NHH

INSTITUTT FOR SAMFUNNSØKONOMI

DEPARTMENT OF ECONOMICS

SAM 5 2010

ISSN: 0804-6824

FEBRUARY 2010

Discussion paper

# Optimal Labor Income Taxation under Maximin: An Upper Bound

BY

**LAURENCE JACQUET**



Norges Handelshøyskole

NORWEGIAN SCHOOL OF ECONOMICS AND BUSINESS ADMINISTRATION

# Optimal Labor Income Taxation under Maximin: An Upper Bound*

Laurence Jacquet[†]
Norwegian School of Economics and Business Administration and CESifo

February 2, 2010

Abstract

This paper assumes the standard optimal income tax model of Mirrlees (Review of Economic Studies, 1971). It gives fairly mild conditions under which the optimal nonlinear labor income tax profile derived under maximin has higher marginal tax rates than the ones derived with welfarist criteria that sum over the population any concave transformation of individual utilities. This strict dominance result is always valid close to the bounds of the skill distribution and almost everywhere (except at the upper bound) when quasilinear-in-consumption preferences are assumed.

Key Words: optimal income taxation, maximin.
JEL Classification: H21.

*I would like to deeply thank Raouf Boucekkine for extremely helpful comments on a preliminary version of this paper. I also wish to thank Agnar Sandmo and Guttorm Schjelderup for their helpful comments and Etienne Lehmann for a stimulating discussion which is at the origin of this paper.

[†]Address: Norwegian School of Economics and Business Administration (NHH), Economics Department, Helleveien 30, 5045 Bergen, Norway. Email: laurence.jacquet@nhh.no

# 1 Introduction

Choné and Laroque (2005) and Laroque (2005) show that in the optimal labor income tax model where labor supply responses are modeled along the extensive margin, the Rawlsian criterion (which maximizes the income tax revenue) gives a benchmark, the Laffer bound. Any optimal allocation corresponds to tax schedules which are below that benchmark. This paper gives a comparable result when dealing with the intensive margin. In the optimal (nonlinear) labor income tax model where labor supply responses are modeled along the intensive margin, the maximin solution gives also an upper bound.[1] Assuming an isoelastic disutility of labor and quasilinear-in consumption preferences, the optimal marginal tax schedule under maximin is a benchmark above which it is suboptimal to tax under the integral over the population of any concave function of individual utilities. Assuming a general separable utility function, this result is still valid close to the bottom and the top of the skill distribution.

# 2 The Model

We use the model that has been employed in much of the literature on optimal labor income taxation since the seminal article of Mirrlees (1971). We assume that all households have the same utility function and an additively separable form as in Mirrlees (1971) and Atkinson and Stiglitz (1980):

$$u(x, \ell) = v(x) - h(\ell) \tag{1}$$

where $x$ is consumption and $\ell$ is labor (so $1 - \ell$ is leisure), with $v' > 0 \geq v''$, $h' > 0$ and $h'' \geq 0$, with either $v'' < 0$ or $h'' > 0$.

Households differ only in skills, which correspond with their wage rates given that aggregate production is linear in labor. Skills $w$ are distributed according to the function $F(w)$ for $w \in W = [\underline{w}, \overline{w}]$, where $0 < \underline{w} < \overline{w} < \infty$. The density function, $f(w) = F'(w)$, is assumed to be differentiable and strictly positive for all $w \in W$.

Households obtain their income from wages, with income denoted by $y \equiv w\ell$. Let $x(w)$, $\ell(w)$ and $y(w)$ be consumption, labor supply and income for a household with skill $w$. The government can observe incomes but not wage rates or labor supplied, so it bases its tax scheme on income. Then, the budget constraint for household $w$ is:

$$x(w) = y(w) - T(y(w)) \tag{2}$$

[1] Assuming a linear income tax, Atkinson (1983) numerically shows that the optimal tax rate is always larger under Maximin.

where $T(y(w))$ is the tax imposed on type-$w$ households. The household maximizes (1) subject to (2), yielding the first-order condition:

$$\frac{h'(\ell(w))}{wv'(x(w))} = 1 - T'(y(w)) \tag{3}$$

If we use $\ell = y/w$ to rewrite the utility function as $v(x) - h(y/w)$, the left-hand side of (3) can be interpreted as the marginal rate of substitution between income $y$ and consumption $x$.

We will compare the optimal tax schedules derived under a maximin criterion and a welfarist criterion that sums over all individuals a transformation $\Phi$ of individuals' utility with $\Phi' > 0$ and $\Phi'' \leq 0$ (hence the government has a non-negative aversion to inequality) and $\Phi$ independent of $w$. Under maximin, the government maximizes the welfare of the least well-off households. Given our information assumptions, the worst-off will be those with wage $\underline{w}$ at the bottom of the skill distribution hence the maximin criterion is

$$u(\underline{w}) \tag{4}$$

The welfarist social preferences are

$$\int_{\underline{w}}^{\overline{w}} \Phi(u(w))dw \tag{5}$$

The government chooses the tax schedule $T(y(w))$ or, equivalently, the consumption-income bundle intended for each household $\{(x(w), y(w), w \in W\}$, to maximize its social welfare function, subject to two sorts of constraints.

The first is the government budget constraint, which takes the form:

$$\int_{\underline{w}}^{\overline{w}} [y(w) - x(w)]f(w)dw \geq R \tag{6}$$

where $R$ is an exogenous revenue requirement. This constraint must be binding at the optimum since $u$ is increasing in $x$.

The second is the set of incentive-compatibility constraints, which require that type-$w$ households choose the consumption-income bundle intended for them, that is,

$$v(x(w)) - h\left(\frac{y(w)}{w}\right) \geq v(x(\widehat{w})) - h\left(\frac{y(\widehat{w})}{w}\right) \quad \forall w, \widehat{w} \in W \tag{7}$$

As Mirrlees (1971) shows, the necessary conditions for (7) to be satisfied are:

$$\dot{u}(w) = h'(.)\frac{y(w)}{w^2} = h'(.)\frac{\ell(w)}{w} \qquad \forall w \tag{8}$$

where, using (1), $u(w) \equiv v(x(w)) - h(\ell(w))$.[2]

[2] These so-called first-order incentive compatibility conditions (FOIC) may not be sufficient. Sufficiency is guaranteed by the second-order incentive compatibility (SOIC) conditions, $\dot{y}(w) \geq 0$ (or equivalently $\dot{x}(w) \geq 0$) (Ebert (1992)). If the SOIC constraints are slack ($\dot{y}(w) > 0$), the first-order approach is appropriate. Where they are binding, we have $\dot{x}(w) = \dot{y}(w) = 0$, so there is bunching of households of different skills.

The problem for the government is to choose $x(w)$, $\ell(w)$ and $u(w)$ to maximize its welfare function subject to the budget constraint (6) and the FOIC conditions (8), where $x(w)$ and $\ell(w)$ are controls and $u(w)$ is a state variable. However, since $u(w) = v(x(w)) - h(\ell(w))$ by (1), we can treat $x(w)$ as an implicit function of $u(w)$ and $\ell(w)$ and write it with some abuse of notation as $x(u(w), \ell(w))$, where by differentiating (1), we obtain:

$$\frac{\partial x(u(w),\ell(w))}{\partial \ell} = \frac{h'(\ell(w))}{v'(x(w))}, \qquad \frac{\partial x(u(w),\ell(w))}{\partial u} = \frac{1}{v'(x(w))} \tag{9}$$

We can then suppress $x(w)$ from the government problem and write it as follows:

$$\underset{\{u(w),\ell(w)\}}{Max} W(u(w)) \text{ s.t. } \int_{\underline{w}}^{\overline{w}} [w\ell(w) - x(u(w),\ell(w))]f(w)dw = R, \;\; \dot{u}(w) = \frac{\ell(w)h'(\ell(w))}{w} \tag{10}$$

where the social welfare function $W(u(w))$ represents either (4) or (5).

The corresponding Lagrangian is:

$$\begin{aligned} \pounds \equiv W(u(w)) + \lambda \int_{\underline{w}}^{\overline{w}} \left[ [w\ell(w) - x(u(w),\ell(w))]f(w) - \frac{R}{\overline{w}-\underline{w}} \right] dw \\ + \int_{\underline{w}}^{\overline{w}} \zeta(w) \left[ \frac{\ell(w)h'(\ell(w))}{w} - \dot{u}(w) \right] dw \end{aligned} \tag{11}$$

where $\lambda$ is the multiplier associated with the binding budget constraint (6) and $\zeta(w)$ is the multiplier associated with the FOIC conditions (8). The necessary conditions are given in Appendix. Under maximin, the first-order conditions reduce to the following:

$$\frac{T'_M(y(w))}{1-T'_M(y(w))} = A(w)\frac{1}{wf(w)}v'(x_M(w)) \int_w^{\overline{w}} \frac{f(t)}{v'(x_M(t))}dt \quad \forall w \in W \tag{12}$$

where the subscript $M$ states for maximin and where

$$A(w) = 1 + \frac{h''(\ell(w))\ell(w)}{h'(\ell(w))}$$

is a measure of the elasticity of labor supply.[3]

Under the social welfare function (5), the marginal tax rate denoted by $T'_\Phi(y(w))$ can be expressed as:[4]

$$\frac{T'_\Phi(y(w))}{1-T'_\Phi(y(w))} = A(w)\frac{1}{wf(w)}v'(x_\Phi(w)) \int_w^{\overline{w}} \left( \frac{1}{v'(x_\Phi(t))} - \frac{\Phi'(u(t))}{\lambda_\Phi} \right) f(t)dt \quad \forall w \in W \tag{14}$$

where the subscript $\Phi$ states for the social objective $\int_{\underline{w}}^{\overline{w}} \Phi(u(w))dw$.

[3] The term $A(w) \equiv [1 + \ell h''(\ell)/h'(\ell)]$ is equal to $[1+e^u(w_n)]/e^c(w_n)$ where $e^c(w_n)$ and $e^u(w_n)$ are the compensated and uncompensated elasticities of labor supply, respectively. More precisely, using (3), $e^c(w_n)$ and $e^u(w_n)$ satisfy

$$e^c(w_n) = \frac{h'(\ell)}{(h''(\ell) - w_n^2 v''(x))\,\ell} > 0 \quad \text{and} \quad e^u(w_n) = \frac{h'(\ell) + v''(x)w_n^2\ell}{(h''(\ell) - w_n^2 v''(x))\ell}$$

where $w_n \equiv w(1 - T'(y(w)))$ is the after-tax wage rate (Saez, 2001).

[4] This writing is similar to the optimal tax formula in Atkinson and Stiglitz (1980).

Our analysis has been conducted as if the first-order approach is valid. This will be the case as long as (12) and (14) yield solutions for $T'(y(w))$ such that $y(w)$ (or $x(w)$) is everywhere increasing in $w$. For simplicity, we assume it satisfied.[5]

Assume, following Diamond (1998), that $h(\ell)$ takes the isoelastic form so $A(w)$ is constant. In order to show that the marginal tax rate under maximin is always above or equal to the one under the more general social welfare function, we have to show that $T'_M(y(w))/\left(1-T'_M(y(w))\right) - T'_\Phi(y(w))/\left(1-T'_\Phi(y(w))\right) \geq 0 \ \forall w$ since it is well established that $0 \leq T'(y(w)) < 1$ (Seade, 1977, 1982). Since $A(w)$ and $wf(w)$ do not depend on the objective function, this reduces to show that

$$\Omega(w) \equiv v'(x_M(w)) \int_w^{\overline{w}} \frac{f(t)}{v'(x_M(t))} dt - v'(x_\Phi(w)) \int_w^{\overline{w}} \left( \frac{1}{v'(x_\Phi(t))} - \frac{\Phi'(u(t))}{\lambda} \right) f(t) dt \geq 0 \quad \forall w \tag{15}$$

First, consider $\Omega(w)$ at $w = \underline{w}$. From (23) (in the Appendix) and the transversality condition $\zeta_\Phi(\underline{w}) = 0$, we have:

$$\Omega(\underline{w}) = v'(x_M(\underline{w})) \int_{\underline{w}}^{\overline{w}} \frac{f(t)}{v'(x_M(t))} dt$$

From (27) (in the Appendix) and the transversality condition $\zeta_M(\underline{w}) = -1$, the latter expression becomes:

$$\Omega(\underline{w}) = \frac{v'(x_M(\underline{w}))}{\lambda_M} > 0 \tag{16}$$

Second, putting $w = \overline{w}$ in (15) gives:

$$\Omega(\overline{w}) = 0 \tag{17}$$

Equation (16) relies on the sharp contrast between the optimal marginal tax rate at the bottom under maximin and under a more general social welfare function. Assuming no bunching at the bottom, $T'_\Phi(y(\underline{w})) = 0$ under the more general welfarist criterion (Seade, 1977). Contrastingly, $T'_M(y(\underline{w})) > 0$ under maximin. Intuitively, increasing the marginal tax rate at a skill level $\widetilde{w}$ distorts the labor supply of those with skill $\widetilde{w}$, implying an efficiency loss. However, it also improves equity when the extra tax revenue can be redistributed towards a positive mass of agents with skills $w \leq \widetilde{w}$. As long as the latter outweighs the former in the welfare criterion, such transfers are positively valued, hence an equity gain appears. Under social preferences $\int_{\underline{w}}^{\overline{w}} \Phi(u(w))dw$, the mass of people at the bottom of the skill distribution is zero hence a positive marginal tax rate would not improve equity but would create an efficiency loss. Even when the aversion to inequality approaches infinity in the social welfare function, the marginal tax rate continues to be zero at the

[5] Boadway and Jacquet (2008) show that a non-increasing marginal tax rate is sufficient to satisfy the SOIC conditions and that the marginal tax rate is decreasing under maximin, under fairly mild conditions.

bottom (Boadway and Jacquet, 2008).[6] Contrastingly, under maximin, everyone in the objective function is at $w = \underline{w}$, so the equity effect is positive hence $T'_M(y(\underline{w})) > 0$. Moreover, as well known since Sadka (1976) and Seade (1977), the optimal marginal tax rate at the top is zero with a bounded skill distribution, i.e. $T'_M(y(\overline{w})) = T'_\Phi(y(\overline{w})) = 0$, which yields (17). These results can be summarized as follows.

**Proposition 1** *At the bottom (top) of the skill distribution, the optimal marginal tax rate under maximin is larger (equal) to the one under criterion* $\int_{\underline{w}}^{\overline{w}} \Phi(u(w))dw$.

From (16) and (17), deriving conditions under which $\Omega(w)$ is monotonically decreasing in $w$ on $(\underline{w}, \overline{w})$ implies (15). In other words, $\Omega(w)$ monotonically decreasing in $w$ on $(\underline{w}, \overline{w})$ ensures that the optimal marginal tax rates under maximin are larger than the ones under the general social welfare function. We differentiate (15):

$$\begin{aligned} \Omega'(w) &= v''(x_M(w))\dot{x}_M(w) \int_w^{\overline{w}} \frac{f(t)}{v'(x_M(t))} dt \\ &-v''(x_\Phi(w))\dot{x}_\Phi(w) \int_w^{\overline{w}} \left( \frac{1}{v'(x_\Phi(t))} - \frac{\Phi'(u(t))}{\lambda_\Phi} \right) f(t)dt - v'(x_\Phi(w)) \frac{\Phi'(u(w))}{\lambda_\Phi} f(w) \end{aligned} \tag{18}$$

**Proposition 2** *With quasilinear-in-consumption preferences and when* $h(\ell)$ *takes the isoelastic form, the marginal tax rate* $T'_M(y(w))$ *derived under maximin is always larger than that under the general social welfare function* $\int_{\underline{w}}^{\overline{w}} \Phi(u(w))dw$, $\forall w \in (\underline{w}, \overline{w})$.

**Proof.** Substituting $v'(x) = 1$ and $v''(x) = 0$ into (18), we obtain:

$$\Omega'(w) = -\frac{\Phi'(u(w))}{\lambda_\Phi} f(w) < 0$$

This completes the proof that $\Omega(w)$ is monotonically decreasing in $w$ under quasilinear-in-consumption preferences. ■

**Proposition 3** *With separable utility, close to the bottom and the top of the skill distribution, the marginal tax rate* $T'_M(y(w))$ *derived under maximin is always larger than that under criterion* $\int_{\underline{w}}^{\overline{w}} \Phi(u(w))dw$.

**Proof.** Evaluating (18) at $w = \underline{w}$, using (21), (23), (25) and (27) (in the Appendix) yields:

$$\Omega'(\underline{w}) = \frac{v''(x_M(\underline{w}))\dot{x}_M(\underline{w})}{\lambda_M} - v'(x_\Phi(\underline{w})) \frac{\Phi'(u(\underline{w}))}{\lambda_\Phi} f(\underline{w}) < 0$$

[6] It is worth noting that a discrete support for the skill distribution, hence social preferences written as $\sum_{\underline{w}}^{\overline{w}} \Phi(u(w))dw$, is a sufficient condition for having a strictly positive marginal tax rate at the bottom. Intuitively, the mass of people at the bottom of the skill distribution is then strictly positive hence a positive marginal tax rate at $w = \underline{w}$ improves equity.

From (18), when $w = \overline{w}$, we have:

$$\Omega'(\overline{w}) = \frac{-v'(x_\Phi(\overline{w}))\Phi(u(\overline{w}))f(\overline{w})}{\lambda_\Phi} < 0$$

Therefore, since Equation (17) states that $\Omega(\overline{w}) = 0$ and Equation (16) states $\Omega(\underline{w}) > 0$ we can conclude that $\Omega(w)$ is monotonically decreasing in $w$ close to $\underline{w}$ and $\overline{w}$, with general additively separable preferences. ■

# 3 Conclusion

The purpose of this paper has been to provide conditions under which maximin entails higher optimal marginal tax rates than other social preferences. Assuming quasilinear-in-consumption preferences and an isoelastic disutility of labor, the optimal marginal tax rates under maximin give an upper bound to the ones we would obtain under welfarist criteria that integrate over the population any concave transformation of individual utilities. With additive preferences, this dominance result is also valid close to the bounds of the skill distribution.

# References

Atkinson, A.B. and J.E. Stiglitz (1980), Lectures on Public Economics, New York: McGraw-Hill.

Atkinson, A.B. (1983), How Progressive Should Income Tax Be?, Chapter 15 in Social Justice and Public Policy, MIT Press (Cambridge, Mass).

Boadway, R. and L. Jacquet (2008), Optimal marginal and average income taxation under maximin, Journal of Economic Theory, 143(1): 425-441.

Choné, P. and G. Laroque (2005), Optimal incentives for labor force participation, Journal of Public Economics, 89(2-3), 395-425.

Diamond, P.A. (1998), Optimal income taxation: An example with a U-shaped pattern of optimal marginal tax rates, American Economic Review, 88(1): 83-95.

Ebert, U. (1992), A reexamination of the optimal nonlinear income tax, Journal of Public Economics, 49(1): 47-73.

Laroque, G. (2005), Income maintenance and labor force participation, Econometrica, 73(2): 341-376.

Mirrlees, J.A. (1971), An exploration in the theory of optimum income taxation, Review Economic Studies, 38(2): 175-208.

Pontryagin, L.S. (1964), The Mathematical Theory of Optimal Processes, The MacMillan Company, New York.

Sadka, E. (1976), On income distribution, incentive effects and optimal income taxation, Review of Economic Studies, 43(2): 261–267.

Saez, E. (2001), Using elasticities to derive optimal income tax rates, Review of Economic Studies, 68(1): 205-229.

Seade, J. (1977), On the Shape of Optimal Tax Schedules, Journal of Public Economics, 7(2): 203 236.

Seade, J. (1982), On the Sign of the Optimum Marginal Income Tax, Review of Economic Studies, 49: 637 643.

## Appendix: First-order conditions

This appendix gives the necessary conditions of (10) under the welfarist objective function (5) and the ones under maximin (4).

Integrating by parts to obtain $\int_{\underline{w}}^{\overline{w}} \zeta(w)\dot{u}(w)dw = \zeta(\overline{w})u(\overline{w}) - \zeta(\underline{w})u(\underline{w}) - \int_{\underline{w}}^{\overline{w}} \dot{\zeta}(w)u(w)dw$, the Lagrangian (11) becomes

$$\begin{aligned} \pounds \equiv W(u(w)) + \lambda \int_{\underline{w}}^{\overline{w}} \left[ [w\ell(w) - x(u(w), \ell(w))] f(w) - \frac{R}{\overline{w} - \underline{w}} \right] dw \\ + \zeta(\underline{w})u(\underline{w}) - \zeta(\overline{w})u(\overline{w}) + \int_{\underline{w}}^{\overline{w}} \left[ \zeta(w) \frac{\ell(w)h'(\ell(w))}{w} + \dot{\zeta}(w)u(w) \right] dw \end{aligned}$$

The rest of this section simplifies the mathematical writing by using the same notation for variables at the optimum under both objective functions. However, in the equations we need for a later demonstration, we add subscripts $\Phi$ or $M$ for social preferences (5) and for maximin, respectively. Under (5), the necessary conditions (assuming an interior solution) are:[7]

$$\frac{\partial \pounds}{\partial \ell(w)} = \lambda \left[ w - \frac{h'(\ell(w))}{v'(x(w))} \right] f(w) + \frac{\zeta(w)h'(\ell(w))}{w} \left( 1 + \frac{\ell(w)h''(\ell(w))}{h'(\ell(w))} \right) = 0 \quad \forall w \in W \tag{19}$$

$$\frac{\partial \pounds}{\partial u(w)} = \Phi'(u(w)) f(w) - \frac{\lambda f(w)}{v'(x(w))} + \dot{\zeta}(w) = 0 \quad \forall w \in (\underline{w}, \overline{w}) \tag{20}$$

$$\frac{\partial \pounds}{\partial u(\underline{w})} = \zeta_\Phi(\underline{w}) = 0 \tag{21}$$

$$\frac{\partial \pounds}{\partial u(\overline{w})} = \zeta_\Phi(\overline{w}) = 0 \tag{22}$$

Integrating $\dot{\zeta}(w)$ in (20) and using the transversality condition $\zeta_\Phi(\overline{w}) = 0$, we obtain:

$$-\frac{\zeta_\Phi(w)}{\lambda_\Phi} = \int_w^{\overline{w}} \left( \frac{1}{v'(x_\Phi(t))} - \frac{\Phi'(u(t))}{\lambda_\Phi} \right) f(t)dt > 0 \tag{23}$$

Using (3), (19) may be rewritten as:

$$\frac{T'_\Phi(y(w))}{1 - T'_\Phi(y(w))} = -\frac{\zeta_\Phi(w)v'(x(w))}{\lambda w f(w)} \left( 1 + \frac{\ell(w)h''(\ell(w))}{h'(\ell(w))} \right) \quad \forall w \in W \tag{24}$$

Finally, combining (23) and (24), the first-order conditions characterizing the optimal marginal tax rates under (5) can be written as (14).

Under maximin, we have the necessary condition (20) and also:

$$\frac{\partial \pounds}{\partial u(w)} = -\frac{\lambda f(w)}{v'(x(w))} + \dot{\zeta}_M(w) = 0 \quad \forall w \in (\underline{w}, \overline{w})$$

$$\frac{\partial \pounds}{\partial u(\underline{w})} = 1 + \zeta_M(\underline{w}) = 0 \tag{25}$$

$$\frac{\partial \pounds}{\partial u(\overline{w})} = \zeta_M(\overline{w}) = 0 \tag{26}$$

Integrating $\dot{\zeta}(w)$ in (20) and using the transversality condition $\zeta_M(\overline{w}) = 0$, we obtain:

$$-\frac{\zeta_M(w)}{\lambda} = \int_w^{\overline{w}} \frac{f(t)}{v'(x_M(t))} dt \tag{27}$$

[7] When we differentiate the Lagrangian, we must do so with respect to the end-points as well as the interior points, which gives the transversality conditions. These necessary conditions can also be derived based on variational techniques using Pontryagin's principle (Pontryagin, 1964). As is usual, we assume that $x(w)$ and $\ell(w)$ are continuous throughout and, in the absence of bunching, differentiable.

Using (3), (19) may be rewritten as:

$$\frac{T'_M(y(w))}{1-T'_M(y(w))} = -\frac{\zeta_M(w)v'\left(x(w)\right)}{\lambda w f(w)}\left(1+\frac{\ell(w)h''(\ell(w))}{h'(\ell(w))}\right) \quad \forall w \in W \tag{28}$$

Finally, combining (27) and (28), the first-order conditions characterizing the optimal marginal tax rates can be written as (12).

**Norges**
**Handelshøyskole**

Norwegian School of Economics
and Business Administration

NHH
Helleveien 30
NO-5045 Bergen
Norway

Tlf/Tel: +47 55 95 90 00
Faks/Fax: +47 55 95 91 00
nhh.postmottak@nhh.no
www.nhh.no