Abstract

The paper formulates a positive theory of nonlinear income taxation. Each member of society, given his ability, has a most preferred income tax schedule, that maximises his utility subject to revenue and incentive constraints. Such tax schedules are characterised and found to display a broad tendency towards a marginal tax rate that increases with income, when compared to tax schedules designed in ”original position”. Conditions are given under which agents’ preferences over such tax schedules are single-peaked, so that majority voting over this limited class would enact the wishes of the median voter.

1. Introduction

This paper characterises the nonlinear income tax schedules that would be chosen by members of a society with different skill levels, such as the median-skilled class, if they were given the power to design the tax structure with only their own welfare in mind. Such “selfish” income tax schedules tend to display a marginal tax rate that increases more markedly with income than the usual social welfare maximising tax. Concerning many of the examples of the latter computed in the literature, Stern (1976, p.126) observes that:

There are two main features of the calculated tax schedules which look different from actual income tax schedules. Marginal rates are not monotonically increasing - most of the population is in the region where they are falling - and the highest marginal rates are low.

*An earlier draft of this paper was circulated under the title ”Individually Optimal Income Taxation”. I would like to thank Tony Atkinson, Peter Bernholz, Tim Besley, Peter Diamond, Avinash Dixit, Martin Hellwig, Eric Maskin, James Mirrlees, Howard Rosenthal and John Weymark, as well as seminar participants at LSE, ECARE, Basel, Princeton and ESSET Gerzensee for useful comments and discussions.
Of course everything depends on the particular parametrized setting considered, and there do exist examples in which marginal tax rates are increasing over a substantial range of incomes. But in this paper, we suggest that in practice, increasing marginal tax rates over a broad middle range of incomes may reflect the wishes of an ego-tistical median voter rather than an inequality averse social welfare maximiser.

We first explore, in Section 2, the properties of selfishly optimal tax schedules designed to maximise the utility of an agent with an arbitrary level of ability within society’s spectrum of ability levels. The chosen tax schedule is characterised and found to be rather extreme in the sense that the poor are treated very badly, and caught in a “poverty trap” where extra earnings are heavily taxed. Beyond that range, marginal tax rates are generally very low - indeed negative - for income levels immediately below that earned by the agent who designs the schedule. Agents with superior ability again face a nonnegative marginal tax rate. The tax schedule thus reflects ”Director’s law”, as described by Stigler (1970): as much income as possible is redistributed towards the social class which designs the schedule, in casu the median skill class.

The notion that income tax schedules are designed by some class such as the median skill class immediately raises the issue of whether majority voting over tax schedules is a well-defined choice procedure. Clearly, any income tax proposed can always be outvoted by many other ones: all that is needed is a modification of the tax that benefits many people, albeit very little, and harms fewer people, but perhaps more. Thus voting over all possible income tax schedules is not a viable decisionmaking procedure. Instead, we will ask whether a majority choice within the set of “selfish” income tax schedules only yields a well-defined outcome. We can think of society voting for one of a set of “partisan” politicians, who, once elected, implement the preferred policy of their constituency, and cannot bind themselves in advance to do otherwise. Alternatively, we can think of voting over tax proposals taking place in a legislature composed of “naive” politicians who
each propose only a plan that is fully optimal from the point of view of their own constituency, without any ornaments designed to garner broader voter support. Either way, the tax plan of the median voter will prevail if preferences over selfish tax schedules are single peaked. Section 3 provides a proof of single peakedness for a special case, that of quasilinear preferences. It is not clear yet to what extent this result may be generalised.

Section 4 concludes.

2. Characterisation of selfishly optimal tax schedules

This section explores the properties of “selfishly optimal” income tax schedules, designed to maximise the utility of one particular group of agents of given ability level. It is assumed that, as in the usual optimal income tax problem, all agents are free to choose the consumption-leisure configuration that maximizes their utility subject to the ruling anonymous tax schedule. This means that the group that designs the schedule cannot, in general, exploit others to such a degree that it draws all of society’s wealth towards itself.

A number of alternative approaches are taken in the literature. Buchanan (1975) assumes that individuals have an exogenous initial endowment of consumable goods, acquired costlessly, and freely disposable. He assumes that constitutional restrictions impose a marginal tax rate that is nonnegative, while free disposal ensures that the marginal tax rate will be no greater than one. Under this approach, if the designer of the tax has an initial income of $\zeta$, after tax income $(x)$ will, under his optimal schedule, take the form:

\[
\begin{align*}
x &= z + (\pi - \zeta) & \text{for } z \leq \pi \\
x &= \pi & \text{for } z \geq \pi
\end{align*}
\]

(1)

where $\pi$ is determined by equating total social income before and after tax. There are also a number of articles [Romer (1975, 1977), Roberts (1977), Metzler and Richard (1981)] which envisage the agent choosing among linear tax schedules.
We will consider the tax system designed by a person of type \( i \), who knows that his wage will be \( w_i \) once the tax is put into effect. A full description of the tax schedule is given by a list of each agent’s income net and gross of tax: \( (x_j, z_j) \) for each type of member of society \( j = 1, \ldots, n \). All tax schedules which induce all types to self-select at these points are equivalent.

We follow Guesnerie and Seade (1982) in assuming that there are only finitely many types of agents \( j = 1, \ldots, n \), with wages \( w_1, \ldots, w_n \). Moreover, all types have identical tastes \( u(x, \ell) \) defined over consumption \( x \) and labour \( \ell \); they differ only in their wage rate. The tax system is required to raise a revenue of at least \( R \) per capita (any arbitrary feasible amount). It is infeasible to force anyone to provide negative labour or to consume less than a minimum amount, taken to be zero.\(^1\) The tax schedule that is optimal for type \( i \) will then solve the following maximization problem.

\[
\max \{ u(x_i, z_i/w_i) \} \quad (2)
\]

subject to

\[
u(x_j, z_j/w_j) \geq u(x_k, z_k/w_k) \quad \forall j, k \quad (3)
\]

\[
u(x_j, z_j/w_j) \geq u(0, 0) \quad \forall j \quad (4)
\]

\[
0 \leq z_j \leq Lw_j \quad \forall j \quad (5)
\]

\(^1\)It is important to note that, unlike in the usual social welfare maximization problem, this constraint is always binding. For without it, any agent of ability \( i > 1 \) can design a tax schedule that gives him arbitrarily high utility by forcing arbitrarily large negative consumption on any agent who cannot reach a threshold pretax income \( \bar{\pi} > Lw_1 \), and redistributing all the money to himself and other more skilled agents who are able to earn at least \( \bar{\pi} \).

Since our problem is, in general, not convex the fact that the constraint always binds does not imply that it necessarily holds with equality at the optimum.
Voting over Nonlinear Income Tax Schedules

\[ \sum_{j=1}^{n} p_j(z_j - x_j) \geq R \]  

(6)

where

- \( w_j \) = wage of type \( j \);
- \( z_j \) = before-tax income of type \( j \);
- \( p_j \) = proportion of population that is of type \( j \);
- \( L \) = endowment of labour time.

The following quite weak assumption ensures that given any tax schedule, types with higher wage rates will select at least as high a value of pre-tax income.

**Assumption U.** Preferences can be represented by a twice differentiable utility function \( u(x, \ell) \) that is strictly increasing in consumption \( (x) \) and strictly decreasing in labour \( (\ell) \). In the absence of taxation, given any non-wage income, the demand for consumer goods is increasing in the real wage rate; that is, \(-\ell u_x/u_x\) is a strictly increasing function of \( \ell \) for all \( x \).

As interpreted by Mirrlees (1971), Assumption U means that the pre-tax incomes selected by agents will be nondecreasing in their wage rate. It also implies that all incentive constraints that do not involve adjacent types are redundant. Moreover, since types with a higher wage rate can always attain at least as high a level of utility as those with a lower wage, all but one of the constraints listed in (4) are also redundant. Choosing the index of types in ascending order of wage rates, and introducing the notation \( u_j(x, z) \equiv u(x, z/w_j) \), the maximization problem can be reduced to the following simplified version:

\[ \max_{\{(x_i, \ldots, x_n, z_1, \ldots, z_n)\}} u_i(x_i, z_i) \]  

(7)

\[ ^2 \text{Alternatively, } \lim_{\ell \to L} u(x, \ell) = -\infty. \]
subject to
\[ u_j(x_j, z_j) \geq u_j(x_{j+1}, z_{j+1}) \quad \text{for } j = 1, \ldots, n - 1 \]  
\[ u_j(x_j, z_j) \geq u_j(x_{j-1}, z_{j-1}) \quad \text{for } j = 2, \ldots, n \]  
\[ u_1(x_1, z_1) \geq u(0, 0) \]  
\[ 0 \leq z_j \leq Lw_j \quad \forall j \]  
\[ \sum_{j=1}^{n} p_j(z_j - x_j) \geq R \]

Let us call the \( n \) points \( \{(x_1, z_1), \ldots, (x_n, z_n)\} \) at which the various types self-select \( \{S_1, \ldots, S_n\} \). It is understood that several adjacent points may well coincide. It is useful, following the approach taken in Röell (1985), to classify all adjacent pairs of points by the identity of the incentive constraint (if any) that is binding between the two. Two points \( S_j \) and \( S_{j+1} \) will be termed upward-linked if \( u_j(x_j, z_j) = u_j(x_{j+1}, z_{j+1}) \), i.e. they are connected by the upward incentive constraint of type \( j \), and downward-linked if \( u_{j+1}(x_j, z_j) = u_{j+1}(x_{j+1}, z_{j+1}) \). Clearly points that coincide are both upward- and downward-linked. And if \( -\ell u_t/u_x \) is strictly increasing in \( \ell \) for any \( x \), any points that are both upward- and downward-linked cannot be distinct.

Several properties of the optimal tax schedule can now be derived. The reader is reminded that the index \( i \) will be reserved for the “target” type, that is the agent whose utility is maximised by the tax schedule.

**Proposition 1** The budget constraint (6) is binding at the optimum.

**Proof.** Unfortunately we cannot appeal to Proposition 1 of Guesnerie and Seade (1982) as they only show that a slack budget constraint means that the welfare of some type(s), not specified in advance, can
be improved. In our case it must be shown that the welfare of a particular type, namely type $i$, can be enhanced.

Suppose, then, that the budget constraint is slack:

$$\sum_{j=1}^{n} p_j(z_j - x_j) = R + \varepsilon$$  \hspace{1cm} \text{where } \varepsilon > 0 \tag{13}$$

We claim that there exists a nonnegative vector $(\varepsilon_1, ..., \varepsilon_n)$, with $\varepsilon_i > 0$, such that the tax schedule $\{(x_j + \varepsilon_j, z_j)_{j=1,...,n}\}$ satisfies all incentive constraints and $\sum_{j=1}^{n} \varepsilon_j \leq \varepsilon$. Then, since utility is strictly increasing in the consumption good, the original scheme cannot have been optimal from the viewpoint of type $i$.

Take any number $\varepsilon_i > 0$ and, given the original schedule $\{(x_j, z_j)_{j=1,...,n}\}$, define the function $f(\varepsilon_i) = (f_1(\varepsilon_i), ..., f_n(\varepsilon_i))$ recursively away from $i$ as follows:

$$f_i(\varepsilon_i) = \varepsilon_i$$

$$u_{i-1}(x_{i-1} + f_{i-1}(\varepsilon_i), z_{i-1}) = \max\{u_{i-1}(x_{i-1}, z_{i-1}), u_{i-1}(x_i + \varepsilon_i, z_i)\}$$

$$u_{i-2}(x_{i-2} + f_{i-2}(\varepsilon_i), z_{i-2}) = \max\{u_{i-2}(x_{i-2}, z_{i-2}), u_{i-2}(x_{i-1} + f_{i-1}(\varepsilon_i), z_i)\}$$

etc.; this procedure can be mirrored for values of the index greater than $i$. $f$ is continuous because utility is strictly increasing in consumption and continuous, and clearly $f$ is nonnegative. Moreover, $f(0) = 0$. Hence for $\varepsilon_i > 0$ small enough,

$$0 < \sum_{j=1}^{n} p_j f_j(\varepsilon_i) \leq \varepsilon. \tag{15}$$

For any such value of $\varepsilon_i$, $f(\varepsilon_i)$ is the nonnegative vector claimed to exist. ■

**Proposition 2** For all $j > i$, $(S_{j-1}, S_j)$ are downward-linked.

**Proof.** Suppose not: for some $j > i$, let

$$u_j(x_{j-1}, z_{j-1}) < u_j(x_j, z_j). \tag{16}$$
Define $\varepsilon_j > 0$ as follows:

$$u_j(x_{j-1}, z_{j-1}) = u_j(x_j - \varepsilon_j, z_j). \quad (17)$$

For all greater values of the type index $k > j$, define $\varepsilon_k$ recursively by:

$$u_k(x_{k-1} - \varepsilon_{k-1}, z_{k-1}) = u_k(x_k - \varepsilon_k, z_k). \quad (18)$$

Then clearly $\varepsilon_k \geq 0$ for all $k$ (if not, there is a $k > j$ such that $\varepsilon_{k-1} \geq 0$ but $\varepsilon_k < 0$, which implies that $u_k(x_{k-1}, z_{k-1}) > u_k(x_k, z_k)$, that is, the original scheme was not incentive compatible). Since $\varepsilon_j > 0$, net tax revenue is strictly increased by adopting a new schedule with $x_k$ replaced by $x_k - \varepsilon_k$ for all $k \geq j$. Then the budget constraint is not binding and, by Proposition 1, it is possible to enhance $u_j$. So the original plan cannot have been optimal. ■

**Proposition 3** If the optimal plan has the property that $u_1(x_1, z_1) > u_1(0, 0)$, then $(S_j, S_{j+1})$ are upward-linked for all $j < i$.

**Proof.** This proof is slightly more involved because it may not be possible to adjust after-tax income downwards in the way suggested in the proof of Proposition 2. For this might result in some types being placed in a situation where their utility is lower than it would be at $(0, 0)$, violating the minimum utility constraint (10).

Suppose the proposition is false. Let $k$ be the smallest value of the index such that $(S_k, S_{k+1})$ are not upward-linked:

$$u_k(x_k, z_k) > u_k(x_{k+1}, z_{k+1}).$$

Thus all pairs $(S_j, S_{j+1})$ are upward-linked for $j < k$. Now let $\delta_1 > 0$ be given by:

$$u_1(x_1 - \delta_1, z_1) = u_1(0, 0).$$

For any $\varepsilon_1 \in [0, \delta_1]$, define the function $g(\varepsilon_1) = (\varepsilon_1, g_2(\varepsilon_1), \ldots, g_k(\varepsilon_1))$ recursively upwards for $j \leq k$ as follows:

$$u_{j-1}(x_j - g_j(\varepsilon_1), z_j) = u_{j-1}(x_{j-1} - g_{j-1}(\varepsilon_1), z_{j-1}).$$
This function is be continuous and nonnegative, with \( g(0) = (0, \ldots, 0) \), since by assumption all pairs of points (if any) below \( S_k \) are upward-linked. Hence there exists a value of \( \varepsilon_1 > 0 \) small enough so that \( u_k(x_k - g_k(\varepsilon_1), z_k) \geq u_k(x_{k+1}, z_{k+1}) \).

Then there is a feasible cost saving of \( \sum_{j=1}^{k} g_j(\varepsilon_1) \) to be realized by setting \( x_j \) to \( x_j - g_j(\varepsilon_1) \) for \( j = 1, \ldots, k \). Hence, by Proposition (1), the plan cannot have been optimal.

The situation is more complicated whenever constraint (10) is tight, that is, \( u_1(x_1, z_1) = u_1(0, 0) \). For then it need no longer be true that all pairs of points \( (S_j, S_{j+1}) \) are upward-linked for all \( j < i \). For example, in one of the two cases depicted in Figure 1 there are pairs of points which are not upward-linked. However, the structure of links is still quite simple: starting from the bottom at \( j = 1 \), there is first a range of downward links, then an unlinked pair of points, and then an upward-linked range until \( i \) is reached.
Proposition 4 For types $j,k < i$:

(i) If $(S_{k}, S_{k+1})$ are not upward-linked, then $(S_{j}, S_{j+1})$ are downward-linked for all $j < k$, and $u_1(x_1, z_1) = u_1(0, 0)$.

(ii) If $(S_{k}, S_{k+1})$ are not downward-linked, then $(S_{j}, S_{j+1})$ are upward-linked for all $j$ such that $k < j < i$.

(iii) At most one pair of adjacent points in the range \{S_1, ..., S_i\} is unlinked.

Proof. (i) Suppose not: let $j$ be the largest $j \leq k$ such that $(S_{j-1}, S_{j})$ are not downward-linked. Then there exists a strictly positive number $\delta_j > 0$ such that:

\[
u_j(x_{j-1}, z_{j-1}) = u_j(x_j - \delta_j, z_j).
\]

Given any $\varepsilon_j \in [0, \delta_j]$, define the function $h(\varepsilon_j) = (\varepsilon_j, h_{j+1}(\varepsilon_j), ..., h_k(\varepsilon_j))$ recursively by:

\[
u_{j+1}(x_j - \varepsilon_j, z_j) = u_{j+1}(x_{j+1} - h_{j+1}(\varepsilon_j), z_{j+1}),
\]

\[
u_{j+2}(x_{j+1} - h_{j+1}(\varepsilon_j), z_{j+1}) = u_{j+2}(x_{j+2} - h_{j+2}(\varepsilon_j), z_{j+2}),
\]

and so on.

Then $h$ is a continuous nonnegative-valued function and, since all pairs of points in the range \{S_j, ..., S_{k+1}\} are downward-linked by assumption, $h(0) = 0$. Hence there is a value of $\varepsilon_j$ strictly positive yet small enough so that $h_k(\varepsilon_j) \leq \delta_k$, where $\delta_k > 0$ is given by:

\[
u_{k}(x_k - \delta_k, z_k) = u_k(x_{k+1}, z_{k+1}).
\]

Hence a cost saving of $\sum_{j=1}^{k} p_j h_l(\varepsilon_j) > 0$ can be attained without altering $u_i$, and the original plan cannot be optimal.

(ii) Analogous to (i).

(iii) Follows directly from (i) and (ii). ■

We can now investigate the structure of marginal tax rates implied by statements concerning links between adjacent points. Since
much of the tax schedule, viewed as a function assigning a post-tax income to every pretax income, is arbitrary, its slope is not well-defined: any curve in $(z, x)$ space that connects the optimal set of points $\{(z_1, x_1), \ldots, (z_n, x_n)\}$ and nowhere lies above the relevant indifference curves attained at these points is an acceptable optimal tax schedule. Two alternative definitions of the marginal tax rate will therefore be adopted:

The **implicit marginal tax rate** facing type $j$ is derived from the slope of type $j$'s indifference curve at $(z_j, x_j)$:

$$\tau_j \equiv 1 + \frac{\partial u_j(x_j, z_j)/\partial z_j}{\partial u_j(x_j, z_j)/\partial x_j} = 1 + \frac{\partial u(x_j, z_j/w_j)/\partial z_j}{w_j \partial u(x_j, z_j/w_j)/\partial x_j}$$

The **explicit marginal tax rate** connecting types $j$ and $k > j$ is defined whenever $S_j$ and $S_k$ are distinct by:

$$T_{j,k} = 1 - \frac{x_k - x_j}{z_k - z_j}$$

We now relate up- and downward-linkedness to the marginal tax rate.

**Proposition 5** Suppose that $S_j$ and $S_{j+1}$ are distinct:

(i) If $S_j$ and $S_{j+1}$ are downward-linked, then $0 \leq T_{j,j+1} \leq 1$. Moreover $0 \leq \tau_{j+1} \leq 1$ provided that $S_{j-1}$ and $S_j$ are not upward-linked.

(ii) If $S_j$ and $S_{j+1}$ are upward-linked, then $T_{j,j+1} \leq 0$. Moreover $\tau_{j+1} \leq 0$ provided that $S_{j-1}$ and $S_j$ are not downward-linked.

(iii) If $S_j$ and $S_{j+1}$ are not linked, then $\tau_j = 0$ provided that $(S_{j-1}, S_j)$ are not upward-linked ($\tau_j \leq 0$ if $z_j = 0$), and $\tau_{j+1} = 0$ provided that $(S_{j+1}, S_{j+2})$ are not downward-linked.

**Proof.** (i) It is clear that the marginal tax rate can never be greater than or equal to 1 under either definition, by monotonicity of preferences over consumption and labour. Thus, $\frac{\partial u/\partial z}{\partial u/\partial x}$ is negative; and clearly no-one would be prepared to work harder and earn more pre-tax income if that leads to a fall in post-tax income.
As for the nonnegativity of the marginal tax rate:

If $T_{j,j+1} < 0$, then $z_j - x_j > z_{j+1} - x_{j+1}$. Since $S_j$ and $S_{j+1}$ are downward-linked, it is possible to reset $S_{j+1}$ to coincide with $S_j$: type $j + 1$ would be indifferent to the change, and the move would enhance tax revenue.

If $\tau_{j+1} < 0$, then by definition the slope of $j + 1$’s indifference curve through $S_{j+1}$ is greater than 1:

$$\frac{-\partial u_{j+1}(x_{j+1}, z_{j+1})/\partial z_{j+1}}{-\partial u_{j+1}(x_{j+1}, z_{j+1})/\partial x_{j+1}} > 1.$$  \hspace{1cm} (21)

Thus it would be feasible, and revenue-enhancing, to shift $S_{j+1}$ down somewhat by reducing $x_{j+1}$ and $z_{j+1}$ along $j + 1$’s indifference curve.

(ii) Analogous to (i).

(iii) If $\tau_j > (\leq) 0$, then it is revenue-enhancing to move $S_j$ up (down) a little along $j$’s indifference curve without violating any incentive constraints, where downward movement is only feasible if $z_j > 0$. The same argument applies to $S_{j+1}$. ■

Proposition 5 enables us to describe the structure of the optimal tax schedule. As depicted in Figure 2, if the minimum consumption constraint is binding there may be an initial segment where after tax income increases more slowly than before tax income. One could perhaps call this a “poverty trap”, where the poorest are held down to their subsistence minimum and the marginal tax rate is positive (though not greater than 1, of course), discouraging effort. Then there comes a segment where marginal tax rates are very low, indeed negative: every extra unit of pretax income is transformed into more than one unit of after tax income. This segment stretches up to the income level at which the type whose welfare is being maximised is to be found. For incomes above that level, there is again a positive explicit marginal tax rate between distinct points. And all types above the target type $i$ who are not bunched with him at $S_i$ face a nonnegative implicit marginal tax rate; indeed the richest group faces an
Voting over Nonlinear Income Tax Schedules

Figure 2:

implicit marginal tax rate \( \tau_n \) of zero (as for the conventional social welfare-maximising income tax schedule).

If one thinks of the median type as the designer of the schedule, then the model predicts a rather extreme form of increasing marginal tax rates: the marginal tax rate is nonnegative for incomes above that selected by the median type, and negative for at least some range directly below that level.

3. Single peakedness of preferences over selfishly optimal tax schedules

We wish to show that the selfishly optimal tax schedule designed by the median-skilled agent defeats all other selfishly optimal tax schedules in majority voting contests. A sufficient condition for this outcome is single peakedness of voters’ preferences over this class of tax schedules: that is, any agent \( i \) prefers the tax schedule proposed by
a closer agent $j$ over that proposed by a more distant agent $k$ (i.e. $i \leq j < k$ or $i \geq j > k$).

Our proof of the single peakedness property will be limited to the special case of quasilinear preferences over consumption and effort. Such preferences have the restrictive property that leisure is not a normal good: labour hours are determined solely by the wage rate and not affected by the wealth of the agent.

**Assumption U'.** Agents have quasilinear preferences over consumption and effort.

$$u(x, \ell) = x - f(\ell), \text{ where } f' > 0, f'' \geq 0$$

(22)

Observe that convexity of $f$ implies that assumption U is satisfied.

Even this strong quasilinearity restriction on preferences is not enough to yield a general proof of the single peakedness property. It is readily shown in full generality that under tax schedules that are selfishly optimal for target types $i$ and $i + 1$ respectively, the utilities of all higher-skilled types $j > i$ satisfy $u_{j+1}^i \geq u_j^i$.

But for $j < i$, complications arise as a result of the minimum utility constraint (4), which puts a floor under all agents’ utility of at least $u(0, 0)$. We are therefore forced to limit our proof of singlepeakedness to special cases where the minimum utility constraint is slack enough at the optimum to ensure that all types below the target are upward-linked anyway.

Whether single peakedness obtains in a fully general setting with quasi-linear preferences remains an open question.

3.1. **The Case of an extremely slack minimum utility constraint:** $u_1(x_i^i, z_i^i) \geq u_1(0, 0)$ for all target types $i = 1, \ldots, n$.

We first show that in general, the higher the target-skill type who designs the tax schedule, the more effort it elicits from all agents.

---

3Superscripts will be used to denote the target type for whom the tax schedule is selfishly optimal. Thus $u_j^i$ is the utility of type $j$ under $i$’s selfishly optimal schedule.
Theorem 6  **Ranking of tax schedules by effort elicited.**

Let \( \{(x_j^i, z_j^i)\}_{j=1,...,n} \) and \( \{(x_j^{i+1}, z_j^{i+1})\}_{j=1,...,n} \) be selfishly optimal tax schedules chosen by types \( i \) and \( i + 1 \) respectively. Then:

\[
z_j^i \leq z_j^{i+1} \quad \text{for all } j > i. \tag{23}
\]

If, in addition, \( u_1(x_{i+1}^{i+1}, z_{i+1}^{i+1}) \geq u_1(0, 0) \), then:

\[
z_j^i \leq z_j^{i+1} \quad \text{for all } j. \tag{24}
\]

**Proof.** Because the two schedules are optimal,

\[
u_i(x_i^i, z_i^i) \geq u_i(x_i^{i+1}, z_i^{i+1}) \geq u_i(x_i^{i+1}, z_i^{i+1}) \tag{25}
\]

\[
u_{i+1}(x_{i+1}^{i+1}, z_{i+1}^{i+1}) \geq u_{i+1}(x_{i+1}^{i+1}, z_{i+1}^{i+1}) \geq u_{i+1}(x_{i+1}^{i+1}, z_{i+1}^{i+1}) \tag{26}
\]

so that by Assumption U':

\[
(x_i^i, z_i^i) \leq (x_{i+1}^{i+1}, z_{i+1}^{i+1}). \tag{27}
\]

**The case of** \( j > i \). Let \( k \) be the lowest type above \( i \) for whom:

\[
z_k^i > z_k^{i+1}. \tag{28}
\]

I claim that such a type cannot exist. For it is feasible to make the upper end \( \{S_k^i, ..., S_n^i\} \) of \( i \)'s tax plan vertically parallel to \( i + 1 \)'s tax plan, by changing the former to to:

\[
\{(x_k^{i+1} - \varepsilon, z_k^{i+1}), ..., (x_n^{i+1} - \varepsilon, z_n^{i+1})\} \tag{29}
\]

where \( \varepsilon \) is the solution to:

\[
u_k(x_k^i, z_k^i) = u_k(x_k^{i+1} - \varepsilon, z_k^{i+1}). \tag{30}
\]

This change does not violate any incentive constraints because by definition of \( k \), \( z_{k-1}^{i+1} \leq z_k^{i+1} \leq z_{k+1}^{i+1} \), so that \( k - 1 \)'s incentive constraint is not violated: \( k \) is indifferent between his old and his new point.
and $k - 1$ still prefers to stay put. And because of quasilinearity of preferences, indifference curves are vertically parallel so that the new scheme is incentive compatible. The change alters the revenue from the tax schedule by:

$$\Gamma = \sum_{j=k}^{n} p_j(z_{j+1}^i - z_j^i) - (x_{j+1}^i - x_j^i) + \varepsilon$$

(31)

Similarly, we can change the upper end $\{S_k^{i+1}, ..., S_n^{i+1}\}$ of $i + 1$’s tax plan to:

$$\{(x_k^i + \delta, z_k^i), ..., (x_n^i + \delta, z_n^i)\}$$

(32)

where $\delta$ is the solution to:

$$u_k(x_{k+1}^i, z_{k+1}^i) = u_k(x_k^i + \delta, z_k^i).$$

(33)

As before, this change is incentive-compatible. This time the tax revenue is:

$$\Lambda = \sum_{j=k}^{n} p_j[(z_{j+1}^i - z_{j}^i) - (x_{j+1}^i - x_j^i) - \delta].$$

(34)

Observe that, by quasilinearity of preferences, (30) and (33) imply that:

$$\varepsilon = \delta = u_k(x_{k+1}^i, z_{k+1}^i) - u_k(x_k^i, z_k^i).$$

(35)

Hence

$$\Gamma = -\Lambda.$$

(36)

Hence, either one or the other of the two schedules can be modified to (weakly) increase its revenue without altering the utilities of agents $i$ and $i + 1$. So one of them is (weakly) suboptimal.

The case of $j \leq i$. Here exactly the mirror image of the argument presented for $j > i$ applies, subject to one proviso. There is a danger that the vertical shifts proposed in our method of proof may drive down the utilities of some low-skilled agents below $u(0, 0)$. The
additional condition that \( u_1(x_{i+1}^i, z_{i+1}^i) \geq u_1(0, 0) \) (implying that also \( u_1(x_i^i, z_i^i) \geq u_1(0, 0) \), by the Spence-Mirrlees assumption U) ensures that this cannot happen, since by definition the shifts are incentive-compatible.

It is important to note that Theorem 6 will generally fail to hold if the minimum utility constraint (4) is tight. For then it is natural to expect that, the higher the target type for whom the schedule is selfishly optimal, the more agents will be pushed into the downward-linked “poverty trap” region. Such agents’ effort is distorted downward, while if they are not in the poverty trap, their effort is generally distorted upward relative to the first-best effort level (i.e. where the marginal disutility of effort is equal to the wage rate). Thus for agents who are pushed from the upward-linked to the downward-linked region, their effort declines.

Single-peakedness follows immediately from Theorem 6. Observe that it suffices to consider the preferences of the various types over the schemes proposed by two adjacent types, \( i \) and \( i + 1 \).

**Theorem 7 Single peakedness.** Let \( \{x_j^i, z_j^i\}_{j=1,...,n} \) and \( \{x_j^{i+1}, z_j^{i+1}\}_{j=1,...,n} \) be selfishly optimal tax schedules chosen by types \( i \) and \( i + 1 \) respectively. Then,

\[
  u_j(x_j^i, z_j^i) \leq u_j(x_j^{i+1}, z_j^{i+1}) \text{ for all } j > i \tag{37}
\]

and, provided that \( u_1(x_{i+1}, z_{i+1}) \geq u_1(0, 0) \),

\[
  u_j(x_j^i, z_j^i) \geq u_j(x_j^{i+1}, z_j^{i+1}) \text{ for all } j \leq i. \tag{38}
\]

**Proof.** The proof is by induction, working outwards from the middle. Consider the case \( j > i \). For any \( j \), the two conditions

\[
  z_j^i \leq z_j^{i+1} \tag{39}
\]

and

\[
  u_j(x_j^i, z_j^i) \leq u_j(x_j^{i+1}, z_j^{i+1}) \tag{40}
\]
together imply:

$$u_{j+1}(x^i_j, z^i_j) \leq u_{j+1}(x^{i+1}_j, z^{i+1}_j)$$  \hspace{1cm} (41)

But by Proposition 2, the tax schedules are downward-linked when

$$\phi > \rho$$, so that

$$u_{j+1}(x^i_{j+1}, z^i_{j+1}) = u_{j+1}(x^i_j, z^i_j) \leq u_{j+1}(x^{i+1}_j, z^{i+1}_j) = u_{j+1}(x^{i+1}_{j+1}, z^{i+1}_{j+1}).$$

The premise (39) is true for all $$\phi > \rho$$, by ??, and (40) holds for

$$j = i + 1$$. Hence, by induction, (40) is valid for all $$j \geq i + 1$$.

The exact converse of this argument applies for $$\phi < \rho$$. Note, however, that we do need to assume that

$$u_1(x_{i+1}, z_{i+1}) \geq u_1(0, 0)$$

so as to ensure that $$z^i_j \leq z^{i+1}_j$$, and that all points below $$S_i$$ are upward-linked. ■

The proof of these theorems required the rather strong assumption that in the selfishly optimal schedule of the higher type (in case, $$i + 1$$), the very lowest type (type 1) is better off at point $$S_{i+1}$$ than he would be at $$S_0 \equiv (0, 0):

$$u_1(x^{i+1}_{i+1}, z^{i+1}_{i+1}) \geq u_1(0, 0).$$

One would expect all selfishly optimal income tax schedules to display this property only if the distribution of skill levels is fairly tightly packed. For if, for example, the highest type $$n$$ has an undistorted earnings level $$z^*_n$$ (at which $$f'(z^*_n/w_n) = w_n$$) which is greater than $$w_1 L$$, the condition must necessarily fail since earnings of $$z^*_n$$ are unattainable for type 1, and we know that the schedule of target type $$n$$ will, if anything; distort his own effort upward above $$z^*_n$$.

### 3.2. The case of a moderately slack minimum utility constraint: $$u_1(x^*_1, z^*_1) > u_1(0, 0)$$ for all target types $$i = 1, \ldots, n$$.

Here we impose stronger conditions on the functional form of utility but weaker conditions on the configuration of the selfishly optimal tax schedule. In particular, we can show that if $$\ell f''(\ell)/f'(\ell)$$ is decreasing in $$\ell$$, it suffices to assume that $$u_1(x^*_1, z^*_1) > u_1(0, 0)$$ for all
tax schedules to be compared: a condition ensuring that all pairs of points are upward-linked all the way from type 1 up to type \( i \). This is generally a much weaker condition than \( u_1(x_i^i, z_i^i) \geq u_1(0,0) \).

**Theorem 8** *Effort ranking and single peakedness (II).*

Let \( \{(x_j^i, z_j^i)_{j=1,...,n}\} \) and \( \{(x_j^{i+1}, z_j^{i+1})_{j=1,...,n}\} \) be selfishly optimal tax schedules chosen by types \( i \) and \( i+1 \) respectively. Then,

\[
z_j^j \leq z_j^{i+1} \quad \text{and} \quad u_j(x_j^i, z_j^i) \leq u_j(x_j^{i+1}, z_j^{i+1}) \quad \text{for all} \quad j > i
\]

(42)

and, provided that \( \ell f''(\ell)/f'(\ell) \) is decreasing in \( \ell \) (for \( 0 < \ell < L \)) and \( u_1(x_1^i, z_1^i) \) and \( u_1(x_1^{i+1}, z_1^{i+1}) \) both strictly exceed \( u_1(0,0) \),

\[
z_j^j \leq z_j^{i+1} \quad \text{and} \quad u_j(x_j^i, z_j^i) \geq u_j(x_j^{i+1}, z_j^{i+1}) \quad \text{for all} \quad j \leq i.
\]

(43)

**Proof.** We focus on the case \( j < i \), as for \( j \geq i \) the result has already been proven in Theorems 6 and 7. The assumption that the minimum utility constraint is slack ensures, by Proposition 3, that all the points \( \{S_1^1, \ldots, S_i^i\} \) and \( \{S_1^{i+1}, \ldots, S_{i+1}^{i+1}\} \) form upward-linked chains. Let \( k \) be the highest type \( k < i \) such that:

\[
z_k^i > z_k^{i+1}.
\]

(44)

It will be shown that such a type does not exist.

Recall that the general method of proof used in Theorem 6 involves rendering the two optimal tax schedules vertically parallel to one another for all types from \( k \) down, by either replacing \( \{z_1^1, \ldots, z_k^1\} \) by \( \{z_k^i, \ldots, z_k^1\} \equiv \{z_1^{i+1}, \ldots, z_k^{i+1}\} \) or conversely, replacing \( \{z_1^{i+1}, \ldots, z_k^{i+1}\} \) by \( \{z_1^{i+1}, \ldots, z_k^i\} \equiv \{z_1^1, \ldots, z_k^i\} \), adjusting the corresponding consumption levels so as to maintain the upward links and arguing that these alterations produced equal and opposite changes in revenue of \( \Lambda \) and \( -\Lambda \), so that one of them is a feasible way of generating (weakly) more revenue. Then, by Proposition 1, one of the original schedules is (weakly) suboptimal in the first place.
The problem with this method of proof is now that one or the other of the proposed adjustments may be infeasible, because it could push the lowest-skilled agent(s) 1, 2,... down to utility levels below \( u(0,0) \), as depicted in Figure 3.

In the setting of Figure 3, \( k = 3 \) since \( z_3 > z_3^{i+1} \). But the proposed adjustment that places \( \hat{S}_1^{i+1}, \hat{S}_2^{i+1} \) and \( \hat{S}_3^{i+1} \) vertically below the corresponding \( S_1^i, S_2^i \) and \( S_3^i \), while preserving the upward links, leads to an infeasible schedule for which \( \hat{u}_1 < u_1(0,0) \).

But if \( \ell f''(\ell)/f'(\ell) \) is a nonincreasing function of \( \ell \), it is possible to construct feasible tax schedules that are intermediate between \( \{S_j^i\}_{j=1,...,n} \) and \( \{\hat{S}_j^i\}_{j=1,...,n} \) (and similarly for \( \{S_j^{i+1}\}_{j=1,...,n} \) and \( \{\hat{S}_j^{i+1}\}_{j=1,...,n} \)) and raise at least as much revenue as one or both, with \( u_1 \) arbitrarily close to \( u_1(S_1^i) \) (or, similarly, \( u_1(S_1^{i+1}) \)), thus satisfying the minimum utility constraint. Then, we again conclude that one or other of the original tax schedules was suboptimal.
It remains to show that such intermediate schedules can be constructed. We will suppress the superscripts $i$ and $i + 1$ since the argument applies for each of the target types’ schedules separately, and make use of the following fact:

**Lemma** If $\ell f''(\ell)/f'(\ell)$ is decreasing (increasing) in $\ell$, then for any $\lambda \in (0, 1)$, $z$ and $\hat{z}$,

$$f(\hat{z}/w) = \lambda f(z/w) + (1 - \lambda)f(\hat{z}/w)$$  \hspace{1cm} (45)

implies that

$$f(\hat{z}/w') \geq (\leq)\lambda f(z/w') + (1 - \lambda)f(\hat{z}/w')$$  \hspace{1cm} (46)

for $w' < w$.

**Proof.** If $\ell f''(\ell)/f'(\ell)$ is decreasing (increasing) in $\ell$, then the lower is $\lambda$, the less convex the function $f_\lambda(z) \equiv f(\lambda = \omega)$.

In effect this lemma allows us to take a weighted average of any two alternative feasible, incentive-compatible tax schedules (where the weighted average is taken over consumption and the disutility of effort rather than consumption and effort itself) without violating any upward incentive constraints. In particular, given any $\lambda \in (0, 1)$ we can define the points $\{\hat{S}_1, \ldots, \hat{S}_k\}$, intermediate between $\{S_1, \ldots, S_k\}$ and $\{\hat{S}_1, \ldots, \hat{S}_k\}$ as follows:

$$\hat{x}_j \equiv \lambda x_j + (1 - \lambda)\hat{x}_j \quad \text{for } j = 1, \ldots, k.$$  \hspace{1cm} (47)

Then utility is a convex combination:

$$u_j(\hat{x}_j, \hat{z}_j) = \lambda u_j(x_j, z_j) + (1 - \lambda)u_j(\hat{x}_j, \hat{z}_j) \quad \text{for } j = 1, \ldots, k.$$  \hspace{1cm} (48)

The revenue from the points $\{\hat{S}_1, \ldots, \hat{S}_k\}$ of the weighted-average tax scheme is at least the $\lambda$-weighted average of the revenues from the original two schemes:

$$\hat{z}_j - \hat{x}_j = w_j f^{-1}(\lambda f(z_j/w_j) + (1 - \lambda)f(\hat{z}_j/w_j))$$

$$- (\lambda x_j + (1 - \lambda)\hat{x}_j) \geq \lambda(z_j - x_j) + (1 - \lambda)(\hat{z}_j - \hat{x}_j),$$  \hspace{1cm} (49)
since $f$ is convex.

Is this an incentive-compatible plan? By the above Lemma we have for all $j = 2, \ldots, k$:

$$u_{j-1}(\tilde{x}_j, \tilde{z}_j) \leq \lambda u_{j-1}(x_j, z_j) + (1 - \lambda)u_{j-1}(\hat{x}_j, \hat{z}_j)$$
$$\leq \lambda u_{j-1}(x_{j-1}, z_{j-1}) + (1 - \lambda)u_{j-1}(\hat{x}_{j-1}, \hat{z}_{j-1})$$
$$= u_{j-1}(\tilde{x}_{j-1}, \tilde{z}_{j-1}), \quad (50)$$

provided that $z_j$ and $\hat{z}_j$ are both attainable earnings levels for type $j - 1$, i.e. below $w_{j-1}L$. This must be the case because, by assumption, all points in this range are upward-linked. Hence no agents will want to switch to a point intended for a higher type (they are welcome to switch to points intended for lower types, as by upward-linkedness, $z_{j-1} - x_{j-1} \geq z_j - x_j$ and $\hat{z}_{j-1} - \hat{x}_{j-1} \geq z_j - x_j$ for all $j \leq i$. Thus any agent $j$ will still yield at least the $\lambda$-weighted average revenue from $S_j$ and $\hat{S}_j$).

Since by assumption $u_1(x_1, z_1) > 0$, by choosing $\lambda \in (0, 1)$ large enough we can ensure that $u_1(\tilde{x}_1, \tilde{z}_1) \geq 0$.

Summarizing, we can design modifications to the original tax schedules that yield revenue increases of at least $(1 - \lambda)\Lambda$ and $(1 - \lambda')(\Lambda)$ for some $\lambda$ and $\lambda' \in [0, 1)$. Whatever the sign of $\Lambda$, one of these numbers is nonnegative. Hence, by Proposition 1, one of the two tax schedules was (weakly) suboptimal in the first place.

This proves that $z_j^* < z_j^{i+1}$ for all $j = 1, \ldots, n$. The utility inequalities of the statement of the theorem now follow directly, as in the proof of Theorem 7. ■

4. Conclusions

This paper has investigated two issues. The first is, the nature of the "selfishly optimal income tax schedule" that an agent maximising the utility of his own type would choose to implement if he had the power to do so. As summarised in Figure 2, the schedule has some notable features that seem quite descriptive of reality: there is a "poverty trap "at the low end of the income scale where effort is discouraged;
then there is a lower-middle income range where effort is strongly encouraged, indeed subsidised by a negative marginal tax rate; while those who are more skilled than the agent in question do face a positive marginal tax rate, which is, as in the classical optimal income tax problem, zero at the very top. These features, and in particular the negative-marginal-tax range, are rather extreme. In practice one would expect an agent to know roughly but not precisely what his type will be; a schedule that is somewhat intermediate between the Benthamite SWF maximising schedule and the one characterised in this paper.

A second issue is whether one can expect the political process to come to a well-defined choice that maximises the median voter’s welfare. The answer is yes, at least for the quasi-linear case, if voters choose a citizen-candidate who cannot commit himself to campaign promises and therefore is expected to implement the tax schedule that suits his own type best.
References


Bierbrauer, Felix, and John Weymark, 2012,

Brett, Craig and John A. Weymark, 2012, Citizen candidates and voting over incentive-compatible nonlinear income tax schedules.


Buchanan, James M. and Gordon Tullock, 19... The Calculus of Consent (The University of Michigan Press, Ann Arbor).

Cukierman-Metzler??

Diamond, Peter A., 1994, Optimal income taxation: an example with a U-shaped pattern of optimal marginal tax rates, working paper 94-14, Department of Economics, M.I.T.


