Optimal Piecewise Linear Income Taxation

PATRICIA APPS
NGO VAN LONG
RAY REES

CESIFO WORKING PAPER NO. 2565
CATEGORY 1: PUBLIC FINANCE
FEBRUARY 2009

CESifo *Working Paper No. 2565*

Optimal Piecewise Linear Income Taxation

Abstract

Given its significance in practice, the piecewise linear tax system seems to have received disproportionately little attention in the literature on optimal income taxation. This paper offers a simple and transparent analysis of its main characteristics.

JEL Code: H21, H31, J22.

Keywords: piecewise linear, income, taxation.

Patricia Apps
Law Faculty
University of Sydney
p.apps@usyd.edu.au

Ngo Van Long
Economics Department
McGill University
ngo.long@mcgill.ca

Ray Rees
Center for Economic Studies at the
University of Munich
Schackstrasse 4
80539 Munich
Germany
ray.rees@lrz.uni-muenchen.de

January 19, 2009
We would like to thank Michael Hoy and Efraim Sadka as well as participants in seminars at IZA and Vanderbilt University for helpful comments on this paper. The usual disclaimer applies.

1 Introduction

The foundations of the theory of optimal income taxation were provided by the theory of nonlinear taxation first developed by James Mirrlees (1971), and the theory of linear taxation formulated by Eytan Sheshinski (1972). In Mirrlees's analysis, the problem is seen as one of 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 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 of a tax system interact to determine the parameters of that system, and in particular its marginal rate structure and degree of progressivity. As Boadway (1998) points out, the optimal nonlinear tax is Pareto superior to a linear tax for any given revenue requirement and set of consumers, implying a superior tradeoff between equity and efficiency. Nevertheless, tax policy makers or "central planners" do not seem to adopt the Mirrlees approach to the design of tax systems in practice.

In reality most tax systems are neither linear in the sense of Sheshinski nor nonlinear in the sense of Mirrlees, but rather piecewise linear. Gross income is divided into (usually relatively few) brackets and marginal tax rates are constant within but vary across these brackets.[1] When we consider *formal* income tax systems, 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 transfers that are paid and withdrawn as a function of gross income 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*

[1] The German tax system is a rare exception to this. It has four brackets and in the second and third of these marginal tax rates increase linearly with income. For further discussion see Apps and Rees (2009a).

case. The reason for planners' preference for piecewise linear as opposed to optimal nonlinear tax systems could be that the former overcome a large part of the inefficiency of a simple linear tax while remaining relatively simple to implement.[2] The present paper is concerned with the realistic case in which policy makers are not trying to solve a mechanism design problem. It can therefore be regarded as an extension of optimal linear taxation, rather than a restricted or approximative form of optimal nonlinear taxation. As we see below, interpretation of the results draws heavily on optimal linear taxation theory.

The problem of the empirical estimation of labour supply functions when a worker/consumer faces a piecewise linear budget constraint has been widely discussed in the econometrics literature.[3] Moreover, the literature[4] on the estimation of the marginal social cost of public funds 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 general problem of optimal piecewise linear income taxation. There are two main papers in the theoretical literature on the continuum-of-types case, by Sheshinski (1989) and Slemrod et al (1994).[5] We believe these papers leave the literature in a rather unfinished state, despite the fact that the paper by Slemrod et al gives a very thorough and insightful discussion of the results of its simulation analysis of the nonconvex case, as well as of the problem of piecewise linear taxation in a model consisting of only two types.

The contribution by Sheshinski was the first to formulate and solve the problem of the optimal two-bracket piecewise linear tax system, including the choice of the bracket threshold, for a continuum of worker/consumer-types. However, he claims to have shown that, under standard assumptions, marginal rate progressivity, the convex case, must *always* hold: in the social

[2] It could also be mentioned that the mechanism design approach does not extend readily to deal with realistic aspects of tax systems, for example two-earner households and multiple time periods. For further discussion of this see Apps and Rees (2009b).

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

[4] See in particular Dahlby (1998).

[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.

optimum the tax rate on the higher income bracket must always exceed that on the lower. Slemrod et al (1996) show that Sheshinski's proof does not hold in general because it ignores the existence of a discontinuity in the tax revenue function in the nonconvex case. They then carry out simulations which, using standard functional forms for the social welfare function, the individual utility function and the distribution of wage rates/productivities, in all cases produce the converse result - the upper-bracket marginal tax rate is optimally always lower.

The result that a nonconvex system could be optimal should not be surprising; for example it is foreshadowed by Sadka (1976), who established the "no distortion at the top" result for optimal nonlinear taxation and provided a strong intuition for why marginal tax rates could be lower at higher levels of income. The fact that the nonconvex case[6] *always* turns out to be optimal 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 possibility that the specific functional forms and parameter values chosen by Slemrod et al for their simulations are biased toward producing the nonconvexity result. Secondly, in practice, in virtually all countries, tax systems do in fact exhibit a substantial degree of marginal rate progressivity. It is as if tax policy makers when designing the formal tax system aim for a basically convex system. However, adjustments are made, possibly by other agencies of government, which relate to specific income ranges, typically at the low-to-middle end of the income distribution, and which have the effect of introducing nonconvexities. Leading examples are earned income tax credit systems or "targeted" cash benefit payments which are withdrawn as a function of gross income.

In this paper, we find it useful to separate the two types of system and examine the conditions that characterise a system when it is optimal. We provide a simple and transparent model which allows the characteristics of each type of tax system, and particularly the optimal bracket thresholds, to be easily seen and compared, and characterise for the first time the optimal marginal tax rates in the nonconvex case.

[6] Which does however exhibit average tax rate progressivity.

2 Individual Choice Problems

We present first the analysis of the choice problems for the individual in the face of respectively convex and nonconvex tax systems. In the next section we discuss the optimal tax structures in each case.

Consumers have identical quasilinear utility functions[7]

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

where x is consumption and l is labour supply. Gross income is $y = wl$, with the wage rate $w \in [w_0, w_1] \subset \mathbf{R}_+$. 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

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

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

We assume a differentiable wage distribution function, $F(w)$, with continuous density $f(w)$, strictly positive for all $w \in [w_0, w_1]$.

2.1 Convex case: $t_1 \leq t_2$

There are three solution possibilities:[8]

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

$$1 - t_1 - D'(\frac{y}{w})\frac{1}{w} = 0 \tag{4}$$

Defining $\psi(.)$ as the inverse function of $D'(.)$, this yields

$$y^* = w\psi((1 - t_1)w) \equiv \phi(t_1, w) \tag{5}$$

[7]Thus we are ruling out income effects. This considerably clarifies the results of the analysis.

[8]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.

giving in turn the indirect utility function

$$v(a, t_1, w) = a + (1 - t_1)\phi(t_1, w) - D(\frac{\phi(t_1, w)}{w}) \tag{6}$$

with derivatives

$$\frac{\partial v}{\partial a} = 1; \quad \frac{\partial v}{\partial t_1} = -\phi(t_1, w); \quad \frac{dv}{dw} = D'(\frac{y}{w})\frac{\phi(t_1, w)}{w^2} > 0 \tag{7}$$

We define the unique value of the wage type $\tilde{w}$ by

$$\hat{y} = \phi(t_1, \tilde{w}) \tag{8}$$

Note that $w < \tilde{w} \Leftrightarrow y^* < \hat{y}$, and $\partial\tilde{w}/\partial\hat{y} > 0$.

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

$$1 - t_2 - D'(\frac{y}{w})\frac{1}{w} = 0 \tag{9}$$

implying

$$y^* = \phi(t_2, w) \tag{10}$$

and the indirect utility

$$v(a, t_1, t_2, \hat{y}, w) = a + (t_2 - t_1)\hat{y} + (1 - t_2)\phi(t_2, w) - D(\frac{\phi(t_1, w)}{w}) \tag{11}$$

with derivatives

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

and $dv/dw > 0$ just as before. We define the unique wage type $\bar{w}$ by

$$\hat{y} = \phi(t_2, \bar{w}) \tag{13}$$

and we have $w > \bar{w} \Leftrightarrow y^* > \hat{y}$, and $\partial\bar{w}/\partial\hat{y} > 0$.

(iii) Optimal income $y^* = \hat{y}$. In that case the consumer's indirect utility is

$$v(a, t_1, \hat{y}, w) = a + (1 - t_1)\hat{y} - D(\frac{\hat{y}}{w}) \tag{14}$$

and the derivatives of the indirect utility function are

$$\frac{\partial v}{\partial a} = 1; \quad \frac{\partial v}{\partial t_1} = -\hat{y}; \quad \frac{\partial v}{\partial \hat{y}} = (1 - t_1) - D'(\frac{\hat{y}}{w})\frac{1}{w} \geq 0 \tag{15}$$

To summarise these results: the consumers can be partitioned into three subsets according to their wage type, denoted $C_0 = [w_0, \tilde{w})$, $C_1 = [\tilde{w}, \bar{w}]$, $C_2 = (\bar{w}, w_1]$, with $C \equiv C_0 \cup C_1 \cup C_2 = [w_0, w_1]$. C_0 consists of consumers in equilibrium at tangencies along the steeper part of the budget constraint, C_1 the consumers at the kink, and C_2 the consumers at tangencies on the flatter part of the budget constraint.[9]

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 $D'(\frac{\hat{y}}{w}) < (1-t_1)w$, 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 $F(w)$, consumers are continuously distributed around this budget constraint, with both maximised utility v and gross income y continuous functions of w. v is strictly increasing in w for all w, and y^* is also strictly increasing in w except over the interval C_1. Finally, note that if $t_1 = t_2$, C_1 shrinks to a point.

2.2 Nonconvex case: $t_1 > t_2$

Here there are again three solution possibilities. Given $\hat{y}$, a, t_1 and t_2, with $t_1 > t_2$, there is a unique consumer type denoted by $\hat{w}$, the solution to the equation

$$(1-t_1)\phi(t_1, w) - D(\frac{\phi(t_1, w)}{w}) = (1-t_2)\phi(t_2, w) - D(\frac{\phi(t_2, w)}{w}) \qquad (16)$$

where $\phi(.)$ has the same meaning as before. In equilibrium this type is indifferent between the two tax brackets. Note that

$$\phi(t_1, \hat{w}) < \hat{y} < \phi(t_2, \hat{w}) \qquad (17)$$

and that $\partial\hat{w}/\partial\hat{y} > 0$. The income of consumers in $[w_0, \hat{w})$ is $\phi(t_1, w)$ and in $(\bar{w}, w_1]$ is $\phi(t_2, w)$. They pay taxes of $t_1\phi(t_1, w)$ and $t_2\phi(t_2, w) + (t_1 - t_2)\widehat{y}$ respectively.

For individuals of type $\hat{w}$, the tax payments at the two local maxima are respectively $t_1\phi(t_1, \hat{w})$ and $[t_2(\phi(t_2, \hat{w}) - \widehat{y}) + t_1\widehat{y}] > t_1\phi(t_1, \hat{w})$. In this case, although maximised utility is a continuous function of w over $[w_0, w_1]$, optimal gross income and the resulting tax revenue are not. There is an upward jump in both at $\hat{w}$. Tax paid by a consumer of type $\hat{w}$ if she chooses

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

to be in the higher tax bracket is always higher than that if she chooses the lower bracket, even though the tax rate in the latter is higher. Since however consumers of type $\hat{w}$ are a set of measure zero, their choice of gross income is of no consequence for social welfare or tax revenue. Nevertheless, this discontinuity will play an important role in the optimal tax analysis, as we see in the next section.

Consumers with wages in $[w_0, \hat{w})$ have indirect utilities

$$v(a, t_1, w) = a + (1 - t_1)\phi(t_1, w) - D(\frac{\phi(t_1, w)}{w}) \tag{18}$$

with

$$\frac{\partial v}{\partial a} = 1; \quad \frac{\partial v}{\partial t_1} = -\phi(t_1, w) \tag{19}$$

and for those in $(\hat{w}, w_1]$,

$$v(a, t_1, t_2, \hat{y}, w) = a + (t_2 - t_1)\hat{y} + (1 - t_2)\phi(t_2, w) - D(\frac{\phi(t_2, w)}{w}) \tag{20}$$

with

$$\frac{\partial v}{\partial a} = 1; \quad \frac{\partial v}{\partial t_1} = -\hat{y}; \quad \frac{\partial v}{\partial t_2} = -(\phi(t_2, w) - \hat{y}); \quad \frac{\partial v}{\partial \hat{y}} = (t_2 - t_1) < 0 \tag{21}$$

We can now turn to the optimal tax analysis.

3 Optimal Taxation

3.1 The optimal convex tax system

We assume that the optimal taxation system is convex. The planner chooses the parameters of the tax system to maximise a social welfare function defined as

$$\int_{C_0} S[v(a, t_1, w)]dF + \int_{C_1} S[v(a, t_1, \hat{y}, w)]dF + \int_{C_2} S[v(a, t_1, t_2, \hat{y}, w)]dF \tag{22}$$

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

$$\int_{C_0} t_1\phi(t_1, w)dF + \int_{C_1} t_1\hat{y}dF + \int_{C_2} [t_2\phi(t_2, w) + (t_1 - t_2)\hat{y}]dF - a - G \geq 0 \tag{23}$$

where $G \geq 0$ is a per capita revenue requirement.

From the first order conditions[10] we derive the following:

Result 1:

$$\int_C \left(\frac{S'(v(w))}{\lambda} - 1\right) dF = 0 \tag{24}$$

where λ is the shadow price of tax revenue.[11] This is essentially the same condition as for linear taxation. The marginal social utility of income averaged across the population is equated to the marginal social cost of public expenditure, implying that the dollar value of this average marginal social utility is equated to the dollar marginal cost of expenditure, which is 1. The first term in the brackets is a money measure of the marginal social utility of income to a consumer of type w.

It is useful for later discussion to have the following

Lemma: $S'(v(w))$ continuous and monotonically decreasing in w implies

$$\int_{w_0}^{\tau} \left(\frac{S'(v(w))}{\lambda} - 1\right) f(w)dw > 0 \text{ for } \tau \in [w_0, w_1) \tag{25}$$

and

$$\int_{\tau}^{w_1} \left(\frac{S'(v(w))}{\lambda} - 1\right) f(w)dw < 0 \text{ for } \tau \in (w_0, w_1] \tag{26}$$

Proof: Since $S'(v(w))$ is continuous and monotonically decreasing in w,(24) implies that there exists $w^* \in (w_0, w_1)$ such that $S'(v(w^*)) - \lambda = 0$, $S'(v(w)) - \lambda > 0$, for $w \in [w_0, w^*)$, and $S'(v(w)) - \lambda < 0$, for $w \in (w^*, w_1]$. Then $f(w) > 0$ on $[w_0, w_1]$ implies, again from (24), that

$$\int_{w_0}^{w^*} \left(\frac{S'(v(w))}{\lambda} - 1\right) f(w)dw = -\int_{w^*}^{w_1} \left(\frac{S'(v(w))}{\lambda} - 1\right) f(w)dw \tag{27}$$

which gives the result.

Assuming an interior solution with $t_1 < t_2$, the condition characterising the optimal lower bracket tax rate yields

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

[11] Or the marginal social cost of public expenditure. Needless to say, if we assume that the planner has optimised the tax system, the problem of estimating this paramer becomes much simpler than it is taken to be in the literature on this problem.

Result 2:

$$t_1^* = \frac{\int_{C_0}(\frac{S'}{\lambda} - 1)[\phi(t_1^*, w) - \hat{y}^*]dF}{\int_{C_0} \frac{\partial\phi(t_1^*, w)}{\partial t_1} dF} \tag{28}$$

The denominator, which is negative, can be interpreted as the efficiency effect of the tax. The numerator is the equity effect. Using the above lemma it can be shown that this numerator is also negative and so the tax rate is positive. Thus, given $C_0 = [w_0, \tilde{w})$ and $\phi(t_1^*, w) - \hat{y}^* < 0$ in this interval, $\tilde{w} \leq w^*$ gives the result immediately. If $\tilde{w} > w^*$, since $(S'(v(w)) - \lambda)f(w) > 0$ and $\phi(t_1^*, w) < \phi(t_1^*, w^*)$ for $w \in [w_0, w^*)$, we have

$$\int_{w_0}^{w^*} [\hat{y}^* - \phi(t_1^*, w](S'(v(w)) - \lambda)f(w)dw > \int_{w_0}^{w^*} [\hat{y}^* - \phi(t_1^*, w^*](S'(v(w)) - \lambda)f(w)dw \tag{29}$$

while $(S'(v(w)) - \lambda)f(w) < 0$ and $\phi(t_1^*, w) > \phi(t_1^*, w^*)$ for $w \in (w^*, \tilde{w})$ implies

$$\int_{w^*}^{\tilde{w}} [\hat{y}^* - \phi(t_1^*, w](S'(v(w)) - \lambda)f(w)dw > \int_{w^*}^{\tilde{w}} [\hat{y}^* - \phi(t_1^*, w^*](S'(v(w)) - \lambda)f(w)dw \tag{30}$$

Adding these two inequalities and applying the lemma with $\tau = \tilde{w}$ then gives

$$\int_{C_0} [\hat{y}^* - \phi(t_1^*, w](S'(v(w)) - \lambda)dF > [\hat{y}^* - \phi(t_1^*, w^*] \int_{C_0} (S'(v(w)) - \lambda)dF > 0 \tag{31}$$

as required.

An interesting aspect of this condition is that t_1 is set in the light of its effects only on efficiency and income distribution within the subset C_0, in a way that is analogous to[12] the setting of a linear tax on a population as a whole, even though it also affects utility and tax burdens of the consumers in C_1 and C_2. The reason for this is that in both these subsets, the effect of a change in t_1 on utility $(-\hat{y}^*)$ is the negative of its effect on tax revenue $(\hat{y}^*)$. The first order condition determining t_1^* can then be written as

$$t_1^* \int_{C_0} \frac{\partial\phi(t_1^*, w)}{\partial t_1} dF = \int_{C_0} (\frac{S'}{\lambda} - 1)y^* dF + \hat{y}^* \int_{C_1 \cup C_2} (\frac{S'}{\lambda} - 1)dF \tag{32}$$

Using (24) then gives Result 2.

[12] But not identical to, since the numerator term is not the covariance of the marginal social utility of income and gross income for the subset C_0.

Result 3:

$$t_2^* = \frac{\int_{C_2}(\frac{S'}{\lambda}-1)[\phi(t_2^*,w)-\hat{y}^*]dF}{\int_{C_2}\frac{\partial\phi(t_2^*,w)}{\partial t_2}dF} \tag{33}$$

Here $\phi(t_1^*, w) - \hat{y}^* > 0$ and is increasing in w, and the numerator here can also be shown to be negative, along the same lines as for t_1^*. Again, the optimal tax rate for the subgroup C_2 is determined (given $\hat{y}^*$) entirely by the characteristics of the wage types in this group. Thus, given the optimal choice of tax brackets, the optimal tax rates are set as optimal linear taxes over the sub-populations strictly within each bracket.

Since in the convex case we must have $t_1^* < t_2^*$, we see that this case is likely to be globally optimal the greater the aggregate deadweight efficiency loss within the C_0 group relative to that in the C_2 group, and the more powerful the marginal tax rate is as a redistributive instrument within the C_2 group relative to the C_0 group, given of course the optimal choices of a and $\hat{y}$. Note that the proportion of the consumers falling into group C_1, $\int_{C_1} dF$, plays no *direct* role in determining these tax rates. It does however of course influence the choice of $\hat{y}$.

Result 4:

The first order condition with respect to $\hat{y}$ can be written as

$$\int_{C_1}\{\frac{S'}{\lambda}v_{\hat{y}} + t_1^*\}dF = -(t_2^* - t_1^*)\int_{C_2}(\frac{S'}{\lambda}-1)dF \tag{34}$$

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 $w \in (\tilde{w}, \bar{w}]$ the marginal utility with respect to a relaxation of the gross income constraint is $v_{\hat{y}} = (1 - t_1) - D'(\frac{\hat{y}}{w})\frac{1}{w} > 0$, 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. A necessary condition for the possibility of this solution is that $[\int_{C_2} dF]^{-1}\int_{C_2}\frac{S'}{\lambda}dF < 1$, so that the marginal social utility of the upper bracket consumers is on average below the population average. The planner then suffers a distributional loss from giving this group a lump sum income increase. Sheshinski argued that if $t_2^* < t_1^*$ the term on the right hand side of (34) 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 can obtain more intuition on these results by considering Figure 1.

Figure 1 about here

Figure 1 illustrates the comparison between the optimal linear tax and the optimal convex two-bracket tax. The line $a_L L$ represents the budget constraint facing all consumers under the optimal linear tax, $a_C CD$ that under the optimal piecewise linear tax. One budget constraint cannot lie entirely above the other over the relevant domain of y-values since this would imply that one of them violates the government budget constraint. Thus there must be at least one intersection point within the domain of y-values. Special cases can be constructed in which there might be only one intersection, but here we consider the case of two such points, at A and B, with $a_C < a_L$.

The essential feature of the illustration is that the piecewise linear tax system redistributes welfare away from the ends towards the middle: all consumers of wage types in the interval (w_A, w_B) are strictly better off under the piecewise linear tax, those in $[w_0, w_A)$ and $(w_B, w_1]$ are strictly worse off. All consumers in the lower tax bracket under the piecewise linear tax will expand their (compensated) labour supplies, all those in the higher bracket will contract theirs, as compared to the linear tax. Conditions that would tend to make the convex piecewise linear tax globally optimal therefore would be a marginal social utility of income that falls slowly over the low and middle income ranges and then then quickly over the upper income range, and substantially higher compensated labour supply elasticities in the middle income range.Note that we cannot rule out the case in which $a_C \geq a_L$, in which the lowest income group would not be made worse off, but clearly there would have to be a compensatingly heavier burden on the upper tax bracket.

3.2 The optimal nonconvex tax system

Assuming that the nonconvex system is optimal we can state the problem as

$$\max_{a,t_1,t_2,\hat{y}} \int_{w_0}^{\hat{w}} S[v(a,t_1,w)]dF + \int_{\hat{w}}^{w_1} S[v(a,t_1,t_2,\hat{y},w)]dF \tag{35}$$

$$s.t. \quad \int_{w_0}^{\hat{w}} t_1\phi(t_1,w)dF + \int_{\hat{w}}^{w_1} [t_2\phi(t_2,w) + (t_1-t_2)\hat{y}]dF - a - R \geq 0 \tag{36}$$

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 (16)-(21) it is easy to see that a change in a does not affect 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

$$\int_{w_0}^{w_1} (\frac{S'}{\lambda} - 1)dF = 0 \tag{37}$$

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

$$\Delta R = [t_2\phi(\hat{w},t_2) - (t_2 - t_1)\widehat{y}] - t_1\phi(\hat{w},t_1) > 0 \tag{38}$$

This is the value of the jump in tax revenue that takes place at $\hat{w}$.

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

Result 5:

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

$$\frac{\partial \hat{w}}{\partial \widehat{y}} \Delta R f(\hat{w}) = (t_2 - t_1) \int_{\hat{w}}^{w_1} \left\{ \frac{S'}{\lambda} - 1 \right\} dF \tag{39}$$

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 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

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

the net income of higher bracket consumers with, on average, below-average marginal social utility 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. More precisely, a discrete increase $\Delta\widehat{y}$ would cause a discrete interval of consumers to jump down into the lower bracket, and, in the limit, as $\Delta\widehat{y} \to 0$, the resulting revenue loss is given by $\Delta Rf(\hat{w})$. Both marginal benefit and marginal cost are positive.

Result 6:

The condition with respect to t_1 is:

$$t_1^* = \frac{\int_{w_0}^{\hat{w}}(\frac{S'}{\lambda} - 1)[y^* - \hat{y}^*]dF + \frac{\partial\hat{w}}{\partial t_1}\Delta Rf(\hat{w})}{\int_{w_0}^{\hat{w}} y_{t_1}(t_1^*, w)dF} \tag{40}$$

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, equal in the limit, as the change in t_1 goes to zero, to $\Delta Rf(\hat{w})$.

Result 7:

The condition with respect to t_2 is:

$$t_2^* = \frac{\int_{\hat{w}}^{w_1}(\frac{S'}{\lambda} - 1)[y^* - \hat{y}^*]dF + \frac{\partial\hat{w}}{\partial t_2}\Delta Rf(\hat{w})}{\int_{\hat{w}}^{w_1} y_{t_2}(t_2^*, w)dF} \tag{41}$$

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 for this term 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.

The existence of the discontinuity terms in these expressions adds a new element to the discussion of the optimal tax parameters. Since $t_1^* > t_2^*$, their presence, other things equal, makes for a widening in the gap between the two tax rates. The influence of the other terms in the conditions is the same as in the discussion of the convex case.

Figure 2 about here

In Figure 2 we compare the optimal linear and nonconvex piecewise linear tax systems. The budget constraint corresponding to the linear tax is again $a_L L$, that of the piecewise linear tax is $a_N EF$. Thus we see that as compared to the linear tax the nonconvex piecewise linear tax redistributes welfare from the middle towards the ends.[14] Lower bracket consumers, who now pay a higher marginal rate, reduce their labour supplies and gross incomes, higher bracket consumers do the reverse.

4 Conclusions

Given its significance in practice, the piecewise linear tax system seems to have received disproportionately little attention in the literature on optimal income taxation. This paper offers a simple and transparent analysis of its main characteristics. We have considered only the two bracket case, but it is easy to see how this can be extended to an arbitrary number of brackets. It is quite possible in this case that some portions of the tax system might be convex and some nonconvex, in a way that depends on the income distributional preferences of the tax policy maker and the way in which labour supply elasticities vary with wage type. The question of the optimal determination of the number of brackets is left open. Note, however, that we are not trying to find the best piecewise linear approximation to a nonlinear tax function that is optimal in the sense of Mirrlees. Rather, we start from the position that it is practical only to pool all consumer types. Given the complexity of the situation which faces the design of tax systems, this may be the only feasible approach.

References

[1] P F Apps and R Rees, 2009a, *Public Economics and the Household*, Cambridge: Cambridge University Press.

[2] P F Apps and R Rees, 2009b, "Two Extensions to the Theory of Optimal Income Taxation", mimeo.

[14] Though we cannot exclude the case in which $a_N < a_L$, with only upper bracket consumers being made better off. This could arise for a sufficiently high welfare weighting and/or sufficiently strong deadweight losses on high income consumers.

[3] R Boadway, "The Mirrlees Approach to the Theory of Economic Policy", *International Tax and Public Finance*, 5, 67-81.

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

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

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

[7] E Sadka, 1976, "On Income Distribution, Incentive Effects and Optimal Income Taxation, *Review of Economic Studies,* 261-267.

[8] 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.

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

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

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

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

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

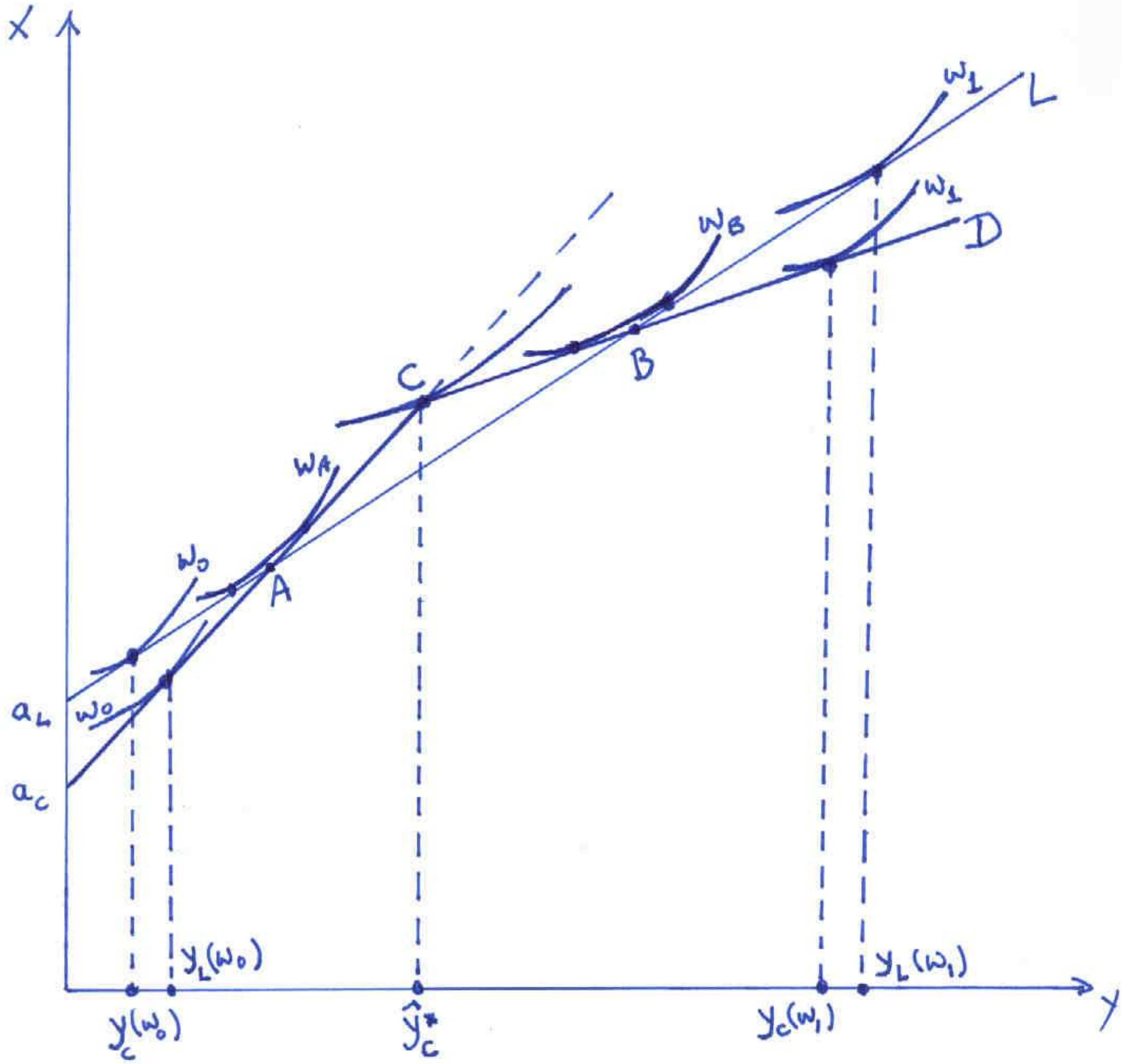

Figure 1. Optimal Convex vs. Optimal Linear Tax Schedules

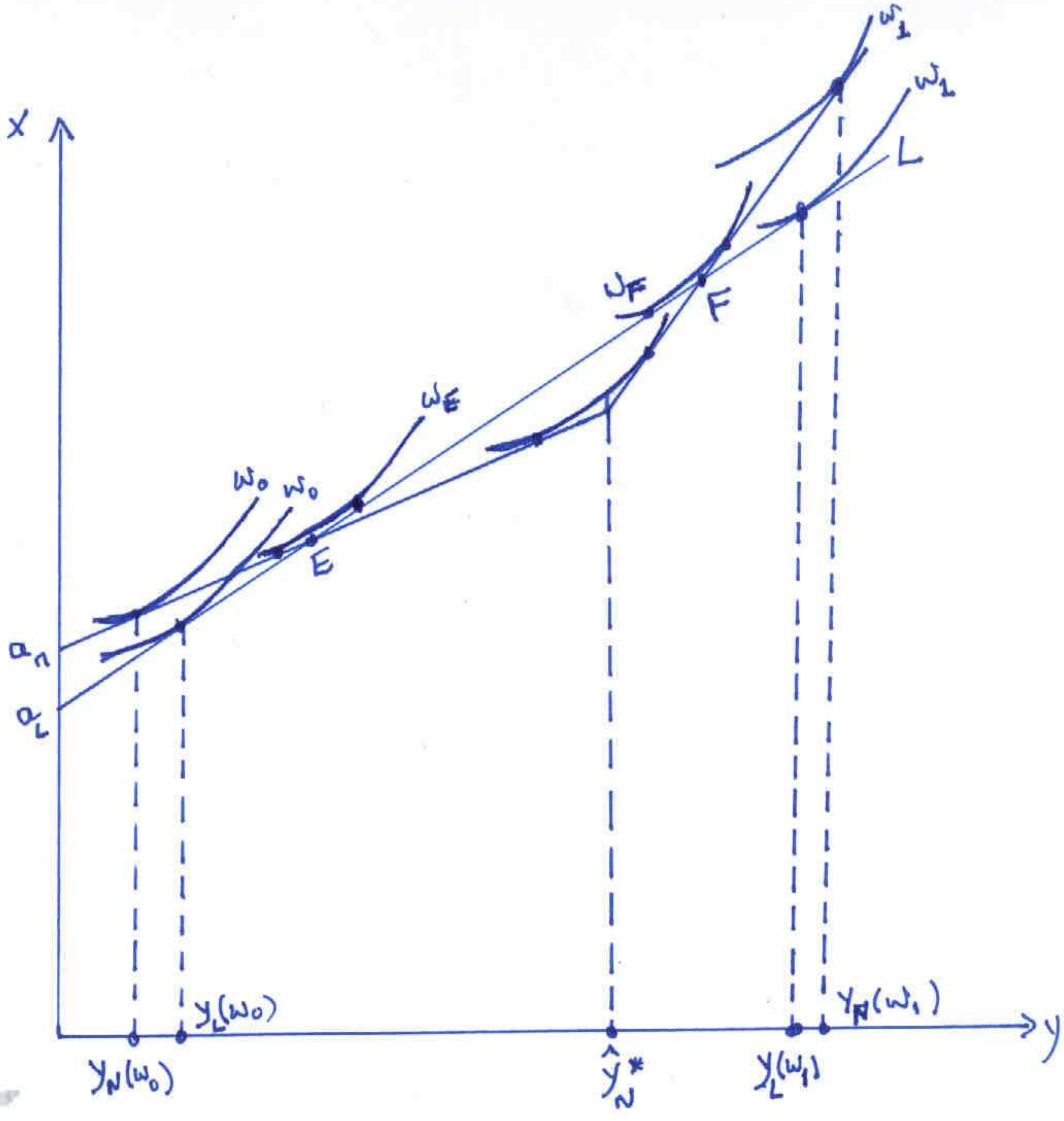

Figure 2 Optimal Nonconvex vs Optimal Linear Tax Schedules

CESifo Working Paper Series

for full list see www.cesifo-group.org/wp
(address: Poschingerstr. 5, 81679 Munich, Germany, office@cesifo.de)

2501 Andrew Clark, Andreas Knabe and Steffen Rätzel, Boon or Bane? Others' Unemployment, Well-being and Job Insecurity, December 2008

2502 Lukas Menkhoff, Rafael R. Rebitzky and Michael Schröder, Heterogeneity in Exchange Rate Expectations: Evidence on the Chartist-Fundamentalist Approach, December 2008

2503 Salvador Barrios, Harry Huizinga, Luc Laeven and Gaëtan Nicodème, International Taxation and Multinational Firm Location Decisions, December 2008

2504 Andreas Irmen, Cross-Country Income Differences and Technology Diffusion in a Competitive World, December 2008

2505 Wenan Fei, Claude Fluet and Harris Schlesinger, Uncertain Bequest Needs and Long-Term Insurance Contracts, December 2008

2506 Wido Geis, Silke Uebelmesser and Martin Werding, How do Migrants Choose their Destination Country? An Analysis of Institutional Determinants, December 2008

2507 Hiroyuki Kasahara and Katsumi Shimotsu, Sequential Estimation of Structural Models with a Fixed Point Constraint, December 2008

2508 Barbara Hofmann, Work Incentives? Ex Post Effects of Unemployment Insurance Sanctions – Evidence from West Germany, December 2008

2509 Louis Hotte and Stanley L. Winer, The Demands for Environmental Regulation and for Trade in the Presence of Private Mitigation, December 2008

2510 Konstantinos Angelopoulos, Jim Malley and Apostolis Philippopoulos, Welfare Implications of Public Education Spending Rules, December 2008

2511 Robert Orlowski and Regina T. Riphahn, The East German Wage Structure after Transition, December 2008

2512 Michel Beine, Frédéric Docquier and Maurice Schiff, International Migration, Transfers of Norms and Home Country Fertility, December 2008

2513 Dirk Schindler and Benjamin Weigert, Educational and Wage Risk: Social Insurance vs. Quality of Education, December 2008

2514 Bernd Hayo and Stefan Voigt, The Relevance of Judicial Procedure for Economic Growth, December 2008

2515 Bruno S. Frey and Susanne Neckermann, Awards in Economics – Towards a New Field of Inquiry, January 2009

2516 Gregory Gilpin and Michael Kaganovich, The Quantity and Quality of Teachers: A Dynamic Trade-off, January 2009

2517 Sascha O. Becker, Peter H. Egger and Valeria Merlo, How Low Business Tax Rates Attract Multinational Headquarters: Municipality-Level Evidence from Germany, January 2009

2518 Geir H. Bjønnes, Steinar Holden, Dagfinn Rime and Haakon O.Aa. Solheim, ‚Large' vs. ‚Small' Players: A Closer Look at the Dynamics of Speculative Attacks, January 2009

2519 Jesus Crespo Cuaresma, Gernot Doppelhofer and Martin Feldkircher, The Determinants of Economic Growth in European Regions, January 2009

2520 Salvador Valdés-Prieto, The 2008 Chilean Reform to First-Pillar Pensions, January 2009

2521 Geir B. Asheim and Tapan Mitra, Sustainability and Discounted Utilitarianism in Models of Economic Growth, January 2009

2522 Etienne Farvaque and Gaël Lagadec, Electoral Control when Policies are for Sale, January 2009

2523 Nicholas Barr and Peter Diamond, Reforming Pensions, January 2009

2524 Eric A. Hanushek and Ludger Woessmann, Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation, January 2009

2525 Richard Arnott and Eren Inci, The Stability of Downtown Parking and Traffic Congestion, January 2009

2526 John Whalley, Jun Yu and Shunming Zhang, Trade Retaliation in a Monetary-Trade Model, January 2009

2527 Mathias Hoffmann and Thomas Nitschka, Securitization of Mortgage Debt, Asset Prices and International Risk Sharing, January 2009

2528 Steven Brakman and Harry Garretsen, Trade and Geography: Paul Krugman and the 2008 Nobel Prize in Economics, January 2009

2529 Bas Jacobs, Dirk Schindler and Hongyan Yang, Optimal Taxation of Risky Human Capital, January 2009

2530 Annette Alstadsæter and Erik Fjærli, Neutral Taxation of Shareholder Income? Corporate Responses to an Announced Dividend Tax, January 2009

2531 Bruno S. Frey and Susanne Neckermann, Academics Appreciate Awards – A New Aspect of Incentives in Research, January 2009

2532 Nannette Lindenberg and Frank Westermann, Common Trends and Common Cycles among Interest Rates of the G7-Countries, January 2009

2533 Erkki Koskela and Jan König, The Role of Profit Sharing in a Dual Labour Market with Flexible Outsourcing, January 2009

2534 Tomasz Michalak, Jacob Engwerda and Joseph Plasmans, Strategic Interactions between Fiscal and Monetary Authorities in a Multi-Country New-Keynesian Model of a Monetary Union, January 2009

2535 Michael Overesch and Johannes Rincke, What Drives Corporate Tax Rates Down? A Reassessment of Globalization, Tax Competition, and Dynamic Adjustment to Shocks, February 2009

2536 Xenia Matschke and Anja Schöttner, Antidumping as Strategic Trade Policy Under Asymmetric Information, February 2009

2537 John Whalley, Weimin Zhou and Xiaopeng An, Chinese Experience with Global 3G Standard-Setting, February 2009

2538 Claus Thustrup Kreiner and Nicolaj Verdelin, Optimal Provision of Public Goods: A Synthesis, February 2009

2539 Jerome L. Stein, Application of Stochastic Optimal Control to Financial Market Debt Crises, February 2009

2540 Lars P. Feld and Jost H. Heckemeyer, FDI and Taxation: A Meta-Study, February 2009

2541 Philipp C. Bauer and Regina T. Riphahn, Age at School Entry and Intergenerational Educational Mobility, February 2009

2542 Thomas Eichner and Rüdiger Pethig, Carbon Leakage, the Green Paradox and Perfect Future Markets, February 2009

2543 M. Hashem Pesaran, Andreas Pick and Allan Timmermann, Variable Selection and Inference for Multi-period Forecasting Problems, February 2009

2544 Mathias Hoffmann and Iryna Shcherbakova, Consumption Risk Sharing over the Business Cycle: the Role of Small Firms' Access to Credit Markets, February 2009

2545 John Beirne, Guglielmo Maria Caporale, Marianne Schulze-Ghattas and Nicola Spagnolo, Volatility Spillovers and Contagion from Mature to Emerging Stock Markets, February 2009

2546 Ali Bayar and Bram Smeets, Economic and Political Determinants of Budget Deficits in the European Union: A Dynamic Random Coefficient Approach, February 2009

2547 Jan K. Brueckner and Anming Zhang, Airline Emission Charges: Effects on Airfares, Service Quality, and Aircraft Design, February 2009

2548 Dolores Messer and Stefan C. Wolter, Money Matters – Evidence from a Large-Scale Randomized Field Experiment with Vouchers for Adult Training, February 2009

2549 Johannes Rincke and Christian Traxler, Deterrence through Word of Mouth, February 2009

2550 Gabriella Legrenzi, Asymmetric and Non-Linear Adjustments in Local Fiscal Policy, February 2009

2551 Bruno S. Frey, David A. Savage and Benno Torgler, Surviving the Titanic Disaster: Economic, Natural and Social Determinants, February 2009

2552 Per Engström, Patrik Hesselius and Bertil Holmlund, Vacancy Referrals, Job Search, and the Duration of Unemployment: A Randomized Experiment, February 2009

2553 Giorgio Bellettini, Carlotta Berti Ceroni and Giovanni Prarolo, Political Persistence, Connections and Economic Growth, February 2009

2554 Steinar Holden and Fredrik Wulfsberg, Wage Rigidity, Institutions, and Inflation, February 2009

2555 Alexander Haupt and Tim Krieger, The Role of Mobility in Tax and Subsidy Competition, February 2009

2556 Harald Badinger and Peter Egger, Estimation of Higher-Order Spatial Autoregressive Panel Data Error Component Models, February 2009

2557 Christian Keuschnigg, Corporate Taxation and the Welfare State, February 2009

2558 Marcel Gérard, Hubert Jayet and Sonia Paty, Tax Interactions among Belgian Municipalities: Does Language Matter?, February 2009

2559 António Afonso and Christophe Rault, Budgetary and External Imbalances Relationship: A Panel Data Diagnostic, February 2009

2560 Stefan Krasa and Mattias Polborn, Political Competition between Differentiated Candidates, February 2009

2561 Carsten Hefeker, Taxation, Corruption and the Exchange Rate Regime, February 2009

2562 Jiahua Che and Gerald Willmann, The Economics of a Multilateral Investment Agreement, February 2009

2563 Scott Alan Carson, Demographic, Residential, and Socioeconomic Effects on the Distribution of 19th Century US White Statures, February 2009

2564 Philipp Harms, Oliver Lorz and Dieter Urban, Offshoring along the Production Chain, February 2009

2565 Patricia Apps, Ngo Van Long and Ray Rees, Optimal Piecewise Linear Income Taxation, February 2009