DISCUSSION PAPER SERIES

No. 8608

OPTIMAL INCOME TAXATION WITH TAX AVOIDANCE

Georges Casamatta

PUBLIC POLICY

ISSN 0265-8003

OPTIMAL INCOME TAXATION WITH TAX AVOIDANCE

Georges Casamatta, Toulouse School of Economics and CEPR

Discussion Paper No. 8608
October 2011

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 8608

October 2011

ABSTRACT

Optimal income taxation with tax avoidance*

We follow the approach of Grochulski (2007), who determines the optimal income tax schedule when individuals have the possibility of avoiding paying taxes. We however modify his setup by considering a convex concealment cost function. This assumption violates the subadditivity property used in Grochulski (2007) and this has strong implications for the design of the tax schedule. This latter indeed shows that, with subadditivity, all individuals should declare their true income. Tax avoidance is thus not optimal. With a convex cost function, we fi nd that a subset of individuals, located in the interior of the income distribution, should be allowed to avoid taxes, provided that the marginal cost of avoiding the fi rst euro is su-ciently small. We also provide a characterization of the optimal income tax curve.

JEL Classification: H21
Keywords: fiscal avoidance and optimal income tax

Georges Casamatta
Toulouse School of Economics
Manufacture des Tabacs (bât. F)
21, Allée de Brienne
31000 Toulouse
France

Email:
georges.casamatta@univ-tlse1.fr

* I thank Helmuth Cremer, Laurence Jacquet and Marie-Louise Leroux for their comments, as well as seminar participants in Toulouse (Political Economy and Public Economics Workshop), Bastia (ECOPA) and Marseille (Xèmes journées Louis-André Gérard-Varet).

Submitted 07 October 2011

1 Introduction

Individual responses to taxation can be classified into two broad categories. On the one hand, individuals react to taxation by changing arguments of the utility function, i.e. leisure and other goods and services. Slemrod (1995) names this effect the *real* response to taxation. Conceptually distinct from real substitution responses are efforts to reduce one's tax liability without modifying economic decisions, such as labor supply or savings. These responses can be legal (*avoidance*) or not (evasion). Slemrod and Yitzhaki (2002), building on the work of Stiglitz (1985), distinguish three basic principles of tax avoidance: retiming, tax arbitrage and income shifting. Retiming occurs when the timing of certain transactions responds to changes in tax rates. The classic example is the anticipation of capital gains realizations following the announcement of the tax rate increase in the Tax Reform Act of 1986 (TRA86). Tax arbitrage denotes all the activities that take advantage of inconsistencies in the tax law. Income shifting arises when the reduction in reported incomes is due to a shift away from taxable individual income toward other forms of taxable income, such as corporate income. An illustration is given by the shift from C corporations into S corporations (which are taxed like partnerships and therefore are not subject to the corporation income tax) following the drop in the top individual rate below the corporate rate in TRA86.

There exists now quite a substantial empirical literature, summarized in Saez et al. (2010), that assesses the extent of avoidance responses to taxation. These studies are mainly based on the natural experiment provided by TRA86. Saez (2004) finds that income shifting can explain most of the rise in Subchapter S and partnership income. Gruber and Saez (2002) estimate and compare the elasticities of taxable and of broad income. They find a much lower value for the former, suggesting that much of the taxable income response comes through deductions, exemptions, and exclusions. Kleven et al. (2011) conduct a field experiment and determine the effects of changes in marginal tax rates on reported income. They conclude that most of the elasticity of reported income with respect to tax rates can be explained by (legal) avoidance rather than (illegal) evasion. Overall there is compelling evidence of strong behavioral responses to taxation. Moreover these responses fall mainly in the avoidance category.

In contrast, the theoretical literature dealing with tax avoidance is quite limited. The optimal taxation literature, initiated by Mirrlees (1971), focuses on the real response to taxation. It aims at identifying the optimal income tax curve when individuals react to the tax by decreasing their labor supply. As argued before, this response is not the empirically most relevant. The taxable income is very sensitive to the tax rate mainly because of tax avoidance and evasion.[1]

Slemrod (2001) studies the effect of income taxation in a model where both real (change in labor supply) and avoidance responses are taken into account. He does however adopt a purely positive standpoint and does not determine the optimal level of taxes. Slemrod and Kopczuk (2002) determine the optimal level of avoidance. Contrarily to labor supply responses, avoidance behaviors can be, at least partly, controlled by the government. This has crucial implications for the design of the tax system. If avoidance responses to taxation are large, the best policy would not be to lower tax rates (as suggested by the standard Mirrleesian approach), but instead to broaden the tax base and eliminate avoidance opportunities.

Quite surprisingly, there are no theoretical studies that address the problem of the optimal nonlinear income tax schedule when individuals try to avoid taxes, with the notable exception of Grochulski (2007). This latter develops a standard optimal taxation model, in which individuals respond to the income tax by hiding part of their income, at a cost, instead of reducing their labor supply, as in the Mirrlees model. He finds 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.

These results are very clear-cut. They are however derived with a subadditive concealment cost function. In this article, we consider the case of convex cost function (that violates subadditivity). It turns out that the no-falsification theorem does not hold anymore. We show that, provided that the marginal cost of concealing the first euro is low enough, individuals belonging to the middle-class should optimally hide

[1] For theoretical studies of the optimal tax schedule when individuals evade taxes, see Cremer and Gahvari (1995) or Chander and Wilde (1998).

part of their income to the fiscal authority. For a marginal cost close to 1 however, all individuals should declare their true income. Finally the first-best (that consists in fully equalizing after-tax incomes) is achieved when the marginal cost is large enough (greater than 1). We also characterize optimal marginal tax rates and thus the shape of the optimal income tax schedule. Marginal tax rates are constant for non-avoiding people. They are greater for individuals who avoid paying taxes. The way they vary with income depend on the shape of the income distribution, as well as the characteristics of the concealment cost function and the preferences of the social planner. We construct an example leading to a bell-shaped curve of optimal tax rates. The corresponding optimal tax schedule is first convex and then concave.

2 Model

2.1 Population and preferences

Individuals differ with respect to income w, distributed according to the cumulative distribution function $F(.)$ and the density $f(.)$ on the support $[w_-, w_+]$; average income is denoted $\overline{w}$. Labor supply is assumed to be inelastic so that income is fixed. 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, where ϕ is continuous and $\phi(0) = 0$, $\phi'(.) > 0$, $\phi''(.) > 0$. Observe that we consider a convex cost function, which does not satisfy the subadditivity property.[2] The income declared by an individual with true income w is denoted $\hat{w}(w)$. It is assumed that individuals cannot declare more than their true income: $\hat{w}(w) \in [w_-, w]$.

Preferences depend only on consumption c, i.e. after-tax income. We assume for simplicity a linear utility function: $u(c) = c$. In the remainder of the paper, we will therefore talk indifferently of utility or consumption.

[2] A function f is subadditive iff $f(x+y) \leq f(x) + f(y)$. One can show that an increasing convex function that passes through the origin is not subadditive.

2.2 Tax policy

The government levies a tax $T^w(\hat{w})$ on declared income. We consider a purely redistributive problem, so that the government budget constraint is:

$$\int T^w(\hat{w}(w))f(w)dw \geq 0.$$

Consumption is equal to net-of-tax income minus the avoidance cost: $c(w) = w - T^w(\hat{w}(w)) - \phi(w - \hat{w}(w))$.

3 The optimal income tax schedule

3.1 Government's problem

The problem of the government consists in finding the tax function on income, $T^w(\hat{w})$, that maximizes a given social welfare function. By the Revelation Principle, this problem can be conveniently addressed by restricting ourselves to direct and revealing mechanisms. In other words, individuals are asked to directly declare their type and are assigned a reported income and a tax levels $\hat{w}(\tilde{w})$ and $T(\tilde{w})$, contingent on their report $\tilde{w}$. The allocation they receive should be designed such that individuals have incentives to reveal truthfully their type: $\tilde{w} = w$. Assuming that the planner maximizes the sum of a concave transformation $G(.)$ of individual utility levels, his program can be written as:

$$\max_{0 \leq \hat{w}(w) \leq w, T(w)} \int G(U(w))dF(w)$$

st

$$U(w) = w - T(w) - \phi(w - \hat{w}(w)),$$
$$\int T(w)f(w)dw \geq 0 \tag{1}$$

and

$$U(w) \geq w - T(w') - \phi(w - \hat{w}(w')). \tag{2}$$

The third constraint is the Government Budget Constraint (GBC) and the last one is the incentive constraint: a type w individual should not want to pretend that he is of type w'.

3.2 The solution without incentive constraints: first-best allocation

Without incentive constraints, there is no cost in making individuals reveal their true income, so that the first-best allocation can be achieved. Solving the previous program without the constraint (2) and denoting μ the Lagrange multiplier of the GBC, we get:

$$G'(U(w)) = \mu.$$

Quite obviously, as the government maximizes a concave transformation of individual consumptions, the first-best allocation consists in giving all individuals the same consumption level. As soon as the marginal cost of avoiding the first euro is not too large (less than 1 precisely), this is not incentive compatible. To implement the first-best, one must have a tax schedule with 100% marginal tax rates. Avoiding 1 euro then increases consumption by the same amount, less the avoidance cost. As soon as this cost is lower than 1, it is thus optimal for individuals to conceal part of their income.

3.3 The optimality of avoidance

The incentive constraint (2) implies that every individual should report truthfully his type. Therefore:

$$w = \arg\max_{w'} w - T(w') - \phi(w - \hat{w}(w')),$$

The first-order condition then implies:

$$-T'(w) + \hat{w}'(w)\phi'(w - \hat{w}(w)) = 0. \tag{3}$$

Noting that $T^w(\hat{w}(w)) = T(w)$, we have $T'(w) = \hat{w}'(w)T^{w'}(\hat{w}(w))$ and thus:

$$T^{w\prime}(\hat{w}(w)) = \phi'(w - \hat{w}(w)). \tag{4}$$

In words, the marginal tax rate should be equal to the marginal avoidance cost. This result, which has already been obtained by Slemrod (2001) and Grochulski (2007) is intuitive: should the marginal tax rate be lower (resp. greater) than the marginal cost, individuals should decrease (resp. increase) the amount of avoidance. A second lesson of this formula is that, because the cost function is assumed to be convex, *individuals who conceal more income face larger marginal tax rates.*

Using standard technique in mechanism design, the second-order condition for a local optimum can be shown to be:

$$\hat{w}'(w)\phi''(w - \hat{w}(w)) > 0.$$

As the cost function is assumed to be convex, the second-order condition is satisfied if and only if $\hat{w}'(w) > 0$, i.e. reported income increases with true income.[3] Violation of this 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. In the remainder of this article, we shall assume that the second-order condition is satisfied.[4]

Recalling that $U(w) = w - T(w) - \phi(w - \hat{w}(w))$ and using (3), we have

$$\frac{dU}{dw} = 1 - \phi'(w - \hat{w}(w)). \tag{5}$$

This condition is intuitive. The social planner, who wants to equalize consumption levels in the first-best, wishes to make the change in utility with respect to income as small as possible. There is however a limit to this, caused by the incentive constraints. If the second-best allocation were to imply $dU/dw < 1 - \phi'(w - \hat{w}(w))$, it would not be incentive compatible as the individual w would want to mimic the individual with a little less income. The change in "private" consumption, $1 - \phi'(w - \hat{w}(w))$, would more than compensate the loss in "public" consumption, dU/dw.

Anticipating on later results, we are not able to say if $\phi'(w - \hat{w}(w))$ is lower or greater than 1 at the optimum, leaving open the possibility that utility be decreasing with income for a subset of the population. This stands in contrast with the Mirrlees model, in which utility is necessarily increasing with productivity; otherwise high productivity individuals would have interest in mimicking low productivity ones. Here this is not guaranteed: if high incomes incur a large marginal cost of avoidance, they do not want to pretend having a lower income, even though they end up with a lower consumption level.

[3] This is the analogous condition to having pre-tax income being increasing with productivity in the optimal taxation literature (Theorem 1 in Mirrlees (1971)).

[4] For a careful treatment of bunching in optimal taxation models, see Lollivier and Rochet (1983), Ebert (1992) or Boadway et al. (2000).

We can thus restate the planner's problem as follows

$$\max_{0\leq \hat{w}(w)\leq w, T(w)} \int G(U(w))dF(w)$$

st

$$U(w) = w - T(w) - \phi(w - \hat{w}(w)),$$
$$\int T(w)f(w)dw \geq 0,$$
$$\frac{dU}{dw} = 1 - \phi'(w - \hat{w}(w)).$$

Taking U as the state variable, we form the Hamiltonian associated to this program:

$$\mathcal{H} = (G(U(w)) + \mu T(w))f(w) + \lambda(w)\frac{dU}{dw} + \beta(w)(w - \hat{w}(w)),$$

where μ and $\lambda(w)$ are the multipliers associated to the GBC and the incentive constraints respectively; $\beta(w)$ is the multiplier on the constraint ensuring that individuals report less than their true income. We did not include the multiplier on the constraint of positive report as this constraint can be shown to be non-binding at the optimum. The first-order conditions are then

$$\frac{\partial \mathcal{H}}{\partial \hat{w}} = 0$$
$$\Leftrightarrow \mu \left.\frac{dT}{d\hat{w}}\right|_U f(w) + \lambda(w)\phi''(w - \hat{w}(w)) - \beta(w) = 0, \tag{6}$$
$$\frac{\partial \mathcal{H}}{\partial U} = -\lambda'(w)$$
$$\Leftrightarrow -\lambda'(w) = (G'(U(w)) + \mu \left.\frac{dT}{dU}\right|_{\hat{w}})f(w). \tag{7}$$

Noting that $dT/d\hat{w}|_U = \phi'(w - \hat{w}(w))$ and $dT/dU|_{\hat{w}} = -1$, conditions (6) and (7) become

$$\mu\phi'(w - \hat{w}(w))f(w) + \lambda(w)\phi''(w - \hat{w}(w)) - \beta(w) = 0, \tag{8}$$
$$-\lambda'(w) = (G'(U(w)) - \mu)f(w).$$

Integrating the second condition and using the endpoint condition $\lambda(w_+) = 0$ yields

$$\lambda(w) = \int_w^{w_+} (G'(U(t)) - \mu)f(t)dt. \tag{9}$$

This multiplier measures the change in social welfare when individuals from w to the top are given one extra euro. On the one hand, the utility of the concerned individuals is increased and this is valued $G'(U(t)))$ by the social planner. On the other hand, this change is costly to society; the corresponding change in social welfare is given by μ, the multiplier of the GBC. Inspecting (8), it should be observed that $\lambda(w)$ is negative for individuals who do avoid taxes (for which $\beta(w) = 0$).

From the endpoint condition $\lambda(w_-) = 0$, we obtain

$$\mu = \int G'(U(w))dF(w). \tag{10}$$

We now argue that, when the marginal cost of hiding the first euro, $\phi'(0)$, is low enough, some individuals will report strictly less than their true income. On the other hand, for $\phi'(0)$ sufficiently close to 1, all individuals report truthfully their income and there is no tax avoidance at the optimum. Suppose that all individuals declare their true income: $\hat{w} = w$, $\forall w$. Then (5) implies:

$$\frac{dU}{dw} = 1 - \phi'(0).$$

Integrating this condition yields

$$U(w) = (1 - \phi'(0))w + k,$$

Recalling that utility is equal to consumption, the GBC can be written:

$$\int U(w)f(w)dw = \overline{w} - \int \phi(w - \hat{w}(w))f(w)dw.$$

As $\hat{w} = w$ and $\phi(0) = 0$, this becomes:

$$\begin{aligned} & \int U(w)f(w)dw = \overline{w} \\ \Leftrightarrow \quad & (1 - \phi'(0))\overline{w} + k = \overline{w} \\ \Leftrightarrow \quad & k = \overline{w}\phi'(0). \end{aligned}$$

As soon as $\phi'(0) < 1$, $U(w)$ is an increasing function of w. From the concavity of $G(.)$, we can conclude that $\lambda(w)$ is everywhere negative (except at w_- and w_+ where it is 0). When $\phi'(0) = 0$, the first term in (8) disappears. Noting that β and ϕ'' are positive, condition (8) is violated for any $w \in (w_-, w_+)$. Therefore *it cannot be* *the case that all individuals declare their true income.* By continuity, this conclusion holds true when $\phi'(0) \to 0$.

When $\phi'(0) = 1$, we have $U(w) = \overline{w}$ and $\lambda(w) = 0$, $\forall w$, so that the first-best allocation is attained. The inspection of (8) makes clear that $\beta(w) = \mu\phi'(0)f(w)$ is positive for all w, meaning that no avoidance is optimal for all individuals. The intuition is clear: when avoidance is too costly, individuals have no better choice than declaring their true income. This conclusion holds true for $\phi'(0) > 1$. When $\phi'(0) \to 1$, the first-best is not attained but a continuity argument allows to conclude that *all individuals declare their true income.* The marginal tax rate in such a case is constant and equal to $\phi'(0)$ (see (4)) but consumption levels, which are $(1-\phi'(0))w + \overline{w}\phi'(0)$, are not fully equalized.

These results suggest that there exists a threshold value for the marginal cost $\phi'(0)$, denoted $\tilde{\phi}$, such that no individual avoids taxation if and only if $\phi'(0) \geq \tilde{\phi}$. From (8), no individual will avoid taxes as soon as:

$$\mu\phi'(0)f(w) + \lambda(w)\phi''(0) \geq 0 \text{ for all } w$$

where

$$\lambda(w) = \int_w^{w_+} (G'((1-\phi'(0))t + \overline{w}\phi'(0)) - \mu)f(t)dt$$

and

$$\mu = \int G'((1-\phi'(0))w + \overline{w}\phi'(0))dF(w).$$

This condition is equivalent to

$$-\frac{\lambda(w)}{f(w)} \leq \mu\frac{\phi'(0)}{\phi''(0)}.$$

The limit value of $\phi'(0)$, $\tilde{\phi}$, is thus implicitly defined by

$$\begin{aligned} \max_w \quad & -\frac{\int_w^{w_+}(G'((1-\tilde{\phi})t + \overline{w}\tilde{\phi}) - \int G'((1-\tilde{\phi})w + \overline{w}\phi'(0))dF(w))f(t)dt}{f(w)} \\ = \quad & \frac{\tilde{\phi}}{\phi''(0)}\int G'((1-\tilde{\phi})w + \overline{w}\phi'(0))dF(w), \end{aligned} \tag{11}$$

where it should be noted that $\tilde{\phi}$ depends on $\phi''(0)$.

We have shown that some individuals will optimally avoid taxation when $\phi'(0) < \tilde{\phi}$. Noting that, as $\lambda(w_-) = \lambda(w_+) = 0$, individuals at the top and the bottom of the income distribution should report their true income, we obtain that there exist two threshold values $w_{\inf} \geq w_-$ and $w_{\sup} \leq w_+$ such that individuals with income $w \leq w_{\inf}$ and $w \geq w_{\sup}$ declare their true income. Moreover individuals located closely to the "right" of $w_{\inf}$ and to the "left" of $w_{\sup}$ understate their income report to the fiscal authority; $w_{\inf}$ and $w_{\sup}$ are solutions to

$$\mu\phi'(0)f(w) + \lambda(w)\phi''(0) = 0. \tag{12}$$

Note that there may exist more than two solutions to this equation, in which case some subsets of individuals located in the interior of the income distribution also declare truthfully.

We summarize in the following proposition the results of this section.

Proposition 1 *1. There exists $\tilde{\phi} \in (0,1)$, implicitly defined by (11), such that*

(i) If $\phi'(0) \geq \tilde{\phi}$, $\hat{w}(w) = w$, $\forall w$;

(ii) If $\phi'(0) < \tilde{\phi}$, $\exists w \in (w_-, w_+)$ such that $\hat{w}(w) < w$.

2. When $\phi'(0) < \tilde{\phi}$, there exist $w_{\inf}$ and $w_{\sup}$, obtained as solutions to (12), such that

(i) $\hat{w}(w) = w$, $\forall w \leq w_{\inf}$ and $w \geq w_{\sup}$;

(ii) There exists $\delta > 0$ such that $\hat{w}(w_{\inf}+\delta) < w_{\inf}+\delta$ and $\hat{w}(w_{\sup}-\delta) < w_{\sup}-\delta$.

Optimal reported incomes and consumption levels are represented on figures 1 and 2 respectively.

We now give the intuition of our main result, namely that some individuals should optimally conceal income when $\phi'(0) \to 0$. Suppose there is no avoidance and make individual $\tilde{w}$ avoid at the margin by perturbing the consumption schedule as represented on figure 3. If this new consumption schedule is both feasible and incentive compatible, it is then socially preferred to the original one (as it allows to "flatten" the consumption curve), meaning that avoidance is optimal.

Making $\tilde{w}$ avoid at the margin ($\hat{w} = \tilde{w} - \varepsilon$) allows to relax incentive constraints: because of convex concealment costs, higher income individuals are less tempted to mimic $\tilde{w}$.[5] This corresponds to the term $-\lambda(w)\phi''(0)$ in (8). But it also has a cost

[5] To see this, consider a discretized version of the model and assume that individuals $\tilde{w} + \delta$ are

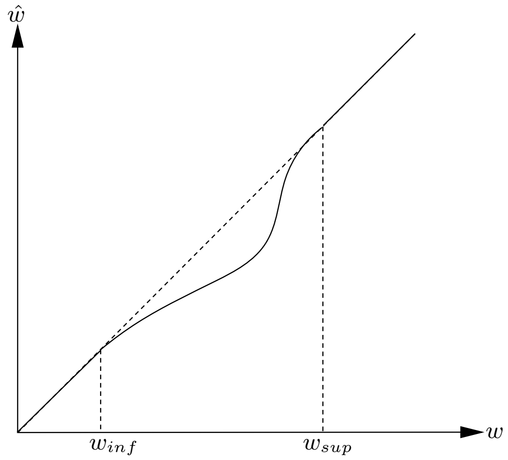

Figure 1: Reported incomes

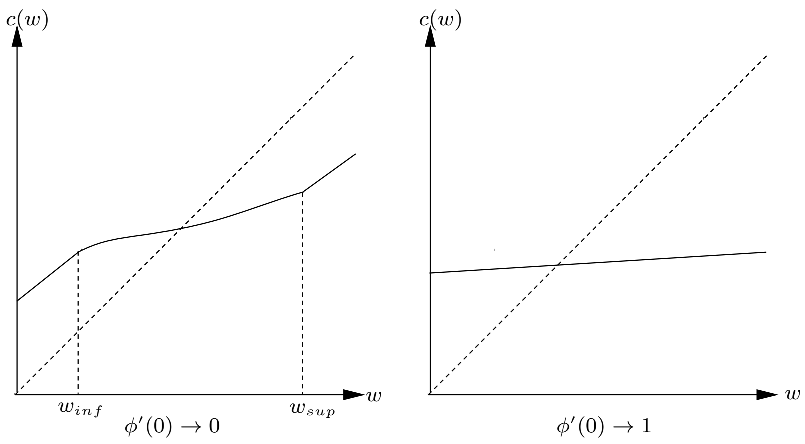

Figure 2: Consumption levels

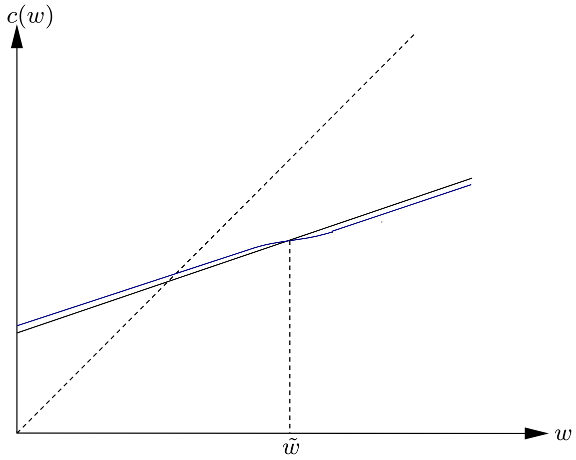

Figure 3: The effect of allowing avoidance

represented by the term $\mu\phi'(0)f(w)$: $\tilde{w}$ must incur a lower tax in order to stay at the same consumption level (to compensate for the cost of avoidance). When $\phi'(0) \to 0$, the benefit outweighs the cost for almost all individuals (not for individuals at the extreme of the distribution as $\lambda(w_-) = \lambda(w_+) = 0$). When $\phi'(0) \to 1$, $\lambda(w) \to 0$ and the cost outweighs the benefit for all individuals. It thus explains why it is optimal to allow for avoidance when the marginal cost of concealing the first euro is low enough. It also helps to explain why it concerns individuals belonging to the middle-class and not the very poor and the very rich.

3.4 Marginal tax rates

From (4), we know that marginal tax rates are equal to marginal avoidance costs and are thus everywhere *positive*. As emphasized previously, we are however not able to conclude about whether they are lower or greater than 1. In the latter case, this would imply that utility decreases with income (see (5)).

---

indifferent between mimicking $\tilde{w}$ or not:

$$\tilde{w} + \delta - T^w(\tilde{w} + \delta) = \tilde{w} + \delta - T^w(\tilde{w}) - \phi(\delta).$$

Now suppose that $\tilde{w}$ avoid taxes by declaring $\tilde{w} - \varepsilon$ instead of $\tilde{w}$. The utility from complying for individuals $\tilde{w} + \delta$ is unchanged. However, the utility when mimicking is now $\tilde{w} + \delta - T^w(\tilde{w} - \varepsilon) - \phi(\delta + \varepsilon)$. The change in the tax paid is thus $T^w(\tilde{w} - \varepsilon) - T^w(\tilde{w})$ while the change in avoidance cost is $\phi(\delta + \varepsilon) - \phi(\delta)$. For ε small enough, they can be approximated by $T^{w\prime}(\tilde{w})$ and $\phi'(\delta)$ respectively. Because the cost function is convex, we have that the increase in the avoidance cost $\phi'(\delta)$ is larger than the save in taxes $T^{w\prime}(\tilde{w}) \approx \phi'(0)$. Therefore $\tilde{w} + \delta$ is not indifferent anymore and strictly prefers not to mimick $\tilde{w}$. In other words, the incentive constraint has been relaxed.

For individuals who declare their true income ($\hat{w} = w$), we thus readily obtain that they face the marginal tax $\phi'(0)$. For the others, we can, using (8) with β set to 0, express the marginal tax rate as follows:

$$T^{w\prime}(\hat{w}(w)) = -\frac{\lambda(w)}{\mu}\frac{1}{f(w)}\phi''(w - \hat{w}(w)).$$

This expression is close to (9) in Diamond (1998) and its interpretation is by now standard in the optimal taxation literature (See, e.g., Saez (2001)). On the one hand, increasing the marginal tax rate at a given income level generates a distortion at this point so that the more there are people at this income level, as measured by $f(w)$, the lower the marginal tax rate should be. The distortion comes from the fact that individuals will react to the increased marginal tax rate by reducing their reported income. The term $1/\phi''(w - \hat{w}(w))$ measures this distortion (it can be obtained by differentiating (4)) and accordingly the lower $\phi''(.)$, the lower should be the marginal tax rate. On the other hand, raising the marginal tax rate locally allows to raise additional taxes on all individuals with higher income, without affecting incentive constraints. The net benefit of doing so is given by $-\lambda(w)$ (it is divided by μ in order to convert it from welfare to monetary units). The larger this benefit, the larger the marginal tax rate.

It is thus quite hard to predict how marginal tax rates should vary with income. It depends on the way $\lambda(w)$, $f(w)$ and $\phi''(w - \hat{w}(w))$ vary with w. We should however notice that *marginal tax rates are always larger for individuals who avoid with respect to non-avoiding people.* This is obtained readily by using (4) and observing that, due to the convexity of ϕ, $\phi'(w - \hat{w}(w)) > \phi'(0)$ whenever $\hat{w}(w) < w$.

4 Numerical illustration

To illustrate the model, we have constructed two numerical examples. In both examples, income is distributed uniformly on the support [0,10]. The cost of avoidance is $\phi(x) = x^2/2 + \alpha x$, so that $\phi'(0) = \alpha$ and $\phi''(x) = 1$ and $G(x) = \ln x$. In the first simulation, $\alpha = 0.4$ and $\alpha = 0.3$ in the second one. We obtain that, in both simulations, some individuals avoid, the threshold values for the avoiding individuals being respectively $w_{\inf}$ =1.28, $w_{\sup}$ =8.66 and $w_{\inf}$ =0.84, $w_{\sup}$ =9.3. Not surprisingly the set of avoiding people expands when the marginal cost $\phi'(0)$ is lowered.

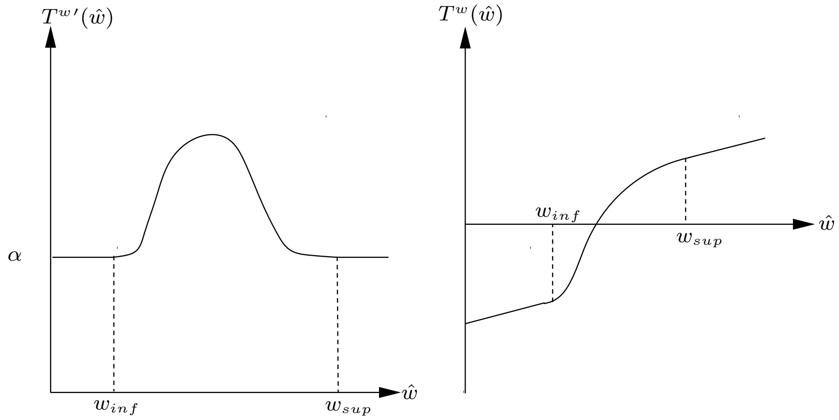

Figure 4: Shape of marginal tax rates and optimal tax scheme in the numerical examples

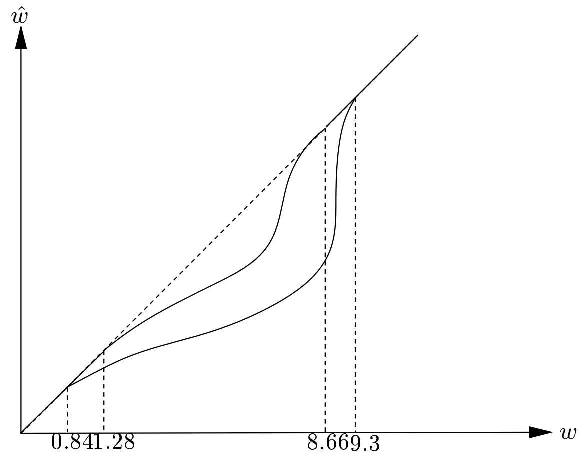

Figure 5: Reported incomes in the numerical examples

We also obtain a bell-shaped curve of marginal tax rates, the corresponding optimal tax schedule being first convex and then concave. This is represented on the figures below.

5 Generalizing the cost function

Two modifications of the cost function can be envisaged. First, we introduce a *fixed* cost in the avoidance technology. Individuals who want to avoid taxes should go through a costly information acquisition process concerning the tax law and have to pay a fixed amount, independently of the amount of income concealed. Second, individuals may have different avoidance opportunities, depending on their income level. In particular richer people may find it easier (meaning incurring a lower total and marginal cost) to avoid taxes than poor individuals. The generalized cost function thus takes the form:

$$\phi(\Delta, w) = \beta(w) + \xi(\Delta, w),$$

where β represents the fixed cost (possibly dependent on the true income level) and ξ the variable cost, that both depends on the amount concealed and the true income level.

With this cost function, formula (4) is modified to:

$$T^{w\prime}(\hat{w}(w)) = \xi_{\Delta}(w - \hat{w}(w), w).$$

Marginal tax rates are still equal to marginal avoidance costs, but these latter now depend on the income level w. If we assume that the marginal cost of avoidance decreases with income ($\xi_{\Delta w} < 0$), this implies that, for a given amount of avoidance, *rich individuals face a lower marginal tax rates than the poor.*

We then turn to the incidence of introducing a fixed cost in the analysis. This modification makes the cost function discontinuous at 0. This in turn implies that the social planner problem is non continuous and cannot be simply solved by analyzing first-order conditions. It is clear that with prohibitive fixed costs, no individuals will be allowed to avoid taxes. With moderate fixed costs however the optimality conditions remain the same as the ones derived above. We conjecture that the main change in the results would be that, intuitively, less individuals avoid taxes. To see this, consider the individuals for which the (unconstrained) solution was $\hat{w}(w) = w$. In such a case, the planner is indifferent between letting these people avoid at the margin or not. With a fixed cost of avoidance however, the planner now strictly prefers that these individuals declare their true income (as the fixed cost is saved when people do not conceal income). This suggests that *the set of avoiding individuals should shrink when avoidance generates a fixed cost.*

6 Conclusion

We have shown that it is optimal for some individuals to conceal income to the fiscal authority when the avoidance cost is convex. This contrasts with the result of Grochulski (2007), who proves a no-falsification theorem in the case of a sub-additive cost function. Our result relies on the idea that permitting avoidance allows to relax incentive constraints as high income individuals are less tempted to mimic lower income ones when these latter avoid taxes. The convexity of the cost function is crucial for this effect to arise and this thus explains the difference in the results between Grochulski (2007) and our approach.

We have assumed a bounded support for the distribution of incomes. A general result in the optimal taxation literature is that, with a bounded support for the distribution of productivities, the marginal tax rate is 0 at the highest sill level (Sadka (1976)). With an unbounded distribution, matters are however different. Diamond (1998) argue that for some utility functions and skill distributions, marginal tax rates may be increasing with productivity and be strictly positive at the limit. It should be noted that in our setting it makes no difference whether the distribution of incomes is bounded or not. One can readily check that all our results go through with an unbounded support. Individuals with income high enough do not conceal income and face a marginal tax rate equal to $\phi'(0)$.

Our results are quite provocative. At the optimal allocation, only the middle-class individuals should avoid taxes. This contrasts with evidence that points to the fact that the richest taxpayers in society are more prone to enter into tax avoidance activities (Agell and Persson (2000), Roine (2006)). Roine (2006) develops a political economy analysis that offers predictions in line with observed behaviors. He indeed shows that the equilibrium tax rates may be supported by a coalition of the poor and the rich. The poor would like to increase the tax rate because they benefit from the redistribution. The rich are also beneficiaries of the tax system as they exploit the avoidance opportunities and thus end up paying relatively small taxes. The middle-class people do not conceal income and are opposed to a further increase in the tax rate. The equilibrium predictions are thus at odds with the normative recommendation arising out of our model.

The fact that the rich do not conceal income in our normative analysis could be thought to be driven by our assumption that all individuals face the same avoidance opportunities, in the sense that they all face the same cost of avoidance. In section 5, we give some arguments why offering to the rich better avoidance opportunities (both with respect to the fixed and the marginal cost of avoidance) would not affect qualitatively the results. These arguments are however derived in an informal way and a more careful analysis is needed.

Finally, we have only considered the avoidance response to taxation. In order to get a better sense of the shape of the optimal tax schedule, it is desirable to incorporate in the model real responses to taxation, that is to allow individuals to choose optimally, together with the amount of reported income, their labor supply.

References

Agell, J. and M. Persson (2000). Tax arbitrage and labor supply. *Journal of Public Economics 78*, 3–24.

Boadway, R., K. Cuff, and M. Marchand (2000). Optimal income taxation with quasi-linear preferences revisited. *Journal of Public Economic Theory 2*(4), 435–60.

Chander, P. and L. L. Wilde (1998). A general characterization of optimum income tax enforcement. *Review of Economic Studies 65*(1), 165–83.

Cremer, H. and F. Gahvari (1995). Tax evasion and the optimum general income tax. *Journal of Public Economics 60*, 235–49.

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*, 47–73.

Grochulski, B. (2007). Optimal nonlinear income taxation with costly tax avoidance. *Economic Quarterly 93*(1), 77–109.

Gruber, J. and E. Saez (2002). The elasticity of taxable income: Evidence and implications. *Journal of Public Economics 84*(1), 1–32.

Kleven, H., M. Knudsen, C. Kreiner, S. Pedersen, and E. Saez (2011). Unwilling or unable to cheat? evidence from a tax audit experiment in denmark. *Econometrica 79*(3), 651–92.

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

Mirrlees, J. A. (1971). An exploration in the theory of optimal income taxation. *Review of Economic Studies 38*, 175–208.

Roine, J. (2006). The political economics of not paying taxes. *Public Choice 126*, 107–34.

Sadka, E. (1976). On income distribution, incentive effects and optimal income taxation. *Review of Economic Studies 43*, 261–67.

Saez, E. (2001). Using elasticities to derive optimal income tax rates. *Review of Economic Studies 68*(1), 205–29.

Saez, E. (2004). Reported incomes and marginal tax rates, 1960-2000: Evidence and policy implications. In J. Poterba (Ed.), *Tax Policy and the Economy*, Volume 18, pp. 117–74. Cambridge, MA: MIT Press.

Saez, E., J. Slemrod, and S. H. Giertz (2010). The elasticity of taxable income with respect to marginal tax rates: a critical review. *Journal of Economic Literature forthcoming.*

Slemrod, J. and S. Yitzhaki (2002). Tax avoidance, evasion, and administration. In A. J. Auerbach and M. Feldstein (Eds.), *Handbook of Public Economics*, Volume 3, pp. 1423–70. Amsterdam: North Holland.

Slemrod, J. B. (1995). Income creation or income shifting? behavioral responses to the tax reform act of 1986. *American Economic Review 85*(2), 175–80.

Slemrod, J. B. (2001). A general model of the behavioral response to taxation. *International Tax and Public Finance 8*(2), 119–28.

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

Stiglitz, J. E. (1985). The general theory of tax avoidance. *National Tax Journal 38*(3), 325–37.