Optimal income taxation with tax avoidance and endogenous labor supply

Georges Casamatta

Optimal income taxation with tax avoidance and endogenous labor supply*

Georges Casamatta

UMR CNRS 6240 LISA

Università di Corsica Pasquale Paoli, Corte, France

and Toulouse School of Economics

casamatta_g@univ-corse.fr

September 7, 2022

Abstract

We determine the optimal income tax schedule when individuals both choose endogenously their labor supply and have the possibility of avoiding paying taxes. Considering a convex concealment cost function, we propose a formula for the optimal marginal tax rate, that generalizes the formula of the standard Mirrlees model to the case of tax avoidance. We also show that the results obtained by Casamatta (2021) in the fixed income case hold true when labor supply is endogenous: with a low enough marginal cost of avoidance, it is optimal to let some taxpayers, located in the interior of the skill distribution, avoid taxes.

Keywords: tax avoidance, optimal income taxation, labor supply.

JEL: H21, H26.

*I thank the editor, Katherine Cuff, as well as two anonymous referees for very useful comments.

1 Introduction

For a long time, the literature on optimal labor income taxation, initiated by Mirrlees (1971) and surveyed in Piketty and Saez (2013), has focused on the labor supply response to taxes. In these models, individuals react to changes in marginal tax rates by adjusting their labor supply and thus the income they earn. More recently, it has been recognized that modifications in marginal tax rates can induce individuals to alter their taxable income, not only through a change in labor supply, but also through a change in declared income (while keeping true income unchanged) (Saez et al., 2012). This latter form of behavioral response to taxation can be illegal (tax evasion) or legal (tax avoidance).

In this article, we focus on tax avoidance, that can be defined, following Piketty et al. (2014), as "changes in reported income due to changes in the form of compensation but not in the total level of compensation". More precisely, we incorporate avoidance responses by taxpayers in the standard Mirrleesian income taxation model, besides the usual labor supply response. This is all the more important given that empirical evidence has pointed to fairly modest effects of taxes on labor supply (Keane, 2011), whereas the avoidance response appears stronger (Slemrod, 1995; Saez et al., 2012).

The first study that derived the optimal tax schedule while accounting for the avoidance margin was proposed by Grochulski (2007). In this model, labor supply is fixed so that individuals respond to taxes along the avoidance margin only. Considering a subadditive concealment cost function, Grochulski (2007) proves two main results. First, at the optimum with taxes, no individuals should hide income. This result is called the no-falsification theorem. Second, the optimal tax schedule is such that marginal tax rates are equal to the marginal falsification costs. Casamatta (2021) later relaxed the subadditivity assumption, by considering a convex cost function, and showed that the no-falsification theorem was not valid anymore in this setup, with part of the taxpayers reporting income lower than their true income.

None of these two articles did incorporate labor supply responses to taxation. Earlier work by Kopczuk (2001) and Slemrod and Kopczuk (2002) consider both tax avoidance and endogenous

labor supply, while focusing on a linear income tax schedule. More recently, Piketty et al. (2014) have proposed a formula for the top-income marginal tax rate when individuals have a quasi-linear utility function (no income effects on labor supply). Huang and Rios (2016) also adopt the quasi-linear framework and characterize the full nonlinear optimal tax curve. Other related work include Selin and Simula (2020), who analyze income shifting, and Doligalski and Rojas (2022), who determine the optimal tax schedule when individuals may decide to work in the informal sector. These two last articles differ mainly from the rest of the literature in that they consider a fixed cost of avoidance.[1]

Our work belongs to this strand of the literature and extends it in two directions. First, we consider a general form for the utility function, hence allowing for income effects on the labor supply. Second, we consider a concealment cost function which may be non-differentiable at 0, implying that some individuals adopt a corner solution for their tax avoidance. We develop a formula for the marginal tax rates in this general framework and contrast it with the formula of the standard Mirrlees model. We also show that individuals located at the extremes of the skill distribution do not avoid taxes, while those located in the interior of the distribution do, thus extending the result obtained by Casamatta (2021) in the fixed income case, to the case of endogenous labor supply.

2 Model

Individuals differ with respect to productivity w, distributed according to the cumulative distribution function $F(.)$ and the density $f(.)$ on the support $[w_-, w_+]$, with $f(w) > 0 \; \forall w \in [w_-, w_+]$. An individual with productivity w generates income $y = wl$, where l denotes his labor supply.

True income is not observable to the fiscal authority and individuals have the possibility to hide (legally) part of it to the government. This action is however costly and we denote $\phi(\Delta)$ the cost of hiding Δ euros, with $\phi(0) = 0$.

We allow for the possibility that individuals declare more than their true income, in which

[1] To be precise, Doligalski and Rojas (2022) consider both a fixed cost and a variable cost of tax avoidance. They however assume that the variable cost of avoidance has a constant rate, while the mechanism of our analysis relies on the strictly convex schedule of the variable cost of avoidance.

case $\Delta < 0$. We assume that ϕ is continuous and strictly convex.[2] Moreover, it is decreasing for $\Delta < 0$ and increasing for $\Delta > 0$. It is differentiable everywhere, except possibly at 0 where the right-hand (resp. left-hand) derivative is positive (resp. negative). We denote these derivatives $\phi'_+(0)$ and $\phi'_-(0)$, respectively. The literature on tax avoidance, with the exception of Casamatta (2021), assumes a differentiable avoidance cost function: $\phi'_+(0) = \phi'_-(0) = 0$. We shall see in what follows that this assumption has important implications for tax avoidance behavior, as well as for the shape of optimal marginal tax rates.

The utility function, $u(c, l)$, is increasing in consumption c and decreasing in labor supply l.[3] Moreover, we assume that leisure is a normal good and that indifference curves are strictly convex and single-crossing in the (c, y) space.

3 The social planner's program

We adopt a mechanism design approach in which each individual directly announces his/her type w to the social planner and is assigned a reported income $r(w)$ and a tax payment $T(w)$. From the revelation principle, individuals should report their type truthfully.

When designing the functions $r(.)$ and $T(.)$, the planner anticipates the optimal reaction of individuals for their true income, y, which is not observable. The income earned by an individual with productivity w who pretends to be of type w' is denoted $\hat{y}(w', w)$. It is obtained by solving the program:

$$\max_y V(y, w', w) \equiv u(y - T(w') - \phi(y - r(w')), y/w). \tag{1}$$

Under truth-telling, the income earned by an individual with productivity w is $y(w) \equiv \hat{y}(w, w)$.

The social planner chooses the functions $r(.)$ and $T(.)$ in order to maximize social welfare, expressed as the sum of a continuous, increasing and concave transformation $G(.)$ of individual

[2] Strict convexity, combined with the fact that $\phi(0) \leq 0$, implies that ϕ violates the subaddivity assumption in Grochulski (2007) and Landier and Plantin (2017), where a function f is subadditive if and only if $f(x + y) \leq f(x) + f(y)$.

[3] Avoidance is assumed to generate a *monetary* cost only, as in Piketty et al. (2014) and Huang and Rios (2016). A more general formulation would allow avoidance to enter directly as an argument of the utility function, recognizing that avoidance potentially also generates non-monetary costs, such as the time spent to optimize the income reporting strategy.

utility levels, under resource and incentive constraints:

$$(P1) \max_{r(.),T(.)} \int G(U(w))dF(w)$$

subject to

$$U(w) = u(c(w), y(w)/w) \tag{2}$$

$$c(w) = y(w) - T(w) - \phi(y(w) - r(w)) \tag{3}$$

$$\int T(w)f(w)dw \geq 0 \tag{4}$$

and

$$U(w) \geq V(\hat{y}(w', w), w', w), \ \forall w, w' \in [w_-, w_+], \tag{5}$$

where equation (4) represents the Government Budget Constraint (GBC) and (5) the incentive constraints: a type w individual should not want to pretend that he is of type w'.

3.1 Optimal labor supply

As stated above, individuals determine their true income by solving program (1). An *interior* solution satisfies the first-order condition:

$$1 - \phi'(\hat{y}(w', w) - r(w')) = MRS_{c,l}(\hat{y}(w', w), w', w), \tag{6}$$

where

$$MRS_{c,l}(y, w', w) \equiv -\frac{1}{w} \frac{u_l(y - T(w') - \phi(y - r(w')), y/w)}{u_c(y - T(w') - \phi(y - r(w')), y/w)}$$

is the marginal rate of substitution between leisure and consumption of an individual with income y and type w who pretends to have type w'.

Three cases are possible: $\hat{y}(w', w) > r(w')$, $\hat{y}(w', w) < r(w')$ or $\hat{y}(w', w) = r(w')$. The first two

cases involve an interior solution. They arise respectively when:

$$1 - \phi'_+(0) > MRS_{c,l}(r(w'), w', w)$$

and

$$1 - \phi'_-(0) < MRS_{c,l}(r(w'), w', w).$$

The corner solution $\hat{y}(w', w) = r(w')$ occurs when:

$$1 - \phi'_+(0) \leq MRS_{c,l}(r(w'), w', w) \leq 1 - \phi'_-(0). \tag{7}$$

3.2 Incentive constraints

From the incentive constraints in program $(P1)$, every individual should report truthfully his type. Therefore:

$$w = \arg\max_{w'} V(\hat{y}(w', w), w', w). \tag{8}$$

In the remainder of this article, *we shall assume that the second-order condition of program (8) is satisfied*, therefore solving the first-order mechanism design approach (Jacquet and Lehmann, 2021). Violation of the second-order condition implies that a subset of individuals should be bunched at the same allocation, declaring the same level of income and paying the same amount of taxes (Lollivier and Rochet, 1983; Ebert, 1992; Boadway et al., 2000).

Using the first-order condition of program (8), we can prove the following lemma.

Lemma 1. *The incentive constraints (5) imply:*

$$\frac{dU}{dw} = -\frac{y(w)}{w^2} u_l. \tag{9}$$

Proof. See appendix A. □

Condition (9), which summarizes the incentive constraints under the first-order approach, is identical to the one encountered in the standard Mirrlees model (Mirrlees, 1971; Piketty and Saez,

2013).

Taking $U(w)$ as the state variable and $r(w)$ as the control variable, the planner's program can then be restated as follows:

$$(P2) \max_{r(.),U(.)} \int G(U(w))dF(w)$$

st

$$\int (y(w) - c(w) - \phi(y(w) - r(w)))f(w)dw \geq 0 \text{ and } (9),$$

where the first constraint is the GBC, that has been re-expressed by using (3). Optimal labor income $y(w)$ and consumption $c(w)$ are determined by inverting the condition $U(w) = u(c(w), y(w)/w)$ and solving program (1) with $w' = w$.

4 First-order conditions of the planner's program

We form the Hamiltonian associated to program $(P2)$ above:

$$\mathcal{H} = (G(U(w)) + \mu(y(w) - c(w) - \phi(y(w) - r(w))))f(w) + \lambda(w)\frac{dU}{dw},$$

where μ and $\lambda(w)$ are the multipliers associated to the GBC and the incentive constraints respectively.

Assuming an interior solution for y ($y(w) \neq r(w)$), the first-order conditions write:

$$\begin{aligned}
&\frac{\partial \mathcal{H}}{\partial r} = 0 \\
\Leftrightarrow &\mu(\phi' + \frac{dy}{dr}(1 - \phi') - \frac{dc}{dr})f(w) + \lambda(w)(\frac{dy}{dr}(-\frac{1}{w^2}u_l - \frac{y(w)}{w^3}u_{ll}) - \frac{dc}{dr}\frac{y(w)}{w^2}u_{cl}) = 0. \qquad (10) \\
&\frac{\partial \mathcal{H}}{\partial U} = -\lambda'(w) \\
\Leftrightarrow &-\lambda'(w) = (G'(U(w)) - \mu\frac{dc}{dU})f(w) - \lambda(w)\frac{y(w)}{w^2}u_{cl}\frac{dc}{dU}. \qquad (11)
\end{aligned}$$

We differentiate (2), to obtain:

$$\frac{dc}{dr} = -\frac{1}{w}\frac{u_l}{u_c}\frac{dy}{dr}$$
$$\frac{dc}{dU} = \frac{1}{u_c}.$$

Substituting these expressions into the first-order conditions (10) and (11), these latter become:

$$\mu(\phi' + \frac{dy}{dr}(1 - \phi' + \frac{1}{w}\frac{u_l}{u_c}))f(w) + \lambda(w)\frac{dy}{dr}(-\frac{1}{w^2}u_l + \frac{y(w)}{w^3}(u_{cl}\frac{u_l}{u_c} - u_{ll})) = 0 \tag{12}$$

$$-\lambda'(w) = (G'(U(w)) - \frac{\mu}{u_c})f(w) - \lambda(w)\frac{y(w)}{w^2}\frac{u_{cl}}{u_c}. \tag{13}$$

In the case of a corner solution ($y(w) = r(w)$), (12) becomes:

$$\mu(1 + \frac{1}{w}\frac{u_l}{u_c}))f(w) + \lambda(w)(-\frac{1}{w^2}u_l + \frac{y(w)}{w^3}(u_{cl}\frac{u_l}{u_c} - u_{ll})) = 0. \tag{14}$$

Integrating (13) yields:

$$\lambda(w) = \int_w^{w_+} (G'(U(m)) - \frac{\mu}{u_c}) \exp(-\int_w^m (y(s)u_{cl})/(s^2 u_c)ds)f(m)dm, \tag{15}$$

where $\lambda(w_-) = \lambda(w_+) = 0$.

5 Reported incomes at the second-best optimum

We now establish a first feature of the optimal allocation, that generalizes a result by Casamatta (2021), obtained in the case of fixed incomes.

Proposition 1. *At the optimal (second-best) allocation:*

1. *Individuals do not declare more than their true income:* $r(w) \leq y(w), \forall w \in [w_-, w_+]$.

2. *Assume that* $\phi'_+(0) > 0$. *Then Individuals located at the extremes of the skill distribution declare their true income. In other words, there exist* $w' > w_-$ *and* $w'' < w_+$ *such that*

$r(w) = y(w), \forall w \in [w_-, w'] \cup [w'', w_+]$.

3. *Assume that the optimal allocation* $\{c(w), l(w)\}$*, as well as the multiplier* $\lambda(w)$*, are continuous with respect to* $\phi'_+(0)$*. Then, for* $\phi'_+(0)$ *low enough, some individuals, located in the interior of the skill distribution, avoid taxes. In other words, there exists* $w \in (w_-, w_+)$ *such that* $r(w) < y(w)$*.*

Proof. See appendix B. □

This proposition says that individuals either declare their true income or understate their income to the fiscal administration. The reason why it may be optimal to let some people avoid taxes is that it allows to relax the incentive constraints.

To see this, consider a discretized version of the model, with a finite number of possible types equally spaced by an amount δ. The utility obtained by individuals $\tilde{w}+\delta$ when mimicking individuals $\tilde{w}$ is:

$$V(\hat{y}(\tilde{w}+\delta, \tilde{w}), \tilde{w}+\delta, \tilde{w}) = u(\hat{y}(\tilde{w}+\delta, \tilde{w}) - \tilde{T} - \phi(\hat{y}(\tilde{w}+\delta, \tilde{w}) - \tilde{r}), \hat{y}(\tilde{w}+\delta, \tilde{w})/(\tilde{w}+\delta)),$$

where $\tilde{T} \equiv T(\tilde{w})$ and $\tilde{r} \equiv r(\tilde{w}) = y(\tilde{w}) - \epsilon$, with ϵ the amount of avoidance by individuals $\tilde{w}$.

Consider now an increase in the amount of avoidance $d\epsilon$ that leaves individuals $\tilde{w}$ with the same level of utility. Using the first-order condition (6), it can be checked that $d\tilde{T}/d\epsilon = -\phi'(\epsilon)$: the increase in the avoidance cost must be compensated by a the reduction in the tax paid by individuals $\tilde{w}$.

It follows, again using (6), that the utility variation of a type $\tilde{w}+\delta$ who mimics a type $\tilde{w}$ is:

$$\begin{aligned}
\frac{dV}{d\epsilon}(\hat{y}(\tilde{w}+\delta, \tilde{w}), \tilde{w}+\delta, \tilde{w}) &= (-\frac{d\tilde{T}}{d\epsilon} - \phi'(\hat{y}(\tilde{w}+\delta, \tilde{w}) - \tilde{r}))u_c(.) \\
&= (\phi'(\epsilon) - \phi'(\hat{y}(\tilde{w}+\delta, \tilde{w}) - \tilde{r}))u_c(.) \\
&= (\phi'(\epsilon) - \phi'(\epsilon + \hat{y}(\tilde{w}+\delta, \tilde{w}) - \tilde{y}))u_c(.).
\end{aligned}$$

With a convex concealment cost function, and provided that $\hat{y}(\tilde{w}+\delta,\tilde{w}) > \tilde{y}$,[4] this expression is negative. This means that individuals $\tilde{w}+\delta$ are less tempted to mimic individuals $\tilde{w}$ when avoidance by these latter is increased.

Note that individuals $\tilde{w}$ must be compensated for this increase in avoidance, through a reduction in their tax payment equal to $\phi'(\epsilon)$. As soon as ϵ and $\phi'_+(0)$ are sufficiently small, however, the cost for the planner in terms of public funds is negligible, implying that the welfare gain from the relaxation of the incentive constraints outweighs the resource cost. This is the case for individuals located in the interior of the skill distribution, for whom the benefit of relaxing the incentive constraint, as measured by the absolute value of the multiplier $\lambda(w)$, is strictly positive.[5] For individuals located at the extremes of the skill distribution, on the other hand, the benefit of relaxing the incentive constraints is nil, implying that they should not avoid taxes. This property can be related to the classical result of no distortion at the top and the bottom in optimal tax analysis (Seade, 1977).[6] By continuity of $\lambda(w)$, individuals located close to the top or the bottom should not avoid taxes either, as long as $\phi'_+(0) > 0$.[7]

To summarize, altering the allocation at a given skill level $\tilde{w}$ by marginally increasing the amount of avoidance makes this allocation less attractive for individuals with a higher skill because, the avoidance cost being strictly convex, these higher skill individuals incur a larger marginal avoidance cost than $\tilde{w}$ when mimicking and therefore require a larger monetary compensation for staying at their initial utility level. Therefore, individuals with skills higher than $\tilde{w}$ are less tempted to mimick $\tilde{w}$ when avoidance by these latter is increased, meaning that incentive constraints are being relaxed.

[4] One can check that this is the case as soon as the agent monotonicity, or single-crossing, condition is satisfied (Mirrlees, 1976; Ebert, 1992).

[5] One cannot guarantee a priori that *all* individuals in one interior range of skills should avoid taxes. This depends on the comparison between the multiplier $\lambda(w)$ and the marginal avoidance cost $\phi'_+(0)$.

[6] Mathematically, this is reflected by the fact that $\lambda(w_-) = \lambda(w_+) = 0$. Relaxing the incentive constraint at a given skill level allows to increase the tax receipts from people with higher skills, and to redistribute this additional income towards people with skills below this level. Therefore, there is no benefit of allowing avoidance neither at the top of the skill distribution, as there are no individuals to be taxed above this level, nor at the bottom, as the mass of people who could benefit from this enhanced redistribution is 0.

[7] It should be noted that, when $\phi'(0) = 0$, the usual assumption in the literature, all individuals in the interior of the skill distribution should misreport their income.

6 Marginal tax rates

In this section, we determine the optimal marginal tax rates, both on individuals who report their true income and on individuals who practice avoidance. In this purpose, we first derive the elasticities of real and taxable income. For a given tax schedule $T_r(r)$, individuals choose real income y and taxable income r in order to maximize $u(c, l)$ subject to the budget constraint $c = y - T_r(r) - \phi(y - r)$. Substituting the budget constraint into the utility function, an individual with productivity w solves the program:

$$\max_{y \geq 0, r \geq 0} u(y - T_r(r) - \phi(y - r), y/w). \tag{16}$$

In the case where both y and r are interior solutions, the first-order conditions write:

$$\begin{aligned} (1 - \phi')u_c + \frac{1}{w}u_l &= 0 \\ (1 - \tau) - 1 + \phi' &= 0, \end{aligned}$$

where $\tau \equiv T_r'(r)$. Substituting the second condition into the first, this system of equation becomes:

$$(1 - \tau)u_c + \frac{1}{w}u_l = 0 \tag{17}$$

$$(1 - \tau) - 1 + \phi' = 0. \tag{18}$$

Differentiating these conditions, and defining Ω as:

$$\Omega \equiv u_{ll} + w^2(1 - \tau)^2 u_{cc} + 2w(1 - \tau)u_{cl},$$

we derive formulas for the elasticities of the true and reported incomes with respect to the net-of-tax rate, which are presented in the next lemma.

Lemma 2. *When an individual chooses an interior solution for both y and r:*

1. *Uncompensated and compensated elasticities of true income with respect to the net-of-tax rate*

are, respectively:

$$e_y = -\frac{1-\tau}{y}\frac{w^2(u_c + r(1-\tau)u_{cc} + \frac{r}{w}u_{cl})}{\Omega} \tag{19}$$

$$e_y^c = \frac{1-\tau}{y}\frac{-w^2 u_c + (y-r)(w^2(1-\tau)u_{cc} + w u_{cl})}{\Omega} \tag{20}$$

2. *Uncompensated and compensated elasticities of reported income with respect to the net-of-tax rate are, respectively:*

$$e_r = \frac{1}{\phi''}\frac{1-\tau}{r} + \frac{y}{r}e_y = \frac{1-\tau}{r}\frac{1}{\phi''}(1 - \phi''\frac{w^2(u_c + r(1-\tau)u_{cc} + \frac{r}{w}u_{cl})}{\Omega}) \tag{21}$$

$$e_r^c = \frac{1-\tau}{r}\frac{1}{\phi''}(1 - \phi''\frac{w^2 u_c}{\Omega}) \tag{22}$$

Proof. See appendix C. □

In the case of a corner solution with no avoidance ($y = r$),[8] the elasticity formulas for true income y, (19) and (20), coincide with the ones obtained in the standard Mirrlees model (formulas (23) and (24) in Saez (2001)). Moreover, the elasticities of reported income, e_r and e_r^c, coincide with the elasticities of true income, e_y and e_y^c. Lemma 2 then implies that the elasticity of reported income, both in its compensated and uncompensated form, is *discontinuous* at the point $r = y$:

$$e_r|_{r \nearrow y} = \frac{1}{\phi''_+(0)}\frac{1-\tau}{y} + e_y > e_r|_{r=y} = e_y \tag{23}$$

$$e_r^c|_{r \nearrow y} = \frac{1}{\phi''_+(0)}\frac{1-\tau}{y} + e_y^c > e_r^c|_{r=y} = e_y^c. \tag{24}$$

The discontinuity arises because altering r for an avoiding individual induces a modification in the marginal avoidance cost (term $\phi''_+(0)$ in the above expressions). This effect is absent for individuals who report their true income.

[8] Other types of corner solutions correspond to violations of the non-negativity constraints $r \geq 0$ and $y \geq 0$.

6.1 The optimal marginal tax rate curve

We now turn to determining the optimal marginal tax rates. When the optimal choice of $r(w)$ by the social planner follows from the first-order conditions (12) or (14), we can, using (17) and (18),[9] express the marginal tax rate as follows:

$$\tau = -\frac{\lambda(w)}{\mu}\frac{1}{wf(w)}\frac{dy}{dr}(-\frac{1}{w}u_l + \frac{y(w)}{w^2}(u_{cl}\frac{u_l}{u_c} - u_{ll})), \tag{25}$$

where we note that $dy/dr = 1$ for individuals choosing $y = r$. For some individuals, however, we will show that the planner's solution for $r(w)$ may arise at a corner. This leads to the next proposition.

Proposition 2. *The optimal marginal tax rate faced by an individual with productivity w is equal to:*

$$T_r'(r(w)) = \min(\tau^*(w), \phi_+'(0)) \qquad \text{if } r(w) = y(w)$$
$$T_r'(r(w)) = \tau^*(w) > \phi_+'(0) \qquad \text{if } r(w) < y(w)$$

where $\tau^*(w)$ satisfies:

$$\frac{\tau^*}{1-\tau^*} = A(w)B(w)C(w) \tag{26}$$

with

$$A(w) = \frac{y(1+e_y) + (y-r)(e_r - e_r^c)}{re_r^c}$$
$$B(w) = \frac{1-F(w)}{wf(w)}$$
$$C(w) = \int_w^{w_+}\left(1 - \frac{G'(U(m))u_c(m)}{\mu}\right)\exp\left[\int_w^m\left(1-\frac{e_r}{e_r^c}\right)\frac{dr(s)}{r(s)}\right]\frac{f(m)}{1-F(w)}dm.$$

Proof. See appendix D. □

[9] It can be checked easily that (17) remains true in the case of a corner solution $y = r$.

The optimal marginal tax rate is lower than or equal to $\phi'_+(0)$ for complying individuals whereas it is larger than $\phi'_+(0)$ for avoiding ones. For complying individuals, two cases are possible. Either the planner solution for r is at a corner and the optimal tax rate is $\phi'_+(0)$, or the marginal tax rate is obtained from the first-order condition (14). For tax avoiding individuals, the optimal marginal tax rates is derived from (12).

In the last two cases (marginal tax rate derived from (12) or (14)), we recover the ABC formula in the optimal income taxation literature (Diamond, 1998; Saez, 2001; Jacquet et al., 2013), where A is the efficiency term, B the skill distribution term and C represents mechanical and income effects. We remind that the parameter μ is the multiplier of the government budget constraint (4). It represents the social welfare gain of giving an additional unit of money to all individuals and therefore allows to express the loss of welfare for individuals above w in terms of revenue.

In the absence of avoidance, the terms A and C become:

$$A(w) = \frac{1+e_y}{e_y^c}$$
$$C(w) = \int_w^{w_+} \left(1 - \frac{G'(U(m))u_c(m)}{\mu}\right) \exp\left[\int_w^m \left(1 - \frac{e_y}{e_y^c}\right) \frac{dy(s)}{y(s)}\right] \frac{f(m)}{1-F(w)} dm,$$

so that the optimal tax solution coincides with the standard ABC formula of the Mirrlees model (equation (15) in Saez (2001)). Recalling that individuals close to the bottom or to the top do not avoid taxes (Proposition 1) and moreover that the distribution of skills is assumed to be bounded, we therefore recover the zero marginal tax rate at the bottom and the top result, provided that there is no bunching at the bottom (Seade, 1977).

In the case of a quasi-linear utility function (no income effects on labor supply), formula (26)

simplifies to:[10]

$$\frac{T_r'(r(w))}{1-T_r'(r(w))} = \frac{y(1+e_y)}{re_r}\frac{1-F(w)}{wf(w)}\int_w^{w_+}\left(1-\frac{G'(U(m))}{\mu}\right)\frac{f(m)}{1-F(w)}dm.$$

6.2 A decomposition of the efficiency term

Formula (26) makes clear that the optimal marginal tax rates not only depend on the elasticity of earned income, as in the standard Mirrlees model, but also on the elasticity of reported income − both in its compensated and uncompensated form − as well as on the size of true income y relatively to reported income r. These differences with the standard optimal taxation setting arise entirely through the efficiency effect A. The next proposition, which is derived from results in Jacquet et al. (2013), allows for a better understanding of this effect.

Proposition 3. *The efficiency term $A(w)$ in the optimal marginal tax rate formula (26) is equal to $\gamma(w)/\varepsilon(w)$, where*

$$\gamma(w) = \frac{y(1+e_y)+(y-r)(e_r-e_r^c)}{r(1+\frac{T_r''}{1-\tau}re_r^c)} \tag{27}$$

is the elasticity of reported income with respect to productivity and

$$\varepsilon(w) = \frac{e_r^c}{1+\frac{T_r''}{1-\tau}re_r^c}$$

is the compensated elasticity of reported income with respect to the net-of-tax rate along the tax schedule.

Proof. See appendix E. □

The efficiency effect A can thus be decomposed into two terms, $\varepsilon(w)$ and $\gamma(w)$.

[10] Denoting sheltered income as $z \equiv y - r$, it can further be shown that this formula can be written as:

$$\frac{T_r'(r(w))}{1-T_r'(r(w))} = \left(\frac{ze_z}{re_r}+1\right)\left(1+\frac{1}{e_y}\right)\frac{1-F(w)}{wf(w)}\int_w^{w_+}\left(1-\frac{G'(U(m))}{\mu}\right)\frac{f(m)}{1-F(w)}dm,$$

which is close to formula (8) in Huang and Rios (2016), with two differences. First, these latter consider an alternative social welfare function, with exogenous weights on the utility of each individual. Second, they allow for a proportional tax t on consumption, in addition to the tax on labor income.

The elasticity $\varepsilon(w)$ represents the substitution effect: a modification of the marginal tax rate at a given income level induces individuals to lower their reported income. Because the tax system should be designed to limit these distortions as much as possible, we obtain that the optimal marginal tax rate is inversely related to this term. Note that $\varepsilon(w)$ differs from the standard elasticity e_r^c as soon as $T_r'' \neq 0$. This is due to the fact that, with a nonlinear tax schedule, a variation in reported income r induces a change in the marginal tax rate $T_r'(r)$. This feedback effect is not taken into account by the standard elasticity concept.[11]

Furthermore, the distortion caused by a modification in the marginal tax rate is increasing with the number of individuals directly affected by it. The income and skill distributions are linked through the following relationship:

$$wf(w) = \gamma(w) r h(r).$$

Therefore, the number of individuals at a given income level, $h(r)$, is inversely related to the elasticity $\gamma(w)$, which explains why the optimal tax rate is increasing with $\gamma(w)$. Inspecting (27), it appears that $\gamma(w)$ is increasing with the elasticity of true income, e_y, as in the standard Mirrlees model. However, when avoidance is possible, it is also increasing with the relative importance of true income, y, with respect to reported income, r, as well as with the size of income effects, $e_r - e_r^c$.

6.3 Sufficient statistics formula

In the next proposition, we express, following Saez (2001), the optimal marginal tax rates as a function of reported incomes.

Proposition 4. *When (26) is satisfied, the marginal tax rate faced by an individual with reported income r^* can be expressed as follows:*

$$\frac{T_r'(r^*)}{1 - T_r'(r^*) + r^* e_r^c T_r''} = \frac{1}{e_r^c} \frac{1 - H(r^*)}{r^* h(r^*)} \int_{r^*}^{r+} \left(1 - \frac{G'(U) u_c}{\mu}\right) \exp\left[\int_{r^*}^{r} \left(1 - \frac{e_r}{e_r^c}\right) \frac{dr'}{r'}\right] \frac{h(r)}{1 - H(r^*)} dr \tag{28}$$

[11] See Jacquet et al. (2013) or Jacquet and Lehmann (2021) for a detailed discussion of this point.

where $H(.)$ and $h(.)$ are the cdf and pdf of reported earnings, respectively.

Proof. See appendix F. □

This formula is identical to equation (14) in Saez (2001).[12] It can be interpreted in the same way, as the combination of mechanical, substitution and income effects. This of course does not imply that the optimal marginal tax rates are the same with and without avoidance, as the distribution of reported incomes depends on the avoidance technology, as emphasized in the discussion above. Nevertheless, this result is reassuring, in that the use of the formula derived by Saez (2001) in applied work remains valid under tax avoidance by individuals, with the usual caveat that elasticities, as well as the distribution of reported incomes, should be estimated at the optimal allocation, not the current one (Kleven, 2021).

7 Numerical examples

We report numerical examples in which the productivities are distributed according to a lognormal distribution on the support $[1, 10]$.

The utility function is assumed to be quasilinear and the disutility of labor takes an isoelastic form:

$$u(c, l) = c - \frac{l^{1+1/e}}{1 + 1/e},$$

where e, the elasticity of labor supply with respect to the net-of-tax-rate (both compensated and uncompensated as there are no income effects on labor supply), is assumed to be equal to 0.25.

We consider both a Utilitarian planner, with a social welfare transformation of utility $G(U) = \ln(U)$, and a Rawlsian planner, who maximizes the utility of the worse-off individuals in society, namely those with productivity w_-.

Finally, the avoidance cost function takes the form $\phi(\Delta) = \Delta^2/10 + |\alpha\Delta|$ where $\alpha \geq 0$, which implies $\phi'_-(0) = -\alpha$ and $\phi'_+(0) = \alpha$.

[12] Saez (2001) gets rid of the term $re_r^c T_r''$ in the denominator of the left-hand side by using the so-called virtual, instead of actual, density of incomes.

Results are displayed in figures 1 and 2, which represent the Utilitarian and Rawlsian cases, respectively.

In the Utilitarian setting (figure 1), we report simulations corresponding to three values of α, $\alpha = 0$, $\alpha = 0.15$ and $\alpha = 0.5$. When $\alpha = 0$, the marginal cost of avoiding the first euro is zero, so that all individuals, except those located at the bottom and the top of the skill distribution, practice avoidance. This is seen in the bottom right panel, that represents concealed income $y - r$ as a function of productivity w: all the individuals in the interior of the skill distribution report strictly less income than their true income. This is not the case anymore when $\alpha = 0.15$: some individuals, located close enough to the extremes of the skill distribution report their true income. When $\alpha = 0.5$, no individual avoids taxes and the optimal tax curve coincides with the solution of the standard Mirrlees model.

We note also that, for all values of α, the optimal marginal tax rate, represented on the top left panel, is zero at the bottom and the top. For individuals located in the interior of the skill distribution, the marginal tax rates correspond with the solution of the standard Mirrlees model when α is high enough ($\alpha = 0.5$), as no individual avoids taxes in such a case. For lower values of α ($\alpha = 0.15$), the marginal tax rate curve exhibits horizontal portions, corresponding to a tax rate equal to $\phi'_+(0) = 0.15$, in line with the result in Proposition 2. When $\alpha = 0$, the optimal marginal tax rates are everywhere given by the ABC formula, although they differ from the ones encountered in the model without tax avoidance.

Numerical simulations with a Rawlsian planner (figure 2) exhibit three main differences with respect to the Utilitarian case. First, marginal tax rates at the bottom are not zero anymore. As explained, e.g. in Boadway and Jacquet (2008), it is optimal to introduce a distortion at the bottom since, contrarily to the Utilitarian case, individuals located at the bottom do not have a zero mass anymore in the objective function (precisely because the Rawlsian planner maximizes the utility of people at the bottom), implying a positive equity effect, and therefore the reappearance of the equity-efficiency trade-off. Second, the existence of positive, and large, marginal tax rates at the bottom of the skill distribution discourages individuals from reporting their true income, leading to some tax avoidance behavior, in contrast to what we obtained with a Utilitarian criterion. Third, we

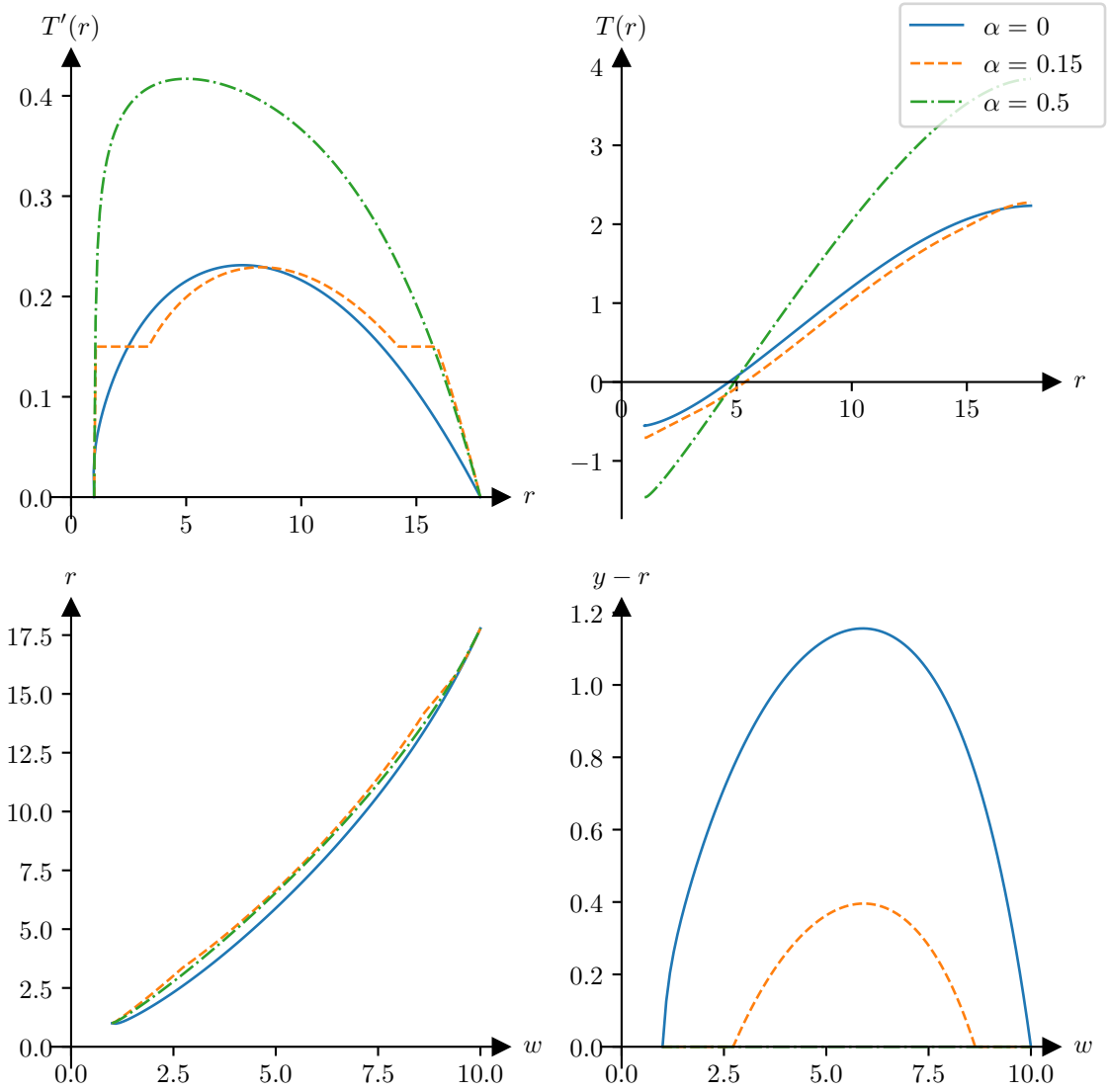

Figure 1: Optimal tax schedule with the Utilitarian criterion

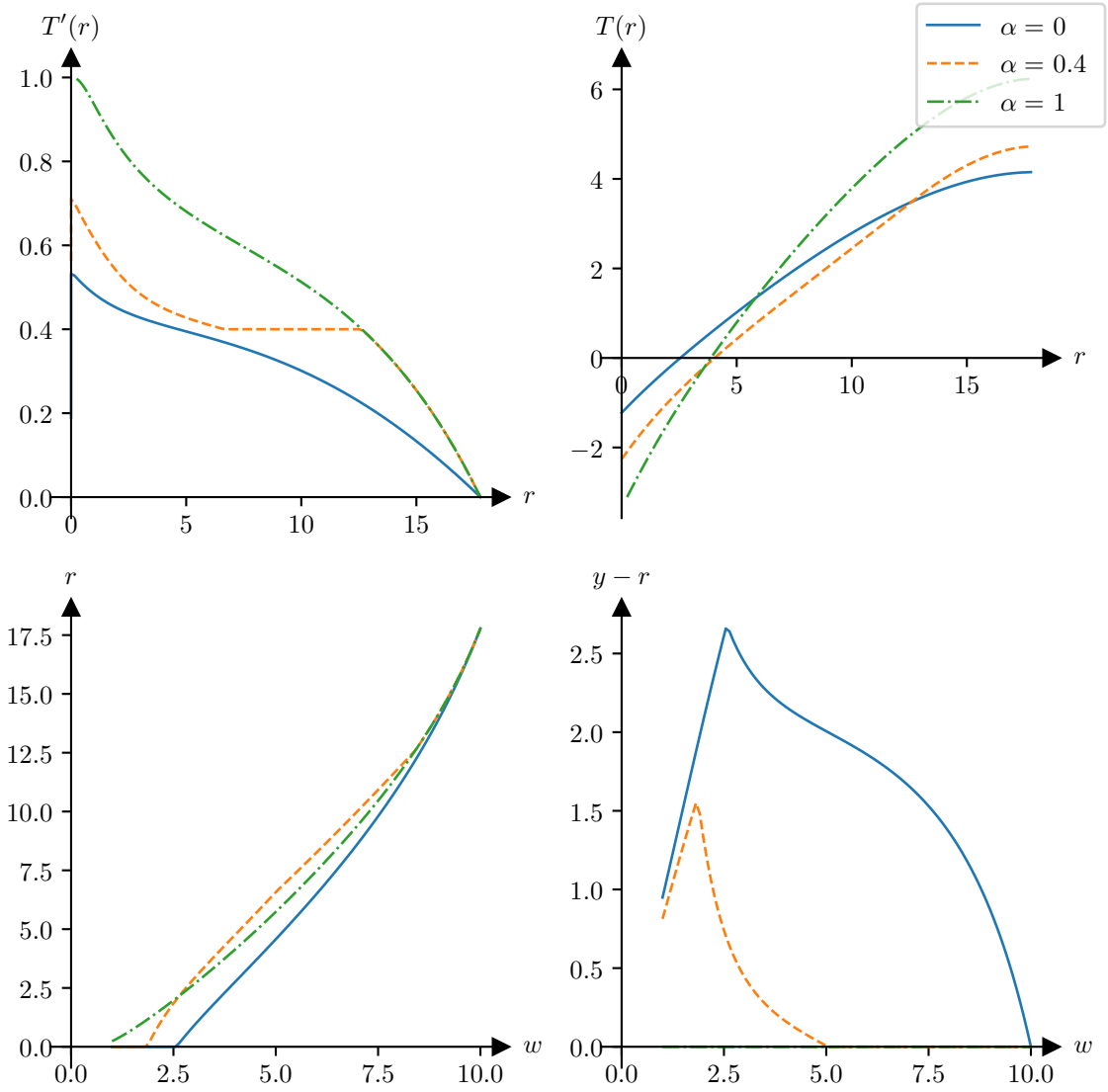

Figure 2: Optimal tax schedule with the Rawlsian criterion

observe that this may lead to corner solutions for reported incomes, individuals with a relatively low productivity declaring to have no income. Apart from these differences, we recover some properties of the optimal tax schedule already obtained in the Utilitarian case. When the marginal cost of avoiding the first euro is zero ($\alpha = 0$ in figure 2), all individuals, except those at the top, practice tax avoidance. For intermediate values of α ($\alpha = 0.4$ in the simulations), part of the individuals avoid taxes (those with the lowest productivities), whereas the others declare their true income. Finally, for a large enough marginal cost of avoidance ($\alpha = 1$ in the simulations), nobody avoids taxes and the optimal tax curve corresponds with the solution of the standard model, that has been analyzed in the Rawlsian case by, e.g., Boadway and Jacquet (2008).

8 Conclusion

We have studied the optimal labor income taxation problem when individuals both determine their labor supply and have the possibility of avoiding paying taxes. We have shown that, with a low enough marginal cost of avoidance, some taxpayers, located in the interior of the skill distribution, choose to conceal part of their income to the tax administration. This result was obtained by Casamatta (2021) in the fixed income case, and is therefore robust to the introduction of endogenous labor supply. For these individuals, the optimal marginal tax rate depends both on the elasticities of true and reported incomes. For the other individuals, located at the extremes of the skill distribution, true and reported incomes coincide.

These results were obtained with a concealment cost function assumed to be continuous and identical for all individuals. These assumptions should be relaxed in future work by considering a more general concealment technology. In particular, it is of interest to introduce a fixed cost of avoidance (Selin and Simula, 2020; Guyton et al., 2021; Doligalski and Rojas, 2022). Also, the case of marginal costs differing with the level of productivity, reflecting the fact that richer individuals may have access to better avoidance opportunities, should receive some attention. We expect that incentive constraints of more productive individuals will be tightened if they have access to a less costly avoidance technology. These latter should therefore face lower marginal tax rates and hide

more income. However, a careful investigation is needed to confirm this conjecture.

Another generalization of this work concerns the possibility of partial avoidance, with sheltered income being taxed at a lower rate than taxable income, which is typically the case when individuals resort to income shifting or income retiming (Piketty and Saez, 2013; Piketty et al., 2014; Selin and Simula, 2020).

Finally, it should be noted that governments have the possibility to reduce avoidance opportunities by modifying the tax environment. This study could therefore be completed by endogenizing the avoidance technology, in line with Kopczuk (2001) or Slemrod and Kopczuk (2002).

References

Boadway, R., K. Cuff, and M. Marchand (2000, October). Optimal Income Taxation With Quasi-Linear Preferences Revisited. *Journal of Public Economic Theory 2*(4), 435–460. 3.2

Boadway, R. and L. Jacquet (2008, November). Optimal marginal and average income taxation under maximin. *Journal of Economic Theory 143*(1), 425–441. 7

Casamatta, G. (2021). Optimal income taxation with tax avoidance. *Journal of Public Economic Theory 23*(3), 534–550. (document), 1, 2, 5, 8

Diamond, P. A. (1998). Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates. *The American Economic Review 88*(1), 83–95. 6.1

Doligalski, P. and L. E. Rojas (2022). Optimal Redistribution with a Shadow Economy. *Theoretical Economics forthcoming*. 1, 8

Ebert, U. (1992, October). A reexamination of the optimal nonlinear income tax. *Journal of Public Economics 49*(1), 47–73. 3.2, 4

Grochulski, B. (2007). Optimal Nonlinear Income Taxation with Costly Tax Avoidance. *Economic Quarterly 93*(1), 77–109. 1, 2

Guyton, J., P. Langetieg, D. Reck, M. Risch, and G. Zucman (2021, March). Tax Evasion at the Top of the Income Distribution: Theory and Evidence. Technical Report w28542, National Bureau of Economic Research, Cambridge, MA. 8

Huang, J. and J. Rios (2016, December). Optimal tax mix with income tax non-compliance. *Journal of Public Economics 144*, 52–63. 1, 3, 10

Jacquet, L. and E. Lehmann (2021, April). Optimal Income Taxation with Composition Effects. *Journal of the European Economic Association 19*(2), 1299–1341. 3.2, 11

Jacquet, L., E. Lehmann, and B. Van der Linden (2013, September). Optimal redistributive taxation with both extensive and intensive responses. *Journal of Economic Theory 148*(5), 1770–1805. 6.1, 6.2, 11

Keane, M. P. (2011, December). Labor Supply and Taxes: A Survey. *Journal of Economic Literature 49*(4), 961–1075. 1

Kleven, H. (2021, August). Sufficient Statistics Revisited. *Annual Review of Economics 13*, 515–538. 6.3

Kopczuk, W. (2001, July). Redistribution when avoidance behavior is heterogeneous. *Journal of Public Economics 81*(1), 51–71. 1, 8

Landier, A. and G. Plantin (2017, July). Taxing the Rich. *Review of Economic Studies 84*(3), 1186–1209. 2

Lollivier, S. and J.-C. Rochet (1983, December). Bunching and second-order conditions: A note on optimal tax theory. *Journal of Economic Theory 31*(2), 392–400. 3.2

Mirrlees, J. A. (1971, April). An Exploration in the Theory of Optimum Income Taxation. *The Review of Economic Studies 38*(2), 175–208. 1, 3.2

Mirrlees, J. A. (1976). Optimal Tax Theory: A Synthesis. *Journal of Public Economics 6*, 327–358. 4

Piketty, T. and E. Saez (2013). Optimal Labor Income Taxation. In *Handbook of Public Economics*, Volume 5, pp. 391–474. Elsevier. 1, 3.2, 8

Piketty, T., E. Saez, and S. Stantcheva (2014, February). Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities. *American Economic Journal: Economic Policy 6*(1), 230–271. 1, 3, 8

Saez, E. (2001). Using Elasticities to Derive Optimal Income Tax Rates. *The Review of Economic Studies 68*(1), 205–229. 6, 6.1, 6.3, 6.3, 12, D, D

Saez, E., J. Slemrod, and S. H. Giertz (2012). The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review. *Journal of Economic Literature 50*(1), 3–50. 1

Seade, J. K. (1977, April). On the shape of optimal tax schedules. *Journal of Public Economics 7*(2), 203–235. 5, 6.1

Selin, H. and L. Simula (2020, February). Income shifting as income creation? *Journal of Public Economics 182*, 104081. 1, 8

Slemrod, J. (1995, August). High-Income Families and the Tax Changes of the 1980s: The Anatomy of Behavioral Response. Technical Report w5218, National Bureau of Economic Research, Cambridge, MA. 1

Slemrod, J. and W. Kopczuk (2002, April). The optimal elasticity of taxable income. *Journal of Public Economics 84*(1), 91–112. 1, 8

Werning, I. (2000). An Elementary Proof of Positive Optimal Marginal Taxes. B

Appendix

A Proof of Lemma 1

Two cases are possible. In the first one, $y(w) = \hat{y}(w, w)$ is an interior solution of program (1) (it is different from $r(w)$). The first-order condition of program (8) then writes:

$$(-T'(w) + r'(w)\phi')u_c + \left.\frac{\partial \hat{y}}{\partial w'}\right|_{(w,w)} ((1-\phi')u_c + \frac{1}{w}u_l) = 0. \tag{29}$$

From the first-order condition (6), the second term in (29) vanishes, so that this condition becomes:

$$-T'(w) + r'(w)\phi'(y(w) - r(w)) = 0. \tag{30}$$

In the second case, $y(w)$ is a corner solution: $y(w) = r(w)$. In such a case, the first-order condition of program (8) is:

$$-T'(w)u_c + \left.\frac{\partial \hat{y}}{\partial w'}\right|_{(w,w)} (u_c + \frac{1}{w}u_l) = 0. \tag{31}$$

Observing that $\partial\hat{y}/\partial w'\big|_{(w,w)} = r'(w)$, condition (31) becomes:

$$(-T'(w) + r'(w))u_c + r'(w)\frac{1}{w}u_l = 0. \tag{32}$$

Let us differentiate (2) with respect to w, again distinguishing whether $y(w)$ is an interior solution or not. In the first case, we have:

$$\frac{dU}{dw} = y'(w)((1-\phi')u_c + \frac{1}{w}u_l) + (-T'(w) + r'(w)\phi')u_c - \frac{y(w)}{w^2}u_l.$$

Condition (6) then implies that the term $(1-\phi')u_c + (1/w)u_l$ on the right-hand side of the above expression is equal to 0. In the case of a corner solution:

$$\frac{dU}{dw} = y'(w)(u_c + \frac{1}{w}u_l) - T'(w)u_c - \frac{y(w)}{w^2}u_l,$$

where we note that $y'(w) = r'(w)$. Using (30) and (32), respectively, we obtain that in both cases:

$$\frac{dU}{dw} = -\frac{y(w)}{w^2}u_l.$$

B Proof of Proposition 1

1. We evaluate the derivative of the Hamiltonian when $r(w)$ is larger than $y(w)$. From (6), we have that $1 - \phi' + (1/w)(u_l/u_c) = 0$ when $r(w) \neq y(w)$. Therefore, (12) implies:

$$\left.\frac{\partial \mathcal{H}}{\partial r}\right|_{r(w)>y(w)} = \mu\phi'_- f(w) + \lambda(w) \left.\frac{dy}{dr}\right|_{r(w)>y(w)} (-\frac{1}{w^2}u_l + \frac{y(w)}{w^3}(u_{cl}\frac{u_l}{u_c} - u_{ll})).$$

We show that this expression is strictly negative.

First, differentiating (6), we have:

$$\frac{dy}{dr} = \frac{\phi'' u_c + (1-\phi')\phi' u_{cc} + \frac{1}{w}\phi' u_{cl}}{-D_y},$$

where D_y is the derivative of (6) with respect to y. Assuming that the second-order condition of program (1) is satisfied, D_y is negative. We use (6) to rewrite the numerator, evaluated at $r(w) > y(w)$, as:

$$\phi'' u_c + \frac{1}{w}\phi'_-(-\frac{u_l}{u_c}u_{cc} + u_{cl}).$$

Indifference curves are assumed to be increasing and convex in the (c, y) space. They are therefore increasing and concave in the (y, c) space, implying:

$$\left.\frac{d^2 y}{dc^2}\right|_u < 0$$
$$\Leftrightarrow -\frac{u_l}{u_c}u_{cc} + u_{cl} < 0.$$

Recalling that, by assumption, $\phi'_- < 0$ and $\phi'' > 0$, we obtain $\partial y/\partial r\big|_{r(w)>y(w)} > 0$.

Second, convexity of the indifference curves in the (c, y) space implies $u_{cl}u_l/u_c - u_{ll} > 0$.

Recalling that $u_l < 0$, this implies that the last term between parenthesis is positive.

Finally, we prove below that $\lambda(w) \leq 0, \forall w$. Combined with the full support assumption ($f(w) > 0 \ \forall w \in [w_-, w_+]$), it follows that $\partial\mathcal{H}/\partial r\big|_{r(w)>y(w)} < 0$, meaning that the social planner has no incentive to increase $r(w)$ above $y(w)$.

2. Let us write the derivative of the Hamiltonian with respect to reported income, when this latter tends to true income from below:

$$\left.\frac{\partial\mathcal{H}}{\partial r}\right|_{r(w)\nearrow y(w)} = \mu\phi'_+(0)f(w) + \lambda(w)\left.\frac{dy}{dr}\right|_{r(w)\nearrow y(w)}(-\frac{1}{w^2}u_l + \frac{y(w)}{w^3}(u_{cl}\frac{u_l}{u_c} - u_{ll})). \tag{33}$$

Under the full support assumption ($f(w) > 0$) and assuming that $\phi'_+(0) > 0$, the first term in the right-hand side is strictly positive. Moreover, recalling that $\lambda(w_-) = \lambda(w_+) = 0$, the second term vanishes when w is close to w_- or w_+. Therefore, $\partial\mathcal{H}/\partial r\big|_{r(w)\nearrow y(w)} > 0$ when $w \to w_-$ or $w \to w_+$, which implies that there exist $w' > w_-$ and $w'' < w_+$ such that $r(w) = y(w), \forall w \in [w_-, w'] \cup [w'', w_+]$.

Observe that this result ($r(w) = y(w)$) is true for all w such that $\lambda(w) = 0$. Noting that $dy/dr = 1$ when $r(w) = y(w)$, (12) implies:

$$-\frac{1}{w}\frac{u_l}{u_c} = 1 \tag{34}$$

whenever $\lambda(w) = 0$. This allows to us to prove, following the argument developed by Werning (2000), that $\lambda(w) < 0 \ \forall w \in (w_-, w_+)$, and thus, recalling that $\lambda(w_-) = \lambda(w_+) = 0$, that $\lambda(w) \leq 0 \ \forall w \in [w_-, w_+]$. if $\lambda(w) > 0$ for some $w \in (w_-, w_+)$, there must be some maximal interval $[w^0, w^1]$ with $w^0 < w^1$ such that $\lambda(w) \geq 0$ for $w \in [w^0, w^1]$, with $\lambda(w^0) = \lambda(w^1) = 0$ and $\lambda'(w^0) \geq 0$ and $\lambda'(w^1) \leq 0$. We note that (9) implies $U(w^0) \leq U(w^1)$. Therefore, recalling that G is a concave function, we have that $G'(U(w^0)) \geq G'(U(w^0))$. Using (13), we then obtain:

$$\frac{1}{u_c(c(w^0), l(w^0))} = \frac{1}{\mu}(G'(U(w^0)) + \frac{\lambda'(w^0)}{f(w^0)}) \geq \frac{1}{\mu}(G'(U(w^1)) + \frac{\lambda'(w^1)}{f(w^1)}) = \frac{1}{u_c(c(w^1), l(w^1))},$$

The rest of the proof, that consists in showing that this inequality cannot be true, provided that

leisure is a normal good and that (34) is satisfied, follows directly from Werning (2000).

3. We prove that $\partial\mathcal{H}/\partial r\big|_{r(w)\nearrow y(w)} < 0$ when $\phi'_+(0) = 0$ and $w \in (w_-, w_+)$.

When $\phi'_+(0) = 0$, the first term on the right-hand side of (33) vanishes. Moreover, from the discussion above (sentence that starts with "Second, convexity of the indifference curves ..." in point 1.), the last term between parenthesis is (strictly) positive. Because λ is strictly negative for individuals with $w \in (w_-, w_+)$, it remains to be shown that $dy/dr\big|_{r(w)\nearrow y(w)}$ is positive:

$$\left.\frac{dy}{dr}\right|_{r(w)\nearrow y(w)} = \frac{\phi''_+(0)u_c + \frac{1}{w}\phi'_+(0)(-\frac{u_l}{u_c}u_{cc} + u_{cl})}{-D_y > 0}.$$

Recalling that $\phi''_+(0) > 0$, this expression is strictly positive when $\phi'_+(0) = 0$.

If the optimal allocation $\{c(w), l(w)\}$ and the multiplier $\lambda(w)$ are continuous with respect to $\phi'_+(0)$, the above reasoning implies that there exist individuals with $w \in (w_-, w_+)$ such that $\partial\mathcal{H}/\partial r\big|_{r(w)\nearrow y(w)} < 0$ when $\phi'_+(0) \to 0$. Therefore, these individuals avoid taxes at the tax optimum.

C Proof of Lemma 2

We consider a linearized tax schedule, such that $c = y - \tau r - \phi(y - r) + I$ with τ independent of r,[13] and differentiate conditions (17) and (18) with respect to y, r, $1 - \tau$ and I:

$$dy\frac{\Omega}{w^2} + d(1-\tau)(u_c + r(1-\tau)u_{cc} + \frac{r}{w}u_{cl}) + dI((1-\tau)u_{cc} + \frac{1}{w}u_{cl}) = 0 \tag{35}$$

$$d(1-\tau) + dy\phi'' - dr\phi'' = 0, \tag{36}$$

where $\Omega \equiv u_{ll} + w^2(1-\tau)^2u_{cc} + 2w(1-\tau)u_{cl}$. Re-arranging these expressions leads to:

$$\begin{aligned} e_y &= -\frac{1-\tau}{y}\frac{w^2(u_c + r(1-\tau)u_{cc} + \frac{r}{w}u_{cl})}{\Omega} \\ e_r &= \frac{1}{\phi''}\frac{1-\tau}{r} + \frac{y}{r}e_y = \frac{1-\tau}{r}\frac{1}{\phi''}(1 - \phi''\frac{w^2(u_c + r(1-\tau)u_{cc} + \frac{r}{w}u_{cl})}{\Omega}), \end{aligned}$$

[13] I represents exogenous income.

where $e_y \equiv (dy/d(1-\tau))((1-\tau)/y)$ and $e_r \equiv (dr/d(1-\tau))((1-\tau)/r)$ are the elasticities of the true and reported incomes with respect to the net-of-tax rate, respectively.

We also determine compensated elasticities. In this purpose, we first compute, using (35) and (36), the response of labor income (true and declared) to a change in exogenous income I:

$$\begin{aligned} \frac{dy}{dI} &= -\frac{w^2(1-\tau)u_{cc} + wu_{cl}}{\Omega} \\ \frac{dr}{dI} &= \frac{dy}{dI}. \end{aligned} \tag{37}$$

We then use the Slutsky equation (expressed in elasticity form):

$$e^c_\varpi = e_\varpi - \eta_\varpi, \ \varpi = y, r \tag{38}$$

where $\eta_\varpi \equiv (1-\tau)(d\varpi/dI)$, to obtain:

$$\begin{aligned} e^c_y &= \frac{1-\tau}{y}\frac{-w^2u_c + (y-r)(w^2(1-\tau)u_{cc} + wu_{cl})}{\Omega} \\ e^c_r &= \frac{1-\tau}{r}\frac{1}{\phi''}(1 - \phi''\frac{w^2u_c}{\Omega}). \end{aligned}$$

D Proof of Proposition 2

When there exists an interior solution for r, the optimal marginal tax rate is obtained by re-arranging the first-order conditions (12) and (14) of the social planner program, which correspond respectively to the case of an interior solution for y ($y(w) > r(w)$) and of a corner solution ($y(w) = r(w)$).

It should, however, be noted that the Hamiltonian function is not differentiable at the point where individuals are just indifferent between avoiding or not, that is when:

$$MRS_{c,l} = -\frac{1}{w}\frac{u_l}{u_c} = 1 - \phi'_+(0). \tag{39}$$

To see this, we rewrite the derivative of the Hamiltonian, as expressed in (12), when r tends to y

from below:

$$\left.\frac{\partial \mathcal{H}}{\partial r}\right|_{r \nearrow y} = \mu(\phi'_+(0) + \frac{dy}{dr}(1 - \phi'_+(0) + \frac{1}{w}\frac{u_l}{u_c}))f(w) + \lambda(w)\frac{dy}{dr}(-\frac{1}{w^2}u_l + \frac{y(w)}{w^3}(u_{cl}\frac{u_l}{u_c} - u_{ll}))$$
$$= \mu\phi'_+(0)f(w) + \lambda(w)\frac{dy}{dr}(-\frac{1}{w^2}u_l + \frac{y(w)}{w^3}(u_{cl}\frac{u_l}{u_c} - u_{ll})),$$

where the second equality follows from (39).

It is shown below that:

$$\frac{dy}{dr} = \frac{re^c_r - re_r + ye_y}{re^c_r}. \tag{40}$$

Therefore, using (23) and (24):

$$\left.\frac{dy}{dr}\right|_{r \nearrow y} = \frac{e^c_y}{\frac{1}{\phi''_+(0)}\frac{1-\tau}{y} + e^c_y} < 1 = \left.\frac{dy}{dr}\right|_{r=y}.$$

Using (14), and reminding that $\lambda(w) < 0 \; \forall w \in (w_-, w_+)$, $u_l < 0$ and, because of the strict convexity of indifference curves, $u_{cl}u_l/u_c - u_{ll} > 0$, we obtain that:

$$\left.\frac{\partial \mathcal{H}}{\partial r}\right|_{r \nearrow y} < \left.\frac{\partial \mathcal{H}}{\partial r}\right|_{r=y}.$$

Three cases are then possible. If $\partial\mathcal{H}/\partial r\big|_{r \nearrow y} < 0$, the solution involves $r(w) < y(w)$ and the formula for the optimal marginal tax rate follows from (12). If $\partial\mathcal{H}/\partial r\big|_{r \nearrow y} \geq 0$ and $\partial\mathcal{H}/\partial r\big|_{r=y} > 0$, the optimal solution involves $r(w) = y(w)$ and the formula for the optimal marginal tax rate follows from (14). Finally, if $\partial\mathcal{H}/\partial r\big|_{r \nearrow y} \geq 0$ and $\partial\mathcal{H}/\partial r\big|_{r=y} \leq 0$, the optimal solution involves $r(w) = y(w)$ and the optimal marginal tax rate is equal to $\phi'_+(0)$, as it is the tax rate that satisfies (17) and (39).

We now analyze the first two cases, in which the optimal tax rate is derived from conditions (14) and (12), respectively.

When individuals adopt a corner solution, $r = y$, the problem faced by the social planner is the same as in the standard Mirrlees model. Re-arranging condition (14), we arrive at the usual formula (Saez, 2001). It should be noted that *the optimal tax rate cannot be larger than* $\phi'_+(0)$,

otherwise individuals would earn a level of income y strictly larger than r. This follows from the resolution of program (16). In the case of a corner solution $y = r$, the first-order condition on y implies:

$$\left.\frac{\partial \mathcal{U}}{\partial y}\right|_{r \nearrow y} = (1 - T_r'(r))u_c + \frac{1}{w}u_l > 0,$$

where $\mathcal{U}$ is the objective function of program (16). A marginal tax rate $T_r'(r)$ larger than $\phi_+'(0)$ would therefore imply a contradiction with (7).

We now turn to the case of an interior solution, $r < y$.

1. We first prove that:

$$\frac{y}{w^2}(u_{cl}\frac{u_l}{u_c} - u_{ll})) = (1-\tau)u_c \frac{1 + e_y - e_y^c}{\frac{r}{y}e_r^c - \frac{r}{y}e_r + e_y}. \tag{41}$$

From (19):

$$\begin{aligned} 1 + e_y &= \frac{u_{ll} + w^2(1-\tau)^2 u_{cc} + 2w(1-\tau)u_{cl} - w^2(1-\tau)\frac{u_c}{y} - \frac{r}{y}w^2(1-\tau)^2 u_{cc} - \frac{r}{y}w(1-\tau)u_{cl}}{\Omega} \\ &= \frac{u_{ll} + (1-\frac{r}{y})w^2(1-\tau)^2 u_{cc} + (2-\frac{r}{y})w(1-\tau)u_{cl} - w^2(1-\tau)\frac{u_c}{y}}{\Omega}. \end{aligned}$$

Using (20):

$$\begin{aligned} 1 + e_y - e_y^c &= \frac{u_{ll} + (1-\frac{r}{y})w^2(1-\tau)^2 u_{cc} + (2-\frac{r}{y})w(1-\tau)u_{cl} - w^2(1-\tau)\frac{u_c}{y}}{\Omega} \\ &\quad - \frac{-w^2(1-\tau)\frac{u_c}{y} + (1-\frac{r}{y})(w^2(1-\tau)^2 u_{cc} + w(1-\tau)u_{cl})}{\Omega} \\ &= \frac{u_{ll} + w(1-\tau)u_{cl}}{\Omega}. \end{aligned} \tag{42}$$

Using (21) and (22):

$$\begin{aligned} \frac{r}{y}e_r^c - \frac{r}{y}e_r + e_y &= \frac{r}{y}e_r^c - \frac{1}{\phi''}\frac{1-\tau}{y} \\ &= -\frac{w^2 u_c}{\Omega}\frac{1-\tau}{y}. \end{aligned} \tag{43}$$

Combining (42) and (43):

$$\begin{aligned}\frac{1+e_y-e_y^c}{\frac{r}{y}e_r^c-\frac{r}{y}e_r+e_y} &= \frac{u_{ll}+w(1-\tau)u_{cl}}{-w^2u_c}\frac{y}{1-\tau}\\ &= \frac{\frac{u_l}{u_c}u_{cl}-u_{ll}}{w^2u_c}\frac{y}{1-\tau}.\end{aligned}$$

Re-arranging the last expression leads to (41).

2. We then prove that:

$$\frac{dy}{dr} = \frac{re_r^c - re_r + ye_y}{re_r^c}. \tag{44}$$

We differentiate (17) and (18) with respect to y and r:

$$\begin{aligned}dy\frac{\Omega}{w^2} - drT_r''u_c &= 0\\ dy\phi'' - dr(\phi''+T_r'') &= 0.\end{aligned}$$

This leads to:

$$\begin{aligned}\frac{dy}{dr} &= T_r''\frac{w^2u_c}{\Omega}\\ \frac{dy}{dr} &= 1+\frac{T_r''}{\phi''}.\end{aligned} \tag{45}$$

Using (43):

$$\frac{dy}{dr} = -T_r''\frac{re_r^c - re_r + ye_y}{1-\tau}.$$

Using (21):

$$\begin{aligned}\frac{1-\tau}{\phi''} &= re_r - ye_y\\ \Leftrightarrow \frac{T_r''}{1-\tau} &= \frac{T_r''}{\phi''}\frac{1}{re_r-ye_y}.\end{aligned}$$

Therefore:

$$\frac{dy}{dr} = -\frac{T_r''}{\phi''}\frac{re_r^c - re_r + ye_y}{re_r - ye_y}. \tag{46}$$

From (18):

$$\frac{T_r''}{\phi''} = \frac{dy}{dr} - 1.$$

Substituting into (46):

$$\begin{aligned} \frac{dy}{dr} &= (1 - \frac{dy}{dr})\frac{re_r^c - re_r + ye_y}{re_r - ye_y}. \\ \Rightarrow \frac{dy}{dr} &= \frac{re_r^c - re_r + ye_y}{re_r^c}. \end{aligned}$$

3. We show that:

$$\frac{T_r'(r(w))}{1 - T_r'(r(w))} = -\frac{u_c \frac{\lambda(w)}{\mu}}{1 - F(w)} \frac{1 - F(w)}{wf(w)} (\frac{y(1 + e_y) + (y - r)(e_r - e_r^c)}{re_r^c}). \tag{47}$$

From (17):

$$-\frac{1}{w} u_l = (1 - \tau)u_c. \tag{48}$$

Starting from (25), and using (41), (44), and (48), we arrive at:

$$\frac{T_r'(r(w))}{1 - T_r'(r(w))} = -\frac{\lambda(w)}{\mu} \frac{u_c}{wf(w)} \frac{re_r^c - re_r + ye_y}{re_r^c} (1 + \frac{1 + e_y - e_y^c}{\frac{r}{y}e_r^c - \frac{r}{y}e_r + e_y}).$$

Moreover, it can be shown easily, using (38) and (37), that

$$e_y - e_y^c = e_r - e_r^c. \tag{49}$$

The previous expression then becomes:

$$\frac{T_r'(r(w))}{1 - T_r'(r(w))} = -\frac{\lambda(w)}{\mu} \frac{u_c}{wf(w)} \frac{y(1 + e_y) + (y - r)(e_r - e_r^c)}{re_r^c}.$$

4. Adapting the method in Saez (2001), we show that:

$$u_c \frac{\lambda(w)}{\mu} = \int_w^{w_+} \left(\frac{G'(U(m))u_c(m)}{\mu} - 1\right) \exp\left[\int_w^m \frac{e_r^c - e_r}{\frac{r}{y}e_r^c - \frac{r}{y}e_r + e_y} \frac{y'(s)}{y(s)} ds\right] f(m) dm$$

From (15),

$$u_c \frac{\lambda(w)}{\mu} = \int_w^{w_+} \left(\frac{G'(U(m))u_c(m)}{\mu} - 1\right) \frac{u_c(w)}{u_c(m)} T_{wm} f(m) dm$$

where

$$T_{wm} \equiv \exp\left(-\int_w^m (y(s)u_{cl})/(s^2 u_c) ds\right).$$

It follows that:

$$\frac{u_c(w)}{u_c(m)} T_{wm} = \exp\left[\int_w^m J(s) ds\right],$$

where

$$J(s) \equiv -\frac{du_c/ds}{u_c} - \frac{y(s)u_{cl}}{s^2 u_c}.$$

We differentiate $U(s) = u(c(s), l(s))$:

$$\frac{dU}{ds} = c'(s)u_c + l'(s)u_l.$$

Using (9), this yields:

$$c'(s)u_c + l'(s)u_l = -\frac{y(s)}{s^2} u_l. \tag{50}$$

Moreover,

$$\frac{du_c}{ds} = c'(s)u_{cc} + l'(s)u_{cl}.$$

Substituting $c'(s)$ from (50) into this equation, and recalling that $y(s) = sl(s)$, we obtain:

$$\frac{du_c}{ds} = -\left(l(s) + sl'(s)\right) \frac{u_l u_{cc}}{s u_c} + l'(s)u_{cl}.$$

It follows that

$$J(s) = \left(l(s) + sl'(s)\right)\frac{u_l u_{cc}}{su_c^2} - \frac{sl'(s)u_{cl}}{su_c} - \frac{l(s)u_{cl}}{su_c}$$
$$= \left(\frac{l(s) + sl'(s)}{sl(s)}\right)\left(\frac{l(s)u_l u_{cc}}{u_c^2} - \frac{l(s)u_{cl}}{u_c}\right).$$

From (21) and (22), we have:

$$e_r^c - e_r = \frac{1-\tau}{\Omega}\left(w^2(1-\tau)u_{cc} + wu_{cl}\right).$$

Condition (17) implies:

$$1 - \tau = -\frac{1}{w}\frac{u_l}{u_c}.$$

Substituting into the previous expression:

$$e_r^c - e_r = \frac{1-\tau}{\Omega}\left(-wu_{cc}\frac{u_l}{u_c} + wu_{cl}\right).$$

We then use (43), to obtain:

$$\frac{e_r^c - e_r}{\frac{r}{y}e_r^c - \frac{r}{y}e_r + e_y} = \frac{lu_l u_{cc}}{u_c^2} - \frac{lu_{cl}}{u_c}.$$

Therefore:

$$J(s) = \frac{y'(s)}{y(s)}\frac{e_r^c - e_r}{\frac{r}{y}e_r^c - \frac{r}{y}e_r + e_y}.$$

This implies:

$$u_c\frac{\lambda(w)}{\mu} = \int_w^{w_+}\left(\frac{G'(U(m))u_c(m)}{\mu} - 1\right)\exp\left[\int_w^m \frac{e_r^c - e_r}{\frac{r}{y}e_r^c - \frac{r}{y}e_r + e_y}\frac{1}{y(s)}dy(s)\right]f(m)dm$$
$$= \int_w^{w_+}\left(\frac{G'(U(m))u_c(m)}{\mu} - 1\right)\exp\left[\int_w^m \left(1 - \frac{e_r}{e_r^c}\right)\frac{dr(s)}{r(s)}\right]f(m)dm,$$

where the second equality follows from (44).

5. The optimal marginal tax rate faced by avoiding individuals ($r(w) < y(w)$) is larger than $\phi'_+(0)$.

This follows readily from the first-order condition (18). The marginal tax rate being equal to $\phi'(y(w) - r(w))$ and the function ϕ being strictly convex, it follows that the marginal tax rate is larger than $\phi'_+(0)$.

E Proof of Proposition 3

We distinguish between the cases $y > r$ and $y = r$.

1. $y > r$

Let us differentiate the first-order conditions (17) and (18) with respect to y, r, $1-\tau$ and w:

$$dy\frac{\Omega}{w^2} - drT''_r u_c + d(1-\tau)u_c + dw(-(1-\tau)\frac{y}{w^2}u_{cl} - \frac{1}{w^2}u_l - \frac{y}{w^3}u_{ll}) = 0 \tag{51}$$

$$dy\phi'' - dr(\phi'' + T''_r) + d(1-\tau) = 0. \tag{52}$$

Substituting the second condition into the first, we obtain:

$$dr\frac{\phi'' + T''_r}{\phi''}\frac{\Omega}{w^2} - drT''_r u_c + d(1-\tau)(u_c - \frac{\Omega}{\phi'' w^2}) - dw((1-\tau)\frac{y}{w^2}u_{cl} + \frac{1}{w^2}u_l + \frac{y}{w^3}u_{ll}) = 0.$$

This implies:

$$\frac{dr}{dw} = \frac{(1-\tau)yu_{cl} + u_l + \frac{y}{w}u_{ll}}{\frac{\phi''+T''_r}{\phi''}\Omega - T''_r w^2 u_c} = \frac{-\frac{y}{w}u_{cl}\frac{u_l}{u_c} - w(1-\tau)u_c + \frac{y}{w}u_{ll}}{\frac{\phi''+T''_r}{\phi''}\Omega - T''_r w^2 u_c},$$

$$\frac{dr}{d(1-\tau)} = \frac{-w^2(u_c - \frac{\Omega}{\phi'' w^2})}{\frac{\phi''+T''_r}{\phi''}\Omega - T''_r w^2 u_c},$$

where the last equality in the first equation follows from (17).

Using (41) and (49), we obtain:

$$\frac{dr}{dw} = \frac{-w(1-\tau)u_c \frac{1+e_y-e_y^c}{\frac{r}{y}e_r^c - \frac{r}{y}e_r + e_y} - w(1-\tau)u_c}{\frac{\phi''+T_r''}{\phi''}\Omega - T_r'' w^2 u_c}$$
$$= \frac{-w(1-\tau)u_c(\frac{y(1+e_y)+(y-r)(e_r-e_r^c)}{re_r^c+ye_y-re_r})}{\frac{\phi''+T_r''}{\phi''}\Omega - T_r'' w^2 u_c}. \tag{53}$$

Moreover, from (43), we have:

$$\frac{\Omega}{w^2} = -\frac{u_c(1-\tau)}{re_r^c - re_r + ye_y},$$

and, from (21):

$$\frac{1-\tau}{\phi''} = re_r - ye_y.$$

These two conditions imply:

$$\frac{dr}{d(1-\tau)} = \frac{-w^2(u_c + u_c\frac{re_r-ye_y}{re_r^c-re_r+ye_y})}{\frac{\phi''+T_r''}{\phi''}\Omega - T_r'' w^2 u_c} = \frac{-w^2 u_c(\frac{re_r^c}{re_r^c-re_r+ye_y})}{\frac{\phi''+T_r''}{\phi''}\Omega - T_r'' w^2 u_c}. \tag{54}$$

We then use (22):

$$e_r^c = \frac{1-\tau}{r}\frac{1}{\phi''} - \frac{1-\tau}{r}\frac{w^2 u_c}{\Omega},$$

which implies

$$T_r'' w^2 u_c = T_r''\Omega(\frac{1}{\phi''} - \frac{re_r^c}{1-\tau})$$

and thus

$$\frac{\phi''+T_r''}{\phi''}\Omega - T_r'' w^2 u_c = \Omega(1 + \frac{T_r''}{1-\tau}re_r^c). \tag{55}$$

Using (43), we also obtain:

$$w(1-\tau)u_c = -\frac{\Omega}{w}(re_r^c + ye_y - re_r). \tag{56}$$

Combining (55) and (56) leads to:

$$\frac{w(1-\tau)u_c}{\frac{\phi''+T_r''}{\phi''}\Omega - T_r''w^2u_c} = -\frac{re_r^c + ye_y - re_r}{w(1+\frac{T_r''}{1-\tau}re_r^c)},$$

which, from (53) and (54), implies:

$$\frac{dr}{dw} = \frac{y(1+e_y)+(y-r)(e_r-e_r^c)}{w(1+\frac{T_r''}{1-\tau}re_r^c)}$$
$$\frac{dr}{d(1-\tau)} = \frac{(1-\tau)re_r^c}{(1+\frac{T_r''}{1-\tau}re_r^c)}.$$

It follows that:

$$\gamma(w) = \frac{y(1+e_y)+(y-r)(e_r-e_r^c)}{r(1+\frac{T_r''}{1-\tau}re_r^c)}$$

and

$$\varepsilon(w) = \frac{e_r^c}{1+\frac{T_r''}{1-\tau}re_r^c}.$$

Therefore:

$$\frac{\gamma(w)}{\varepsilon(w)} = \frac{y(1+e_y)+(y-r)(e_r-e_r^c)}{re_r^c} = A(w).$$

2. $y = r$

Let us differentiate the first-order conditions (17) with respect to r, $1-\tau$ and w:

$$dr(\frac{\Omega}{w^2} - T_r''u_c) + d(1-\tau)u_c + dw(-(1-\tau)\frac{r}{w^2}u_{cl} - \frac{1}{w^2}u_l - \frac{y}{w^3}u_{ll}) = 0.$$

This implies:

$$\frac{dr}{dw} = \frac{(1-\tau)ru_{cl}+u_l+\frac{r}{w}u_{ll}}{\Omega - T_r''w^2u_c} = \frac{-\frac{r}{w}u_{cl}\frac{u_l}{u_c} - w(1-\tau)u_c + \frac{r}{w}u_{ll}}{\Omega - T_r''w^2u_c},$$
$$\frac{dr}{d(1-\tau)} = \frac{-w^2u_c}{\Omega - T_r''w^2u_c},$$

where the last equality in the first equation follows from (17).

Using (41) and (49), we obtain:

$$\frac{dr}{dw} = \frac{-w(1-\tau)u_c(\frac{r(1+e_r)}{re_r^c})}{\Omega - T_r''w^2u_c}.$$

Moreover, from (43), we have:

$$\frac{\Omega}{w^2} = -\frac{u_c(1-\tau)}{re_r^c}.$$

This implies:

$$\frac{dr}{dw} = \frac{r(1+e_r)}{w(1+\frac{T_r''re_r^c}{1-\tau})},$$
$$\frac{dr}{d(1-\tau)} = \frac{(1-\tau)re_r^c}{(1+\frac{T_r''}{1-\tau}re_r^c)}.$$

It follows that:

$$\gamma(w) = \frac{1+e_r}{1+\frac{T_r''}{1-\tau}re_r^c}$$

and

$$\varepsilon(w) = \frac{e_r^c}{1+\frac{T_r''}{1-\tau}re_r^c}.$$

Therefore:

$$\frac{\gamma(w)}{\varepsilon(w)} = \frac{1+e_r}{e_r^c} = A(w).$$

F Proof of Proposition 4

Using the fact that:

$$h(r(w))dr = f(w)dw \Rightarrow f(w) = \frac{dr}{dw}h(r),$$

the term $B(w)$ can be rewritten as:

$$\begin{aligned} B(w) &= \frac{1-H(r)}{w\frac{dr}{dw}h(r)} \\ &= \frac{1-H(r)}{h(r)} \frac{1+\frac{T_r''}{1-\tau}re_r^c}{y(1+e_y)+(y-r)(e_r-e_r^c)}. \end{aligned}$$

This leads to:

$$A(w)B(w) = \frac{1+\frac{T_r''}{1-\tau}re_r^c}{re_r^c}\frac{1-H(r)}{h(r)},$$

and therefore, from (26):

$$\begin{aligned} \frac{T_r'(r(w))}{1-T_r'(r(w))} &= \frac{1+\frac{T_r''}{1-\tau}re_r^c}{re_r^c}\frac{1-H(r)}{h(r)}C(w) \\ \Rightarrow \frac{T_r'}{1-T_r'+re_r^cT_r''} &= \frac{1}{e_r^c}\frac{1-H(r)}{rh(r)}C(w). \end{aligned}$$

Noting finally that $C(w)$ can be written as:

$$\int_{r^*}^{r+}\left(1-\frac{G'(U)u_c}{\mu}\right)\exp\left[\int_{r^*}^{r}\left(1-\frac{e_r}{e_r^c}\right)\frac{dr'}{r'}\right]\frac{h(r)}{1-H(r^*)}dr,$$

we have demonstrated that the optimal marginal tax rate satisfies (28).