Optimal Piecewise Linear Income Taxation

Patricia Apps
University of Sydney

Ngo Van Long
McGill University

Ray Rees
University of Munich

August 14, 2008

Abstract

This paper first sets out to clarify the existing literature on optimal piecewise linear income taxation. It then extends the analysis to the taxation of two-earner households.

JEL classification H21, H31, J22.

Keywords: piecewise linear; taxation; singles; couples.

Please note: this is a report of work in progress. If you should wish to refer to it, we would be grateful if you would ask us for the latest version of the paper.

1 Introduction

The point of departure of the theory of optimal income taxation is the proposition that, ideally, a tax should be levied on an individual's innate productivity endowment, which determines the utility level he can achieve on the labour market. Since this is however unobservable, a tax is instead levied on money income. The underlying model of behaviour, whether in the theory of nonlinear taxation first developed by Mirrlees (1971), or in the theory of linear taxation formulated by Sheshinski (1972), is that of a utility maximizing individual who divides his time optimally between market labour supply and leisure, given his net wage. The gross wage measures his productivity. There is a given distribution of wage rates over the population, and the problem is to maximize some social welfare function defined on individual utilities. In Mirrlees's nonlinear tax analysis, the problem is seen as one in mechanism design. An optimally chosen menu of marginal tax rates and lump sum tax/subsidies is offered, and individuals select from this menu in a way that reveals their productivity type. As well as the government budget constraint therefore, a key role is played by incentive compatibility or self selection constraints. In Sheshinski's linear tax analysis on the other hand, there is no attempt to solve the mechanism design problem. All individuals are pooled, and the problem is to find the optimal marginal tax rate and lump sum subsidy (sometimes called the demogrant) over the population as a whole, subject only to the government budget constraint. In each case, the theory provides an analysis of how concerns with the equity and efficiency effects of a tax system interact to determine the parameters of that system, and in particular its marginal rate structure and degree of progressivity.

In reality most tax systems are neither linear in the sense of Sheshinski nor nonlinear in the sense of Mirrlees,[1] but rather piecewise linear. Gross income is divided into (usually relatively few) brackets and marginal tax rates vary across these brackets. When we consider *formal* income tax systems, narrowly defined, the marginal tax rates are typically strictly increasing with the income levels defining the brackets. We refer to this case of strict marginal rate progressivity as *the convex case*, since it defines for an income earner a convex budget set in the space of gross income-net income/consumption. However, when we widen the definition of the tax system to include cash

[1] The German tax system is the main example of a nonlinear system, with marginal tax rates increasing with income in a piecewise linear way, up to a maximum rate which is then constant with respect to income. For further discussion see Apps and Rees (2008).

benefit transfers that are paid and withdrawn as a function of gross income, which we refer to as the *effective* income tax system, we see that typically this may lead marginal tax rates to fall over some range as gross income increases. Since this introduces nonconvexities into the budget set income earners actually face, we refer to this as the *nonconvex case*.

The problem of the empirical estimation of labour supply functions when a worker/consumer faces a piecewise linear budget constraint has been extensively discussed in the econometrics literature.[2] Moreover, the recent literature[3] on the estimation of the marginal social cost of public funds (MCPF) has been concerned with the deadweight losses associated with raising a marginal unit of tax revenue in the context of some given piecewise linear tax system, which is assumed *not* to represent an optimal tax system. Yet there is surprisingly little analysis of the problem of *optimal* piecewise linear income taxation, even though this cannot be thought of as a simple adaptation of either the linear or nonlinear analyses.[4] There are two main papers in the theoretical literature on this subject, by Sheshinski (1989) and Slemrod et al (1994).[5] We believe these papers leave the literature in a rather unsatisfactory state.

The contribution by Sheshinski first formulated and solved the problem of the optimal two-bracket piecewise linear tax system for an individual worker/consumer. Unfortunately, he claims to have proved that, under standard assumptions, marginal rate progressivity, the convex case, must always hold: in the social optimum, the tax rate on the higher income bracket must always exceed that on the lower. However, Slemrod et al (1996) show that

[2] For a very extensive discussion see in particular Pudney (1989).

[3] See in particular Dahlby (1998)....

[4] It is possible to analyse this problem from the point of view of imposing a piecewise linearity constraint on the optimal tax function which is found by solving a Mirrlees type of of mechanism design problem. For a general theory of this class of problems, see Gjesdal (1988). The present paper is however concerned with the more realistic case in which policy makers are not trying to solve this problem. It can therefore be regarded as an extension of optimal linear taxation, rather than a restricted form of optimal nonlinear taxation. As we see below, interpretation of the results draws on optimal linear taxation theory rather than on Mirrlees' analysis.

[5] Strawczynski (1998) also considers the optimal piecewise linear income tax, but gross income in his model is exogenous and attention is focussed, as in Varian (1980) on income uncertainty, where taxation essentially becomes social insurance. Kesselman and Garfinkel (1978) compare linear and piecewise linear tax systems in a two-type economy, taking however the tax brackets as fixed. Sadka et al (1982) extend this to the case of a continuum of types.

there is a mistake in Sheshinski's proof, in that he ignores the existence of a discontinuity in the tax revenue function in the nonconvex case. They then carry out simulations which, again on standard assumptions, in all cases produce the converse result - the upper-bracket marginal tax rate is optimally always lower. This is however also somewhat problematic, for two reasons. First, in general non-parameterised models there is no reason to rule out the convex case, and there is the suspicion that the specific functional forms and parameter values chosen by Slemrod et al for their simulations are biassed toward nonconvexity. Secondly, in practice in virtually all countries tax systems do in fact exhibit a substantial degree of marginal rate progressivity. It is as if policy makers aim for a basically convex system, but make adjustments to it over particular income ranges which have the effect of introducing nonconvexities. After reading these papers we are not left with a clear idea of the conditions under which we might expect the different cases to occur. The first aim of this paper therefore is to provide a unified analysis to clarify this issue.

A second issue is that, when we come to consider debates on actual tax policy, it becomes clear that a central problem, that of how *couples* should be taxed, is not directly addressed by the literature discussed so far. Thus this paper considers the extension of the analysis of piecewise linear tax systems to the case of two person households.

A small literature has developed on the extension of the linear and nonlinear tax models to the case of two-earner households. In the case of linear taxation, Boskin and Sheshinski (1983) derived the result that optimally, women should be taxed at a lower rate than men. This builds on the observation, dating back to Munnell (1980) and Rosen (1977), that since women have higher compensated labour supply elasticities, standard Ramsey arguments would imply, other things equal, lower tax rates. This is not however a conclusive argument. The optimal tax rate in a linear tax model depends not only on the efficiency effects of taxation, but also on the distributional effects, and it is *a priori* possible that the tax rate on women should optimally be higher, despite the higher elasticities, if this tax rate were a sufficiently better instrument for redistribution than that of men. This depends on the covariance between the marginal social utility of income and gross income of, respectively, men and women. Boskin and Sheshinski use a model calibrated with parameter values meant to be representative of the empirical estimates to derive the result that, when distributional effects are taken into account, the optimal tax rate on women is indeed below that on men. However, this is still just an example, and there has been little further work to test its robustness, though Apps and Rees (1999), (2008) show that, both in the tax reform and optimal linear tax cases, this result, hailed as the "conventional wisdom" in this area, can be put on a firmer foundation.[6]

It is almost a trivial result that male and female tax rates should differ. Equalising their marginal tax rates, as is done in a joint taxation or income splitting system such as those in the US and Germany, amounts to imposing a constraint on the optimal tax problem which cannot increase, and in general will reduce, the optimised value of social welfare. Less trivial is the argument that women should be taxed at lower rates than men with the same gross income.

In the case of nonlinear taxation, Schroyen (2003), Apps and Rees (2008), Brett (2007) and Kleven, Kreiner and Saez (2007) consider the problem of the extension to two-earner households. General results are hard to find, essentially because of the complexity of the two-dimensional screening problem that arises when the productivity of each household member is the household's private information. Even in the relatively simple case of two wage types and therefore four household types, the multiplicity of potentially binding incentive compatibility constraints gives rise to a wide range of possible solutions. Perhaps the main general result is that the tax rates on men and women will vary with their productivity type, so individual taxation is still in general optimal, but the tax rate on a given individual of one of the two types will also depend on the type of his or her partner. In this sense, the tax unit consists of both the individual and the couple.[7]

A simple linear tax system seems to be too constrained, a fully nonlinear system too complex, to be satisfactory approaches to applicable tax systems. Piecewise linear tax systems therefore are of interest on both theoretical and practical grounds. As far as we are aware, this is the first paper which tries to extend the analysis of piecewise linear taxation to the two-earner case. However, it seems to us to be immediately clear that the basic result of the linear tax model - individual taxation for men and women with a lower tax rate for the latter - will carry over to the piecewise linear case. Therefore the useful simplification of assuming that, even under individual taxation, men and women will face the same piecewise linear tax schedule, will also

[6] See also Feldstein and Feenberg (1996).

[7] This suggests that the question that has often been posed in the literature: "should the individual or the couple be the appropriate tax unit" is wrongly formulated, or at least only makes sense in the context of linear taxation.

allow us to focus on the role that progressivity of the tax system plays in the comparison of joint *vs.* individual taxation.

The paper is structured as follows. In the next two sections we consider optimal piecewise linear income taxation of single-earner households, in an attempt to clarify the exisitng literature. We then go on to extend the analysis to taxation of two-earner households.

2 Single Person Households

In this section we present an in-depth analysis of the problem of optimal piecewise linear taxation of single person households, in an attempt to clarify the existing literature. We present first the analysis of the choice problems for the individual in the face of respectively convex and nonconvex tax systems, and then discuss the optimal tax structures in each case.

2.1 Solution to the Consumer Choice Problem

We assume consumers have identical quasilinear utility functions[8]

$$u = x - D(l) \qquad D' > 0, D'' > 0 \tag{1}$$

where x is consumption and l is labour supply. Given a two-bracket tax system with parameters $(a, t_1, t_2, \hat{y})$, with a the lump sum payment to all households, t_1 and t_2 the marginal tax rates in the first and second brackets respectively, and $\hat{y}$ the income level determining the upper limit of the first bracket, the consumer faces the budget constraint

$$x \leq a + (1 - t_1)y \qquad y \leq \hat{y} \tag{2}$$

$$x \leq a + (t_2 - t_1)\hat{y} + (1 - t_2)y \qquad y > \hat{y} \tag{3}$$

where $y = wl$. In analysing the consumer's choice problem, it is useful to work in the (y, x)-space rather than the (l, x)-space used by Sheshinski and Slemrod et al, because in the former the budget constraint is the same for all consumers, in the latter it varies with the wage. We just have to redefine the utility function:

$$u = x - D(\frac{y}{w}) \equiv x - \psi(y, w) \qquad \psi_y > 0, \psi_{yy} > 0, \psi_{yw} < 0 \tag{4}$$

[8] Thus we are ruling out income effects. This considerably simplifies the analysis without, we would argue, losing too much of interest.

The slope of an indifference curve in the (y, x)-space is $\psi_y(y, w) > 0$ and it decreases continuously with w, the consumer's type. We assume a differentiable wage distribution function, $F(w)$, with continuous density $f(w)$, strictly positive for all $w \in [w_0, w_1]$.

In discussing the solutions to the consumer's problem we have to distinguish between the convex and nonconvex cases:

Convex case: $t_1 < t_2$

Here there are three solution possibilities:[9]

(i) Optimal income $y^* < \hat{y}$. In that case we have

$$\psi_y(y^*, w) = 1 - t_1 \Rightarrow y^* = y^*(t_1, w) \tag{5}$$

$$x^* = a + (1 - t_1)y^*(t_1, w) \tag{6}$$

$$u = a + (1 - t_1)y^* - \psi(y^*(t_1, w), w) \equiv v(a, t_1, w) \tag{7}$$

and the derivatives of the indirect utility function are

$$\frac{\partial v}{\partial a} = 1; \qquad \frac{\partial v}{\partial t_1} = -y^*(t_1, w) \tag{8}$$

(ii) Optimal income $y^* > \hat{y}$. In that case we have

$$\psi_y(y^*, w) = 1 - t_2 \Rightarrow y^* = y^*(t_2, w) \tag{9}$$

$$x^* = a + (t_2 - t_1)\hat{y} + (1 - t_2)y^*(t_2, w) \tag{10}$$

$$u = a + (t_2 - t_1)\hat{y} + (1 - t_2)y^*(t_2, w) - \psi(y^*(t_2, w), w) \equiv v(a, t_1, t_2, \hat{y}, w) \tag{11}$$

and the derivatives of the indirect utility function are

$$\frac{\partial v}{\partial a} = 1; \quad \frac{\partial v}{\partial t_1} = -\hat{y}; \quad \frac{\partial v}{\partial t_2} = -(y^*(t_2, w) - \hat{y}); \quad \frac{\partial v}{\partial \hat{y}} = (t_2 - t_1) \tag{12}$$

(iii) Optimal income $y^* = \hat{y}$. In that case we have

$$\psi_y(\hat{y}, w) \leq 1 - t_1 \tag{13}$$

$$x^* = a + (1 - t_1)\hat{y} \tag{14}$$

[9] It is assumed throughout that all consumers have positive labour supply in equilibrium. It could of course be the case that for some lowest sub interval of wage rates consumers have zero labour supply. We do not explicitly consider this case but it is not difficult to extend the discussion to take it into account.

$$u = a + (1 - t_1)\hat{y} - \psi(\hat{y}, w) \equiv v(a, t_1, \hat{y}, w) \tag{15}$$

and the derivatives of the indirect utility function are

$$\frac{\partial v}{\partial a} = 1; \quad \frac{\partial v}{\partial t_1} = -\hat{y}; \quad \frac{\partial v}{\partial \hat{y}} = (1 - t_1) - \psi_y(\hat{y}, w) \geq 0 \tag{16}$$

To summarise these results: the consumers can be partitioned into three groups according to their wage type. Thus let $\tilde{w}$ and $\hat{w}$ be defined by

$$\psi_y(\hat{y}, \tilde{w}) = 1 - t_1 \tag{17}$$

$$\psi_y(\hat{y}, \hat{w}) = 1 - t_2 \tag{18}$$

Then the groups correspond to the subsets of wages[10]

$$C_0 = \{w; w \in [w_0, \tilde{w})\} \tag{19}$$

$$C_1 = \{w; w \in [\tilde{w}, \hat{w}]\} \tag{20}$$

$$C_2 = \{w; w \in (\hat{w}, w_1]\} \tag{21}$$

C_0 consists of consumers in equilibrium at tangencies along the steeper part of the budget constraint, C_1 are the consumers at the kink, and C_2 consists of the consumers at tangencies on the flatter part of the budget constraint. Note that the consumers in C_1, with the exception of type $\tilde{w}$, are effectively constrained at $\hat{y}$; in the sense that they would prefer to earn extra gross income if it could be taxed at the rate t_1, since $\psi_y(\hat{y}, w) < 1 - t_1$, but since it would in fact be taxed at the higher rate t_2, they prefer to stay at $\hat{y}$. Given the continuity of $G(w)$, consumers are continuously distributed around this budget constraint, with both maximised utility v and gross income y continuous functions of w. Finally note that (17) and (18) can be solved to derive the differentiable functions $\tilde{w}(t_1, \hat{y})$, $\hat{w}(t_2, \hat{y})$.

Nonconvex case: $t_1 > t_2$

Here there are again three solution possibilities. First, there is a unique consumer type, denoted by $\hat{w}$, which is in equilibrium indifferent between being in either of the two tax brackets. There are two local maxima that yield the same utility. This type is characterised by the conditions

$$\psi_y(y_1^*, \hat{w}) = 1 - t_1 \tag{22}$$
$$\psi_y(y_2^*, \hat{w}) = 1 - t_2 \tag{23}$$

[10] We assume that the tax parameters are such that none of these subsets is empty.

$$x_1^* = a + (1 - t_1)y_1^*(t_1, \hat{w}) \tag{24}$$

$$x_2^* = a + (t_2 - t_1)\hat{y} + (1 - t_2)y_2^*(t_2, \hat{w}) \tag{25}$$

$$x_1^* - \psi(y_1^*, \hat{w}) = x_2^* - \psi(y_2^*, \hat{w}) \tag{26}$$

which yield as solutions in particular the functions , $y_1^*(\hat{w}, t_1)$ and $y_2^*(\hat{w}, t_2)$, $\hat{w}(t_1, t_2, \hat{y})$. Then, for consumers with wages in $[w_0, \hat{w})$, we have only tangency solutions on the first, flatter segment of the budget constraint

$$\psi_y(y^*, w) = 1 - t_1 \Rightarrow y^*(t_1, w) \tag{27}$$

while for those in $(\hat{w}, w_1]$, we have only tangencies on the second, steeper segment

$$\psi_y(y^*, w) = 1 - t_2 \Rightarrow y^*(t_2, w) \tag{28}$$

It is straightforward to show that $y^*(t_1, w) < y_1^* < \hat{y} < y_2^* < y^*(t_2, w)$. For individuals of type $\hat{w}$, the tax payments at the two local maxima are $t_1 y_1^*(\hat{w}, t_1)$ and $[t_2 y_2^*(\hat{w}, t_2) - (t_2 - t_1)\widehat{y}] > 0$. In this case, although maximised utility is a continuous function of w, optimal gross income and the resulting tax revenue are not: There is an upward jump in both at $\hat{w}$. It is this fact that seems to have been overlooked by Sheshinski.

2.2 The optimal convex tax system

As just shown, in the convex case, consumers are distributed around the same convex budget set, with some in equilibrium on the first, steeper line segment, some in a constrained equilibrium at the kink, and some in equilibrium on the second, flatter line segment. We can derive the optimal piecewise linear tax system for this case as follows. The planner chooses the parameters of the tax system to maximise a social welfare function defined as

$$\int_{C_0} S[v(a, t_1, w)]dF + \int_{C_1} S[v(a, t_1, \hat{y}, w)]dF + \int_{C_2} S[v(a, t_1, t_2, w)]dF \tag{29}$$

where $S(.)$ is a strictly concave and increasing social welfare function. The government budget constraint is

$$\int_{C_0} t_1 y(t_1, w)dF + \int_{C_1} t_1 \hat{y} dF + \int_{C_2} [t_2 y(t_2, w) + (t_1 - t_2)\hat{y}]dF - a - G \geq 0 \tag{30}$$

where $G \geq 0$ is a per capita revenue requirement. From the first order conditions characterizing a maximum of social welfare subject to the government budget constraint[11] we derive the following:

Result 1:

$$\sigma \equiv \int_{C_0 \cup C_1 \cup C_2} \frac{S'}{\lambda} dF = 1 \tag{31}$$

where σ is the average marginal social utility of income over the entire population and λ is the shadow price of tax revenue. Since the same lump sum a is paid to each consumer, this is essentially the same condition as for linear taxation. However, since it implies

$$\int_{C_0 \cup C_1} (\frac{S'}{\lambda} - 1) dF = -\int_{C_2} (\frac{S'}{\lambda} - 1) dF \tag{32}$$

and S'/λ falls with the wage, the left hand side must be positive and so the value of the integral on the right hand side negative. That is, the consumers in C_2, the higher tax bracket, on average have marginal social utilities of income below the population average, and the converse is true for consumers in the lower tax bracket. This is of course what we would expect.

The conditions characterising the optimal marginal tax rates yield[12]

Result 2:

$$t_1^* = \frac{\int_{C_0} (\frac{S'}{\lambda} - 1)[y^* - \hat{y}^*] dF}{\int_{C_0} y_{t_1}(t_1^*, w) dF} \tag{33}$$

$$t_2^* = \frac{\int_{C_2} (\frac{S'}{\lambda} - 1)[y^* - \hat{y}^*] dF}{\int_{C_2} y_{t_2}(t_2^*, w) dF} \tag{34}$$

The denominator, the average (compensated) derivative of gross income with respect to the marginal tax rate, which is negative, can be interpreted as the efficiency effect of the tax. The numerator is the equity effect. Since $y^* < \hat{y}^*$ for the subset C_0, while marginal social utilities of income are above the average, the numerator will also be negative. Likewise $y^* > \hat{y}^*$ for the subset

[11] In deriving these conditions, it must of course be taken into account that the limits of integration $\tilde{w}$ and $\hat{w}$ are functions of the tax parameters. Because of the continuity of optimal gross income and tax revenue in w, these effects all cancel and the first order conditions reduce to those shown here. Note that we are assuming an interior optimum.

[12] For convenience we change the partial derivative notation, writing y_{t_1} for $\partial y / \partial t_1$ and so on. Note also we have used (32) in deriving (33).

C_2, while marginal social utilities will be below average, and so the numerator here is also negative.

Note the strong formal similarity with the results for optimal linear taxation. The welfare gain of piecewise linear over linear taxation arises out of the fact that the marginal tax rates t_1^*, t_2^* reflect more closely the covariation of income with the marginal social utility of income, and the average compensated gross income derivatives, a measure of deadweight loss, within the respective subgroups.

The wholly new element of course is the determination of the optimal income threshold at which the tax brackets change, $\hat{y}^*$.The condition for optimal choice of $\hat{y}$ is:

Result 3:

$$\int_{C_1} \{\frac{S'}{\lambda} v_{\hat{y}} + t_1^*\} dF = -(t_2^* - t_1^*) \int_{C_2} (\frac{S'}{\lambda} - 1) dF \tag{35}$$

The left hand side gives the marginal social benefit of a relaxation of the constraint on the consumer types in C_1 who are effectively constrained by $\hat{y}$. First, for $w \in (\tilde{w}, \hat{w}]$ the marginal utility with respect to a relaxation of the gross income constraint is $v_{\hat{y}} = (1 - t_1) - \psi_y(\hat{y}, w) > 0$, as shown earlier. This is weighted by the marginal social utility of income to these consumer types. Moreover, since they increase their gross income, this increases tax revenue at the rate t_1^*. The right hand side is positive and gives the marginal social cost of increasing $\hat{y}$, thus, since $t_2^* > t_1^*$, reducing the tax burden on the higher income group. This can be thought of as equivalent to giving a lump sum payment to higher rate taxpayers proportionate to the difference in marginal tax rates, and this is weighted by a term reflecting the net marginal social utilities of income to consumers in this group, which is negative, as we just showed. The planner suffers a distributional loss from giving this group a lump sum income increase. Sheshinski argued that if $t_2^* < t_1^*$ this term on the right hand side of () must be negative, yielding a contradiction, and therefore ruling out the possibility of nonconvex taxation. However, because of the discontinuity in the tax revenue function in the nonconvex case, this is not the appropriate necessary condition in that case, as was pointed out by Slemrod et al. We now go on to derive formally the relevant necassary conditions.

2.3 The optimal nonconvex tax system

In this case we can state the optimal tax problem as

$$\max_{a,t_1,t_2,\hat{y}} \int_{w_0}^{\hat{w}} S[v(a,t_1,w)]dF + \int_{\hat{w}}^{w_1} S[v(a,t_1,t_2,\hat{y},w)]dF \tag{36}$$

$$s.t. \quad \int_{w_0}^{\hat{w}} t_1 y(t_1,w)dF + \int_{\hat{w}}^{w_1} [t_2 y(t_2,w) + (t_1-t_2)\hat{y}]dF - a - R \geq 0 \tag{37}$$

where it has to be remembered that indirect utility is continuous in w, but that there is a discontinuity in tax revenue at $\hat{w}$.

From (24)-(26) it is easy to see that a change in a does not effect the value[13] of $\hat{w}$, and so the first order condition with respect to a is just as before, and can be written again as

$$\int_{w_0}^{w_1} (\frac{S'}{\lambda} - 1)dF = 0 \tag{38}$$

However, for each of the remaining tax parameters the discontinuity in gross income will be relevant, because a change will cause a change in $\hat{w}$, the type that is just indifferent to being in either of the tax brackets.

Now define

$$\Delta R = [t_2 y_2^*(\hat{w},t_2) - (t_2-t_1)\widehat{y}] - t_1 y_1^*(\hat{w},t_1) > 0 \tag{39}$$

This is the value of the jump in tax revenue that takes place at $\hat{w}$. Note that tax revenue if a consumer chooses to be in the higher tax bracket is always higher than that if she chooses the lower bracket, even the the tax rate in the latter is higher.

From the first order conditions for the above problem,[14] we then have

Result 4:

The condition with respect to the optimal bracket value $\widehat{y}$ is

$$\frac{\partial \hat{w}}{\partial \widehat{y}} \Delta R f(\hat{w}) = (t_2 - t_1) \int_{\hat{w}}^{w_1} \left\{ \frac{S'}{\lambda} - 1 \right\} dF \tag{40}$$

It can be shown that $\partial\hat{w}/\partial\widehat{y} > 0$, and, on the same arguments as used before, but with $(t_2 - t_1) < 0$, the right hand side is also positive. Thus, there is

[13] Note the usefulness of the quasilinearity assumption in this respect.

[14] Note that equilibrium utilities, and therefore social utility, still vary continuously with w. The only discontuity is the upward jump in the tax revenue function at $\hat{w}$.

nothing *a priori* to rule this case out, contrary to Sheshinski's assertion. The intuition is straightforward. The right hand side now gives the marginal benefit of an increase in $\widehat{y}$ to the planner, namely a lump sum reduction in the net income of higher bracket consumers with below-average marginal social benefit of income. The marginal cost of this is a jump downward in tax revenue from consumers who now find the first tax bracket better than the second. Both marginal benefit and marginal cost are positive.

Result 5:

The condition with respect to t_1 is:

$$t_1^* = \frac{\int_{w_0}^{\hat{w}} (\frac{S'}{\lambda} - 1)[y^* - \hat{y}^*]dF + \frac{\partial \hat{w}}{\partial t_1} \Delta R f(\hat{w})}{\int_{w_0}^{\hat{w}} y_{t_1}(t_1^*, w)dF} \tag{41}$$

The new element here, as compared to the convex case, is the second term in the numerator, which, since $\partial \hat{w}/\partial t_1 < 0$, is also negative. Thus this term acts to increase the absolute value of the numerator, and therefore the value of t_1, as compared to the convex case. The intuition for this term is simply that an increase in t_1 expands the subset of consumers who prefer to be in the upper tax bracket (with the lower tax rate) and so causes an upward jump in tax revenue.

Result 6:

The condition with respect to t_2 is:

$$t_2^* = \frac{\int_{\hat{w}}^{w_1} (\frac{S'}{\lambda} - 1)[y^* - \hat{y}^*]dF + \frac{\partial \hat{w}}{\partial t_2} \Delta R f(\hat{w})}{\int_{\hat{w}}^{w_1} y_{t_2}(t_2^*, w)dF} \tag{42}$$

Again the new element here is the second term in the numerator, which, since $\partial \hat{w}/\partial t_2 > 0$, is positive. Thus this tends to reduce the tax rate in the upper bracket as compared to the convex case. The intuition is that an increase in t_2 widens the subset of consumers who prefer to be in the lower bracket, and so causes a downward jump in tax revenue. This then makes for a lower tax rate in the upper income bracket.

All other terms in these conditions have the same interpretations as in the convex case.

3 Two-earner Households

The previous section summarised the theoretical analysis of optimal two-bracket piecewise linear taxation in the context of the standard model where a single consumer divides his time between leisure and market work. We now want to extend the analysis to the case of two-earner households. First, we construct a simple model of such a household, then we go on to analyse optimal taxation, focussing on the convex case, since this is a simpler context in which to show the implications of the change in household model.

3.1 Household Models

To formulate the household's budget constraint we have to specify the tax system. We can define three possible types of piecewise linear tax system which, in order of increasing restrictiveness, are:

Gender based taxation[15]*:* the household faces two separate two-bracket piecewise linear tax systems, one for each individual, implying that its budget constraint takes the form, for given tax parameters $a_i, t_1^i, t_2^i, \hat{y}_i$

$$x \leq \sum_i [a_i + (1 - t_1^i)y_i] \quad y_i \leq \hat{y}_i \quad i = f, m \tag{43}$$

$$x \leq a_i + a_j + (1 - t_1^i)y_i + (t_2^j - t_1^j)\hat{y}_j + (1 - t_2^j)y_j \quad y_i \leq \hat{y}_i, y_j > \hat{y}_j, \quad i, j = f, m, i \neq j \tag{44}$$

$$x \leq \sum_i [a_i + (t_2^i - t_1^i)\hat{y}_i + (1 - t_2^i)y_i] \quad y_i > \hat{y}_i, \qquad i = f, m \tag{45}$$

Thus (43) refers to the case in which both spouses have gross incomes within the first bracket of their respective tax schedules, (44) to the cases in which one is in the lower and the other in the higher bracket, and (45) the case in which both are in the higher bracket.

Individual taxation: in this case the tax schedules faced by f and m are constrained to be the same, but they are separately applied to the individual incomes y_i. The threshold $\hat{y}$ is a value of individual income. Thus the household budget constraint is:

$$x \leq a + (1 - t_1)\sum_i y_i \quad y_i \leq \hat{y} \quad i = f, m \tag{46}$$

$$x \leq a + (1 - t_1)y_i + (t_2 - t_1)\hat{y} + (1 - t_2)y_j \quad y_i \leq \hat{y}, y_j > \hat{y}, \quad i, j = f, m, i \neq j \tag{47}$$

[15] This useful term was introduced by Alesina et al (2007). This present fornmulation is the extension to the piecewise linear case of the case for linear taxation analysed by Boskin and Sheshinski.

$$x \leq a + 2(t_2 - t_1)\hat{y} + (1 - t_2)\sum_i y_i \quad y_i > \hat{y}, \qquad i = f, m \tag{48}$$

Again we have the three possibilities: both partners are in the same tax bracket, high or low, or they are in different brackets.

Joint taxation: the tax rates t_1, t_2 are applied to total income $y = \sum_i y_i$, and the threshold $\hat{y}$ is a value of total income. Thus the household budget constraint is

$$x \leq a + (1 - t_1)\sum_i y_i \quad \sum_i y_i \leq \hat{y} \quad i = f, m \tag{49}$$

$$x \leq a + (t_2 - t_1)\hat{y} + (1 - t_2)\sum_i y_i \quad \sum_i y_i > \hat{y} \tag{50}$$

Since marginal tax rates are equlised, both partners are always in the same bracket.

We extend the individual quasilinear utility function used previously to the case of two-person households in the simplest possible way. Thus the household solves

$$\max_{x, y_i} x + \psi^f(y_f; w_f) + \psi^m(y_m; w_m) \tag{51}$$

subject to the budget constraints defined above for the respective three types of tax system. We consider the solutions in turn.

Gender-based taxation: We can draw here on the results presented above for the single consumer household. For each individual there are three solution possibilities:

$$-\psi_y^i = (1 - t_1^i) \quad y_i \leq \hat{y}_i, \;\; i = f, m \tag{52}$$

$$(1 - t_1^i) > -\psi_y^i \geq (1 - t_2^i) \quad y_i = \hat{y}_i, \;\; i = f, m \tag{53}$$

$$-\psi_y^i = (1 - t_2^i) \quad y_i > \hat{y}_i, \;\; i = f, m \tag{54}$$

The first is a tangency solution in the lower tax bracket, the second a solution at the kink in the budget constraint, and the third a tangency in the upper tax bracket. As we saw earlier, in the second case the individual is effectively constrained by the tax system, in the sense that she would like to choose a gross income greater than $\hat{y}_i$ as long as she could continue to be taxed at the lower marginal tax rate. It follows that a marginal increase in $\hat{y}_i$ increases utility of all individuals at the kink, while leaving the utility of those satisfying () unchanged. In general the two individuals in a household may be at different types of equilibria, and so there are nine possible cases, defined by the possible pairs of tax brackets in which the household members may choose to be.

Individual taxation: There are the same nine possible cases, and the formal conditions are as those in (52)-(54), but with the i-subscripts removed from the marginal tax rates.

Joint taxation: Each household faces the same piecewise linear tax system based on total household income The first order conditions for the solution, as in the case of single individuals analysed in the preceding section, yield three possible cases:

$$-\psi_y^f = -\psi_y^m = (1-t_1) \qquad \sum_{i=f,m} y_i \leq \hat{y} \tag{55}$$

$$-\psi_y^f = -\psi_y^m = (1-t_1) - \mu \qquad \sum_{i=f,m} y_i = \hat{y} \tag{56}$$

$$-\psi_y^f = -\psi_y^m = (1-t_2) \qquad \sum_{i=f,m} y_i > \hat{y} \tag{57}$$

where $\mu > 0$ is a Lagrange multiplier associated with the constraint $\sum_{i=f,m} y_i \leq \hat{y}$, which in this case is binding. By the Envelope Theorem, we have that

$$\mu = (1-t_1) + \psi_y^m = (1-t_1) + \psi_y^f > 0 \tag{58}$$

gives the derivative of maximised household utility with respect to the constraint $\hat{y}$. To summarise these results: the households can be partitioned into three groups according to their wage type and the total joint income that arises from their optimal choice of labour supply. Thus:

$$\Omega_0 = \{(w_f, w_m); \sum y_i^*(t_1, w_i) \equiv y(t_1, w_f, w_m) < \hat{y}\} \tag{59}$$

$$\Omega_1 = \{(w_f, w_m); \sum y_i^*(t_1, w_i) = \hat{y}\} \tag{60}$$

$$\Omega_2 = \{(w_f, w_m); \sum y_i^*(t_2, w_i) \equiv y(t_2, w_f, w_m) > \hat{y}\} \tag{61}$$

Given the continuity of $f(w_f, w_m)$, consumers are continuously distributed around the budget constraint, with both maximised utility v and gross incomes y_i^*, y increasing, continuous functions of the w_i. This means that household gross income and utility are both strictly increasing functions of each wage rate.

3.2 The optimal convex joint taxation system

As just shown, in the convex case, consumers are distributed around the same convex budget set, with some in equilibrium on the first, steeper line segment, some in a constrained equilibrium at the kink, and some in equilibrium on the second, flatter line segment. We can derive the optimal piecewise linear tax system for this case as follows. The planner chooses the parameters of the tax system to maximise a social welfare function defined as

$$\int\int_{\Omega_0} S[v(a, t_1, w_f, w_m)]f(w_f, w_m)dw_f dw_m \tag{62}$$

$$+\int\int_{\Omega_1} S[v(a, t_1, \hat{y}, w_f, w_m)]f(w_f, w_m)dw_f dw_m \tag{63}$$

$$+\int\int_{\Omega_2} S[v(a, t_1, t_2, w_f, w_m)]f(w_f, w_m)dw_f dw_m \tag{64}$$

where $S(.)$ is a strictly concave and increasing social welfare function. The government budget constraint is

$$\int\int_{\Omega_0} t_1 y(t_1, w_f, w_m)f(w_f, w_m)dw_f dw_m \tag{65}$$

$$+\int\int_{\Omega_1} t_1\hat{y}f(w_f, w_m)dw_f dw_m \tag{66}$$

$$+\int\int_{\Omega_2} [t_2 y(t_2, w_f, w_m) + (t_1 - t_2)\hat{y}]f(w_f, w_m)dw_f dw_m - a \tag{67}$$

$$\geq G \tag{68}$$

where $G \geq 0$ is a per capita revenue requirement. From the first order conditions characterizing a maximum of social welfare subject to the government budget constraint[16] we derive the following:

Result 7:

$$\sigma \equiv \int\int_{\Omega_0\cup\Omega_1\cup\Omega_2} \frac{S'}{\lambda}f(w_f, w_m)dw_f dw_m = 1 \tag{69}$$

[16] In deriving these conditions, it must of course be taken into account that the limits of integration are functions of the tax parameters. Because of the continuity of optimal gross income in w_f, w_m, these effects all cancel and the first order conditions reduce to those shown here.

where σ is the average marginal social utility of income over the entire population and λ is the shadow price of tax revenue. Since the same lump sum a is paid to each household, this is essentially the same condition as for linear taxation. However, since it implies

$$\int\int_{\Omega_0\cup\Omega_1}(\frac{S'}{\lambda}-1)f(w_f,w_m)dw_f dw_m = -\int\int_{\Omega_2}(\frac{S'}{\lambda}-1)f(w_f,w_m)dw_f dw_m \tag{70}$$

and S'/λ falls with each wage, the left hand side must be positive and so the value of the integral on the right hand side negative. That is, the consumers in Ω_2, the higher tax bracket, on average have marginal social utilities of income below the population average, and the converse is true for consumers in the lower tax bracket. This is of course what we would expect.

The conditions characterising the optimal marginal tax rates yield

Result 8:

$$t_1^* = \frac{\int\int_{\Omega_0}(\frac{S'}{\lambda}-1)[y^*-\hat{y}^*]f(w_f,w_m)dw_f dw_m}{\int\int_{\Omega_0} y_{t_1}(t_1^*,w)f(w_f,w_m)dw_f dw_m} \tag{71}$$

$$t_2^* = \frac{\int\int_{\Omega_2}(\frac{S'}{\lambda}-1)[y^*-\hat{y}^*]f(w_f,w_m)dw_f dw_m}{\int\int_{\Omega_2} y_{t_2}(t_2^*,w)f(w_f,w_m)dw_f dw_m} \tag{72}$$

The denominator, the average (compensated) derivative of gross household income with respect to the marginal tax rate, which is negative, can be interpreted as the efficiency effect of the tax. The numerator is the equity effect. Since $y^* < \hat{y}^*$ for the subset Ω_0, while marginal social utilities of income are above the average, the numerator will also be negative. Likewise $y^* > \hat{y}^*$ for the subset Ω_2, while marginal social utilities will be below average, and so the numerator here is also negative.

Note the strong formal similarity with the results for optimal linear taxation. The welfare gain of piecewise linear over linear taxation arises out of the fact that the marginal tax rates t_1^*, t_2^* reflect more closely the covariation of income with the marginal social utility of income, and the average compensated gross income derivatives, a measure of deadweight loss, within the respective subgroups.

The wholly new element of course is the determination of the optimal income threshold at which the tax brackets change, $\hat{y}^*$.The condition for optimal choice of $\hat{y}$ is:

Result 9:

$$\int\int_{\Omega_1}\{\frac{S'}{\lambda}v_{\hat{y}}+t_1^*\}f(w_f,w_m)dw_fdw_m=-(t_2^*-t_1^*)\int\int_{\Omega_2}(\frac{S'}{\lambda}-1)f(w_f,w_m)dw_fdw_m \tag{73}$$

The left hand side gives the marginal social benefit of a relaxation of the constraint on the consumer types in Ω_1 who are effectively constrained by $\hat{y}$. First, for $w_f, w_m \in \Omega_1$ the marginal utility with respect to a relaxation of the gross income constraint is $v_{\hat{y}} > 0$, as shown earlier. This is weighted by the marginal social utility of income to these household types. Moreover, since they increase their gross income, this increases tax revenue at the rate t_1^*. The right hand side is positive and gives the marginal social cost of increasing $\hat{y}$. Since $t_2^* > t_1^*$, this reduces the tax burden on the higher income group. This can be thought of as equivalent to giving a lump sum payment to higher rate taxpayers proportionate to the difference in marginal tax rates, and this is weighted by the sum of net marginal social utilities of income to consumers in this group, which is negative, as we just showed.

References

[1] A Alesina, A. Ichino and L.Karabarbounis, 2007, "Gender based taxation and the division of family chores", Harvard University Discussion Paper.

[2] P F Apps and R Rees, 1988, "Taxation and the Household", *Journal of Public Economics*, 35, 355-369.

[3] P F Apps and R Rees, 1999a, "On the Taxation of Trade within and between Households", *Journal of Public Economics*, 75, 241-263.

[4] P F Apps and R Rees, 1999b, "Joint vs Individual Taxation in Models of Household Production", *Journal of Political Economy*, 107, 178-190.

[5] P F Apps and R Rees, 2008, *Public Economics and the Household*, Cambridge: Cambridge University Press.

[6] B Dahlby, 1998, "Progressive Taxation and the Marginal Social Cost of Public Funds", *Journal of Public Economics*, 67, 105-122.

[7] M J Boskin and E Sheshinski, 1983, "Optimal Tax Treatment of the Family: Married Couples", *Journal of Public Economics,* 20, 281-297.

[8] M Feldstein and D R Feenberg, 1996, “The Taxation of Two-Earner Families”, in M Feldstein and J M Poterba (eds), *Empirical Foundations of Household Taxation,* University of Chicago Press, 39-76.

[9] H J Kleven, C T Kreiner and E Saez, 2007, "The Optimal Income Taxation of Couples as a Multi-Dimensional Screening Problem", CESifo Working Paper 2092.

[10] J A Mirrlees, 1971, “An Exploration in the Theory of Optimum Income Taxation”, *Review of Economic Studies*, 38, 175-208.

[11] A Munnell, 1980, “The Couple versus the Individual under the Federal Personal Income Tax”, in H Aaron and M Boskin (eds),*The Economics of Taxation,* The Brookings Institution, 247-280.

[12] S Pudney, 1989, *Modelling Individual Choice: The Econometrics of Corners, Kinks and Holes,* Basil Blackwells, Oxford.

[13] H Rosen, 1977, “Is It Time to Abandon Joint Filing”, *National Tax Journal,*XXX, 423-428.

[14] E Sadka, I Garfinkel and K Moreland, (1982), "Income Testing and Social Welfare: An Optimal Tax-Transfer Model", in I Garfinkel (ed), *Income Tested Transfer Programs,: The Case For and Against*, New York: Academic Press.

[15] A Sandmo, 1990, “Tax Distortions and Household Production”, *Oxford Economic Papers*, 42, 78-90.

[16] F Schroyen, 2003, ”Redistributive taxation and the household: The case of individual filings”, *Journal of Public Economics, 87, 2527-2547.*

[17] E Sheshinski, 1972, “The Optimal Linear Income Tax”, *Review of Economic Studies*, 39, 297-302.

[18] E Sheshinski, 1989, "Note on the Shape of the Optimum Income Tax Schedule", *Journal of Public Economics*, 40, 201-215.

[19] J Slemrod, S Yitzhaki, J Mayshar and M Lundholm, 1994, "The Optimal Two-Bracket Linear Income Tax", *Journal of Public Economics*, 53, 269-290.Individual optimization.

[20] M Strawczinsky, 1988, Social Insurance and the Optimum Piecewise Linear Income Tax", *Journal of Public Economics*,, 69, 371-388.

[21] H R Varian, 1980, "Redistributive Taxes as Social Insurance", *Journal of Public Economics*,, 141(1), 49-68.