Wage endogeneity, tax evasion and optimal nonlinear income taxation*

Firouz Gahvari
Department of Economics
University of Illinois at Urbana-Champaign, USA

Luca Micheletto
Department of Law, University of Milan, Italy,
UCFS Uppsala, Sweden, and
CESifo Germany

December 2019

*We thank an anonymous referee and the Editor in charge, Helmuth Cremer, for their helpful comments and advice.

Abstract

This paper studies the interaction between tax evasion and wage endogeneity within a Mirrleesian optimal tax framework. It characterizes the optimal marginal income tax rates on the skilled and the unskilled workers and the optimal amount of resources to be spent on deterring tax evasion. It shows that tax evasion weakens the incentives for the government to manipulate the marginal tax rates for the purpose of exploiting general equilibrium effects on wages. Moreover, the extent of this depends on the curvature of the evasion cost function. It also argues that marginal income tax rates are likely to be higher when the government attempts to deter evasion.

JEL classification: H21, H26.

Keywords: Nonlinear income taxation; endogenous wages; income-misreporting; tax administration; redistribution.

1 Introduction

Nearly five decades has past since Allingham and Sandmo (1972) formulated the first formal model of tax evasion in the economics discipline. A vast literature has since emerged expanding the subject into numerous directions. [1] One strand, to which this paper belongs, has sought to integrate tax evasion into the optimal income tax literature with variable labor supply. What distinguishes the current paper from the previous papers is that it allows for imperfect substitutability between skilled and unskilled workers and the endogeneity of their wages. Previous studies have followed the traditional optimal income tax models that assume wages are exogenously determined.[2]

Endogeneity of wages matters and not just for its realism; it opens up a second channel for redistribution. The change in the relative *gross-of-tax* wages, arising from government tax policies, plays as important a factor in income distribution as taxation itself (which aims to change the relative net-of-tax wages). These general equilibrium changes can very well reduce the burden of distortionary redistributive taxes. The point is that highly increasing marginal income tax rates leads to decreasing labor supplies of the higher-skilled workers. In turn, this will lead to an *increase* in the gross-of-tax wage rate of the higher-skilled relative to lower-skilled workers. By opting for a less increasing marginal income tax rate schedule, one can mitigate this perverse redistributive effect of taxation.

[1] See, e.g., Cremer *et al.* (1990), Cremer and Gahvari (1994, 1996, 2000), and Slemrod and Yitzhaki (2002) for a survey.

[2] Even though Stiglitz (1982) included a section on wage endogeneity in his two-group formulation of the optimum general income tax, the vast body of this general literature has focused on the original Mirrleesian assumption of fixed wages. Stiglitz (1985), Naito (1999), Huber (1999), Pirttilä and Tuomala (2002), Blackorby and Brett (2004), Micheletto (2004), Gaube (2005), Gahvari (2014) and Bastani *et al.* (2015) are among the exceptions.

Starting with Sandmo (1981), the contributors to the optimal income taxation literature in the presence of tax evasion have generally simplified the problem by restricting the income tax schedule to be linear. Cremer and Gahvari (1996) and Schroyen (1997) have reconsidered the problem within the Mirrleesian general income tax framework. However, all writers in this area have assumed that wages are constant. Thus far, to the best of our knowledge, nobody has considered tax evasion within the general income tax framework that allows for wage endogeneity.

The difference that wage endogeneity makes for the optimal marginal income tax rates was demonstrated by Stiglitz (1982, 1987) in his two-group model. One was that a marginal income subsidy on the high-skilled agents replaces the no-distortion-at-the-top recommendation. Secondly, while the marginal tax rate on the low-skilled is always positive, its magnitude depends on the elasticity of substitution between the two types of labor. The smaller the elasticity of substitution, the larger should be the marginal tax rate—with the government increasingly relying on the changes in the gross-of-tax wages to redistribute income. These modification to the characterization of the optimal marginal income tax rates are meant to bring about an increase in the labor supply ratio of the high-skilled relative to the low-skilled workers and a decrease in their relative wages.

This paper focuses on the implications of tax evasion for the optimal marginal income tax rates, particularly the component that reflects the change in the relative gross-of-tax wages. Prior studies on optimal taxation and tax evasion have neglected this specific question. To keep things simple, we adopt the riskless approach to evasion introduced in the literature by Usher (1986) and has since been used in a number of subsequent contributions.[3] It assumes taxpayers are able to fully avoid detection by incurring a cost that depends on the amount they misreport. In this way, we can focus on the implication of tax evasion for the marginal income tax rates that are exclusively due to the endogeneity of wages.

We append a production function, with skilled and unskilled workers as inputs, to the two-group model of Stiglitz. The inputs are imperfect substitutes so that wages of the two inputs are determined endogenously. The economic environment is competitive and wages reflect the value of marginal product. Specifying the evasion cost as a smooth convex function of the amount of evaded income, we characterize the optimal marginal income tax rates faced by the two types of workers. The first lesson we learn from this

[3]See, e.g., Mayshar (1991), Boadway *et al.* (1994), Kopczuk (2001), Slemrod (2001), Christiansen and Tuomala (2008), Chetty (2009), Gahvari and Micheletto (2014), and Blomquist *et al.* (2016).

exercise is that tax evasion weakens the incentives for the government to manipulate marginal income tax rates for the purpose of exploiting general equilibrium effects on wages. The second lesson we learn is that the effect of tax evasion on the optimal marginal income tax rates depends on the convexity of the concealment cost function. The tax rate decreases as the cost function becomes less convex. Put differently, the less convex is the concealment cost function (resulting in a higher amount of evasion), the less labor supplies should be distorted. Loosely speaking, this arises because the easier it is for agents to misreport their true income to the tax authority, the less effective the marginal income tax rate becomes as an instrument for changing labor supplies.

As an extension, we allow the concealment costs to also depend on government expenditures to deter evasion. We characterize the optimal amount that the government should spend on deterrence and discuss how this changes our earlier results under a passive government. Numerical exercises show that optimally spending resources to deter evasion strengthens the redistributive potential of taxation by allowing the government to rely further on the general-equilibrium effects on wages for redistribution. And, to do so, it further raises marginal income tax rates (in absolute value). Moreover, and in accordance with the theoretical results established in the tax evasion literature, the evasion-deterrence efforts of the government are primarily targeted towards agents reporting a low level of income.

Interestingly, the characterization of the optimal deterrence cost also highlights the role that endogeneity of wages play in determining this cost. Specifically, wage endogeneity generates an additional source of gain from marginally raising the amount of resources spent on deterrence. However, whether this implies that wage endogeneity strengthens the importance of spending resources to deter evasion will ultimately depend on the extent to which redistribution is made easier for the government when it can operate both on an ex-ante (wage channel) and on an ex-post basis (tax-transfer channel).[4]

[4] To avoid misunderstanding, note that this is a different question from the preceding one. The former

2 The setting

Consider an economy consisting of two types of agents, low-skilled (denoted by superscript ℓ) and high-skilled workers (denoted by superscript h). The number of workers of type j, with $j = \ell, h$, is denoted by N^j. All workers have identical quasi-linear preferences defined over consumption c and labor supply L. This is represented by $U = c - v(L)$ where $v(L)$ is strictly increasing and convex.[5]

The single consumption good c is produced through a linear homogeneous production function using labor hours supplied by each of the two types of agents

$$Q = P\left(N^\ell L^\ell, N^h L^h\right).$$

Labor markets are assumed competitive so that the equilibrium wage rate for an agent of skill-type j, w^j, is equal to the value of his marginal product

$$\begin{aligned} w^\ell &= \frac{\partial P\left(N^\ell L^\ell, N^h L^h\right)}{\partial N^\ell L^\ell} \equiv P_\ell, \\ w^h &= \frac{\partial P\left(N^\ell L^\ell, N^h L^h\right)}{\partial N^h L^h} \equiv P_h. \end{aligned}$$

Wage rates are thus determined endogenously (as long as the skill types are not perfect substitutes). As usual, the linear homogeneity of $P(\cdot)$ allows a simpler characterization. Define

$$n \equiv \frac{N^h L^h}{N^\ell L^\ell}. \tag{1}$$

concerns the effect of deterrence expenditures on the optimal marginal income tax rates (assuming wages are determined endogenously and we have tax evasion). The latter asks whether wage endogeneity (in comparison to fixed wages) calls for higher deterrence costs.

[5] Quasi-linearity is often invoked in the optimal income tax literature; see, e.g., Diamond (1998) or Salanié (2011). The ensuing implications of no income effects on labor supply and a constant marginal utility of income make the optimal tax formulas less "heavy" allowing one to gain a sharper insight into the elements that determine the optimal marginal income tax rates. We have nevertheless worked out the formulas for non-quasi-linear preferences which give us essentially the same picture, albeit with some nuances, in Gahvari and Micheletto (2017).

Then rewrite the production function as

$$Q = N^{\ell} L^{\ell} P\left(1, \frac{N^h L^h}{N^{\ell} L^{\ell}}\right) \equiv N^{\ell} L^{\ell} p\left(n\right), \tag{2}$$

and the equilibrium wage rates as[6]

$$\begin{aligned} w^{\ell} &= p\left(n\right) - np'\left(n\right), &(3)\\ w^{h} &= p'\left(n\right). &(4) \end{aligned}$$

3 The optimal tax problem

3.1 Design

The government intends to design a Pareto-efficient tax policy that allows it to achieve its revenue raising and redistributive goals. The informational structure of the problem includes the standard Mirrleesian assumption that the government knows only the distribution of types in the population and not "who is who". This rules out the possibility of using first-best type-specific lump-sum taxes/subsidies. However, in contrast to what is commonly assumed in optimal taxation models, we shall assume that earned incomes are not publicly observable either. This opens up the possibility of tax evasion.

To model tax evasion, we follow the riskless approach introduced by Usher (1986); specifically, once agents have incurred a cost that is increasing in the amount they misreport, they face no risk of detection. This simple structure allows the government to achieve its objectives through a general income tax function $T\left(\cdot\right)$ levied on *reported incomes*, M. To see this, recall that in a two-group model without evasion one needs only to determine the two groups' allocations which can be done by a direct mechanism

[6] We have,

$$\begin{aligned} w^{\ell} &= \frac{\partial Q}{\partial\left(N^{\ell}L^{\ell}\right)} = p\left(n\right) + \left(N^{\ell}L^{\ell}\right)p'\left(n\right)\frac{\partial n}{\partial\left(N^{\ell}L^{\ell}\right)} = p\left(n\right) - \left(\frac{N^h L^h}{N^{\ell}L^{\ell}}\right)p'\left(n\right), \\ w^{h} &= \frac{\partial Q}{\partial\left(N^{h}L^{h}\right)} = \left(N^{\ell}L^{\ell}\right)p'\left(n\right)\frac{\partial n}{\partial\left(N^{h}L^{h}\right)} = \left(\frac{N^{\ell} L^{\ell}}{N^{\ell}L^{\ell}}\right)p'\left(n\right). \end{aligned}$$

consisting of two bundles each specifying a particular amount of income (earned equal to reported), and consumption. The same procedure works in our model if the two bundles, the one intended for the low-skilled agents and the other for the high-skilled ones, are specified in terms of reported incomes M and tax payments T (equivalently, incomes taxpayers may keep: $B = M - T$). The reason is that, as we show below, an (M, B)-bundle corresponds to a (c, L)-bundle; notwithstanding the fact that reported incomes will likely differ from actual incomes.

Consider the optimization problem of an agent who is to choose between bundles (M^ℓ, B^ℓ) and (M^h, B^h). Regardless of which bundle he chooses, he would want to maximize his utility $U = c - v(L)$. Denote the amount he evades, or more precisely misreports, by a.[7] Evasion is costly. A taxpayer who evades a incurs $f(a)$ in costs where $f(\cdot)$ is assumed to be non-negative, increasing in a, strictly convex, and that $f(0) = f'(0) = 0$. Thus the consumption level of a taxpayer who selects (M, B), and subsequently evades a, is equal to

$$c = a + B - f(a),$$

where we have dropped the superscripts for ease in notation. Now, with the taxpayer's true earnings being equal to wL, the amount he evades must be equal to $a = wL - M$. Substituting this in the above equation for c and the resulting expression in $U = c - v(L)$ yields

$$U = wL - M + B - f(wL - M) - v(L). \tag{5}$$

The taxpayer's optimization problem is thus choosing L to maximize (5). This results

[7] With wage endogeneity, high-skilled agents may end up receiving a subsidy on their marginal income. This will indeed occur in a two-group model when there are no incentives reasons to tax high-skilled types at the margin. These agents then have an incentive to over-report rather than under-report their income. Either way though, the modeling is identical. The only difference is that while with under-reporting and tax evasion $a > 0$, over-reporting implies $a < 0$. Under this latter case, the agents still need to incur a concealment cost to get away with this cheating. The concealment cost will then be assumed to be a function of the absolute value of a with the same properties. In principle, then, the problem is better referred to as income misreporting rather than tax evasion. We shall use the two terms interchangeably.

in the first-order condition

$$v'(L) = w\left[1 - f'(wL - M)\right], \tag{6}$$

which determines L and subsequently c.

Having shown that a taxpayer's choice of (M, B) determines his consumption bundle (c, L), we now discuss how the mechanism designer determines (M^{ℓ}, B^{ℓ}) and (M^h, B^h). There are two types of constraints that must be considered in this problem. One is that the bundles must satisfy the economy's resource constraint. Secondly, with the taxpayers being ultimately free to determine their allocations when given a tax function, the bundles must be incentive compatible (requiring each ability type to prefer the (M, B)-bundle intended for him to that intended for the other type). However, to reduce the number of the incentive compatible constraints that must be imposed, assume that in equilibrium $w^h \geq w^{\ell}$ so that the redistribution is from high- to low-skilled workers.[8] This allows one to ignore the self-selection constraint corresponding to low-skilled "mimicking" high-skilled workers (whom, following the literature, we shall refer to as "mimickers").

Let $V(M, B; w)$ denote the maximum value function of problem (5). A Pareto-efficient tax structure is characterized as the solution to the following problem:

$$\max_{M^h, B^h, M^{\ell}, B^{\ell}} V\left(M^h, B^h; w^h\right)$$

subject to,

$$\begin{aligned} V\left(M^{\ell}, B^{\ell}; w^{\ell}\right) &\geq \overline{U}^{\ell}, \\ V\left(M^h, B^h; w^h\right) &\geq V\left(M^{\ell}, B^{\ell}; w^h\right), \\ P\left(N^{\ell}\frac{M^{\ell} + a^{\ell}}{w^{\ell}}, N^h\frac{M^h + a^h}{w^h}\right) &\geq \left(B^{\ell} + a^{\ell}\right)N^{\ell} + \left(B^h + a^h\right)N^h, \end{aligned}$$

[8] Gaube (2005) has demonstrated that wage endogeneity may lead to second-best solutions other than the traditional "redistributive" and "regressive" cases. We nevertheless concentrate on the $w^h \geq w^{\ell}$ alone in light of the paper's goals.

where $(M^j + a^j)/w^j = L^j$. The first constraint prescribes a minimum utility for the low-skilled agents; the second constraint is the binding self-selection constraint requiring high-skilled agents not to behave as mimickers, and the last constraint is the resource constraint of the economy.

Introduce the notation

$$g(n) \equiv p(n) - np'(n), \tag{7}$$

and treat w^ℓ as an artificial control variable for the government. This allows one to formulate the Lagrangian of the government's problem above as

$$\begin{aligned}\mathcal{L} \;=\;& V\left(M^h, B^h, w^h\right) + \mu V\left(M^\ell, B^\ell, w^\ell\right) \\ &+\gamma\left\{P\left(N^\ell \frac{M^\ell + a^\ell}{w^\ell}, N^h \frac{M^h + a^h}{w^h}\right) - \left(B^\ell + a^\ell\right) N^\ell - \left(B^h + a^h\right) N^h\right\} \\ &+\lambda\left[V\left(M^h, B^h, w^h\right) - V\left(M^\ell, B^\ell, w^h\right)\right] \\ &+\phi\left[w^\ell - g(n)\right],\end{aligned} \tag{8}$$

where μ, γ, λ, and ϕ are the Lagrange multipliers associated with the constraint requiring a given level of utility for the low-skilled, the economy's resource balance constraint, the self-selection constraint, and the constraint capturing the equilibrium value of the wage rate of low-skilled agents. The optimal values of M^h, B^h, M^ℓ, B^ℓ are given by the first-order conditions to this problem presented in the Appendix [equations (A30)–(A31) and (A36)–(A37)].

3.2 Properties

We can now specify the properties of the optimal tax schedule $T(M)$ that implements (M^j, B^j), and thus (c^j, L^j), $j = \ell, h$. Faced with the tax schedule $T(M)$, or alternatively $B(M) \equiv M - T(M)$, an agent with the utility function $V(M, B; w)$ chooses M to maximize $V(M, B(M); w)$. Assuming $T(M)$ and thus $B(M)$ is differentiable, the first-order condition for this problem is,

$$\frac{dV}{dM} = \frac{\partial V}{\partial M} + \frac{\partial V}{\partial B}\frac{dB}{dM} = 0 \implies \frac{dB}{dM} = -\frac{\partial V/\partial M}{\partial V/\partial B}$$

We also have, from differentiating $B(M) \equiv M - T(M)$, that $dB/dM = 1 - T'(M)$. Hence one can rewrite the above condition as $1 - T'(M) = -(\partial V/\partial M)/(\partial V/\partial B)$. This property allows us to implicitly define the marginal income tax faced by an agent, $T'(M)$, as

$$T'(M) = 1 + \frac{\partial V/\partial M}{\partial V/\partial B} \equiv 1 - MRS_{MB}, \tag{9}$$

where MRS_{MB} denotes the marginal rate of substitution between M and B.[9] Now recall that $V(M, B; w)$ represents the maximum value function of problem (5). Invoking the envelope theorem, $\partial V/\partial M = -[1 - f'(wL - M)]$ and $\partial V/\partial B = 1$. Thus, in equilibrium, we will have,[10]

$$T'(M) = f'(wL - M). \tag{10}$$

In order to establish the properties of $T'(M)$, we begin by presenting two results. Lemma 1 proves that at any point in the (M, B)-space, the indifference curves of low-skilled agents are steeper than those of the high-skilled agents. Lemma 2 proves that ϕ, the multiplier associated with the last constraint in the government's problem summarized by the Lagrangian (8), is negative.

Lemma 1 *Assume that the Mirrleesian agent-monotonicity condition in income and consumption (wL, c)-space is satisfied. Then:*

(i) The same property holds in the (M, B)-space so that

(ii) The high-skilled worker pretending to be low-skilled will evade more than the low-skilled worker. That is, $a^{h\ell} > a^{\ell}$.

Proof. See the Appendix. ■

[9] As is well known (see, e.g. Stiglitz, 1982, p. 217, and Stiglitz, 1987, p. 1003), with two groups of individuals, the implementing tax function $T(M)$ will not in general be differentiable at one point. In our setting, the problem occurs at M^{ℓ}. Nevertheless, there always exists an implementing tax structure whose left-hand derivative at M^{ℓ} will be given by (9).

[10] The first-order conditions hinge on the assumptions about the evasion cost function $f(\cdot)$; in particular $f(0) = f'(0) = 0$. Observe also that the equilibrium condition (10) does not imply that tax evasion affects the marginal income tax rate only through the shape of $f(\cdot)$. Evasion also affects wL; thus the argument of $f'(\cdot)$ and with it $T'(M)$.

Lemma 2 *The multiplier ϕ, as specified in the Lagrangian (8), is negative.*

Proof. See the Appendix. ■

We are now in a position to give a characterization for the optimal marginal income tax rates of the skilled and unskilled workers. These characterizations are the basis for the main results of the paper.

Proposition 1 *Consider a two-skill-type Mirrleesian general income tax problem wherein wages w^j, $j = h, \ell$, are determined endogenously. Preferences are quasi-linear in consumption c with the labor supply L as the other argument. Let M denote reported income, $M - B$ income tax paid, MRS_{MB} the marginal rate of substitution between M and B, N^j is the number of j-type workers with $n \equiv N^h L^h / N^\ell L^\ell$, $g(n)$ by (7) with $g'(\cdot) > 0$, and γ, λ, ϕ the Lagrange multipliers associated with the economy's resource balance constraint, the self-selection constraint, and the constraint capturing the equilibrium value of the wage rate of low-skilled agents.*

If taxpayers can misreport their earning wL and avoid detection by incurring a cost $f(a)$ which is increasing in a and convex, the optimal marginal income tax rates faced by the high- and the low-skilled agents are given by:

$$T'\left(M^h\right) = \Delta^h \frac{\phi n g'}{\gamma w^h N^h L^h} \tag{11}$$

$$T'\left(M^\ell\right) = \frac{\lambda}{\gamma N^\ell}\left[MRS^\ell_{MB} - MRS^{h\ell}_{MB}\right] - \Delta^\ell \frac{\phi n g'}{\gamma w^\ell N^\ell L^\ell}, \tag{12}$$

where,

$$0 < \Delta^j \equiv \frac{\left(w^j\right)^2 f''\left(a^j\right)}{v''\left(L^j\right) + \left(w^j\right)^2 f''\left(a^j\right)} \leq 1. \tag{13}$$

Proof. See the Appendix. ■

The first point to note about Proposition 1 is that the optimal marginal income tax rate on the low-skilled type consists of two expressions. The first expression on the right-hand side of (12) is standard and reflects the incentive term that characterizes the optimal marginal tax rate faced by low-skilled agents in a setting where wages are fixed.

Given the agent-monotonicity condition in the (M, B)-space, $MRS^{\ell}_{MB} - MRS^{h\ell}_{MB} > 0$ so that this term continues to be positive with tax evasion. That there is no such term in (11) is of course due to the so-called "no-distortion-at-the-top" result.

Secondly, the second expression on the right-hand side of (12), as well as the sole expression on the right-hand side of (11), arises because of the endogeneity of wage rates. With $g' = p' - p' - np'' = -np'' > 0$, and $\phi < 0$, *this calls for an additional tax on the marginal income of the low-skilled and a subsidy on the marginal income of high-skilled* both of which result in increasing the relative wage of the low-skilled to the high-skilled. This result is in keeping with previous findings in the literature in the absence of tax evasion (see, e.g., Stiglitz 1982, 1987).

3.3 Evasion versus no evasion

The two points we made above explain the components of the optimal marginal income tax rates in our model. However to understand the import of tax evasion *per se*, in a setting where wage rates are endogenously determined, one must compare these expression with the corresponding optimal marginal income tax rates in the absence of evasion. Thus assume earnings are publicly observable and cannot be misreported. Then $a^j = 0$, $M^j = w^j L^j$, and $B^j = c^j$. We show in the Appendix that the optimal marginal income tax rates in the absence of tax evasion are characterized by

$$T'\left(M^h\right) = \frac{\phi n g'}{\gamma w^h N^h L^h} \tag{14}$$

$$T'\left(M^\ell\right) = \frac{\lambda}{\gamma N^\ell}\left[MRS^{\ell}_{MB} - MRS^{h\ell}_{MB}\right] - \frac{\phi n g'}{\gamma w^\ell N^\ell L^\ell}. \tag{15}$$

Comparing (14)–(15) with (11)–(12) reveals, firstly, that tax evasion leaves the incentive terms in the optimal marginal income tax rate intact but mitigates the impact coming from the channel due to general-equilibrium effect on the relative wage rates. This follows because of these latter terms are multiplied by a factor Δ^j, as defined in (13), with the property $0 < \Delta^j \leq 1$. Consequently, *tax evasion points to the desirability of a "lower" tax rate on the low-skilled and a "lower" subsidy rate on the high-skilled* (smaller positive distortion on the labor supply of high-skilled and smaller negative distortion on the labor supply of low-skilled).[11]

We must emphasize here that "lower" and "higher" in this discussion refers only to what the "tax rules" suggest and not the final equilibrium "levels" of the tax. The common variables appearing in the tax formulas with and without tax evasion assume different values in the two cases and do not allow a comparison between tax levels. Our terminology is based on the distinction Atkinson and Stern (1974) make between the "rule" for and the "level" of optimal provision of public goods in a first-best versus a second-best world.

The second observation about the impact of tax evasion can be gleaned from the expression for Δ^j in (13). It tells us that the effect of tax evasion on the optimal marginal income tax rates depends on the convexity of the concealment cost function. The reduction becomes greater (i.e. the tax rate decreases) as $f(a)$ becomes less convex ($f''(a)$ becomes smaller). In words, *the less convex is the concealment cost function (resulting in a higher amount of evasion), the less one should distort labor supplies.*[12] Loosely speaking, this arises because the easier it is for agents to misreport their true income to the tax authority, the less effective the marginal income tax rate becomes as an instrument to affect labor supplies.

4 Deterrence

The riskless approach to tax evasion assumes that taxpayers can get away with cheating by spending a given amount of resources (for book-keeping, lawyers, etc.). This makes the tax authority a passive agent. A more realistic approach assigns a role to the tax

[11] As pointed out earlier, in a two-group model, the high-skilled agents face a negative marginal income tax rate. Tax evasion for them takes the form of over-stating their true income to the tax authority to get a better tax treatment.

[12] In deriving our results, we rely on condition (10), which in turn hinges on the assumed properties of the f-function, and in particular the fact that it is strictly convex, and that $f(0) = f'(0) = 0$. These assumptions are standard in the literature (see references in footnote 2) that adopts the riskless approach to modeling tax evasion. Their inclusion though implies that one cannot apply equations (11)–(12) to the case when $f'' = 0$.

authority and recognizes its power to make cheating more difficult for taxpayers. One way to model this, while retaining the riskless property of our approach, is to assume the tax authority can make taxpayers' cheating more costly by spending resources of its own to deter tax evasion. Of course, given the assumption that taxpayers are always able to conceal their cheating, the tax authority's attempt at enforcement can only be useful if it weakens the self-selection constraint $V\left(M^h, B^h; w^h\right) \geq V\left(M^\ell, B^\ell; w^h\right)$ that would otherwise be binding. If not, deterrence attempts will be pure waste (to the government as well as the economy).

Specifically, assume that the concealment cost of evasion is given by the non-negative function $f(a, R)$ where R denotes the (per capita) amount of resources the tax authority spends to prevent evasion. Assume also that $f(0, R) = f_a(0, R) = 0$, $f(a, R)$ is increasing in R and a (in absolute value), $f_{aR}(a, R) > 0$ (a in absolute value), $f_{RR}(a, R) > 0$, and $f_{aa}(a, R) > 0$ (where one subscript on f denotes a first partial derivative and two subscripts a second partial derivative).[13]

With R as an additional policy tool, the Lagrangian of the government's problem becomes:

$$\begin{aligned}
\mathcal{L} \;=\; & V\left(M^h, B^h, R, w^h\right) + \mu V\left(M^\ell, B^\ell, R, w^\ell\right) \\
& + \gamma \left\{ P\left(N^\ell \frac{M^\ell + a^\ell}{w^\ell}, N^h \frac{M^h + a^h}{w^h}\right) - \left(B^\ell + a^\ell - R\right) N^\ell - \left(B^h + a^h - R\right) N^h \right\} \\
& + \lambda \left[V\left(M^h, B^h, R, w^h\right) - V\left(M^\ell, B^\ell, R, w^h\right) \right] \\
& + \phi \left[w^\ell - g(n) \right].
\end{aligned}$$

It is clear from this representation that the characterization of the optimal marginal income tax rates provided by (11)–(12) remains valid even after expanding the armory of government's policy tools to include R as an additional instrument. The obvious

[13]The assumption $f_{aR} > 0$ ensures that R will have an interior solution. Intuitively, the assumption means that increasing the expenditure on deterrence increases the individual's marginal cost of evasion.

reason for this result is that the optimal marginal income tax rates are obtained by properly combining the first-order condition to the government's problem with respect to M^h, B^h, M^ℓ, and B^ℓ; *for a given value for* R, these first-order conditions take exactly the same form as in the problem considered in the previous section.[14]

This leaves us with the task of characterizing the optimal policy with respect to the choice of R. The first-order condition of the government's problem with respect to this policy variable is given by:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial R} &= (1+\lambda)\frac{\partial V^h}{\partial R} + \mu \frac{\partial V^\ell}{\partial R} - \lambda \frac{\partial V^{h\ell}}{\partial R} \\ &\quad + \gamma \left[P_\ell \frac{N^\ell}{w^\ell}\frac{\partial a^\ell}{\partial R} + P_h \frac{N^h}{w^h}\frac{\partial a^h}{\partial R} - N^\ell \left(\frac{\partial a^\ell}{\partial R}+1\right) - N^h\left(\frac{\partial a^h}{\partial R}+1\right)\right] - \phi g' \frac{\partial n}{\partial R} \leq 0,\end{aligned}$$

with the associated Kuhn-Tucker condition $R\partial\mathcal{L}/\partial R = 0$ to allow for the possibility that the optimal value for R is zero. Setting $P_\ell = w^\ell$ and $P_h = w^h$, this is simplified to

$$\frac{\partial \mathcal{L}}{\partial R} = (1+\lambda)\frac{\partial V^h}{\partial R} + \mu \frac{\partial V^\ell}{\partial R} - \lambda \frac{\partial V^{h\ell}}{\partial R} - \left(N^\ell + N^h\right)\gamma - \phi g' \frac{\partial n}{\partial R} \leq 0. \tag{16}$$

Focusing on an interior solution for R (i.e. $R > 0$), problem (16) provides us with a characterization for the optimal value for R.

Proposition 2 *Assume Problem (16) has an interior solution (i.e.* $R > 0$*) and let*

$$\Psi^j \equiv v''\left(L^j\right) + \left(w^j\right)^2 f_{aa}\left(a^j, R\right), \quad j = h, \ell.$$

[14] Of course, this does not mean that the numerical value of the marginal income tax rates characterized by (11)–(12) will be identical in the two cases (unless the optimal policy for the government is to set $R = 0$). Moreover, with the f-function now depending both on a and R, one needs to replace $f''\left(a^h\right)$ in (11) and $f''\left(a^\ell\right)$ in (12) with $f_{aa}\left(a^h, R\right)$ and $f_{aa}\left(a^\ell, R\right)$, and to properly redefine Δ^j as

$$\Delta^j \equiv \frac{\left(w^j\right)^2 f_{aa}(a^j, R)}{v''\left(L^j\right) + \left(w^j\right)^2 f_{aa}(a^j, R)}, \; j = h, \ell.$$

The optimal amount of resources spent to deter evasion is implicitly characterized by

$$\underbrace{\lambda\left[f_R\left(a^{h\ell},R\right)-f_R\left(a^{\ell},R\right)\right]}_{\Gamma>0}+\underbrace{\left[\frac{w^h f_{aR}\left(a^h,R\right)}{L^h\Psi^h}-\frac{w^\ell f_{aR}\left(a^\ell,R\right)}{L^\ell\Psi^\ell}\right]\phi n g'}_{\Omega>0}=$$

$$\underbrace{\gamma\left[N^\ell\left(1+f_R\left(a^\ell,R\right)\right)+N^h\left(1+f_R\left(a^h,R\right)\right)\right]}_{>0}. \tag{17}$$

Proof. See the Appendix (for both the expression and the signs). ■

To interpret condition (17) it is useful to consider the effects of a policy reform that marginally raises R while at the same time adjusting B^ℓ and B^h in such a way as to leave the utility of all non-mimicking agents unaffected. This requires accompanying the marginal increase in R with an upward adjustment in B^ℓ and B^h given by $dB^\ell = f_R\left(a^\ell,R\right)$ and $dB^h = f_R\left(a^h,R\right)$. The right-hand side of condition (17) represents the overall revenue cost of this reform.

The left hand side of (17) captures the gains of the proposed reform. They come from two sources. First, there is a gain associated with relaxing the binding self-selection constraint. This is given by the first term on the left-hand side of condition (17), the term labeled Γ. We have already shown in the Appendix, when the concealment cost is independent of R, that $a^{h\ell} > a^\ell > 0$. However, it is clear from the proof that it also works when $f(\cdot)$ depends on R. This implies that $f_R\left(a^{h\ell},R\right) > f_R\left(a^\ell,R\right)$; hence, while the reform leaves the utility of the low-skilled agents intact, it lowers the utility of a high-skilled mimicker.

Second, there is a gain, working via general-equilibrium effects, emanating from a reduction in the wage gap between high- and low-skilled workers. This comes about because, as shown in the proof of Proposition 2 in the Appendix, $\partial n/\partial R > 0$ so that the proposed reform induces an increase in n, i.e. in the ratio of the labor supply of high-skilled workers to that of low-skilled workers. This increases the relative wage of low-skilled to high-skilled workers and represents a gain through its impact on the ϕ-

constraint.[15] The second term on the left-hand side of condition (17), the term labeled Ω, captures this effect.[16]

Notice that in a setting with exogenous wages the counterpart of (17) would be the following simplified condition:

$$\lambda\left[f_R\left(a^{h\ell},R\right)-f_R\left(a^{\ell},R\right)\right]=\gamma\left[N^{\ell}\left(1+f_R\left(a^{\ell},R\right)\right)+N^{h}\left(1+f_R\left(a^{h},R\right)\right)\right], \tag{18}$$

i.e. the term labeled Ω in (17), representing one of the two sources of gains from raising R, would vanish. That $\Omega > 0$ suggests that spending resources to deter tax evasion might be relatively more valuable in a setting with endogenous wages than in a setting where wages are exogenous. Intuitive as this may seem, such a rule-type inference is not quite warranted. The gain to combating tax evasion comes not just from Ω but Γ as well. Now when different types of labor are imperfect substitutes, the fact that redistribution can operate both on an ex-ante (wage channel) and on an ex-post basis (tax-transfer channel) tends to weaken the tightness of the self-selection constraint and to lower the equilibrium value of the Lagrange multiplier λ in (18) as compared to (17).

4.1 Targeted deterrence

The way we have introduced the policy variable R in above suggests that the evasion-deterrence efforts of the government are not targeted towards specific levels of reported income. Suppose instead that the resources devoted to fight evasion can be made income-dependent so that an agent reporting an income M^j will face a concealment cost $f\left(a,R^j\right)$, where R^j denotes the (per capita) resources devoted to deter evasion by agents reporting an income M^j. The Lagrangian of the government's problem would

[15] What is behind this change in relative wages is that, with two factors of production, they are necessarily Edgeworth complements.

[16] The gain represented by the term Ω in (17) is, ultimately, a gain due to mimicking-deterring effects. This is apparent once one notices that, as implied by (A29), the value of the shadow price ϕ is related to the value of the Lagrange multiplier λ, and that $\phi = 0$ when $\lambda = 0$.

become:

$$
\begin{aligned}
\mathcal{L} \;=\; & V\left(M^h, B^h, R^h, w^h\right) + \mu V\left(M^\ell, B^\ell, R^\ell, w^\ell\right) \\
& +\gamma\left\{P\left(N^\ell \frac{M^\ell + a^\ell}{w^\ell}, N^h \frac{M^h + a^h}{w^h}\right) - \left(B^\ell + a^\ell - R^\ell\right) N^\ell - \left(B^h + a^h - R^h\right) N^h\right\} \\
& +\lambda\left[V\left(M^h, B^h, R^h, w^h\right) - V\left(M^\ell, B^\ell, R^\ell, w^h\right)\right] \\
& +\phi\left[w^\ell - g\left(n\right)\right].
\end{aligned}
$$

Focusing on the case of an interior solution for both R^ℓ and R^h, it is straightforward to show that their optimal values would be implicitly characterized by the two following conditions, which respectively refer to R^ℓ and R^h:

$$
\underbrace{\lambda\left[f_R\left(a^{h\ell}, R^\ell\right) - f_R\left(a^\ell, R^\ell\right)\right]}_{>0} - \underbrace{\frac{w^\ell}{L^\ell} \frac{f_{aR}\left(a^\ell, R^\ell\right)}{\Psi^\ell} \phi n g'}_{<0} = \underbrace{\gamma N^\ell\left[1 + f_R\left(a^\ell, R^\ell\right)\right]}_{>0}, \tag{19}
$$

$$
\underbrace{\frac{w^h}{L^h} \frac{f_{aR}\left(a^h, R^h\right)}{\Psi^h} \phi n g'}_{>0} = \underbrace{\gamma N^h\left[1 + f_R\left(a^h, R^h\right)\right]}_{>0}. \tag{20}
$$

In both equations (19) and (20), the left-hand side provides a measure of the social gains arising from a marginal increase in the value of the relevant policy variable, and the right hand side provides a measure of the associated social cost. Eq. (19) highlights that the evasion-deterrence efforts targeted at the agents reporting M^ℓ generate two sources of gains, a direct mimicking-deterring effect due to the fact that $a^{h\ell} > a^\ell$, and an indirect mimicking-deterring effect that operates through a compression of the wage distribution. On the other hand, from eq. (20) one can see that only the latter source of gains can be reaped by the evasion-deterrence efforts targeted at agents reporting M^h. At least in terms of policy rules, this observation leads support to the idea that it is more effective to target the evasion-deterrence efforts towards low-income earners. At the same time, however, the fact that wages are endogenously determined implies that one should not refrain from allocate resources to deter evasion by agents reporting a high level of income.

5 Two extensions

5.1 Commodity taxes

If labor income could be costlessly observed by the government, there would clearly be no role for supplementing a nonlinear income tax with a commodity tax in our setting. With one single consumption good in our model, the effect of any commodity tax could simply be replicated by a proper adjustment of the nonlinear labor income tax. However, when agents can misreport income for tax purposes, commodity taxation will be a non-redundant policy instrument if aggregate sales of commodities are publicly observable and are taxed (at a uniform rate to all consumers). These taxes will not be subject to evasion. Under this circumstance, one can show that all the expressions for the optimal marginal income tax rates would have to be amended to include an additional negative term that depends on the rate at which the commodity tax is levied. Moreover, although the additional negative term takes a different value for each skill type in the population because of their different consumption levels, for each skill type its absolute value is decreasing in the curvature of the f-function.[17]

5.2 Many-skill-types

We now show that the main message of our paper, namely, the possibility that tax evasion mitigates the kind of tax-induced general equilibrium effects originally emphasized by Stiglitz (1982, 1987), does not rely on a two-group specification. Consider a setting where the number of skill types is larger than two, say, K. Output is then produced through a generalized version of the constant-returns-to-scale production function (2) according to,

$$Q = P\left(N^0L^0, N^1L^1, ..., N^KL^K\right) = N^0L^0 p\left(\frac{N^1L^1}{N^0L^0}, \frac{N^2L^2}{N^0L^0}, ..., \frac{N^KL^K}{N^0L^0}\right),$$

[17] The detailed analytical results are available upon request.

where the set of skill types is given by $\{0, ..., K\}$ with a higher index corresponding to a higher skill type. Let

$$n^j \equiv \frac{N^j L^j}{N^0 L^0}, \quad j = 1, 2, \ldots, K.$$

The labor markets being competitive, wage rates are given by,

$$\begin{aligned}
w^0 &= p + N^0 L^0 \sum_{i=1}^{K} \frac{\partial p}{\partial n^j} \frac{\partial n^j}{\partial \left(N^0 L^0\right)} = p\left(n^1, n^2, ..., n^K\right) - \sum_{i=1}^{K} n^j p'_j \left(n^1, n^2, ..., n^K\right), \\
w^1 &= N^0 L^0 \frac{\partial p}{\partial n^1} \frac{\partial n^1}{\partial \left(N^1 L^1\right)} = p'_1 \left(n^1, n^2, ..., n^K\right), \\
w^2 &= N^0 L^0 \frac{\partial p}{\partial n^2} \frac{\partial n^2}{\partial \left(N^2 L^2\right)} = p'_2 \left(n^1, n^2, ..., n^K\right), \\
&\vdots \\
w^K &= N^0 L^0 \frac{\partial p}{\partial n^K} \frac{\partial n^K}{\partial \left(N^K L^K\right)} = p'_K \left(n^1, n^2, ..., n^K\right).
\end{aligned}$$

Next define,

$$\begin{aligned}
g_0\left(n^1, n^2, ..., n^K\right) &\equiv p\left(n^1, n^2, ..., n^K\right) - \sum_{i=1}^{K} n^j p'_j \left(n^1, n^2, ..., n^K\right), \\
g_i\left(n^1, n^2, ..., n^K\right) &\equiv p'_i \left(n^1, n^2, ..., n^K\right), \quad i = 1, 2, \ldots, K-1.
\end{aligned}$$

The constrained Pareto-efficient allocations can then be characterized by the first-order conditions of a government problem summarized by the Lagrangian:

$$
\begin{aligned}
\mathcal{L} \;=\;& \sum_{i=0}^{K}\alpha^{i}V\left(M^{i},B^{i},w^{i}\right) \\
&+\gamma\left\{P\left(N^{0}\frac{M^{0}+a^{0}}{w^{0}},N^{1}\frac{M^{1}+a^{1}}{w^{1}},...,N^{K}\frac{M^{K}+a^{K}}{w^{K}}\right)-\sum_{i=0}^{K}\left(B^{0}+a^{0}\right)N^{0}\right\} \\
&+\sum_{i=1}^{K}\lambda_{i,i-1}\left[V\left(M^{i},B^{i},w^{i}\right)-V\left(M^{i-1},B^{i-1},w^{i}\right)\right] \\
&+\sum_{i=0}^{K-1}\phi_{i}\left[w^{i}-g_{i}\left(n^{1},n^{2},...,n^{K}\right)\right],
\end{aligned}
$$

where in writing the set of self-selection constraints we have assumed that the binding ones are those linking downward pair of adjacent types.[18] Following the same approach we used to derive (11)-(12), and denoting $V\left(M^{i},B^{i},w^{i}\right)$ and $V\left(M^{i-1},B^{i-1},w^{i}\right)$ by V^{i} and $V^{i,i-1}$, one would end up with the following expressions for $T'\left(M^{K}\right)$, $T'\left(M^{0}\right)$, and $T'\left(M^{j}\right)$ for $j\in\{1,2,...,K-1\}$:[19]

$$
\begin{aligned}
T'\left(M^{0}\right) =& \frac{\lambda_{1,0}}{\gamma N^{0}}\left(MRS_{MB}^{0}-MRS_{MB}^{1,0}\right) \\
&-\frac{\Delta^{0}}{\gamma w^{0}N^{0}L^{0}}\sum_{i=0}^{K-1}\phi_{i}\left[\sum_{j=1}^{K}\frac{\partial g_{i}\left(n^{1},n^{2},...,n^{K}\right)}{\partial n^{j}}n^{j}\right], \qquad (21)\\
T'\left(M^{j}\right) =& \frac{\lambda_{j+1,j}}{\gamma N^{j}}\left(MRS_{MB}^{j}-MRS_{MB}^{j+1,j}\right) \\
&+\frac{\Delta^{j}}{\gamma w^{j}N^{j}L^{j}}\left[\sum_{i=0}^{K-1}\phi_{i}\frac{\partial g_{i}\left(n^{1},n^{2},...,n^{K}\right)}{\partial n^{j}}n^{j}\right], \quad j=1,2,...,K-1, \qquad (22)\\
T'\left(M^{K}\right) =& \frac{\Delta^{K}}{\gamma w^{K}N^{K}L^{K}}\left[\sum_{i=0}^{K-1}\phi_{i}\frac{\partial g_{i}\left(n^{1},n^{2},...,n^{K}\right)}{\partial n^{K}}n^{K}\right]. \qquad (23)
\end{aligned}
$$

One can also easily show that when tax evasion is not possible the optimal marginal income tax rates are again characterized by (21)–(23) except that $\Delta^{j}=1$ for all

[18] See Guesnerie and Seade (1982) for a justification of this assumption.

[19] The detailed derivations are available upon request.

$j = 0, 1, \ldots, K$. And, with Δ^j being defined as previously by (13), the optimal marginal income tax rates tend to decrease more and more as $f(a)$ becomes less and less convex ($f''(a)$ becomes smaller). In words, the extent to which the government uses the marginal income tax rates as an instrument to affect redistribution via general equilibrium effects depends on the curvature of the evasion cost function: a lower curvature weakens the incentives for the government to manipulate the marginal income tax rates to induce general equilibrium effects on wages. One should nevertheless point out here that with more than two factors of production the sign of expressions like $\partial g_i/\partial n^j$ will depend on the complementarity/substitutability relationship between factors of production. This also prevents one to find a determinate sign for all ϕ_is. Nevertheless the fact that $\Delta^j < 1$, our conclusions about the effect of tax evasion remain valid.

Finally, observe that in the case when the government can affect the concealment costs, one can also find a generalized expression corresponding to (17). This is given by,

$$\sum_{i=1}^{K}\lambda_{i,i-1}\left[f_R\left(a^{i,i-1},R\right)-f_R\left(a^{i-1},R\right)\right]+\sum_{i=0}^{K-1}\phi_i\left[\sum_{j=1}^{K}\frac{\partial g_i}{\partial n_j}n_j\left(\frac{w^j f_{aR}\left(a^j,R\right)}{\Psi^j L^j}-\frac{w^0 f_{aR}\left(a^0,R\right)}{\Psi^0 L^0}\right)\right]$$
$$=\gamma\sum_{i=0}^{K}N^i\left(1+f_R\left(a^i,R\right)\right).$$

and has the same interpretation as previously.

6 Numerical illustrations

This section illustrates our results through a number of numerical examples based on four types. There are low-skilled and high-skilled agents who are imperfect substitutes with varying wages. At the same time, within each skill type, there are two different types of workers who are perfect substitutes but have different productivities: one type produces more than the other for one hour of labor but at a fixed rate. Specifically, assume that the production function is of the CES type and given by,

$$Q=\left[\frac{1}{10}\left(N^{\ell 1}L^{\ell 1}+1.1N^{\ell 2}L^{\ell 2}\right)^{\frac{\sigma-1}{\sigma}}+\frac{9}{10}\left(N^{h1}L^{h1}+1.4N^{h2}L^{h2}\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}, \qquad (24)$$

where σ denotes the elasticity of substitution between high- and low-skilled workers, the superscripts $\ell 1$ and $\ell 2$ denote the less and the more productive workers among the low-skilled agents, and the superscripts $h1$ and $h2$ denote the less and the more productive workers among the high-skilled agents.[20] We assume that there is an equal proportion of agents of each skill type and set $N^{\ell 1} = N^{\ell 2} = N^{h1} = N^{h2} = 1$.

All agents have identical preferences represented by:

$$U = \ln c - \frac{L^2}{2}. \tag{25}$$

Regarding the government's objective function we will consider both the case of a purely utilitarian objective and the case of a maxi-min social welfare function. Finally, assume that the evasion function is given by:

$$f(a, R) = \beta a^2 + 25a^2 R, \tag{26}$$

where β is a parameter affecting the curvature of $f(a, R)$, and R is per capita government expenditures to fight evasion (set equal to such expenditures/4 in our computations).

6.1 Utilitarian objective

We begin by solving for the case of a purely utilitarian social welfare function, i.e. the case when the government maximizes $W = \Sigma_{j=\ell,h}\Sigma_{i=1,2}N^{ji}U^{ji}$. We set $\sigma = 1/2$ in (24) and initially assume that R is not a policy instrument (i.e. we set $R = 0$). Table 1 reports, for the same model specification, the optimal marginal income tax rates, $MITR$, the resulting wages, and the associated utility levels when the possibility of tax evasion does not exist (the first three columns) and when it does (the last three columns with $\beta = 1$ in (26)).

[20] As usual, $\sigma = 0$ corresponds to the Leontief case and perfect substitutability is attained when σ approaches ∞.

Comparing the results, one observes that the optimal marginal tax rates are more compressed when agents have the possibility to evade. Specifically, tax evasion lowers the marginal income tax rates for all types (in absolute value). Additionally, tax evasion leads to lower wage rates and lower utility levels for low-skilled workers ($\ell 1$- and $\ell 2$ types), and higher wage rates and higher utility levels for high-skilled workers ($h1$- and $h2$ types). This reverse redistribution lowers the utilitarian social welfare. The outcome corroborates the first point we made in subsection 3.3 regarding the effects of tax evasion in our model. *The possibility of tax evasion weakens the incentive for the government to manipulate the marginal tax rates for the purpose of exploiting general equilibrium effects on wages.*

Second, we examine the effects of the curvature of concealment cost on the optimal marginal income tax rates (the second point stated in subsection 3.3). To do so, we lower β from 1 to $\beta = 0.1$ making the concealment cost function less convex. The results for this case are reported in the first three columns of Table 2. Comparing these results with those in the last three columns of Table 1, indicates that the lowering of the curvature of the concealment cost function leads to drastic reductions in the values of the marginal income tax rates faced by $\ell 1$-, $\ell 2$, and $h1$-types. This is also in line with our theoretical result that the less convex is the f-function the weaker will be the incentive for the government to manipulate the marginal tax rates. Additionally, a less convex concealment cost function results in lower wages and utility levels for the low-skilled workers ($\ell 1$- and $\ell 2$ types) and higher wages and utility levels for the high-skilled workers ($h1$- and $h2$ types). The utilitarian social welfare too is further reduced.

Third, we consider how government expenditure to deter tax evasion, R, affects our results. The only thing we could say about this theoretically was that for R to be positive, it must be possible to increase redistribution from the high-skilled towards the low-skilled group of workers. The figures in the last three columns of Table 2 report our findings when R is a policy instrument. These figures show that R has indeed an interior solution. Moreover, comparisons with the corresponding figures in the first three columns show that spending resources to deter tax evasion enables the government to raise social welfare by strengthening redistribution from the high-skilled towards the low-skilled group of workers (in line with the observation we made in Section 4). By making it more costly for the agents to evade, the government finds it more effective to raise the marginal income tax rates (in absolute value). Which allows it to exploit, to a larger extent, the tax-induced general-equilibrium effects on wages.

Finally, to illustrate how results are affected when R is allowed to be income-dependent, we use the same identical evasion cost function (26) while replacing per capita deterrence cost R by $R\left(M^j\right)$, the per capita value for the group reporting income M^j. The results are reported in Table 3. They are to be contrasted with the corresponding results in Table 2 in the case where the governments attempts to deter evasion in a non-targeted way.

The first thing to note is the fact that the utilitarian social welfare has now increased from -1.101 to -1.079. This accords with one's expectation in that allowing $R\left(M^j\right)$ to vary across income types can never decrease social welfare. The relevant point to highlight, however, is the finding that the evasion-deterrence efforts of the government are primarily targeted towards agents reporting a low level of income. Moreover, no resources are to be spent in this case on thwarting the evasion possibilities of agents reporting the highest level of income (i.e. we obtain a corner solution for $R\left(M^{h2}\right)$). This is in accordance with the theoretical results established in the literature [see, e.g., Cremer and Gahvari (1996)].

6.2 Maxi-min objective

To further illustrate our findings we redo our computations for a maxi-min social welfare function, i.e. the case when the government maximizes $W = U^{\ell 1}$. We again set $\sigma = 1/2$ and start first with the case when R is not a policy instrument ($R = 0$). The resulting optimal marginal income tax rates, $MITR$, wages, and the associated utility levels are reported in Table 3 (with the first three columns corresponding to the no tax evasion case and the last three columns to when there is evasion with $\beta = 1$). The results show an even higher compression of marginal income tax rates as compared to the utilitarian case thus corroborating our theoretical comparisons in subsection 3.3. The absolute value of the marginal income tax rates decrease drastically for all types (particularly the low-skilled types). Additionally, the possibility of tax evasion lowers the wages and the utility levels of $\ell 1$- and $\ell 2$-types while increasing the wages and utility levels of the $h1$- and $h2$-types. With the utility level of the low-skilled workers decreasing, the possibility of tax evasion lowers social welfare.

Second, to re-examine the effects of the curvature of concealment cost, we again make the function less convex by lowering β from 1 to $\beta = 0.1$ and reporting the results in the first three columns of Table 4. Comparing these results with those in the last three columns of Table 3 indicates that lowering the curvature lowers the marginal income tax rates quite substantially for all type (in absolute values). Once again this corroborates the second point we made in subsection 3.3. In particular, the less convex is the f-function and the weaker is the incentive for the government to manipulate the marginal tax rates for the purpose of exploiting general equilibrium effects on wages. Additionally, the wages and utility levels of $\ell 1$- and $\ell 2$- types diminish but the wages and the utility levels of the $h1$- and $h2$-types improve. Social welfare decreases.

Third, we again consider the effects of government attempting to deter tax evasion. The figures in the last three columns of Table 4 report our findings when R is a policy instrument. Comparisons with the corresponding figures in the first three columns show that spending resources to fight tax evasion enables the government to raise social welfare by strengthening redistribution towards the low-skilled group of workers. By making it more costly for the agents to evade, the government finds it more effective to raise the marginal income tax rates (in absolute value). This allows it to exploit, to a larger extent, the tax-induced general-equilibrium effects on wages and increase social welfare.

Finally, we again allow for R to be income-dependent and use the same identical evasion cost function (26) to determine how the tax authority will spend its resources. The results are reported in Table 6. They are to be contrasted with the corresponding results in Table 5 in the case where the governments attempts to deter evasion in a non-targeted way.

The maxi-min social welfare now increases from -1.284 to -1.252 as a result of allowing $R\left(M^j\right)$ to vary across income types. Again note that the evasion-deterrence efforts of the government are primarily targeted towards agents reporting a low level of income. Moreover, no resources are spent in this case too on thwarting the evasion possibilities of agents reporting the highest level of income.[21]

7 Conclusion

This paper has studied, within a Mirrleesian optimal tax framework, the interaction between tax evasion and wage endogeneity (due to the effects of taxation on gross-of-tax wages). Prior studies on optimal taxation with endogenous wages have neglected tax evasion; while prior studies of tax evasion have neglected the endogeneity of wages. We appended a production function, that has skilled and unskilled workers as inputs, to the two-group model of Stiglitz (1982). Using this setup, and in conjunction with the riskless approach of Usher (1986) to tax evasion, we have derived a characterization for the optimal marginal income tax rates on the skilled- and the unskilled workers and for the optimal amount of resources that the government should spend to deter tax evasion. The results have then been generalized to a many-skilled group model. Finally, we have run a few numerical simulations to illustrate the magnitude of the phenomena that we have studied.

Among the findings of the paper are: (i) the possibility of tax evasion weakens the

[21] Running many simulations, we find that our result on the government targeting lower-income types, in its attempt to deter evasion, is robust for the maxi-min criterion as well as the utilitarian case. However, the zero expenditure on the highest income type can change. For example, if the coefficient of a^2R in the evasion cost function changes from 25 to 250, implying that the evasion-deterrence efforts of the government prove to be more effective in raising the private marginal cost of evasion, $R\left(M^j\right)$ in the maxi-min case will decline from 25.94 to 20.88 to 15.61 and finally to 2.23 as income increases.

incentives for the government to manipulate the marginal tax rates for the purpose of exploiting general equilibrium effects on wages; (ii) the effect of tax evasion on the optimal marginal income tax rates depends on the convexity of the concealment cost function: the tax rate decreases as the cost function becomes less convex; (iii) optimally spending resources to deter evasion strengthens the redistributive potential of taxation via the general-equilibrium effects on wages resulting in higher marginal income tax rates (in absolute value); and the evasion-deterrence efforts of the government are primarily targeted towards agents reporting a low level of income.

Appendix

Proof of Lemma 1. *Part (i).* Recall that the agent monotonicity condition in (wL, c)-space requires

$$\frac{d}{dw}\left(\frac{\partial c}{\partial wL}\right) < 0.$$

But $c = a + B - f(a)$ and $wL = M + a$. Consequently, given any point (M, B), we have

$$\frac{d}{dw}\left(\frac{\partial\left(a + B - f(a)\right)}{\partial\left(M + a\right)}\right) = \frac{d}{dw}\left(1 - f'(a)\right) < 0. \tag{A1}$$

Recall, from (9)–(10), that $1 - MRS_{MB} = f'(a)$. It thus follows from inequality (A1) that

$$\frac{d}{dw} MRS_{MB} < 0,$$

which is the single-crossing property in the (M, B)-space.

Part (ii). The single-crossing property proved in Part (i) implies

$$MRS^{h}_{MB}\left(M^{\ell}, B^{\ell}\right) < MRS^{\ell}_{MB}\left(M^{\ell}, B^{\ell}\right)$$

On the basis of this inequality and the fact that $1 - MRS_{MB} = f'(a)$, we have $1 - f'\left(a^{h\ell}\right) < 1 - f'\left(a^{\ell}\right)$ or $f'\left(a^{h\ell}\right) > f'\left(a^{\ell}\right)$. It then follows from the convexity of $f(\cdot)$ that,

$$a^{h\ell} > a^{\ell}. \tag{A2}$$

Proof of Lemma 2. We first derive a number of algebraic expressions that are subsequently used in the proof. These expressions also simplify the first-order conditions of the government's problem summarized by the Lagrangian (8) and the proof of Proposition 1.

Step (i): Algebraic derivations to be used in the proofs.

- *The effects of a change in B^j, M^j, and w^j on a j-type's choice of L^j.* Implicitly differentiating equation (6) partially with respect to B^j, M^j, and w^j yields, for all $j = h, \ell$,

$$\begin{aligned} \frac{\partial L^j}{\partial B^j} &= 0, & \text{(A3)} \\ \frac{\partial L^j}{\partial M^j} &= \frac{w^j f''}{v'' + (w^j)^2 f''} > 0, & \text{(A4)} \\ \frac{\partial L^j}{\partial w^j} &= \frac{1 - f' - w^j L^j f''}{v'' + (w^j)^2 f''}. & \text{(A5)} \end{aligned}$$

  Observe also that in the case of a mimicker the first-order condition (6) takes the form

$$v'\left(L^{jk}\right) = w^j \left[1 - f'\left(w^j L^{jk} - M^k\right)\right]$$

  so that its implicit differentiation yields

$$\frac{\partial L^{jk}}{\partial w^j} = \frac{1 - f'\left(a^{jk}\right) - w^j L^{jk} f''\left(a^{jk}\right)}{v''\left(L^{jk}\right) + (w^j)^2 f''\left(a^{jk}\right)}. \tag{A6}$$

- *The effects of a change in B^j, M^j, and w^j on a j-type's choice of a^j.* Recall that individuals' earnings consist of reported and not-reported incomes. That is, $w^j L^j = M^j + a^j$. Partially differentiating this relationship with respect to B^j and M^j and substituting from (A3)–(A5) for $\partial L^j/\partial B^j, \partial L^j/\partial M^j$, and $\partial L^j/\partial w^j$ yields,

$$\begin{aligned} \frac{\partial a^j}{\partial B^j} &= 0, & \text{(A7)} \\ \frac{\partial a^j}{\partial M^j} &= \frac{(w^j)^2 f''}{v'' + (w^j)^2 f''} - 1 = \frac{-v''}{v'' + (w^j)^2 f''} < 0. & \text{(A8)} \\ \frac{\partial a^j}{\partial w^j} &= L^j + w^j \frac{\partial L^j}{\partial w^j} = \frac{v'' L^j + w^j (1 - f')}{v'' + (w^j)^2 f''} & \text{(A9)} \end{aligned}$$

  As to the mimicker, it follows from partial differentiation of $w^j L^{jk} = a^{jk} + M^k$ with respect to w^j that

$$\frac{\partial a^{jk}}{\partial w^j} = L^{jk} + w^j \frac{\partial L^{jk}}{\partial w^j} = \frac{L^{jk} v''\left(L^{jk}\right) + w^j \left[1 - f'\left(a^{jk}\right)\right]}{v''\left(L^{jk}\right) + (w^j)^2 f''\left(a^{jk}\right)}. \tag{A10}$$

- *The effects of a change in B^j and M^j on n.* Partially differentiate $n \equiv N^h L^h / N^\ell L^\ell$ with respect to B^h, M^h, B^ℓ, and M^ℓ. We have,

$$\begin{aligned}
\frac{\partial n}{\partial B^h} &= \frac{N^h}{N^\ell L^\ell}\frac{\partial L^h}{\partial B^h} = 0, && \text{(A11)}\\
\frac{\partial n}{\partial M^h} &= \frac{N^h}{N^\ell L^\ell}\frac{\partial L^h}{\partial M^h} = \frac{n}{L^h}\frac{w^h f''\left(a^h\right)}{v''\left(L^h\right) + \left(w^h\right)^2 f''\left(a^h\right)}, && \text{(A12)}\\
\frac{\partial n}{\partial B^\ell} &= \frac{N^h L^h}{N^\ell}\frac{\partial}{\partial B^\ell}\left(\frac{1}{L^\ell}\right) = 0, && \text{(A13)}\\
\frac{\partial n}{\partial M^\ell} &= \frac{N^h L^h}{N^\ell}\frac{\partial}{\partial M^\ell}\left(\frac{1}{L^\ell}\right) = \frac{N^h L^h}{N^\ell}\left[\frac{-1}{\left(L^\ell\right)^2}\right]\frac{w^\ell f''\left(a^\ell\right)}{v''\left(L^\ell\right) + \left(w^\ell\right)^2 f''\left(a^\ell\right)},\\
&= \left(\frac{-n}{L^\ell}\right)\frac{w^\ell f''\left(a^\ell\right)}{v''\left(L^\ell\right) + \left(w^\ell\right)^2 f''\left(a^\ell\right)}, && \text{(A14)}
\end{aligned}$$

where we have substituted the expressions for $\partial L^h / \partial B^h$ and $\partial L^h / \partial M^h$ from (A3)–(A4) and the expressions for $\partial L^\ell / \partial B^\ell$ and $\partial L^\ell / \partial M^\ell$ from (A3)–(A4).

- *The effect of a change in w^ℓ on n.* Partially differentiate $n \equiv N^h L^h / N^\ell L^\ell$ with respect to w^ℓ to get,

$$\frac{\partial n}{\partial w^\ell} = \frac{N^h}{N^\ell}\frac{\partial}{\partial w^\ell}\left(\frac{L^h}{L^\ell}\right) = \frac{N^h}{N^\ell}\left[\frac{L^\ell \frac{\partial L^h}{\partial w^h}\frac{\partial w^h}{\partial w^\ell} - L^h \frac{\partial L^\ell}{\partial w^\ell}}{\left(L^\ell\right)^2}\right] = n\left(\frac{1}{L^h}\frac{\partial L^h}{\partial w^h}\frac{\partial w^h}{\partial w^\ell} - \frac{1}{L^\ell}\frac{\partial L^\ell}{\partial w^\ell}\right)$$

Substitute the expressions for $\partial L^h / \partial w^h$ and $\partial L^\ell / \partial w^\ell$ from (A5) into above and simplify:

$$\begin{aligned}
\frac{\partial n}{\partial w^\ell} &= n\left[\frac{1}{L^h}\frac{1 - f'\left(a^h\right) - w^h L^h f''\left(a^h\right)}{v''\left(L^h\right) + \left(w^h\right)^2 f''\left(a^h\right)}\frac{\partial w^h}{\partial w^\ell} - \frac{1}{L^\ell}\frac{1 - f'\left(a^\ell\right) - w^\ell L^\ell f''\left(a^\ell\right)}{v''\left(L^\ell\right) + \left(w^\ell\right)^2 f''\left(a^\ell\right)}\right]\\
&= n\frac{1}{w^h L^h}\frac{w^h\left[1 - f'\left(a^h\right)\right] - \left(w^h\right)^2 L^h f''\left(a^h\right)}{v''\left(L^h\right) + \left(w^h\right)^2 f''\left(a^h\right)}\frac{\partial w^h}{\partial w^\ell}\\
&\quad -n\frac{1}{w^\ell L^\ell}\frac{w^\ell\left[1 - f'\left(a^\ell\right)\right] - \left(w^\ell\right)^2 L^\ell f''\left(a^\ell\right)}{v''\left(L^\ell\right) + \left(w^\ell\right)^2 f''\left(a^\ell\right)}. && \text{(A15)}
\end{aligned}$$

Introducing the notation

$$\Psi^j \equiv v''\left(L^j\right) + \left(w^j\right)^2 f''\left(a^j\right) > 0, \quad \text{for } j = \ell, h \qquad \text{(A16)}$$

and recalling the definition of Δ^j (for $j = \ell, h$) provided by (13), one can rewrite (A15) as

$$\begin{aligned}
\frac{\partial n}{\partial w^\ell} &= n\left\{\frac{1}{w^h L^h}\left[\frac{w^h\left[1-f'\left(a^h\right)\right]}{\Psi^h} - L^h\Delta^h\right]\frac{\partial w^h}{\partial w^\ell} - \frac{1}{w^\ell L^\ell}\left[\frac{w^\ell\left[1-f'\left(a^\ell\right)\right]}{\Psi^\ell} - L^\ell\Delta^\ell\right]\right\} \\
&= n\left\{\left[\frac{1-f'\left(a^h\right)}{L^h\Psi^h} - \frac{\Delta^h}{w^h}\right]\frac{\partial w^h}{\partial w^\ell} - \left[\frac{1-f'\left(a^\ell\right)}{L^\ell\Psi^\ell} - \frac{\Delta^\ell}{w^\ell}\right]\right\}. \qquad \text{(A17)}
\end{aligned}$$

To further simplify (A17), observe that,

$$\frac{\partial w^h}{\partial w^\ell} = \frac{\partial w^h/\partial n}{\partial w^\ell/\partial n} = \frac{p''(n)}{p'(n) - p'(n) - np''(n)} = \frac{-1}{n}, \qquad \text{(A18)}$$

where $\partial w^h/\partial n$ and $\partial w^\ell/\partial n$ are found from differentiating relationships (4)–(3). Substituting in (A17) and further simplification results in

$$\frac{\partial n}{\partial w^\ell} = -\left[\frac{1-f'\left(a^h\right)}{L^h\Psi^h} + n\frac{1-f'\left(a^\ell\right)}{L^\ell\Psi^\ell} - \frac{\Delta^h}{w^h} - \frac{n\Delta^\ell}{w^\ell}\right]. \qquad \text{(A19)}$$

- *The effects of a change in B^j, B^k, M^j, M^k, w^j, and w^k on $V^j \equiv V\left(M^j, B^j; w^j\right)$ and $V^{jk} \equiv V\left(M^k, B^k; w^j\right)$.* Recall that $V(M, B; w)$ denotes the maximum value function of problem (5) with the associated notation $V^j \equiv V\left(M^j, B^j; w^j\right)$ and $V^{jk} \equiv V\left(M^k, B^k; w^j\right)$. Invoking the envelop theorem, we have, for all $j = h, \ell$,

$$\begin{aligned}
\frac{\partial V^j}{\partial B^j} &= 1, & \text{(A20)} \\
\frac{\partial V^{jk}}{\partial B^k} &= 1, & \text{(A21)} \\
\frac{\partial V^j}{\partial M^j} &= -\left[1-f'\left(a^j\right)\right] & \text{(A22)} \\
\frac{\partial V^{jk}}{\partial M^k} &= -\left[1-f'\left(a^{jk}\right)\right], & \text{(A23)} \\
\frac{\partial V^j}{\partial w^j} &= L^j\left[1-f'\left(a^j\right)\right], & \text{(A24)} \\
\frac{\partial V^{jk}}{\partial w^j} &= L^{jk}\left[1-f'\left(a^{jk}\right)\right]. & \text{(A25)}
\end{aligned}$$

Step (ii): Proof of the negativity of ϕ. We have,

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial w^\ell} &= (1+\lambda)\frac{\partial V^h}{\partial w^h}\frac{\partial w^h}{\partial w^\ell} + \mu\frac{\partial V^\ell}{\partial w^\ell} - \lambda\frac{\partial V^{h\ell}}{\partial w^h}\frac{\partial w^h}{\partial w^\ell} \\
&\quad +\gamma\left\{w^\ell N^\ell \frac{\partial L^\ell}{\partial w^\ell} + w^h N^h \frac{\partial L^h}{\partial w^h}\frac{\partial w^h}{\partial w^\ell} - N^\ell\frac{\partial a^\ell}{\partial w^\ell} - N^h\frac{\partial a^h}{\partial w^h}\frac{\partial w^h}{\partial w^\ell}\right\} + \left(1 - g'\frac{\partial n}{\partial w^\ell}\right)\phi, \\
&= (1+\lambda) L^h\left[1 - f'\left(a^h\right)\right]\frac{-1}{n} + \mu L^\ell\left[1 - f'\left(a^\ell\right)\right] - \lambda L^{h\ell}\left[1 - f'\left(a^{h\ell}\right)\right]\frac{-1}{n} \\
&\quad +\gamma\left\{w^\ell N^\ell\frac{\partial L^\ell}{\partial w^\ell} + w^h N^h \frac{\partial L^h}{\partial w^h}\frac{-1}{n} - N^\ell\left(L^\ell + w^\ell\frac{\partial L^\ell}{\partial w^\ell}\right) - N^h\left(L^h + w^h\frac{\partial L^h}{\partial w^h}\right)\frac{-1}{n}\right\} \\
&\quad + \left(1 - g'\frac{\partial n}{\partial w^\ell}\right)\phi, \\
&= (1+\lambda) L^h\left[1 - f'\left(a^h\right)\right]\frac{-1}{n} + \mu L^\ell\left[1 - f'\left(a^\ell\right)\right] - \lambda L^{h\ell}\left[1 - f'\left(a^{h\ell}\right)\right]\frac{-1}{n} \\
&\quad +\gamma\left\{-N^\ell\left(L^\ell\right) - N^h\left(L^h\right)\frac{-1}{n}\right\} + \left(1 - g'\frac{\partial n}{\partial w^\ell}\right)\phi, \\
&= (1+\lambda) L^h\left[1 - f'\left(a^h\right)\right]\frac{-1}{n} + \mu L^\ell\left[1 - f'\left(a^\ell\right)\right] - \lambda L^{h\ell}\left[1 - f'\left(a^{h\ell}\right)\right]\frac{-1}{n} \\
&\quad + \left(1 - g'\frac{\partial n}{\partial w^\ell}\right)\phi. \qquad \text{(A26)}
\end{aligned}$$

The expressions involving λ and μ in (A26) can also be simplified. Multiply (A32) by $\left[1 - f'\left(a^h\right)\right] L^h/n$, multiply (A38) by $L^\ell\left[1 - f'\left(a^\ell\right)\right]$, and add the resulting equations together:

$$\left[1 + \lambda - \gamma N^h\right]\frac{-L^h}{n}\left[1 - f'\left(a^h\right)\right] + \left[\mu - \gamma N^\ell - \lambda\right] L^\ell\left[1 - f'\left(a^\ell\right)\right] = 0.$$

Moving all the expressions involving γ to the right-hand side,

$$\begin{aligned}
(1+\lambda)\frac{-L^h}{n}\left[1 - f'\left(a^h\right)\right] + (\mu - \lambda) L^\ell\left[1 - f'\left(a^\ell\right)\right] &= \gamma N^\ell L^\ell\left[1 - f'\left(a^\ell\right)\right] \\
&\quad -\gamma\frac{N^h L^h}{n}\left[1 - f'\left(a^h\right)\right] \\
&= \gamma N^\ell L^\ell\left[f'\left(a^h\right) - f'\left(a^\ell\right)\right]
\end{aligned}$$

Or

$$(1+\lambda)\frac{-L^h}{n}\left[1 - f'\left(a^h\right)\right] + \mu L^\ell\left[1 - f'\left(a^\ell\right)\right] = \lambda L^\ell\left[1 - f'\left(a^\ell\right)\right] + \gamma N^\ell L^\ell\left[f'\left(a^h\right) - f'\left(a^\ell\right)\right]$$

Then substitute it in (A26) to get

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial w^{\ell}} &= \lambda L^{\ell}\left[1-f'\left(a^{\ell}\right)\right]+\gamma N^{\ell}L^{\ell}\left[f'\left(a^{h}\right)-f'\left(a^{\ell}\right)\right]+\lambda\frac{L^{h\ell}}{n}\left[1-f'\left(a^{h\ell}\right)\right]+\left(1-g'\frac{\partial n}{\partial w^{\ell}}\right)\phi \\ &= \lambda\left\{L^{\ell}\left[1-f'\left(a^{\ell}\right)\right]+\frac{L^{h\ell}}{n}\left[1-f'\left(a^{h\ell}\right)\right]\right\}+\left(1-g'\frac{\partial n}{\partial w^{\ell}}\right)\phi+\gamma N^{\ell}L^{\ell}\left[f'\left(a^{h}\right)-f'\left(a^{\ell}\right)\right]\end{aligned} \tag{A27}$$

Now observe, from (10), that $f'\left(a^{j}\right)=T'\left(M^{j}\right)$ where the values of $T'\left(M^{h}\right)$ and $T'\left(M^{\ell}\right)$ are given by (11) and (12). Substituting these values in the last expression of (A27), we have

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial w^{\ell}} &= \lambda\left\{L^{\ell}\left[1-f'\left(a^{\ell}\right)\right]+\frac{L^{h\ell}}{n}\left[1-f'\left(a^{h\ell}\right)\right]\right\}+\left(1-g'\frac{\partial n}{\partial w^{\ell}}\right)\phi \\ &\quad+\gamma N^{\ell}L^{\ell}\left[\Delta^{h}\frac{\phi n g'}{\gamma w^{h}N^{h}L^{h}}-\frac{\lambda}{\gamma N^{\ell}}\left(MRS_{MB}^{\ell}-MRS_{MB}^{h\ell}\right)+\Delta^{\ell}\frac{\phi n g'}{\gamma w^{\ell}N^{\ell}L^{\ell}}\right]\end{aligned}$$

Substituting the expression for $\partial n/\partial w^{\ell}$ from (A19) into above and simplifying,

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial w^{\ell}} &= \lambda\left\{L^{\ell}\left[1-f'\left(a^{\ell}\right)\right]+\frac{L^{h\ell}}{n}\left[1-f'\left(a^{h\ell}\right)\right]\right\}+ \\ &\quad\left\{1+g'\left[\frac{1-f'\left(a^{h}\right)}{L^{h}\Psi^{h}}+n\frac{1-f'\left(a^{\ell}\right)}{L^{\ell}\Psi^{\ell}}-\frac{\Delta^{h}}{w^{h}}-n\frac{\Delta^{\ell}}{w^{\ell}}\right]\right\}\phi+ \\ &\quad\left[\Delta^{h}\frac{\phi g'}{w^{h}}-\lambda L^{\ell}\left(MRS_{MB}^{\ell}-MRS_{MB}^{h\ell}\right)+\Delta^{\ell}\frac{\phi n g'}{w^{\ell}}\right] \\ &= \lambda\left\{L^{\ell}\left[1-f'\left(a^{\ell}\right)\right]+\frac{L^{h\ell}}{n}\left[1-f'\left(a^{h\ell}\right)\right]\right\}+ \\ &\quad\left\{1+g'\left[\frac{1}{L^{h}}\frac{1-f'\left(a^{h}\right)}{\Psi^{h}}+n\frac{1}{L^{\ell}}\frac{1-f'\left(a^{\ell}\right)}{\Psi^{\ell}}\right]\right\}\phi-\lambda L^{\ell}\left(MRS_{MB}^{\ell}-MRS_{MB}^{h\ell}\right).\end{aligned}$$

Finally, substituting MRS_{MB}^{ℓ} for $1-f'\left(a^{\ell}\right)$, $MRS_{MB}^{h\ell}$ for $1-f'\left(a^{h\ell}\right)$, and $MRS_{MB}^{h\ell}$ for $1-f'\left(a^{h\ell}\right)$ in above results in

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial w^{\ell}} &= \lambda\left(L^{\ell}MRS_{MB}^{\ell}+\frac{L^{h\ell}}{n}MRS_{MB}^{h\ell}\right)+\left[1+g'\left(\frac{MRS_{MB}^{h}}{L^{h}\Psi^{h}}+n\frac{MRS_{MB}^{\ell}}{L^{\ell}\Psi^{\ell}}\right)\right]\phi \\ &\quad-\lambda L^{\ell}\left(MRS_{MB}^{\ell}-MRS_{MB}^{h\ell}\right).\end{aligned} \tag{A28}$$

Setting $\partial\mathcal{L}/\partial w^\ell = 0$ in (A28), one ends up with the following expression for ϕ,

$$\phi = -\frac{\lambda\left(\frac{L^{h\ell}}{n} + L^\ell\right) MRS^{h\ell}_{MB}}{1 + \left(\frac{MRS^h_{MB}}{L^h\Psi^h} + \frac{MRS^\ell_{MB}}{L^\ell\Psi^\ell} n\right) g'}, \tag{A29}$$

where $\Psi^j > 0$ was defined in (A16).

With $g' = p' - p' - np'' = -np'' > 0$, it follows that $\phi < 0$.

Proof of Proposition 1: We do this in two steps.

Step (i): Proof of equation (11). The first-order conditions of the government's problem, summarized by the Lagrangian (8), with respect to B^h and M^h are:

$$\frac{\partial V^h}{\partial B^h} + \gamma\left[P_h \frac{N^h}{w^h}\frac{\partial a^h}{\partial B^h} - N^h - N^h\frac{\partial a^h}{\partial B^h}\right] + \lambda\frac{\partial V^h}{\partial B^h} - \phi g'\frac{\partial n}{\partial B^h} = 0, \tag{A30}$$

$$\frac{\partial V^h}{\partial M^h} + \gamma\left[P_h \frac{N^h}{w^h}\left(1 + \frac{\partial a^h}{\partial M^h}\right) - N^h\frac{\partial a^h}{\partial M^h}\right] + \lambda\frac{\partial V^h}{\partial M^h} - \phi g'\frac{\partial n}{\partial M^h} = 0. \tag{A31}$$

Set $P_h = w^h$ and substitute from (A7)–(A8), (A11)–(A12), (A20) and (A22) in (A30)–(A31). Simplify and rearrange the terms to rewrite equations (A30)–(A31) as

$$1 + \lambda = \gamma N^h, \tag{A32}$$

$$(1+\lambda)\frac{\partial V^h}{\partial M^h} = \frac{n}{L^h}\frac{w^h f''\left(a^h\right)}{v''\left(L^h\right) + \left(w^h\right)^2 f''\left(a^h\right)}\phi g' - \gamma N^h. \tag{A33}$$

Substitute for $1+\lambda$ from (A32) into (A33) to get

$$\gamma N^h\left(1 + \frac{\partial V^h}{\partial M^h}\right) = \frac{n}{L^h}\frac{w^h f''\left(a^h\right)}{v''\left(L^h\right) + \left(w^h\right)^2 f''\left(a^h\right)}\phi g'. \tag{A34}$$

Using the definition of the marginal income tax rate in (9), one can rewrite (A34) as,

$$T'\left(M^h\right) = \frac{w^h f''\left(a^h\right)}{v''\left(L^h\right) + \left(w^h\right)^2 f''\left(a^h\right)}\frac{\phi n g'}{\gamma N^h L^h}, \tag{A35}$$

which, using the definition of Δ^h, is identical to (11).

Step (ii): Proof of equation (12). The first-order conditions of the government's problem with respect to B^ℓ and M^ℓ are:

$$\mu\frac{\partial V^\ell}{\partial B^\ell}+\gamma\left[P_\ell\frac{N^\ell}{w^\ell}\frac{\partial a^\ell}{\partial B^\ell}-N^\ell-N^\ell\frac{\partial a^\ell}{\partial B^\ell}\right]-\lambda\frac{\partial V^{h\ell}}{\partial B^\ell}-\phi g'\frac{\partial n}{\partial B^\ell}=0, \tag{A36}$$

$$\mu\frac{\partial V^\ell}{\partial M^\ell}+\gamma\left[P_\ell\frac{N^\ell}{w^\ell}\left(1+\frac{\partial a^\ell}{\partial M^\ell}\right)-N^\ell\frac{\partial a^\ell}{\partial M^\ell}\right]-\lambda\frac{\partial V^{h\ell}}{\partial M^\ell}-\phi g'\frac{\partial n}{\partial M^\ell}=0. \tag{A37}$$

set $P_\ell=w^\ell$ and substitute from (A7)–(A8), (A13)–(A14), and (A20)–(A23) into (A36)–(A37). Simplify and rearrange the terms to rewrite equations (A36)–(A37) as,

$$\mu=\gamma N^\ell+\lambda, \tag{A38}$$

$$\mu\frac{\partial V^\ell}{\partial M^\ell}=-\gamma N^\ell+\lambda\frac{\partial V^{h\ell}}{\partial M^\ell}-\frac{\phi n g'}{L^\ell}\frac{w^\ell f''\left(a^\ell\right)}{v''\left(L^\ell\right)+\left(w^\ell\right)^2 f''\left(a^\ell\right)}. \tag{A39}$$

Substitute for μ from (A38) into (A39) and rearrange the terms to get,

$$\gamma N^\ell\left(1+\frac{\partial V^\ell}{\partial M^\ell}\right)=\lambda\left(\frac{\partial V^{h\ell}}{\partial M^\ell}-\frac{\partial V^\ell}{\partial M^\ell}\right)-\frac{\phi n g'}{L^\ell}\frac{w^\ell f''\left(a^\ell\right)}{v''\left(L^\ell\right)+\left(w^\ell\right)^2 f''\left(a^\ell\right)}.$$

Using the definition of the marginal income tax rate in (9), one can rewrite the above equation as

$$\begin{aligned}T'\left(M^\ell\right)&=\frac{1}{\gamma N^\ell}\left[\lambda\left(\frac{\partial V^{h\ell}}{\partial M^\ell}-\frac{\partial V^\ell}{\partial M^\ell}\right)-\frac{\phi n g'}{L^\ell}\frac{w^\ell f''\left(a^\ell\right)}{v''\left(L^\ell\right)+\left(w^\ell\right)^2 f''\left(a^\ell\right)}\right]\\&=\frac{\lambda}{\gamma N^\ell}\left(MRS^\ell_{MB}-MRS^{h\ell}_{MB}\right)-\frac{w^\ell f''\left(a^\ell\right)}{v''\left(L^\ell\right)+\left(w^\ell\right)^2 f''\left(a^\ell\right)}\frac{\phi n g'}{\gamma N^\ell L^\ell},\end{aligned} \tag{A40}$$

which, using the definition of Δ^ℓ, is identical to (12).

Derivation of equations (14)–(15). In the absence of tax evasion, the government's problem continues to be summarized by the Lagrangian (8) except that in this case $a=0$. Nor is there the taxpayer's optimization problem (5) and the associated first-order condition (6). The first-order conditions of the government's problem with respect to B^h, M^h, B^ℓ, and M^ℓ, equations (A30)–(A31) and (A36)–(A37), simplify to

$$\frac{\partial V^h}{\partial B^h} - \gamma N^h + \lambda \frac{\partial V^h}{\partial B^h} - \phi g' \frac{\partial n}{\partial B^h} = 0, \tag{A41}$$

$$\frac{\partial V^h}{\partial M^h} + \gamma N^h + \lambda \frac{\partial V^h}{\partial M^h} - \phi g' \frac{\partial n}{\partial M^h} = 0,$$

$$\mu \frac{\partial V^\ell}{\partial B^\ell} - \gamma N^\ell - \lambda \frac{\partial V^{h\ell}}{\partial B^\ell} - \phi g' \frac{\partial n}{\partial B^\ell} = 0, \tag{A42}$$

$$\mu \frac{\partial V^\ell}{\partial M^\ell} + \gamma N^\ell - \lambda \frac{\partial V^{h\ell}}{\partial M^\ell} - \phi g' \frac{\partial n}{\partial M^\ell} = 0. \tag{A43}$$

As previously with tax evasion, the expressions for $\partial V^j/\partial B^j, \partial V^{h\ell}/\partial B^\ell, \partial V^j/\partial M^j$, and $\partial V^{h\ell}/\partial M^\ell$, for $j = h, \ell$, are given by equations (A20)–(A23). Similarly, we continue to have $\partial n/\partial B^h = \partial n/\partial B^\ell = 0$ but the expressions for $\partial n/\partial M^h$ and $\partial n/\partial M^\ell$ now simplify to:

$$\begin{aligned}
\frac{\partial n}{\partial M^h} &= \frac{\partial}{w^h \partial L^h}\left(\frac{N^h L^h}{N^\ell L^\ell}\right) = \frac{N^h}{w^h N^\ell L^\ell} = \frac{n}{w^h L^h}, \\
\frac{\partial n}{\partial M^\ell} &= \frac{N^h L^h}{N^\ell} \frac{\partial}{w^\ell \partial L^\ell}\left(\frac{1}{L^\ell}\right) = \frac{N^h L^h}{w^\ell N^\ell} \frac{-1}{(L^\ell)^2} = \frac{-n}{w^\ell L^\ell}.
\end{aligned}$$

Substituting all these expressions in (A41)–(A43) they simplify to,

$$\begin{aligned}
& 1 + \lambda = \gamma N^h, \\
& (1+\lambda) \frac{\partial V^h}{\partial M^h} + \gamma N^h = \phi g' \frac{n}{w^h L^h}, \\
& \mu = \gamma N^\ell + \lambda, \\
& \mu \frac{\partial V^\ell}{\partial M^\ell} + \gamma N^\ell - \lambda \frac{\partial V^{h\ell}}{\partial M^\ell} - \phi g' \left(\frac{-n}{w^\ell L^\ell}\right) = 0.
\end{aligned}$$

Eliminating $1 + \lambda$ between the first two equations and μ between the last equations yield,

$$\begin{aligned}
\gamma N^h \left(1 + \frac{\partial V^h}{\partial M^h}\right) &= \phi g' \frac{n}{w^h L^h}, \\
\gamma N^\ell \left(1 + \frac{\partial V^\ell}{\partial M^\ell}\right) &= \lambda \left(\frac{\partial V^{h\ell}}{\partial M^\ell} - \frac{\partial V^\ell}{\partial M^\ell}\right) - \phi g' \left(\frac{n}{w^\ell L^\ell}\right).
\end{aligned}$$

Then, from the definitions of the marginal income tax rates and the marginal rate of substitution between consumption and income, one can rewrite the above equations as equations (14)–(15) in the text.

Proof of Proposition 2. With R being endogenously determined, the optimization problem (5) is amended as

$$V(M,B,R;w) \equiv \left\{ \max_L \quad U = wL - M + B - f(wL - M, R) - v(L) \right\},$$

so that, from the envelope theorem, we have

$$\frac{\partial V^j}{\partial R} = -f_R\left(a^j, R\right), \tag{A44}$$

$$\frac{\partial V^{jk}}{\partial R} = -f_R\left(a^{jk}, R\right). \tag{A45}$$

Substituting for $\partial V^h/\partial R, \partial V^\ell/\partial R$, and $\partial V^{h\ell}/\partial R$, from equations (A44)–(A45), into the first-order condition (16) in the text, we have

$$-(1+\lambda) f_R\left(a^h, R\right) - \mu f_R\left(a^\ell, R\right) + \lambda f_R\left(a^{h\ell}, R\right) - \left(N^\ell + N^h\right)\gamma - \phi g' \frac{\partial n}{\partial R} \leq 0, \tag{A46}$$

From (A32), substituting γN^h for $(1+\lambda)$, and from (A38) substituting $\left(\gamma N^\ell + \lambda\right)$ for μ, one gets

$$-\gamma N^h f_R\left(a^h, R\right) - \left(\gamma N^\ell + \lambda\right) f_R\left(a^\ell, R\right) + \lambda f_R\left(a^{h\ell}, R\right) - \left(N^\ell + N^h\right)\gamma - \phi g' \frac{\partial n}{\partial R} \leq 0.$$

Rearranging the terms and focusing on the case of an interior optimum (i.e. the case where the optimal value of R is non-zero), we have

$$\lambda\left[f_R\left(a^{h\ell}, R\right) - f_R\left(a^\ell, R\right)\right] - \phi g' \frac{\partial n}{\partial R} = \gamma\left[N^h\left(1 + f_R\left(a^h, R\right)\right) + N^\ell\left(1 + f_R\left(a^\ell, R\right)\right)\right]. \tag{A47}$$

To show that relationship (A47) reduces to condition (17) in the text, we have to prove that

$$\frac{\partial n}{\partial R} = -\left[\frac{w^h f_{aR}\left(a^h, R\right)}{L^h \Psi^h} - \frac{w^\ell f_{aR}\left(a^\ell, R\right)}{L^\ell \Psi^\ell}\right] n. \tag{A48}$$

To do so, differentiate $n \equiv N^h L^h / N^\ell L^\ell$ with respect to R to get

$$\frac{\partial n}{\partial R} = \frac{N^h}{N^\ell}\left[\frac{L^\ell \frac{\partial L^h}{\partial R} - L^h \frac{\partial L^\ell}{\partial R}}{\left(L^\ell\right)^2}\right] = n\left[\frac{1}{L^h}\frac{\partial L^h}{\partial R} - \frac{1}{L^\ell}\frac{\partial L^\ell}{\partial R}\right]. \tag{A49}$$

To determine $\partial L^h/\partial R$ and $\partial L^\ell/\partial R$, observe that the equivalent of equation (6) when R is a choice variable by the government is given by,

$$v'(L) = w\left[1 - f_a(wL - M, R)\right].$$

Partially differentiate this relationship with respect R and rearrange the terms to get,

$$\frac{\partial L}{\partial R} = \frac{-wf_{aR}(a, R)}{v''(L) + w^2 f_{aa}(a, R)} = \frac{-wf_{aR}(a, R)}{\Psi}. \tag{A50}$$

Substituting from (A50) for $\partial L^h/\partial R$ and $\partial L^\ell/\partial R$ into (A49) results in (A48).

Finally, to determine the sign of $\Omega = -(\partial n/\partial R)\,\phi g'$, recall that $\phi < 0$ and $g' > 0$. As to the sign of $\partial n/\partial R$, our assumptions on $v(\cdot)$ and $f(\cdot,\cdot)$ ensure that $v''(L) + w^2 f_{aa}(a, R) > 0$. Moreover, the assumption that $f_R(a, R)$ is increasing in the absolute value of a, and the fact that $a^h < 0$ and $a^\ell > 0$, imply $f_{aR}\left(a^h, R\right) < 0$ and $f_{aR}\left(a^\ell, R\right) > 0$. Hence the bracketed expression on the right-hand side of (A48) is negative implying that $\partial n/\partial R > 0$. Consequently, $\Omega > 0$.

References

[1] Allingham, Michael G., and Agnar Sandmo. 1972. Income tax evasion: a theoretical analysis. Journal of Public Economics 1, 323–338.

[2] Atkinson, Anthony B., and Nicholas H. Stern. 1974. Pigou, taxation and public goods. Review of Economic Studies 41, 119–128.

[3] Bastani, Spencer, Blumkin, Tomer, and Luca Micheletto, 2015. Optimal wage redistribution in the presence of adverse selection in the labor market. Journal of Public Economics 131, 41–57.

[4] Blackorby, Charles, and Craig Brett, 2004. Production efficiency and the direct-indirect tax mix. Journal of Public Economic Theory 6, 165-180.

[5] Blomquist, Sören, Christiansen, Vidar, and Luca Micheletto, 2016. Public provision of private goods, self-selection and income tax avoidance. Scandinavian Journal of Economics 118, 666-692.

[6] Boadway, Robin, Marchand, Maurice, and Pierre Pestieau, 1994. Towards a theory of the direct-indirect tax mix. Journal of Public Economics 55, 71–88.

[7] Chetty, Raj, 2009. Is the taxable income elasticity sufficient to calculate deadweight loss? The implications of evasion and avoidance. American Economic Journal: Economic Policy 1 (2), 31–52.

[8] Christiansen, Vidar and Matti Tuomala, 2008. On taxing capital income with income shifting. International Tax and Public Finance 15, 527–545.

[9] Cremer, Helmuth, and Firouz Gahvari, 1994. Tax evasion, concealment, and the optimal linear income tax. Scandinavian Journal of Economics 96, 219–239.

[10] Cremer, Helmuth, and Firouz Gahvari, 1996. Tax evasion and the optimum general income tax. Journal of Public Economics 60, 235–249.

[11] Cremer, Helmuth, and Firouz Gahvari, 2000. Tax evasion, fiscal competition, and economic integration. European Economic Review 44, 1633–1657.

[12] Cremer, Helmuth, Marchand, Maurice, and Pierre Pestieau, 1990. Evading, auditing and taxing: The equity-compliance tradeoff. Journal of Public Economics 43, 67–92.

[13] Diamond, Peter A., 1998. Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates. American Economic Review 88, 83–95.

[14] Gahvari, Firouz, 2014. Second-best taxation of incomes and non-labor inputs in a model with endogenous wages, Journal of Public Economic Theory 16, 917–935.

[15] Gahvari, Firouz, and Luca Micheletto, 2014. The Friedman rule in an overlapping-generations model with nonlinear taxation and income misreporting. Journal of Public Economics 119, 10–23.

[16] Gahvari, Firouz, and Luca Micheletto, 2017. Optimal nonlinear income taxation with endogenous wages and income misreporting. Unpublished Working Paper.

[17] Gaube, Thomas, 2005. Income taxation, endogenous factor prices and production efficiency. Scandinavian Journal of Economics 107, 335–352.

[18] Guesnerie, Roger, and Jesus Seade, 1982. Nonlinear pricing in a finite economy. Journal of Public Economics 17, 157–179.

[19] Huber, Bernd, 1999. Tax competition and tax coordination in an optimum income tax model. Journal of Public Economics 71, 441-458.

[20] Kopczuk, Wojciech, 2001. Redistribution when avoidance behavior is heterogeneous. Journal of Public Economics 81, 51–71.

[21] Kopczuk, Wojciech, and Joel Slemrod, 2002. The optimal elasticity of taxable income. Journal of Public Economics 84, 91–112.

[22] Mayshar, Joram, 1991. Taxation with costly administration. Scandinavian Journal of Economics 93, 75–88.

[23] Micheletto, Luca, 2004. Optimal redistributive policy with endogenous wages. Finanzarchiv 60, 141–159.

[24] Naito, Hisahiro, 1999. Re-examination of uniform commodity taxes under a non-linear income tax system and its implication for production efficiency. Journal of Public Economics 71, 165–188.

[25] Pirttilä, Jukka, and Matti Tuomala, 2002. Publicly provided private goods and redistribution: a general equilibrium analysis. Scandinavian Journal of Economics 104, 173–188.

[26] Salanié, Bernard, 2011. The Economics of Taxation. Second edition. The MIT Press, Cambridge, Massachusetts.

[27] Sandmo, Agnar, 1981. Tax evasion, labour supply and the equity-efficiency tradeoff. Journal of Public Economics 16, 265-288.

[28] Schroyen, Fred, 1997. Pareto efficient income taxation under costly monitoring. Journal of Public Economics 65, 343-366.

[29] Slemrod, Joel, 2001. A general model of the behavioral response to taxation. International Tax and Public Finance 8, 119–128.

[30] Slemrod, Joel, and Shlomo Yitzhaki, 2002. Tax avoidance, evasion, and administration. In: Handbook of Public Economics vol. 3 (edited by Alan Auerbach and Martin Feldstein), ch. 22, 1423–1470, North-Holland.

[31] Stiglitz, Joseph E., 1982. Self-selection and Pareto efficient taxation. Journal of Public Economics 17, 213–240.

[32] Stiglitz, Joseph E., 1987. Pareto efficient and optimal taxation and the new new welfare economics. In: Handbook of Public Economics vol. 2 (edited by Alan Auerbach and Martin Feldstein), ch. 15, 991–1042, North-Holland.

[33] Usher, Dan, 1986. Tax evasion and the marginal cost of public funds. Economic Inquiry 24, 563–86.

Table 1. Optimal marginal income tax rates with and without evasion

	Without evasion			With evasion: $\beta = 1$		
Skill-type	$MITR$	wage	utility	$MITR$	wage	utility
$\ell 1$	31.01%	0.355	–1.115	28.09%	0.305	–1.283
$\ell 2$	42.98%	0.391	–1.082	32.01%	0.336	–1.243
$h1$	18.82%	0.732	–0.970	14.01%	0.757	–0.947
$h2$	–3.52%	1.020	–0.737	–2.62%	1.059	–0.657
	Utilitarian Social Welfare = –0.976			Utilitarian Social Welfare = –1.032		

Table 2. Optimal marginal income tax rates with and without government fighting evasion

	$\beta = .1$: Evasion is not fought			$\beta = .1$: Evasion is fought		
Skill-type	$MITR$	wage	utility	$MITR$	wage	utility
$\ell 1$	2.86%	0.163	–2.103	29.11%	0.295	–1.384
$\ell 2$	3.18%	0.179	–2.028	32.42%	0.325	–1.341
$h1$	4.11%	0.846	–0.702	13.49%	0.762	–0.993
$h2$	–2.03%	1.184	–0.377	–2.65%	1.067	–0.681
	Utilitarian Social Welfare = –1.302			Utilitarian Social Welfare = –1.101		

Table 3. Income-dependent R with $\beta = .1$

Skill-type	$MITR$	wage	utility	R^j
$\ell 1$	31.31%	0.308	–1.343	3.987
$\ell 2$	34.73%	0.339	–1.303	4.265
$h1$	7.09%	0.755	–1.016	0.905
$h2$	–0.05%	1.058	–0.657	0
	Utilitarian Social Welfare = –1.079			

Table 4. Optimal marginal income tax rates with and without evasion

	Without evasion			With evasion: $\beta = 1$		
Skill-type	$MITR$	wage	utility	$MITR$	wage	utility
$\ell 1$	81.44%	0.582	–1.054	42.80%	0.361	–1.229
$\ell 2$	65.65%	0.640	–1.047	46.75%	0.397	–1.196
$h1$	35.39%	0.640	–1.047	25.71%	0.729	–1.012
$h2$	–20.86%	0.896	–0.877	–11.79%	1.020	–0.754
	Maxi-min Social Welfare = –1.054			Maxi-min Social Welfare = –1.229		

Table 5. Optimal marginal income tax rates with and without government fighting evasion

	$\beta = .1$: Evasion is not fought			$\beta = .1$: Evasion is fought		
Skill-type	$MITR$	wage	utility	$MITR$	wage	utility
$\ell 1$	2.87%	0.164	–2.102	62.97%	0.459	–1.284
$\ell 2$	3.19%	0.180	–2.027	66.96%	0.504	–1.258
$h1$	5.23%	0.845	–0.702	31.16%	0.686	–1.187
$h2$	–2.62%	1.183	–0.381	–8.12%	0.961	–0.927
	Maxi-min Social Welfare = –2.102			Maxi-min Social Welfare = –1.284		

Table 6. Income-dependent R with $\beta = .1$

Skill-type	$MITR$	wage	utility	R^j
$\ell 1$	65.03%	0.476	–1.252	6.29
$\ell 2$	69.13%	0.523	–1.228	6.73
$h1$	28.07%	0.679	–1.174	3.61
$h2$	–1.32%	0.951	–0.901	0
	Maxi-min Social Welfare = –1.252			