

Volume 30, Issue 4

Optimal labor income taxation under maximin: An upper bound

Laurence Jacquet

Norwegian School of Economics and Business Administration and CESifo

Abstract

This paper compares marginal income tax rates for the maximin versus a welfarist criterion in the standard Mirleesian optimal income tax problem. It derives fairly mild conditions under which the former is higher than the latter. This strict dominance result is always valid close to the bounds of the skill distribution and almost everywhere, except at the upper bound, if preferences are quasilinear in consumption.

I would like to deeply thank Raouf Boucekkine, Etienne Lehmann, Agnar Sandmo, Guttorm Schjelderup and the Editor, Firouz Gahvari, for their helpful comments. I take sole responsibility for possible errors in the paper.

Citation: Laurence Jacquet, (2010) "Optimal labor income taxation under maximin: An upper bound", *Economics Bulletin*, Vol. 30 no.4 pp. 3152-3160.

Submitted: Sep 22 2010. **Published:** November 27, 2010.

1. Introduction

Choné and Laroque (2005) and Laroque (2005) have shown, in an optimal income tax model wherein the labor supply response is along the extensive margin, that the Rawlsian criterion provides a benchmark: the Laffer bound. All optimal allocations correspond to tax schedules that are below this benchmark. This note gives a comparable result when the labor supply is at the intensive margin as in the standard optimal income tax problem a la Mirrlees (1971). Here too the maximin solution provides a benchmark. Assuming that preferences are quasilinear in consumption with an isoelastic disutility for labor, maximin criterion gives an upper bound for the optimal marginal income tax schedule. All such schedules derived under a welfarist criterion—aggregating a concave transformation of individual utilities over the entire population—must lie below this benchmark. With a general separable utility function, this result remains 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 individuals have the same utility function and the latter takes an additively separable form as in Mirrlees (1971) and Atkinson and Stiglitz (1980):

$$U(x, \ell) = v(x) - h(\ell)$$

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

Agents 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 $w \in W = [\underline{w}, \bar{w}]$, where $0 < \underline{w} < \bar{w} < \infty$. The density function, $f(w) = F'(w)$, is assumed to be differentiable and strictly positive for all $w \in W$. Individuals obtain their income from wages, with labor income denoted by $y \equiv w\ell$. Therefore, we can use $\ell = y/w$ to rewrite the utility function as

$$v(x) - h(y/w) \tag{1}$$

The government can observe incomes but not skills or labor supplied, so it bases its tax scheme $T(\cdot)$ on income y . $T(\cdot)$ is assumed differentiable in y . The budget constraint for individual w is:

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

where $T(y(w))$ is the tax imposed on type- w individuals. Each agent therefore chooses her income by solving:

$$\max_y v(y(w) - T(y(w))) - h(y/w)$$

Let $u(w)$ be the value of this program. The first-order condition associated to this program implies

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

where the left-hand side is the marginal rate of substitution between income and consumption.

For later use, consumption $x(w)$ can be treated as an implicit function of $u(w)$ and $y(w)$ and denoted by $X(u(w), y(w))$, where by differentiating (1), we obtain:

$$\frac{\partial X(u(w), y(w))}{\partial y} = \frac{h'(y(w)/w)}{wv'(x(w))}, \quad \frac{\partial X(u(w), y(w))}{\partial u} = \frac{1}{v'(x(w))} \quad (4)$$

We will compare the optimal tax schedules derived under a maximin criterion and a welfarist criterion that sums over all individuals a transformation Φ of individuals' utility with $\Phi' > 0$ and $\Phi'' \leq 0$ (hence the government has a non-negative aversion to inequality) and Φ 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 skill \underline{w} at the bottom of the skill distribution hence the maximin criterion is

$$u(\underline{w}) \quad (5)$$

The welfarist social preferences are

$$\int_{\underline{w}}^{\bar{w}} \Phi(u(w))f(w)dw \quad (6)$$

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

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

$$\int_{\underline{w}}^{\bar{w}} [y(w) - X(u(w), y(w))]f(w)dw \geq R \quad (7)$$

where R is an exogenous revenue requirement. This constraint must be binding at the optimum since utility is increasing in consumption.

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

$$u(w) \equiv v\left(x\left(\frac{y(w)}{w}\right)\right) - h\left(\frac{y(w)}{w}\right) \geq v\left(x\left(\frac{y(\hat{w})}{\hat{w}}\right)\right) - h\left(\frac{y(\hat{w})}{\hat{w}}\right) \quad \forall (w, \hat{w}) \in W^2 \quad (8)$$

We assume that $w \mapsto y(w)$ is continuous on $[\underline{w}, \bar{w}]$ and differentiable everywhere, except for a finite number of skill levels and that $w \mapsto u(w)$ is differentiable. Hence, $w \mapsto x(w)$ is also continuous everywhere and differentiable almost everywhere. These assumptions are made for reasons of tractability and have been standard since Guesnerie and Laffont (1984).

Our individual preferences ensure that the strict-single crossing (Spence-Mirrlees) condition holds. Hence, constraints (8) are equivalent to imposing the following differential equation (see Mirrlees 1971) that is called first-order incentive compatibility conditions (FOIC):

$$\dot{u}(w) \stackrel{\text{a.e.}}{=} h'(\cdot) \frac{y(w)}{w^2} > 0 \quad \forall w \quad (9)$$

and the monotonicity requirement that the earnings level $y(w)$ be a nondecreasing function of the skill level w .¹

The problem for the government is to choose $y(w)$ and $u(w)$ to maximize its welfare function subject to the budget constraint (7) and the FOIC conditions (9):

$$\underset{\{u(w), y(w)\}}{\text{Max}} W(u(\cdot)) \text{ s.t. } \int_{\underline{w}}^{\bar{w}} [y(w) - X(u(w), y(w))] f(w) dw = R, \quad \dot{u}(w) = h' \left(\frac{y(w)}{w} \right) \frac{y(w)}{w^2} \quad (10)$$

where the social welfare function $W(u(\cdot))$ represents either (5)² or (6).

The corresponding Lagrangian is:

$$\begin{aligned} \mathcal{L} \equiv & W(u(\cdot)) + \lambda \int_{\underline{w}}^{\bar{w}} \left[[y(w) - X(u(w), y(w))] f(w) - \frac{R}{\bar{w} - \underline{w}} \right] dw \\ & + \int_{\underline{w}}^{\bar{w}} \zeta(w) \left[h' \left(\frac{y(w)}{w} \right) \frac{y(w)}{w^2} - \dot{u}(w) \right] dw \end{aligned} \quad (11)$$

where λ is the multiplier associated with the binding budget constraint (7) and $\zeta(w)$ is the multiplier associated with the FOIC conditions (9). The necessary conditions are given in the Appendix.

Under maximin, the first-order conditions reduce to the following:

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

¹In the core of the paper, for simplicity, we follow the (usual) first-order approach and ignore the monotonicity requirement $\dot{y}(w) \geq 0$ (or equivalently $\dot{x}(w) \geq 0$) (Ebert 1992). If the second-order incentive compatibility (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 agents of different skills. The appendix gives the necessary conditions for the government's problem when bunching occurs.

²The maximin solution can also be obtained from an equivalent revenue-maximizing problem as follows. Take \underline{u} as given and consider the tax profiles that will generate it, given the incentive conditions. Clearly, \underline{u} can be supported by a large number of tax profiles such that tax revenues are no greater than R , that is,

$$\int_{\underline{w}}^{\bar{w}} [y(w) - X(u(w), y(w))] f(w) dw \leq R$$

As long as the incentive constraints are satisfied for all w , we know from the above problem that achieving $u(\underline{w}) = \underline{u}$ requires that the tax revenue generated cannot exceed R , so the above inequality must be satisfied. In fact, if we maximize the amount of tax revenue that will yield utility \underline{u} for the worst-off agents, that level of revenue will be precisely R . Therefore, maximizing tax revenue subject to $u(\underline{w}) \geq \underline{u}$ and the incentive conditions is equivalent maximizing $u(\underline{w})$ subject to the revenue and incentive constraints.

where the subscript M states for maximin and where

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

is a measure of the elasticity of labor supply.³

Under the social welfare function (6), the marginal tax rate denoted by $T'_\Phi(y(w))$ can be expressed as:⁴

$$\frac{T'_\Phi(y(w))}{1 - T'_\Phi(y(w))} = A(w) \frac{1}{wf(w)} v'(x_\Phi(w)) \int_w^{\bar{w}} \left(\frac{1}{v'(x_\Phi(t))} - \frac{\Phi'(u(t))}{\lambda_\Phi} \right) f(t) dt \quad \forall w \in W \quad (13)$$

where the subscript Φ states for the social objective $\int_w^{\bar{w}} \Phi(u(w))f(w) dw$.

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 $T'_M(y(w))/(1 - T'_M(y(w))) - T'_\Phi(y(w))/(1 - T'_\Phi(y(w))) \geq 0 \forall w$ 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

$$\Omega(w) \equiv v'(x_M(w)) \int_w^{\bar{w}} \frac{f(t)}{v'(x_M(t))} dt - v'(x_\Phi(w)) \int_w^{\bar{w}} \left(\frac{1}{v'(x_\Phi(t))} - \frac{\Phi'(u(t))}{\lambda} \right) f(t) dt \geq 0 \quad \forall w \quad (14)$$

First, consider $\Omega(w)$ at $w = \underline{w}$. From (22) (in the Appendix) and the transversality condition $\zeta_\Phi(\underline{w}) = 0$, we have:

$$\Omega(\underline{w}) = v'(x_M(\underline{w})) \int_{\underline{w}}^{\bar{w}} \frac{f(t)}{v'(x_M(t))} dt > 0 \quad (15)$$

Second, putting $w = \bar{w}$ in (14) gives:

$$\Omega(\bar{w}) = 0 \quad (16)$$

Equation (15) 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 \tilde{w} distorts the labor supply of those with skill \tilde{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 $w \leq \tilde{w}$. As long as the latter outweighs the former in the welfare criterion, such transfers are

³The term $A(w)$ that can be rewritten as $[1 + \ell h''(\ell)/h'(\ell)]$ 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).

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

positively valued, hence an equity gain appears. Under social preferences $\int_{\underline{w}}^{\bar{w}} \Phi(u(w))f(w) dw$, 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 bottom (Boadway and Jacquet 2008).⁵ 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. $T'_M(y(\bar{w})) = T'_\Phi(y(\bar{w})) = 0$, which yields (16). These results can be summarized as follows.

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

From (15) and (16), deriving conditions under which $\Omega(w)$ is monotonically decreasing in w on (\underline{w}, \bar{w}) implies (14). In other words, $\Omega(w)$ monotonically decreasing in w on (\underline{w}, \bar{w}) ensures that the optimal marginal tax rates under maximin are larger than the ones under the general social welfare function. We differentiate (14):

$$\begin{aligned} \Omega'(w) = & v''(x_M(w))\dot{x}_M(w) \int_w^{\bar{w}} \frac{f(t)}{v'(x_M(t))} dt \\ & - v''(x_\Phi(w))\dot{x}_\Phi(w) \int_w^{\bar{w}} \left(\frac{1}{v'(x_\Phi(t))} - \frac{\Phi'(u(t))}{\lambda_\Phi} \right) f(t) dt - v'(x_\Phi(w)) \frac{\Phi'(u(w))}{\lambda_\Phi} f(w) \end{aligned} \quad (17)$$

Proposition 1 *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 $\int_{\underline{w}}^{\bar{w}} \Phi(u(w))f(w) dw$, $\forall w \in (\underline{w}, \bar{w})$.*

Proof. Substituting $v'(x) = 1$ and $v''(x) = 0$ into (17), 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 2 *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 $\int_{\underline{w}}^{\bar{w}} \Phi(u(w))f(w) dw$.*

⁵It is worth noting that a discrete support for the skill distribution, hence social preferences written as $\sum_{\underline{w}}^{\bar{w}} \Phi(u(w))dw$, 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.

Proof. Evaluating (17) at $w = \underline{w}$, using (20), (22), (25) and (27) (in the Appendix) yields:

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

From (17), when $w = \bar{w}$, we have:

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

Therefore, since Equation (16) states that $\Omega(\bar{w}) = 0$ and Equation (15) states $\Omega(\underline{w}) > 0$ we can conclude that $\Omega(w)$ is monotonically decreasing in w close to \underline{w} and \bar{w} , with general additively separable preferences. ■

The work undertaken in this note identifies the following extension: What happens when the utility is not linear in consumption or not isoelastic in hours of work? It would be interesting to derive conditions under which, at least over some range of the skill distribution, our upper bound result does not hold, i.e. under which the second term (which is positive) in (14) offsets the other two terms (which are negative). This is left for future research.

3. Conclusion

The purpose of this note has been to provide conditions under which maximin entails higher optimal marginal tax rates than other social preferences, at any skill level. 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.
- 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.
- Guesnerie, R. and J-J. Laffont (1984) “A complete solution to a class of principal-agent problems with an application to the control of a self-managed firm”, *Journal of Public Economics* **25**(3), 329-69.

- 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(4)**, 637 643.

Appendix: First-order conditions

This appendix gives the necessary conditions of (10) under the welfarist objective function (6) and the ones under maximin (5).

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

$$\begin{aligned} \mathcal{L} \equiv & W(u(\cdot)) + \lambda \int_{\underline{w}}^{\bar{w}} \left[[y(w) - X(u(w), y(w))]f(w) - \frac{R}{\bar{w} - \underline{w}} \right] dw \\ & + \zeta(\underline{w})u(\underline{w}) - \zeta(\bar{w})u(\bar{w}) + \int_{\underline{w}}^{\bar{w}} \left[\zeta(w)h' \left(\frac{y(w)}{w} \right) \frac{y(w)}{w^2} + \dot{\zeta}(w)u(w) \right] dw \end{aligned}$$

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 Φ or M for social preferences (6) and for maximin, respectively. Under (6), the necessary conditions (assuming an interior solution) are:⁶

$$\frac{\partial \mathcal{L}}{\partial y(w)} = \lambda \left[1 - \frac{h'(\cdot)}{wh'(\cdot)} \right] f(w) + \frac{\zeta(w)h'(\cdot)}{w^2} \left(1 + \frac{y(w)h''(\cdot)}{wh'(\cdot)} \right) = 0 \quad \forall w \in W \quad (18)$$

$$\frac{\partial \mathcal{L}}{\partial u(w)} = \Phi'(u(w)) f(w) - \frac{\lambda f(w)}{v'(\cdot)} + \dot{\zeta}(w) = 0 \quad \forall w \in (\underline{w}, \bar{w}) \quad (19)$$

$$\frac{\partial \mathcal{L}}{\partial u(\underline{w})} = \zeta_{\Phi}(\underline{w}) = 0 \quad (20)$$

$$\frac{\partial \mathcal{L}}{\partial u(\bar{w})} = -\zeta_{\Phi}(\bar{w}) = 0 \quad (21)$$

Integrating $\dot{\zeta}(w)$ in (19) and using the transversality condition $\zeta_{\Phi}(\bar{w}) = 0$, we obtain:

$$-\frac{\zeta_{\Phi}(w)}{\lambda_{\Phi}} = \int_w^{\bar{w}} \left(\frac{1}{v'(x_{\Phi}(t))} - \frac{\Phi'(u(t))}{\lambda_{\Phi}} \right) f(t)dt \quad (22)$$

⁶When 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).

Using (3), (18) may be rewritten as:

$$\frac{T'_\Phi(y(w))}{1 - T'_\Phi(y(w))} = -\frac{\zeta_\Phi(w)v'(x(w))}{\lambda w f(w)} \left(1 + \frac{y(w)h''(y(w)/w)}{wh'(y(w)/w)}\right) \quad \forall w \in W \quad (23)$$

Finally, combining (22) and (23), the first-order conditions characterizing the optimal marginal tax rates under (6) can be written as (13).

Under maximin, we have the necessary condition (18) and also:

$$\frac{\partial \mathcal{L}}{\partial u(w)} = -\frac{\lambda f(w)}{v'(\cdot)} + \zeta_M(w) = 0 \quad \forall w \in (\underline{w}, \bar{w}) \quad (24)$$

$$\frac{\partial \mathcal{L}}{\partial u(\underline{w})} = 1 + \zeta_M(\underline{w}) = 0 \quad (25)$$

$$\frac{\partial \mathcal{L}}{\partial u(\bar{w})} = -\zeta_M(\bar{w}) = 0 \quad (26)$$

Integrating $\dot{\zeta}(w)$ in (24) and using the transversality condition $\zeta_M(\bar{w}) = 0$, we obtain:

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

Using (3), (18) may be rewritten as:

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

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

When the monotonicity constraint $\dot{y}(w) \geq 0$ binds over $[w_0, w_1]$, there is bunching over this interval. Equation (18) is then modified as follows (see, for instance, Guesneries and Laffont 1984):

$$\int_{w_0}^{w_1} \left\{ \lambda \left[1 - \frac{h'(\cdot)}{wh'(\cdot)} \right] f(w) + \frac{\zeta(w)h'(\cdot)}{w^2} \left(1 + \frac{y(w)h''(\cdot)}{wh'(\cdot)} \right) \right\} dw = 0$$

whereas Equation (19) (derived under (6)) and Equation (24) (derived under maximin) still hold.