more difficult to design.

9. $\bar{w}$ and $\bar{w}'$ are uniquely defined since $\ell_2 > 0$.

Proof: The second inequality follows from the assumption that $\ell_2 > 0$ and the fact that $\bar{y} > 0$. From eqs. (2.2) and (2.3), we have

$$\begin{aligned} \bar{w}\ell(\alpha, \beta\bar{w}) &= \bar{w}'\ell(\alpha', \beta'\bar{w}') \\ &> \bar{w}'\ell(\alpha, \beta'\bar{w}') \\ &> \bar{w}'\ell(\alpha, \beta\bar{w}'), \end{aligned}$$

since $\alpha > \alpha'$, $\beta < \beta'$, $\ell_1 < 0$ and $\ell_2 > 0$. Now the function $w\ell(x, yw)$ is strictly increasing in w for all positive x and y. Therefore, $\bar{w} > \bar{w}'$.[10]

We can now demonstrate

Proposition 1 *Low productivity consumers ($w \leq \bar{w}'$) choose tax schedule τ. High productivity consumers ($w \geq \bar{w}$) choose tax schedule τ'.*

Proof: Let (c^*, ℓ^*) and $(c^{*\prime}, \ell^{*\prime})$ denote the preferred bundles at the income levels and relative prices implied, respectively, by tax schedules τ and τ' (they are computed along the lines suggested in section 1).

Consider first a low productivity consumer ($w \leq \bar{w}'$). Since pre-tax labor income rises with productivity, and given the definition of $\bar{w}'$ in eq. (2.3), we have $w\ell^{*\prime} \leq \bar{y}$. But then, from eq. (2.1),

$$-\alpha + (1-\beta)w\ell^{*\prime} \leq -\alpha' + (1-\beta')w\ell^{*\prime},$$

which implies that

$$\alpha + \beta w\ell^{*\prime} \geq \alpha' + \beta' w\ell^{*\prime} = c^{*\prime}. \tag{2.4}$$

Hence the bundle $(c^{*\prime}, \ell^{*\prime})$ is affordable at the income and relative price level implied by schedule τ. But (c^*, ℓ^*) is the preferred bundle for that schedule. Since preferences are strictly quasi-concave, we must have $u(c^*, \ell^*) > u(c^{*\prime}, \ell^{*\prime})$, so that a low productivity consumer always chooses the τ schedule.

The proof for high productivity consumers ($w > \bar{w}$) is symmetrical: it obtains by circular permutation of primed and non-primed α, β and ℓ^* variables.

Productivity therefore provides the criterion according to which consumers sort themselves between the two tax schedules: workers with very low (high) productivity never (always) find it worthwhile to "buy" a reduction in their marginal tax rate.

In the absence of tighter restrictions on the utility function, we cannot say precisely which tax schedule will be chosen by which consumer with intermediate

10. Notice that the proof relies heavily on our assumption that the wage elasticity of labor supply is positive; were ℓ_2 equal to zero, we would have $\bar{w}' = \bar{w}$.

productivity (those for whom $\bar{w}' < w < \bar{w}$). This is because, under the new tax schedule, these consumers both work and consume more. Depending on the degree of substitutability between consumption and leisure, the latter effect may or may not dominate the former. For consumers at either end of the income distribution, this ambiguity can be lifted, as established above. What we however know, by continuity, is that there is an odd number of consumers who are indifferent between the two tax schedules.[11]

The possibility, which cannot be ruled out,[12] that there are many consumers indifferent between the two schedules, in general prohibits implementing our menu scheme as a standard simple piecewise linear income tax function. As we shall see below, this multiplicity is however no impediment to the determination of the revenue effects of the introduction of a second linear tax schedule.

2.2 Government revenues

To study the effect of the introduction of the second tax schedule τ' on government revenues, it is useful—as suggested by the foregoing analysis—to separately consider low, intermediate and high productivity workers.

Low productivity workers We have shown above that low productivity consumers ($w \leq \bar{w}'$) choose the τ schedule. Tax collection from these workers is thus not affected by the introduction of the second tax schedule.

Intermediate productivity workers While we do not precisely know who, among intermediate productivity workers ($\bar{w} < w < \bar{w}'$), will opt for the new tax schedule, we can guarantee that anybody who chooses τ' ends up paying more taxes than what he had paid under τ:

Proposition 2 *Any consumer with intermediate productivity ($\bar{w}' \leq w \leq \bar{w}$) pays more taxes under tax schedule τ' than under tax schedule τ.*

Proof: Under tax schedule τ, intermediate consumers have labor income $w\ell^* \leq \bar{y}$, and thus pay, from eq. (2.1), taxes no greater than $\bar{t}$, with consumer $\bar{w}$ paying $\bar{t}$. Under tax schedule τ',

11. Consumer $\bar{w}'$ prefers schedule τ , consumer $\bar{w} > \bar{w}'$ prefers schedule τ', and indirect utility functions are continuous under our assumptions.
12. See the appendix for an example of a sufficient condition guaranteeing the existence of a unique indifferent consumer.

intermediate consumers have labor income $w\ell^{*\prime} \geq \bar{y}$, and thus pay taxes no smaller than $\bar{t}$, with consumer $\bar{w}'$ paying $\bar{t}$.

Therefore, the existence of intermediate productivity consumers who choose the τ' tax schedule increases the government's tax revenues.

High productivity workers We have shown above that all consumers with productivity higher than $\bar{w}$ opt for tax schedule τ'. The impact on tax revenues of the adoption of this tax schedule by high productivity consumers is best understood by decomposing it into two parts. First, at their old labor supply ℓ^*, high productivity consumers would pay less taxes under τ' than under τ since $w\ell^* < \bar{y}$—an effect which raises, *ceteris paribus*, their tax bill and which is the stronger the more productive the worker. Second, labor supply and thus labor income increase under τ' because this new schedule offers a lower marginal tax rate—an effect which increases the tax bill and which is the stronger the more elastic labor supply is relative to the wage rate.[13]

There will therefore be two types of high productivity consumers: those for whom the second effect dominates the first (and who will consequently pay more taxes), and those for whom the first effect dominates the second (and who will pay less taxes).[14]

Designing, for any given τ, a menu of linear income tax schedules which is not revenue-decreasing therefore requires that tax authorities make sure that the possible decrease in taxes levied from the some of the high productivity agents does not offset the revenue gains from intermediate workers and from those high productivity workers who end up paying more taxes.

The following argument, based on the possible existence of an upper-limit to the productivity distribution, shows that this goal can always be attained by designing the tax system appropriately:

---

13. The second effect is clearly absent if, contrary to our assumptions, labor supply were inelastic ($\ell_2 = 0$). In that case, as follows from footnote 10, tax collections from consumers with income below $\bar{y}$ would be unchanged under our scheme, while those from income above $\bar{y}$ would go down—an outcome incompatible with budget balance. The feasibility of our scheme thus crucially hinges on a positive wage elasticity of labor supply.

14. If we were willing to restrict preferences further to ensure that income effects are well-behaved, we would be able to identify the first group of consumers with "upper-middle class" workers who have a productivity larger than, but close to $\bar{w}$, and the second group with "upper class" consumers with incomes at the high end of the income distribution.

Proposition 3 *If there is a highest productivity level $w_{\max}$, then it is always possible to design schedule τ' in such a way that its introduction results in an increase in government revenues: for any desired decrease in the marginal tax rate, it suffices to pick $\bar{y}$ equal to the pre-tax income under τ of the largest taxpayer.*

Proof: Pick an arbitrary tax schedule τ and a $\beta' > \beta$ (i.e., select a lower marginal tax rate). If

$$\bar{y} = w_{\max}\ell(\alpha, \beta w_{\max}),$$

we have, using (2.2),

$$\bar{w} = w_{\max}.$$

This construction guarantees, using the results of propositions 1 and 2, that government revenues will increase: because $\bar{w} = w_{\max}$, the only consumers who choose τ' are "intermediate" workers (with productivity between $\bar{w}'$ and $w_{\max}$) who pay more taxes.

The condition of the proposition is of course overly strong, for it is a sufficient but not necessary condition for our scheme to be successful. An alternative sufficient condition—which we view as uninteresting—would of course be that the economy be on the decreasing side of the Laffer curve, on which tax revenues increase when the marginal rate goes down. It must however be emphasized that the revenue-increasing features of our plan do not hinge, as the condition of proposition 3 makes clear, on such circumstances. Instead, the feasibility of our scheme proceeds from the desirability for high productivity tax payers of a tax package offering reduced marginal tax rates in exchange of a larger lump-sum component of the tax bill.

One must note at this juncture, that our scheme provides a global implementation of the well-known (Sadka [1977]) but local result that optimal tax schedules must feature "no distortion at the top" by setting the marginal tax rate of the most able taxpayer, if she exists, equal to zero. By offering a menu of linear income tax schedules, our plan provides the opportunity to many workers to "purchase" a low marginal tax rate—a Pareto-improving approximation, relative to the single τ scheme, of the (perhaps too complex or unknown, and in general non-linear) optimal income tax function which must be flat at its endpoint. As the computations presented below will show, a large fraction of the population may in practice choose to take advantage of the low marginal rate tax schedule. Our menu scheme, because it is implementable even in the presence of multiple indifferent consumers, has thus much wider applicability that Seade's [1977] previous attempt to explore the non-local implications of the "no distortion at the top." His implementation, in contrast with ours, is of little practical guidance because, although not strictly confined to the top taxpayer, it only affects a minuscule fraction of the population and is thus extremely sensitive to changes in the top of the income distribution.

As our goal—establishing that the introduction of a menu of linear income tax schedules enables the government to raise more revenue—is attained by the foregoing analysis, and as additional results cannot be derived without tighter specification of the utility function, we do not pursue further the analysis of the general case. We instead construct an example which, although it does not satisfy all the assumptions of our theoretical analysis, illustrates in a simple manner the issues at hand, how our scheme may in practice be implemented, and its effects on the tax bill of various sections of the population.

3 An example

Assume that the utility function of our consumers is:

$$u(c,\ell) = ac - \frac{1}{1+\theta}\ell^{1+\theta} \quad a,\theta > 0,$$

so that optimal consumption, labor supply and indirect utility of an agent of type w facing only schedule τ are

$$\begin{aligned} c^* &= \alpha + a^{1/\theta}(\beta w)^{(1+\theta)/\theta}, \\ \ell^* &= (a\beta w)^{1/\theta}, \\ U^* &= a\alpha + \frac{\theta}{1+\theta}(a\beta w)^{(1+\theta)/\theta}. \end{aligned}$$

Thus, the wage elasticity of labor supply is equal to $1/\theta$.

Similarly, optimal consumption, labor supply and indirect utility of an agent of type w facing schedule τ' only are

$$\begin{aligned} c^{*\prime} &= \alpha' + a^{1/\theta}(\beta' w)^{(1+\theta)/\theta}, \\ \ell^{*\prime} &= (a\beta' w)^{1/\theta}, \\ U^{*\prime} &= a\alpha' + \frac{\theta}{1+\theta}(a\beta' w)^{(1+\theta)/\theta}. \end{aligned}$$

It is straightforward to show that when offered a choice between the two tax schedules τ and τ', any agent with productivity above (below) a critical level $\tilde{w}$ will choose schedule τ' (τ)[15], where, using eq. (2.1),

$$\tilde{w} = \left[\frac{1+\theta}{\theta}\frac{\alpha-\alpha'}{a^{1/\theta}(\beta'^{(1+\theta)/\theta}-\beta^{(1+\theta)/\theta})}\right]^{\frac{\theta}{1+\theta}} \tag{3.1}$$

3.1 Necessary and sufficient condition

We now derive a necessary and sufficient condition for the introduction of a second tax schedule to be revenue-increasing. For any given α, β and $\beta' > \beta$, let $z(w, \bar{y})$ denote the change in the tax bill of consumer w if he opts for τ' instead of τ when the income level at which the two schedules intersect is $\bar{y}$—i.e, when all consumers with productivity larger than $\tilde{w}$ choose schedule τ'. Using equations (2.1) and (3.1), it is easy to show that

$$\begin{aligned} z(w,\tilde{w}) = a^{1/\theta}\Big[&\frac{\theta}{1+\theta}[\beta'^{(1+\theta)/\theta}-\beta^{(1+\theta)/\theta}]\tilde{w}^{(1+\theta)/\theta} \\ &+[(1-\beta')\beta'^{1/\theta}-(1-\beta)\beta^{1/\theta}]w^{(1+\theta)/\theta}\Big]. \end{aligned} \tag{3.2}$$

Assume, to make the analysis interesting, that β and β' are greater than $(1+\theta)^{-1}$ (otherwise, the economy is on the downward-sloping side of the Laffer curve).[16] The change in tax revenues stemming from the introduction of τ' is thus

$$\Delta R(\tilde{w}) = \int_{\tilde{w}}^{\infty} z(w,\tilde{w})\, dF(w),$$

since only consumers with productivity larger than $\tilde{w}$ choose tax schedule τ'. Since $\beta' - \beta > 0$ by assumption, we immediately find that $\Delta R(\tilde{w}) > 0$ if and only if

$$\frac{(a\beta)^{1/\theta}\tilde{w}^{(1+\theta)/\theta}}{\Phi(\tilde{w})} > \frac{1+\theta}{\theta}\frac{(1-\beta)\beta^{1/\theta}-(1-\beta')\beta'^{1/\theta}}{\beta'^{(1+\theta)/\theta}-\beta^{(1+\theta)/\theta}}, \tag{3.3}$$

where

$$\Phi(x) = \frac{\int_x^{\infty}(a\beta)^{1/\theta}w^{(1+\theta)/\theta}\, dF(w)}{\int_x^{\infty} dF(w)}.$$

15. Consumer $\tilde{w}$ is indifferent between the two schedules since, for him, $U^* = U^{*\prime}$.

16. For any given α, a and w, the tax revenue function $-\alpha + (1-\beta)w(a\beta w)^{1/\theta}$ reaches a maximum—the top of the Laffer curve, at $\beta = (1+\theta)^{-1}$.

denotes the average pre-tax income, prior to the introduction of the τ' schedule, of consumers with productivity larger than x.

According to equation (3.3), government revenues will thus increase if and only if, measured prior to the introduction of the menu of tax schedules, the pre-tax income of the consumer of type $\tilde{w}$ relative to the average income of consumers with productivity higher than $\tilde{w}$ exceeds the value on the right-hand side of the inequality.

One can always construct the schedule τ' to satisfy this inequality. Since $\beta' > \beta > 1$, the right-hand side of (3.3) is strictly smaller than 1,[17] while the left-hand side converges to 1 as $\tilde{w}$ tends to ∞. Therefore, for any given α, β and β' such that $\beta' > \beta > 1$, one can always find a $\tilde{w}$ (or, equivalently by (3.1), an α') which satisfies (3.3).

In particular, following the suggestion of proposition 3, it is easy to show that selecting $\bar{y}$ equal to the income of the largest taxpayer—when such an individual exists—satisfies inequality (3.3) whichever the shape $F(.)$.

3.2 Numerical simulations

We consider two stylized cases. First, a "U.S" economy with an original marginal tax rate of 33%, and then a "Swedish" economy (or a U.S. economy of yesteryear) with a 60% marginal tax rate.

We assume that productivities follow a $\Gamma(a, b)$ distribution, and choose its parameters to fit as closely as possible the empirical distribution of income in each country for plausible wage elasticities of labor supply $1/\theta$.[18] The resulting model parameters are reported in Table 1.

Our numerical exercises illustrate three basic properties of the model:

- The larger the original marginal tax rate, the larger the fraction of the population which will select the new tax schedule with lower marginal rate but higher intercept which we introduce.

---

17. Let $g(x) \equiv x^{(1+\theta)/\theta}/(1+\theta) - x^{1/\theta}$. It is straightforward to show that $\beta' > \beta > 1$ and $\theta > 0$ imply that $g(\beta') < g(\beta)$—which is a necessary and sufficient condition for the righthand side of eq. (3.3) to be smaller than 1 when $\beta' > \beta$.
18. See Salem and Mount [1976] for a justification of the use of the Γ distribution to model the income distribution, and McDonald and Rensom [1979] for details on the estimation technique.

- The larger the reduction in the marginal rate offered by the second schedule relative to the first, the larger the increase in the intercept required to maintain revenue neutrality, and the fewer the number of taxpayers who will choose the new schedule. In other terms, the lump-sum "price" which must be paid to take advantage of the new schedule rises when the reduction in the marginal rate becomes larger;
- The smallest the wage elasticity of labor supply, the smaller the fraction of the population choosing the new tax schedule.

The last observation suggests that our scheme has significant implications only for economies with relatively high labor elasticities. One should however note that when the labor income elasticity is small, the distortionary costs of taxation are small—so that one should not be overly worried about tax distortions in the first place!

In the "U.S." case, Table 2 shows that it is possible to offer a moderate (5%) Pareto-improving and revenue-neutral decrease in the marginal tax rate, and still have, if the wage elasticity of labor supply is 0.8, 11% of the population choosing to take advantage of the new, lower tax rate. If a more drastic cut (18%) in the marginal tax rate is desired, however, the revenue neutrality constraint imposes that the "price" to be paid by consumers to take advantage of this new low rate be so high that only the top 2% of the taxpayers will participate in the scheme—as illustrated in Table 3. The more elastic labor supply is, the larger the fraction of consumers who will choose to take advantage of the lower marginal rate schedule.

The results are of course more spectacular in the "Swedish" case when the original marginal tax rate is close to (but still to the left of) the top of the Laffer curve. A revenue neutral decrease in the marginal rate from 60% to 50% would affect more than half of the population if the wage elasticity of labor supply is 0.6, and close to a quarter when it is 0.4 (Table 4). A schedule offering a decrease in the marginal tax rate from 60% to 30% (while increasing the average tax rate enough to maintain revenue neutrality) would still be selected by 20% of the population if the elasticity of labor is 0.6 (Table 5).[19]

For both the U.S. and Swedish cases, it is straightforward to show that the effect

19. These values slightly overstate our case, as we are implicitly assuming that there is initially only one 60% tax bracket from which all the consumers choosing the new schedule originate. In practice, however, it is often true (in particular in the "Swedish" case) that all consumers but the very poorest ones are subject to the highest marginal tax rate.

of our scheme is to shift the pre-tax income distribution to the right.

Although these numbers are only illustrative, they emphasize i) that our proposal is most relevant for countries with high marginal tax rates, as it is in those countries that the welfare gains of offering to sell the income tax are greatest; and ii) that large reductions in marginal rates for a large fraction of the population are achievable without decreasing government revenue even if the economy is on the upward-sloping side of the Laffer curve if the fisc "sells" its right of distortionary taxation to the public.

4 Discussion

4.1 Implementation as a piecewise linear tax schedule

An interesting question is whether the Pareto improvement achieved by a menu of two tax schedules can be obtained by modifying the original tax schedule, without offering to taxpayers the *choice* between two schedules. As we showed above, workers with "intermediate" productivity levels do not sort themselves monotonically; that is, there exists workers who choose the new tax schedules who are less productive than some of the workers who choose the original schedule. Thus, unless one is willing to design extremely complex (and fragile) tax rules, our scheme is best implemented—and the proposed Pareto improvement achieved—not by designing a piecewise linear income tax schedule but instead by using a menu of schedules and letting consumers sort themselves out, often in a non-monotonic fashion, according to their best self-interest.

4.2 A menu of N tax schedules

A menu of more than two schedules can provide further welfare improvements. One could introduce a third tax schedule, with a lower marginal rate and a higher intercept (i.e., a lower lump sum transfer) than the second schedule. An appropriate choice of parameters, along the lines described in the previous section, improves upon the scheme with two schedules. This argument can be repeated for N schedules. Note, in particular, that nothing rules out a zero marginal rate for the Nth (last) schedule chosen by the most productive workers. Needless to say, the magnitude of N is constrained by the administrative costs of an excessively complicated tax system.

4.3 Piecewise linear functions

An immediate generalization of our results implies that any piecewise linear tax schedule, such as those studied by Sheshinski [1989], can be Pareto dominated, without reducing tax revenues, by introducing one or more new tax schedules. The second tax schedule has a marginal rate lower than the highest one of the original piecewise linear schedule and the new schedule intersect the old one only once. [20] In this case the existence of a kink in the original tax schedule does not affect in any way the proofs of the results shown in Sections 2 and 3.[21]

These results lead to a comment on the optimality of discontinuous piecewise linear schedules. Our preceding analysis suggests that the optimal piecewise linear tax schedule is *dis*continuous for any non-degenerate social welfare functions—in contradiction with Sheshinski's [1989] results on the shape of the optimal piecewise linear income tax schedule.[22] The discontinuity of the tax schedule offers the planner the Pareto-improving opportunity of decreasing the marginal tax rate of high productivity consumers. A social planner will thus in general find it optimal to design, within the class of piecewise linear schedules, a discontinuous income tax function—or, equivalently, a menu of linear income tax schedules.

5 Conclusion

Menus of multiple income tax schedules can Pareto-improve upon traditional linear or piecewise linear income tax schedules without reducing tax revenues. By offering a choice between different tax schedules the government "sells" the right of a lower marginal rate in exchange for a higher lump sum contribution. The more productive taxpayers find it in their interest to buy this reduction of marginal tax rates. Multiple tax schedules thus introduce discontinuities into the pre-tax distribution of income.

---

20. If the two tax schedules intersect twice, the second schedule can be defined in such a way that is available only for tax payers with an income higher than the lower intercept.
21. If this is not the case, the proofs have to be slightly generalized with no qualitative changes in the results. Also, note that, in principle, it is not impossible that the conditions for revenue neutrality could be satisfied even if $\bar{y} < y^*$. However, the range of parameter values for which this is possible is likely to be small.
22. This observation has been confirmed recently in independent work by Slemrod et al. [1991], who point out the mistake in Sheshinski's proof.

This result is related to the "no distortions at the top" principle of the optimal taxation literature, which implies that the derivative of the optimal tax schedule should be zero for the most productive taxpayer. Our results generalize this principle: it applies more globally, and does not rely upon the existence of a well defined and identified "most productive tax payer." In addition, our scheme is relatively easy to implement—particularly if the number of schedules is not too high—and unlikely to produce perverse results in the presence of gradual changes in the income distribution.

Numerical simulations suggest that for economies with relatively high marginal rates at the top (but still in the upward sloping part of the Laffer curve) the marginal rates of a relatively large fraction of the population could be cut without reducing tax revenues. In economies with relatively low marginal rates at the top, such as the United States, our scheme would affect the marginal rates of a small but not trivial fraction of the population. Our scheme is more likely to be successful and affect a larger fraction of the population the more elastic is the labor supply to the after tax wage. If the trend of increasing women participation in the labor force continues, this elasticity is likely to be increasing.

Finally, it should be noted that our scheme makes the income distribution more unequal, which might be an undesirable feature politically. However, our proposed scheme is a Pareto-improvement: the rich are getting richer—but not at the expense of the poor.

References

Atkinson, A. and Stiglitz, J., 1980, *Lectures in public economics.*

Laffont, J.-J. and Tirole, J., 1986, Using cost observation to regulate firms, *Journal of Political Economy* 94, 614–641.

McDonald, J. B. and Rensom, M. R., 1979, Functional forms, estimation techniques and the distribution of income, *Econometrica* 47(6), 1513–1525.

Mirrlees, J., 1977, An exploration in the theory of optimal income taxation, *Review of Economic Studies* 38, 175–208.

Sadka, E., 1977, A theory of uniform taxation, *Journal of Public Economics* 7, 397–399.

Salem, A. B. Z. and Mount, T. D., 1976, A convenient descriptive model of income distribution: The Gamma density, *Econometrica* 42, 1115–1127.

Seade, J. K., 1977, On the shape of the optimal tax schedule, *Journal of Public Economics* 7, 203–235.

Sheshinski, E., 1972, The optimal linear income tax, *Review of Economic Studies* 39, 297–302.

Sheshinski, E., 1989, A note on the shape of the optimum income tax schedule, *mimeo*, Harvard University Cambridge, MA.

Slemrod, J., Yitzhaki, S., and Mayshar, J., 1991, *The Optimal Two-Bracket Linear Income Tax*, Working Paper 3847, NBER.

Appendix

Under additional assumptions on the strength of income effects, one can lift any ambiguity as to the behavior of intermediate productivity workers. A sufficient condition is given by

Proposition A *If $u_1^2 - u_2 u_{11}\ell > 0$, there exists a unique $\tilde{w}$, $\bar{w}' < \tilde{w} < \bar{w}$, such that i) consumer $\tilde{w}$ is indifferent between schedules τ and τ'; ii) consumers with productivity lower (higher) than $\tilde{w}$ strictly prefer schedule τ (τ').*

Proof: Let $v[\alpha, \beta w] = u[\alpha + \beta w \ell(\alpha, \beta w), \ell(\alpha, \beta w)]$ denote the indirect utility of a consumer of type w facing schedule τ. Let

$$\Delta(w) = v[\alpha', \beta' w] - v[\alpha, \beta w]$$

denote the utility gain (or loss, if negative) of choosing schedule τ' over schedule τ. Applying the envelope theorem, we have

$$\begin{aligned} \Delta'(w) \equiv \frac{d\Delta}{dw} &= \beta' \ell^{*\prime} u_1^{*\prime} - \beta \ell^* u_1^* \\ &\equiv f(\alpha', \beta', w) - f(\alpha, \beta, w). \end{aligned}$$

Dropping the * superscripts for ease of notation, it is straightforward to show that

$$\begin{aligned} f_\alpha &= \beta \ell u_{11} + \frac{\beta \ell_1}{u_1}\{u_1^2 + (u_{12}u_1 - u_2 u_{11})\ell\} < 0 \\ f_\beta &= \frac{1}{u_1}\{(u_1^2 - u_2 u_{11}\ell)\ell + \beta w \ell_2 [u_1^2 + (u_{12}u_1 - u_2 u_{11})\ell]\} > 0. \end{aligned}$$

The property that $f_\alpha < 0$ follows from assumptions 3 and 4, and while the result that $f_\beta > 0$ requires, in addition, the condition of the proposition.

Thus, since $\alpha' < \alpha$ and $\beta' > \beta$ by construction, $\Delta'(w) > 0$ for all w. Since $\Delta'(.)$ does not change sign, the $\tilde{w}$ which solves $\Delta(w) = 0$ is unique. Existence is guaranteed by proposition 1, which shows that $\Delta(w) < 0$ for $w < \bar{w}'$ and $\Delta(w) > 0$ for $w > \bar{w}$. The property $\Delta(w) > 0$ for all $w > \tilde{w}$ follows from $\Delta' > 0$.

Notice that, from assumption 4,

$$u_1^2 - u_2 u_{11}\ell > u_1 u_{12}\ell,$$

so that the condition of the proposition is satisfied as soon as income effects are "weak." This occurs, in particular, for all utility function for which $u_{12} \leq 0$.[23] Similarly, utility

23. Sheshinski [1989] assumes, to prove that the optimal continuous piecewise tax schedule is convex, that income effects are non-increasing with productivity, and that (in our notation) $\bar{w}'\ell_1(\alpha', \beta'\bar{w}') \leq \bar{w}\ell_1(\alpha, \beta\bar{w})$. Regularity conditions similar to these would lead to an alternative sufficient condition for the existence of a unique $\tilde{w}$.

functions with curvature in consumption $C \equiv -cu_{11}/u_1 < 1$ satisfy proposition A, since the condition of the proposition is equivalent, using eq. (1.6), to

$$C < 1 + \frac{\alpha}{\beta w \ell}.$$

Table 1: Model Parameters

Labor Elasticity ($1/\theta$)	0.2	0.4	0.6	0.8	1.0

Top of Laffer Curve	0.833	0.714	0.625	0.556	0.500

United States (1988) Mean Income = $24,054					
γ_α	1.65	2.15	2.66	3.21	3.82
γ_β	0.70	0.80	0.90	1.05	1.10
mean w	2.35	2.69	2.95	3.06	3.47
median w	1.90	2.29	2.59	2.74	3.17
variance of w	3.36	3.36	3.28	2.91	3.15

Sweden (1985) Mean Income = 77,900 Kr					
γ_α	0.99	1.18	1.39	1.61	1.85
γ_β	0.30	0.30	0.35	0.40	0.40
mean w	3.29	3.95	3.98	4.04	4.62
median w	2.27	2.91	3.08	3.24	3.82
variance of w	10.96	13.15	11.37	10.09	11.55

Table 2: "U.S." Tax Cut from 33% to 28%

Labor Elasticity ($1/\theta$)	0.2	0.4	0.6	0.8	1.0
% of agents affected	0.0†	0.4	3.6	10.8	21.4
"price"	14,735	7,443	4,766	3,261	2,290
equiv. consumption gain	0.0†	0.0†	0.3	1.0	1.9
critical income	292,556	146,706	93,260	63,336	44,151
revenue-neutral income	300,287	150,993	97,375	67,090	47,446
% who pay more tax	0.0†	0.0†	0.0†	0.0†	0.0†
% who pay less tax	0.0†	0.5	3.6	10.8	21.4
% affected who pay more tax	11.6	0.0†	0.0†	0.0†	0.0†
% affected who pay less tax	88.4	99.9	99.9	99.9	99.9

NOTES:

† indicates a number rounded down to 0.

Labor elasticity is the wage elasticity of labor supply.

"Price" is the lump sum tax the agent would have to 'pay' to be allowed to take advantage of the lower marginal tax rate.

Equivalent consumption gain is the fraction of additional consumption (in $) which would make the average agent indifferent to the existence of the new tax schedule.

Critical income is the income level above which agents opt for the new tax schedule.

Revenue-neutral income is the income below which agents who have opted for the new schedule pay more tax, while those above pay less.

% who pay more/less tax shows the breakdown of agents affected by the new schedule as a fraction of the entire population.

% *affected* who pay more/less tax shows the breakdown only of agents who opt for the new schedule.

Table 3: "U.S." Tax Cut from 33% to 15%

Labor Elasticity ($1/\theta$)	0.2	0.4	0.6	0.8	1.0
% of agents affected	0.0†	0.0†	0.5	2.3	5.9
"price"	79,748	44,024	30,769	23,001	17,741
equiv. consumption gain	0.0†	0.0†	0.2	0.9	2.4
critical income	432,180	232,683	158,576	115,572	86,891
revenue-neutral income	454,825	255,995	182,913	139,806	110,235
% who pay more tax	0.0†	0.0†	0.0†	0.0†	0.0†
% who pay less tax	0.0†	0.0†	0.5	2.3	5.9
% affected who pay more tax	5.0	0.2	0.0†	0.0†	0.0†
% affected who pay less tax	95.0	99.8	99.9	99.9	99.9

NOTES:

† indicates a number rounded down to 0.

Labor elasticity is the wage elasticity of labor supply.

"Price" is the lump sum tax the agent would have to 'pay' to be allowed to take advantage of the lower marginal tax rate.

Equivalent consumption gain is the fraction of additional consumption (in $) which would make the average agent indifferent to the existence of the new tax schedule.

Critical income is the income level above which agents opt for the new tax schedule.

Revenue-neutral income is the income below which agents who have opted for the new schedule pay more tax, while those above pay less.

% who pay more/less tax shows the breakdown of agents affected by the new schedule as a fraction of the entire population.

% *affected* who pay more/less tax shows the breakdown only of agents who opt for the new schedule.

Table 4: "Swedish" Tax Cut from 60% to 50%

Labor Elasticity ($1/\theta$)	0.2	0.4	0.6
% of agents affected	2.1	21.2	52.2
"price"	40,654	14,834	5,104
equiv. consumption gain	0.8	8.1	18.0
critical income	397,202	141,585	47,578
revenue-neutral income	425,601	154,805	54,394
% who pay more tax	0.2	0.0†	0.0†
% who pay less tax	1.9	21.2	52.2
% affected who pay more tax	7.6	0.0†	0.0†
% affected who pay less tax	92.4	99.9	99.9

NOTES:

† indicates a number rounded down to 0.

Labor elasticity is the wage elasticity of labor supply.

"Price" is the lump sum tax the agent would have to 'pay' to be allowed to take advantage of the lower marginal tax rate.

Equivalent consumption gain is the fraction of additional consumption (in Kr.) which would make the average agent indifferent to the existence of the new tax schedule.

Critical income is the income level above which agents opt for the new tax schedule.

Revenue-neutral income is the income below which agents who have opted for the new schedule pay more tax, while those above pay less.

% who pay more/less tax shows the breakdown of agents affected by the new schedule as a fraction of the entire population.

% *affected* who pay more/less tax shows the breakdown only of agents who opt for the new schedule.

Table 5: "Swedish" Tax Cut from 60% to 30%

Labor Elasticity ($1/\theta$)	0.2	0.4	0.6
% of agents affected	0.2	5.9	19.3
"price"	215,856	104,796	59,213
equiv. consumption gain	0.3	8.4	28.0
critical income	676,490	308,472	163,541
income of tax burden switch	761,025	385,901	228,796
% who pay more tax	0.0†	0.0†	0.0†
% who pay less tax	0.2	5.9	19.3
% affected who pay more tax	2.9	0.0†	0.0†
% affected who pay less tax	97.1	99.9	99.9

NOTES:

† indicates a number rounded down to 0.

Labor elasticity is the wage elasticity of labor supply.

"Price" is the lump sum tax the agent would have to 'pay' to be allowed to take advantage of the lower marginal tax rate.

Equivalent consumption gain is the fraction of additional consumption (in Kr.) which would make the average agent indifferent to the existence of the new tax schedule.

Critical income is the income level above which agents opt for the new tax schedule.

Revenue-neutral income is the income below which agents who have opted for the new schedule pay more tax, while those above pay less.

% who pay more/less tax shows the breakdown of agents affected by the new schedule as a fraction of the entire population.

% *affected* who pay more/less tax shows the breakdown only of agents who opt for the new schedule.