OPTIMAL INCOME TAXATION WITHOUT TAX EVASION

Thomas Aronsson* and Fei Xu†

Abstract: This paper is the first to integrate corruption with respect to tax collection in a Mirrleesian model of optimal redistributive taxation. The analysis starts with a simple two-type model, showing that the optimal marginal tax structure resembles that of the original Stiglitz (1982) model, albeit for different reasons. We also extend the analysis to a framework with many types and present policy rules for marginal taxation over the whole ability distribution. The marginal income tax rates are all non-negative and can be expressed in terms of two key determinants: the distributional weights attached to taxpayers, and how the private cost to evade taxes varies with the taxpayers’ income. Finally, we consider the role of government expenditures directly targeting the incentives of the tax collector, such that these public expenditures and the optimal tax structure are implemented simultaneously.

JEL Classification: H21, H26, D73.
Keywords: Redistribution, taxation, corruption.

*Umeå University, Department of Economics, Umeå School of Business, Economics and Statistics, Umeå University, SE-901 87 Umeå, Sweden. E-mail: thomas.aronsson@umu.se
†Umeå University, Department of Economics, Umeå School of Business, Economics and Statistics, Umeå University, SE-901 87 Umeå, Sweden. E-mail: fei.xu@umu.se
1. Introduction

Tax evasion is a serious problem that undermines redistribution policies and reduces the funds available for public services. Although its magnitude varies among countries, it seems, nevertheless, to be present everywhere. For instance, Pissarides and Weber (1989) estimate that unreported incomes correspond to roughly 5.5 percent of GDP in Britain, which is a considerable number, and in Sweden the National Tax Agency estimates the “tax gap” to be around 8-9 percent of tax revenues. Based largely on the seminal contribution of Allingham and Sandmo (1972), and the extension of their analysis by Yitzhaki (1974), there is now a large literature dealing with the nature, determinants, and implications of tax evation. This literature focuses on, among other things, the roles of marginal taxation for the incentives to evade (e.g., Clothfelter, 1983; Crane and Nourzad, 1990), auditing (e.g., Beron, Tauchen, and Witte, 1992; Blumenthal, Christian, and Slemrod, 2001), and the potential corruption among tax collectors (e.g., Besley and McLaren, 1993). The present study deals with the implications of tax evasion for optimal redistributive taxation and, in particular, how an optimal nonlinear income tax ought to respond to corruption in the tax administration. Our approach will be explained and motivated in greater detail below.

A central restriction in the study of optimal redistributive taxation refers to information: in the Mirrleesian (1971) model, productivity and actions are private information, whereas income is observable at the individual level. This rules out type-specific lump-sum taxes (and thus the first-best resource allocation) while allowing for nonlinear taxes on income. Tax evasion reinforces the information problem in the sense that income is no longer fully observable at the individual level, i.e., there is a distinction between the actual and the reported income.

There are two main approaches to tax evasion in the literature on redistributive taxation: one in which tax evasion can be detected through costly auditing, and the other where the evader can avoid detection by incurring a cost. To our knowledge, Cremer and Gahvari (1996) were the first to introduce tax evasion in a two-type version of the Mirrleesian model. Their approach is to examine the optimal mix of nonlinear income taxation and nonlinear penalties, and the results suggest that high-ability individuals should never be audited and should face a zero marginal income tax, while low-ability individuals should be audited with a probability less than one and should either face a positive or zero marginal income
tax. Schroyen (1997) examines a similar problem, albeit with a linear (and predetermined) penalty function, and derives results on marginal income taxation that are qualitatively similar to those of Cremer and Gahvari. Chander and Wilde (1998) extend the analysis to a continuum of taxpayers, where the government is unable to observe income at the individual level without costly auditing. Their study provides a general characterization of the income tax schedule, auditing, and enforcement policy, and they find, among other things, a regressive tax structure with non-increasing average tax rates.

Gahvari and Micheletto (2019) analyze the implications of tax evasion for optimal general income taxation in an economy where the before-tax wage rates are endogenous. Their study is based on the riskless approach to tax evasion introduced by Usher (1986), in which a tax evader can avoid detection by incurring a cost that depends on the amount of income misreported. Gahvari and Micheletto find that tax evasion leads to a modification of the policy rule for marginal income taxation through the mechanism associated with endogenous before-tax wage rates. More specifically, the possibility of tax evasion weakens the policy incentive to modify the marginal income tax rates in response to wage endogeneity. The intuition is that tax evasion makes income taxation a less effective instrument for influencing the labor supply behavior.

Yet, none of the studies referred to above focuses on corruption in the tax administration and its implications for optimal redistributive taxation, which is the key issue of the present study. According to Klitgaard (1985), there are different means of corruption in the tax administration; in particular, payments for speeding up the administration of a particular errand, and bribing tax collectors to approve a lower tax payment. Whereas the former typically involves relatively small amounts of money, the latter can be substantial. Klitgaard exemplifies with the Philippines, where the government estimates that without the latter type of corruption, tax revenues would be 50 percent higher, ceteris paribus. “If a tax examiner discovered these errors [too many deductions, understated income, or incorrect tax computations], he would bring them to the taxpayer’s attention. At this point the arrangement frequently occurred. Typically, the taxpayer paid half of the extra taxes he owed and kept the other half. On the other half, perhaps two-thirds was a bribe to the tax examiner and the other third was paid where it should have been, to the government
coffers.” (Klitgaard, 1985, p6). This suggests that corruption in the tax administration and, in particular, when it comes to tax collection can be very costly.

To our knowledge, our study is the first to integrate corruption with respect to tax collection in a Mirrlesian (1971) model of optimal taxation. In doing so, we simplify by using a discrete formulation of the optimal tax problem, first in terms of the two-type model pioneered by Stiglitz (1982) and later in terms of a more general multi-type model. The model comprises three agent-types: individuals (or taxpayers), a tax collector (referred to as the “taxman”), and a government. Productivity is private information, and the government wants to redistribute from individuals with high productivity to individuals with lower productivity. However, contrary to standard models of optimal taxation, we assume that individual income is not directly observable to the government, while it is observable to the (more or less corrupt) taxman. Thus, there are two ways for any individual to mimic the tax payment of individuals with lower incomes: either by reducing their labor supply in order to reach this particular individual’s before-tax income, as in conventional models of optimal redistributive taxation, or by bribing the taxman such that the tax payment corresponds to this level of income. Since the implications for optimal taxation of the former option are well-known, we focus primarily on the latter option and how the government can design a tax system satisfying incentive compatibility. In other words, there is no corruption at the second-best optimum.

Section 2 presents a two-type model, where the high-ability type has an incentive to bribe the taxman in order to gain from the redistribution policy. Although the underlying channels of mimicking differ from the conventional optimal tax problem, we show in Section 3 that the qualitative results are, nevertheless, similar: whereas the optimal marginal income tax rate implemented for the low-ability type is positive, the marginal income tax implemented for the high-ability type is zero. In Section 4, this model is extended to a multi-type framework, where we examine different possibilities regarding potential mimicking and binding self-selection constraints. The marginal income tax is always non-negative and can be expressed in terms of two key determinants: the government’s preferences for redistribution and how a change in the taxpayers’ income affects the private cost of evading taxes.
Although the second-best optimum characterized in Sections 3 and 4 implies truthful income reporting, it is not necessarily efficient to induce this behavior solely through the tax system. The direct incentives facing the tax collector are also relevant for the functioning of the tax administration. In Section 5, therefore, we extend the analysis by allowing the government to directly affect the salary of the taxman through public expenditures. This means that the optimal tax structure and the public expenditure to avoid corruption are implemented simultaneously. The results show how the marginal income tax structure depends on the cost of tax evasion facing taxpayers which, in turn, depends on the amount potentially evaded, and on the public expenditures used to incentivize the tax administration. In addition, the optimal policy mix completely neutralizes the incentives to evade taxes.

Finally, conclusions are provided in Section 6, while background calculations and proofs can be found in the Appendix.

2. The Model

We consider a model with three types of agents: the taxpayers, a taxman, and the government. We assume that taxpayers' income is not observable to the government but verifiable by the taxman through auditing. Therefore, the formally stipulated duty of a taxman is to report the taxpayer’s true income so that the corresponding taxes would be imposed. A taxpayer chooses the consumption, labor supply, and whether to report their true income and pay the correct amount of tax, or to bribe the taxman in exchange for the illicit service of approving a lower tax payment, i.e., a tax payment corresponding to a lower income level. The role of the incentive constraints in the government’s problem is then to make such tax mimicking undesirable for the taxpayers.

2.1. The Taxman. The taxman receives salary $I_0 > 0$ and is assigned the duty of reporting the taxpayers’ true income $y$ to the government. We assume that the taxman can verify the taxpayers’ true incomes through auditing. Let $y$ denote the taxpayer’s true income and $Y$ the reported income. For each taxpayer, the taxman can take one of two possible actions: either to perform his duty or not, i.e., either truthfully report the taxpayers’ true income $Y = y$ or untruthfully report the taxpayers’ true income $Y \neq y$. Corruption happens when
the taxman receives a bribe in order to switch his duty to untruthfully report the taxpayer’s true income. In this way corruption decisions are viewed as binary: the taxman is either corrupt or not (see e.g., Amir and Burr, 2015; Banerjee, Mullainathan, and Hanna, 2012), and a bribe can be understood as a monetary payoff associated with a certain distortion, i.e., (non)performance of duty.

The model also admits an important dimension in terms of the degree of corruption, i.e., the magnitude by which the taxman distorts his duties (Santos, Koutsougeras, and Xu, 2023). Specifically, corruption decisions should also incorporate the idea that a taxman can be corrupt to a variable degree. Correspondingly, the size of the bribe should also take into account how seriously the taxman distorts his duty. In our model, the degree of (non)performance of duty is understood as the difference between a taxpayer’s true income and the reported income. Since the taxpayers’ true incomes are not verifiable to the taxman without auditing, the (non)performance of duty can be represented by the probability that a taxpayer’s true income is higher than the reported one. To simplify the analysis so that the (non)performance of duty of a taxman depends only on the reported income, we assume that the taxman has a theoretical continuous distribution of taxpayers’ true income $y$ with cumulative distribution $F_y(Y) = P(y \leq Y)$, which has a compact support $[\bar{Y}, Y] \subset \mathbb{R}$. So the probability that the true income $y$ is higher than a reported value $Y$ is given by $\phi(Y) = 1 - F_y(Y)$.\footnote{The analysis is reconciled with the assumption that the taxman can verify the taxpayers’ true incomes through auditing. If the taxman audits the taxpayer, $\phi(Y) = 1$.} Hence, the degree of corruption is assumed to increase when the probability that the taxpayer’s true income is higher than the reported one increases.

Let us define $p = \phi(Y)$. The choice of action by the taxman is governed by a utility function that depends on the performance of duty, as represented by the probability that a taxpayer’s true income is higher than the reported value, as well as on his own income: $v : [0, 1] \times \mathbb{R}_+ \to \mathbb{R}$, which is thus decreasing in $p$, increasing in $I$, and strictly quasi-concave. We also assume that $\partial^2 v / \partial I \partial p \leq 0$.

Misreporting the taxpayer’s income entails a utility cost to the taxman, i.e., $\forall p \in (0, 1]$, $v(p, I) < v(0, I)$. Furthermore, a more serious infraction of duty (i.e., when the probability of misreport is higher) decreases utility. The simplest justification of our assumption is that more severe misreporting entails a higher risk of detection and punishment, or a drop
of professional ethics due to the acknowledgment that misbehavior results in revenue loss. 

We do not model explicitly the probability that the taxman confronts legal sanctions, e.g., being exposed and directly punished, but this can be done easily in which case our assumption would be valid in terms of expected utility.

In particular, given the income of the taxman, $I$, for a reported income $Y$:

- If the taxman reports truthfully, then $\phi(Y) = 0$ and his utility becomes: $v(0, I)$.
- If the taxman reports untruthfully, then $\phi(Y) > 0$ such that the utility becomes: $v(\phi(Y), I)$.

Note that since the probability that the true income is misreported depends on the reported value $Y$, the utility of the taxman can be defined on the support of the income distribution via the utility function $u : R_+ \times R_+ \to R$ defined as $u(Y, I) \equiv v(\phi(Y), I)$. We assume that $u(\cdot, \cdot)$ is twice continuously differentiable. Obviously, $u(Y, I)$ is increasing in each argument and strictly quasi-concave.

Since misreporting the taxpayer’s income (non performance of duty) entails a utility cost for the taxman, a monetary incentive is required for the taxman to take such an action. This is what we understand as ”bribe”. In order to (mis)report income as $Y$, the bribe $B \geq 0$ acceptable by a taxman who receives the salary $I_0$ should satisfy:

(1) \[ u(Y, B + I_0) \geq v(0, I_0) \]

It follows that for each $Y$ the reservation bribe is $B(Y, I_0)$ such that\(^2\)

(2) \[ u(Y, I_0 + B(Y, I_0)) \equiv v(0, I_0). \]

where $B(\cdot, \cdot)$ is twice continuously differentiable. Since $\phi(Y) = 1 - F_Y(Y)$, the reservation bribe function satisfies the following comparative statics property with respect to $Y$:

(3) \[ \frac{\partial B(Y, I_0)}{\partial Y} = \frac{\partial u/\partial Y}{\partial u/\partial I} = -\frac{\partial v/\partial p(-F'_Y(Y))}{\partial v/\partial I} < 0. \]

---

\(^2\)Note that we have taken the point of view that a bribe is a perfect substitute for salary income. In case one wishes to capture the idea of moral inhibitions about receiving bribes, this can be incorporated by a suitable adjustment instead of adding the two sources of monetary payoff.
We can also derive\(^3\)

\[
\frac{\partial B(Y, I_0)}{\partial I_0} = \frac{\partial v(0, I_0)/\partial I - \partial u(Y, I_0 + B)/\partial I}{\partial u(Y, I_0 + B)/\partial I - \partial u(Y, I_0 + B)/\partial I - 1} \geq 0. \tag{4}
\]

While we shall generally cast our analysis in terms of bribing the taxman, this setup is equally applicable to other possible services related to tax evasion such as wealth concealment services, with only a simple relabeling of the variables \(I_0\) (legal consultant fees) and \(B(Y, I_0)\) (fees for wealth concealment services). The intuition is that when third-party reporting plays a role in the tax system, e.g., domestic banks, accounting organizations, family trusts and private foundations, the model captures the potential costs of tax evasion combining the basic service fees \(I_0\) and the value of the wealth involved. The analysis could be further extended to the study of migration behavior in response to wealth taxes (Agrawal et al., 2022; Jakobsen et al., 2021) in which the cost of migration should take into consideration both the residence permit fees or a lump sum payment required to maintain investment, and the transaction cost to immigrate assets which is determined by the value of these assets.

### 2.2. The Taxpayers

We have \(N = \{1, \ldots, n\}\) types of taxpayers who differ in ability but have the same utility function. Each type \(i\), where \(i = 1, \ldots, n\), faces a before-tax wage rate per unit of labor of \(w_i \in W = \{w_1, \ldots, w_n\}\). We assume that the levels of ability represented by the before-tax wage rates are totally ordered. In the absence of taxation, each taxpayer of type \(i\) faces the budget constraint

\[
C_i = w_i L_i, \tag{5}
\]

where \(C\) and \(L\) denote the consumption and labor supply, respectively. The ability (productivity) of the taxpayer is private information, and the labor supply is not observable. In addition to these standard assumptions, we assume that the before-tax income facing a taxpayer of any type \(i\)

\[
Y_i = w_i L_i, \tag{6}
\]

\(^3\)Note that \(v(p, I)\) is increasing and concave in \(I\), \(p\) is decreasing in \(Y\), and \(\partial^2 v/\partial I \partial p \leq 0 \iff \partial v(0, I_0 + B)/\partial I \geq \partial v(p, I_0 + B)/\partial I\). We obtain \(\partial v(0, I_0)/\partial I \geq \partial v(Y, I_0 + B)/\partial I\).
is not directly observable to the government but verifiable to the taxman by auditing. Each taxpayer derives utility from consuming goods and disutility from labor/effort. The utility function can be written as follows:

\[ U^i = U(C_i, Y_i) \equiv V^i(C_i, Y_i; w_i), \]

which is increasing in \( C \), decreasing in \( Y \) for a given before-tax wage rate, and strictly quasi-concave.

Suppose that the government imposes a tax as a function of the income reported by the taxman. In the absence of tax evasion, the reported income is, of course, the taxpayer’s true income. The tax imposed on a taxpayer of any type \( i \) is then defined as \( T_i = T(Y_i) \), while the consumption is given by the taxpayer’s income minus the tax payment

\[ C_i = Y_i - T(Y_i). \]

In the absence of any tax evasion, the utility maximization problem facing the taxpayer can be written as

\[ \max_{C_i, Y_i} V^i(C_i, Y_i; w_i), \text{ s.t., } C_i \leq Y_i - T(Y_i), \]

yielding the first-order condition

\[ \frac{V^i_{Y_i}}{V^i_{C_i}} + 1 = T'(Y_i). \]

Tax evasion enables each taxpayer of type \( i \) to pay a bribe in order to evade taxes. Specifically, the taxpayer of type \( i \) pays the bribe \( B(Y_j, I_0) \) in exchange for the illicit service provided by the taxman to report income \( Y_j \) instead of \( Y_i \) to the government, where \( Y_j < Y_i \) and \( Y_j \) is the income of type \( j \in N \). A taxpayer of any type \( i \) who pays a bribe \( B(Y_j, I_0) \) and is imposed the tax payment \( T(Y_j) \) thus solves

\[ \max_{C_i, Y_j} V^i(C_i, Y_i; w_i), \]

\[ \text{s.t. } C_i \leq Y_i - T(Y_j) - B(Y_j, I_0). \]

where \( T(Y_j) = Y_j - C_j, \ i, j \in N, \ j \neq i \). The first-order condition of the problem (11) can be written as

\[ \frac{V^i_{Y_i}}{V^i_{C_i}} = 0 \]

\[ ^4 \text{The case in which the size of the bribe is fixed is a special example of our analysis.} \]
Let $\bar{C}_i(C_j, Y_j)$, and $\bar{Y}_i(C_j, Y_j)$ be the solution to problem (11). Clearly, since the tax payment and the size of the bribe are both determined by the income of taxpayers of type $j$, the utility of a taxpayer of type $i$, evaluated at the optimum choice $(\bar{C}_i, \bar{Y}_i)$, can be rewritten as a function of $(C_j, Y_j)$. Let $\bar{V}^i(C_j, Y_j) \equiv V_i(C_i(C_j, Y_j), I_i(C_j, Y_j); w_i)$ and $B_j = B(Y_j, I_0)$ be the maximum utility and bribe, respectively, associated with such tax evasion. By the Envelope Theorem, we obtain

\begin{align}
(13) & \quad \bar{V}^i_{C_i} = V^i_{C_i} > 0. \\
(14) & \quad \bar{V}^i_{Y_j} = -V^i_{C_i} (1 + \partial B_j / \partial Y_j).
\end{align}

Equation (13) shows that, as the consumption of type $j$ increases, the maximum utility that type $i$ can achieve by mimicking type $j$ through tax evasion increases as well. Similarly, equation (14) implies that a higher before-tax income of type $j$ has a positive effect, $-V^i_{C_i}$, and a negative effect, $-V^i_{C_i}$, on the maximum utility the type $i$ evader can achieve.\(^5\) Therefore, the overall effect of $Y_j$ on the maximized utility of the type $i$ mimicker is ambiguous.

3. A Two-Type Problem

We start with a basic version of the model, in which there are only two types of taxpayers, indexed by $i = 1, 2$. Type 2 (the high-ability type) is more productive than type 1 (the low-ability type) in the sense that $w_2 > w_1$. Focus will be on a normal case, where the government wants to redistribute from the high-ability to the low-ability type. This means that the high-ability type is the one who would potentially evade taxes and can afford the costs to evade taxes as well.\(^6\) Corruption happens in our model when the taxman accepts the bribe offered by taxpayers of type 2 and untruthfully reports a taxpayer of type 2’s income as $Y_1$. In Section 4, we extend the analysis to a case with multiple ability-types and show that the insights from the two-type model naturally extend to the multi-type case.

\(^5\)Recall that $V^i_{C_i} > 0$ and $\partial B_j / \partial Y_j < 0$.

\(^6\)Alstadsæter, Johannesen and Zucman (2019) find that the 0.01 percent richest households in Scandinavia evade about 25 percent of their taxes. Leenders, et al. (2023) document that in the Netherlands, top 0.01 percent evade around 8 percent of their true tax liability, and that there is substantial evasion among the "merely rich" (90 percent - 99.9 percent) who own around 67 percent of the hidden wealth.
We begin with a purely redistributive tax system, which does not raise any net tax revenue for public consumption. A more general version of the model will be examined in Section 5, where the government also uses public expenditures to fight corruption by incentivizing the taxman. This case means that the government alleviates corruption both through the tax system and via public expenditures.

In our framework, and in the absence of any concerns for incentive compatibility, high-ability individuals could potentially mimic the low-ability type either through adjustments of their labor supply (in order to reach the same before-tax income as the low-ability type) or through tax evasion by bribing the taxman (allowing them to pay the same tax as the low-ability type). Thus, two self-selection constraints must be imposed on the high-ability type. The social decision-problem can then be written as follows:7

\[
\begin{align*}
\max_{C_1, Y_1, C_2, Y_2} & \quad \sum_{i=1}^{2} \beta_i N_i V^i(C_i, Y_i), \\
\text{s.t.} & \quad V^2(C_2, Y_2) \geq V^2(C_1, Y_1) \\
& \quad V^2(C_2, Y_2) \geq V^2(C_1, Y_1), \\
& \quad (Y_1 - C_1) N_1 + (Y_2 - C_2) N_2 = 0,
\end{align*}
\]

where \(N_i\) is the number of taxpayers of type \(i, i = \{1, 2\}\).

The first self-selection constraint serves to prevent taxpayers of the high-ability type from mimicking the low-ability type through tax evasion, and the second serves to prevent mimicking through labor supply adjustments. It is important to emphasize that both self-selection constraints cannot bind simultaneously, except in the unlikely case where both types of mimicking give rise to exactly the same utility for the potential mimicker. Intuitively, if a high-ability individual reduces her labor supply in order to mimic the before-tax income of the low-ability type, there would be no incentives for this individual to pay a positive bribe. Similarly, if the maximized utility of tax mimicking through evasion is greater than the utility of mimicking through labor supply adjustments, there would be no incentive to undertake such adjustments.

\footnote{Since the low-ability type pays a lower tax than the high-ability type, individuals of the low-ability type cannot evade taxes by mimicking the high-ability type, i.e., a self-selection constraint imposed on the low-ability type preventing such evasion would not bind. See the Appendix. For a similar reason, a self-selection constraint preventing low-ability individuals from mimicking the high-ability type through labor supply adjustments would not bind either.}
Since the implications of preventing mimicking through labor supply adjustments are wellknown from earlier research, this paper focuses on the case in which "tax mimicking" by bribing the taxman is more desirable for the taxpayers, i.e., the case where the first self-selection constraint in problem (15) is binding. The Lagrangean corresponding to problem (15) can then be written as follows:

\[ L = \sum_{i=1}^{2} \beta_i N_i V^i(C_i, Y_i) \]

\[ + \lambda [V^2(C_2, Y_2) - \bar{V}^2(C_1, Y_1)] \]

\[ + \gamma [(Y_1 - C_1)N_1 + (Y_2 - C_2)N_2] \]

where \( \lambda \) and \( \gamma \) denote the Lagrange multipliers attached to the self-selection and resource constraint, respectively.

3.1. The Optimal Marginal Tax Structure. The social first-order conditions are presented in the Appendix. Let us now use these conditions together with (10), (13), and (14) to examine the implications for marginal taxation. Consider Proposition 1.

Proposition 1. The optimal marginal tax structure that prevents tax evasion can be characterized as follows:

\[ T'(Y_1) = \frac{\lambda \bar{V}^2_{C_1}}{\gamma N_1} \left( \frac{V^2_{Y_1}}{V^2_{C_1}} - \frac{V^1_{Y_1}}{V^1_{C_1}} \right) = -\frac{\lambda \bar{V}^2_{C_1}/\gamma N_1}{1 + \lambda \bar{V}^2_{C_1}/\gamma N_1} \frac{\partial B_1}{\partial Y_1} > 0 \]

\[ T'(Y_2) = 0. \]

Proof: see the Appendix.

Proposition 1 shows that if the government attempts to redistribute from the high-ability type and at the same time prevent any tax evasion, the optimal tax policy includes a zero marginal tax rate for the high-ability type and a positive marginal tax rate for the low-ability type. Although this accords well with the original two-type model, in which the self-selection constraint prevents mimicking through labor supply adjustments, the marginal tax implemented for the low-ability type takes a different form here. The intuition is, of course, that the government attempts to prevent the high-ability type from mimicking the low-ability type through tax evasion (instead of via the before-tax income).

\[ \text{Note that the bribe, } B(Y_1; I_0), \text{ is endogenous to the government. In the special case where the bribe is fixed, first-best taxation would be fully revealing, i.e., } \lambda = 0 \text{ and the government can attain the first-best through lump-sum taxation.} \]
The expression after the first equality in (17) is the difference in marginal rate of substitution of $C$ for $Y$ between the potential mimicker and the low-ability type (the potentially mimicked agent). By using (13) and (14), we show in the expression after the second equality that this difference in marginal rates of substitution can be rewritten in terms of how the reservation bribe (characterizing the taxman) reacts to an increase in the before-tax income of the low-ability type. The intuition is that a lower reported before-tax income leads to an increase in the reservation bribe (see equation (3)), which motivates the government to increase the marginal tax of the low-ability type’s income in order to make tax evasion more costly.

Finally, note that the redistribution from the high-ability to the low-ability type implies that the tax payment of the high-ability type exceeds the tax payment of the low-ability type. In this case with two types of taxpayers and no direct public expenditure, we obtain $N_2T(Y_2) = -N_1T(Y_1)$. On the one hand, tax evasion motivates higher marginal taxation of low-income earners (the non-evaders). On the other hand, the low-ability type would be secured to benefit from the redistribution achieved by such a tax scheme and receive a transfer corresponding to the tax payment of the high-ability type.

4. Multi-type Model

It is straightforward to extend the results derived above to a multi-type framework and show that all individuals will face non-negative marginal income tax rates. Consider an economy comprising $N = \{1, \ldots, n\}$ types of taxpayers. As before, we assume that the redistribution favors people with lower productivities and will formulate the social welfare function accordingly. Taxpayers of type $n$ could in principle mimic any of the $n-1$ types with lower productivity than type $n$, type $n-1$ could in principle mimic any of the $n-2$ types with lower productivity than type $n-1$, and so on. In such a framework, it is not obvious which self-selection constraint that is binding for a given type. We will, therefore, begin with a general characterization, where the effects of all potential self-selection constraints are described.

The government maximizes a generalized utilitarian social welfare function by offering $n$ packages of private consumption and labor/effort, i.e., $\{C, Y\}$ in which $\{C_i, Y_i\}$ will be chosen for individuals of type $i$. The social decision-problem can then be written as
\[
\begin{align*}
\text{max} \quad & \sum_{i=1}^{n} \beta_i N_i V_i(C_i, Y_i), \\
\text{s.t.} \quad & V^n(C_n, Y_n) \geq \tilde{V}^n(C_{n-1}, Y_{n-1}), \\
& V^n(C_n, Y_n) \geq \tilde{V}^n(C_{n-2}, Y_{n-2}), \\
& \vdots \\
& V^n(C_n, Y_n) \geq \tilde{V}^n(C_1, Y_1), \\
& V^{n-1}(C_{n-1}, Y_{n-1}) \geq \tilde{V}^{n-1}(C_{n-2}, Y_{n-2}), \\
& \vdots \\
& V^{n-1}(C_{n-1}, Y_{n-1}) \geq \tilde{V}^{n-1}(C_1, Y_1), \\
& \vdots \\
& V^2(C_2, Y_2) \geq \tilde{V}^2(C_1, Y_1), \\
& \sum_{i=1}^{n} (Y_i - C_i) N_i = 0,
\end{align*}
\]

where \(N_i\) is the number of taxpayers of type \(i\), \(i = 1, \ldots, n\). The Lagrangean corresponding to problem (19) becomes

\[
L = \sum_{i=1}^{n} \beta_i N_i V_i(C_i, Y_i)
\]

\[
+ \sum_{j=1}^{n-1} \lambda_{n,j} [V^n(C_n, Y_n) - \tilde{V}^n(C_j, Y_j)]
\]

\[
+ \sum_{j=1}^{n-2} \lambda_{n-1,j} [V^{n-1}(C_{n-1}, Y_{n-1}) - \tilde{V}^{n-1}(C_j, Y_j)]
\]

\[
+ \sum_{i=1}^{i-1} \lambda_{i,j} [V^i(C_i, Y_i) - \tilde{V}^i(C_j, Y_j)]
\]

\[
+ \lambda_{2,1} [V^2(C_2, Y_2) - \tilde{V}^2(C_1, Y_1)]
\]

\[
+ \gamma \sum_{i=1}^{n} (Y_i - C_i) N_i,
\]
where \( j \in \mathbb{N} \) and \( j < i \), and \( \lambda_{i,j} \) denotes the Lagrange multiplier on the self-selection constraint that serves to prevent type \( i \) from mimicking type \( j \) through tax evasion. The social first-order conditions are presented in the Appendix. Let us begin with a general characterization of the optimal marginal tax structure, which applies regardless of the structure of binding self-selection constraints. Consider the following Proposition 2.

**Proposition 2.** In a multi-type economy, the optimal marginal tax structure that prevents tax evasion satisfies \( T'(Y_n) = 0 \) and

\[
T'(Y_j) = -\frac{\sum_{i=j+1}^{n} \lambda_{i,j} V_i^{C_j} / \gamma N_j}{1 + \sum_{i=j+1}^{n} \lambda_{i,j} V_i^{C_j} / \gamma N_j} \frac{\partial B_j}{\partial Y_j} \geq 0
\]

for \( j = 1, \ldots, n-1 \).

**Proof:** see the Appendix.

The right-hand side of equation (21) just summarizes the effects of all potentially binding self-selection constraints on the marginal income tax rate implemented for any type \( j \). As such, it reflects and neutralizes the incentives of all types higher than \( j \) to mimic type \( j \). However, except in the extremely unlikely case where several mimicking-options give rise to exactly the same utility for a potential mimicker, we would expect that only one such constraint binds for each distinct agent-type. In Corollary 1, we consider two special cases directly following from Proposition 2. In one of these special cases, each individual potentially mimics the adjacent type with lower productivity (which is the conventional assumption in continuous type models of optimal taxation), and in the other each individual would potentially mimic the type with the lowest productivity. We can interpret the former as a case where additional evasion is very costly to the individual, such that a small decrease in the reported income necessitates a relatively large increase in the bribe, and the latter as a case where it is less costly.

**Corollary 1.** If each individual would potentially mimic the adjacent type with lower productivity, the optimal marginal tax structure in Proposition 2 simplifies to \( T'(Y_n) = 0 \) and

\[
T'(Y_j) = -\frac{\lambda_{j+1,j} V_j^{C_j} / \gamma N_j}{1 + \lambda_{j+1,j} V_j^{C_j} / \gamma N_j} \frac{\partial B_j}{\partial Y_j} > 0
\]

for \( j = 1, \ldots, n-1 \). Instead, if each individual of type \( j > 1 \) would potentially mimic type 1, then \( T'(Y_j) = 0 \) for \( j = 2, \ldots, n \) and
Proof: see the Appendix.

Policy rule (22) is fully analogous to policy rule (17) in the two-type model and thus interpretable in the same general way. To be more specific, it is designed to prevent individuals of any type \( j+1 \) to mimic type \( j \) through tax evasion. In turn, this motivates marginal taxation of type \( j \)'s income. Thus, when each individual would potentially mimic the adjacent type with lower productivity, the marginal income taxes are positive along the whole income distribution (except at the very top). Policy rule (23) is more extreme, as it corresponds to the (somewhat unlikely) scenario where all individuals would mimic type 1 in the absence of self-selection constraints preventing such behavior. In this case, therefore, only type 1 would face a positive marginal tax rates (designed to prevent all higher types from mimicking type 1), while the other marginal income tax rates would be zero.

The two special cases in the corollary can be interpreted as polar cases in terms of how sensitive the bribe is to the size of the amount of taxes evaded. If all individuals of abilities higher than type 1 would prefer to potentially mimic this type, i.e., if mimicking type 1 gives higher utility than the other mimicking options, then the effect on the bribe of the amount of taxes evaded is likely to be relatively small. Otherwise, it would be too costly for the highest types (who would need to evade substantial amounts) to prefer this particular option. On the other hand, if the bribe is very sensitive to the amount evaded, a potential mimicker is likely to prefer less evasion, which is exemplified by the case where individuals would mimic the adjacent type with lower ability in the absence of constraints ensuring incentive compatibility.

Returning finally to the general policy rules in Proposition 2, it is straightforward to rewrite the marginal income tax rates in terms of distributional weights, which reflect the social desire to redistribute from individuals with abilities higher than any type \( j \) to individuals with abilities lower than this particular type. In fact, such a policy rule would always apply regardless of which self-selection constraints that bind. Let \( \delta_j = \beta j V_{C_j} / \gamma \geq 1 \)
denote the distributional weight the government attaches to individuals of type $j$. This weight reflects the marginal social value of consumption relative to the marginal cost of public funds. Corollary 2 below is an immediate consequence of the calculations behind Proposition 2.

**Corollary 2.** The policy rules in Proposition 2 can be reformulated to read

\[
T' (Y_j) = 1 - \delta_j \frac{\partial B_j}{\partial Y_j} \geq 0
\]

for $j = 1, \ldots, n$.

**Proof:** see the Appendix.

Note once again that equation (24) is a reformulation of equation (21) in Proposition 2 and, therefore, also consistent with special cases (22) and (23) in Corollary 1. Thus, as long as we can estimate the distributional weights and the functional relationship between income and the reservation bribe, there is no need to assume which self-selection constraint that is actually binding for each type at the social optimum. To interpret Corollary 2, we assume that the welfare weight is declining in consumption. This means that the higher the distributional weight attached to type $j$, the more the government would like to redistribute income to people with abilities lower than or equal to $j$, ceteris paribus. To accomplish this redistribution, the government raises revenue from people with abilities higher than $j$ through marginal taxation of type $j$’s income. Given the distributional weight, the second factor determines the level of marginal taxation such that the optimal redistribution does not lead people with abilities higher than $j$ to mimic type $j$.

5. **More on the Government’s Effort Against Tax Evasion**

In this section, we extend the analysis to include public expenditures directed at anti-corruption efforts. More specifically, the government can now raise net tax revenue in order to increase the salary of the taxman, which provides another channel through which to improve the quality of the tax administration. This is interesting for at least two reasons. First, earlier research shows that the direct incentives facing tax collectors are important for the functioning of the tax administration (e.g., Besley and Mclaren, 1993). In our framework, public expenditures targeting the incentives facing the taxman contribute to raise the the cost of tax evasion which, in turn, relaxes the self-selection constraints
and opens up for more redistribution. Second, the introduction of direct anti-corruption measures will affect the policy incentives underlying the optimal marginal tax policy.

Let $E$ denote the public expenditure the government uses to incentivize the tax administration; in our case, through the salary of the taxman. We can then rewrite the reservation bribe to read $B(Y, E)$, since the salary of the taxman is now a direct decision-variable of the government. Thus, the reservation bribe is now a function of both the reported income of the taxpayer and the public expenditures targeting the incentives of the tax administration. The analysis below is based on, and extends, the model with multiple types of taxpayers developed in Section 4.

The optimization problem facing the true taxpayers, i.e., those who truthfully report their income, is given by problem (9) in Section 2, the solution of which satisfies equation (10). On the other hand, if a taxpayer of any type $i$ were to engage in tax evasion, by reporting a lower before-tax income, she now solves the problem

\[
\max_{C_i, Y_i} V^i(C_i, Y_i), \text{ s.t. } C_i \leq Y_i - (Y_j - C_j) - B(Y_j, E),
\]

where $j, i \in \mathbb{N}$ and $j < i$. Recall that individuals of any type $i > 1$ could potentially mimic any other type with lower productivity through tax evasion and would in that case choose the option giving the highest utility. The first-order condition of problem (25) is

\[
V^i_{C_i} + V^i_{Y_i} = 0.
\]

Let $\bar{C}_i(C_j, Y_j, E)$, and $\bar{Y}_i(C_j, Y_j, E)$ be the solution of problem (25), and let $\bar{V}(C_j, Y_j, E) \equiv V^i(\bar{C}_i(C_j, Y_j, E), \bar{Y}_i(C_j, Y_j, E))$ denote the corresponding value function. We show in the Appendix that this function satisfies the following comparative statics property with respect to $E$:

\[
\frac{\partial \bar{V}^i(C_j, Y_j, E)}{\partial E} = -\frac{\partial \bar{V}^i(C_j, Y_j, E)}{\partial C_j} \frac{\partial B(Y_j, E)}{\partial E} \leq 0.
\]

Thus, by increasing the (publicly funded) salary to the taxman, in order to strengthen the quality of the tax administration, the utility of being a mimicker decreases, ceteris paribus. The intuition is, of course, that this policy increases the reservation bribe at any income level.
We begin with a general case in which each type could in principle mimic any of the types with lower productivity. As we explained in Section 4, it is not obvious which self-selection constraints that are binding at the social optimum, since the cost of tax evasion increases with the amount evaded. More specifically, the more income the individual would like to evade, the higher will be bribe. It is, therefore, useful to start with a general characterization of the marginal tax structure, which applies regardless of which self-selection constraints are binding, and then turn to special cases.

The social decision-problem takes the same form as in Section 4 with two important modifications: (i) the public expenditure directed at anti-corruption, $E$, is now a direct decision-variable of the government/social planner, and (ii) $E$ directly affects all self-selection constraints and thus the incentives underlying tax evasion. The Lagrangean of the social decision-problem can then be written as follows:

\[
L = \sum_{i=1}^{n} \beta_i N_i V^i(C_i, Y_i) \\
+ \sum_{j=1}^{n-1} \lambda_{n,j} [V^n(C_n, Y_n) - \bar{V}^n(C_j, Y_j, E)] \\
+ \sum_{j=1}^{n-2} \lambda_{n-1,j} [V^{n-1}(C_{n-1}, Y_{n-1}) - \bar{V}^{n-1}(C_j, Y_j, E)] \\
+ \sum_{j=1}^{i-1} \lambda_{i,j} [V^i(C_i, Y_i) - \bar{V}^i(C_j, Y_j, E)] \\
+ \lambda_{2,1} [V^2(C_2, Y_2) - \bar{V}^2(C_1, Y_1, E)] \\
+ \gamma \left[ \sum_{i=1}^{n} (Y_i - C_i) N_i - E \right],
\]

which takes the same general form as (20). The social first-order conditions are presented in the Appendix. Consider Proposition 3.

**Proposition 3.** In a multi-type economy where the government can incentivize the tax administration by choosing the salary of the taxman, the optimal marginal tax structure that prevents tax evasion satisfies $T'(Y_n, E) = 0$ and
(29) \[ T'(Y_j, E) = -\frac{\sum_{i=j+1}^{n} \lambda_{i,j} \bar{V}_{C_i}^{j}/\sum_{i=1}^{n} \lambda_{i,j} \bar{V}_{C_i}^{j} \frac{\partial B(Y_{j}, E)}{\partial E} N_{j}}{1 + \sum_{i=j+1}^{n} \lambda_{i,j} \bar{V}_{C_i}^{j}/\sum_{i=1}^{n} \lambda_{i,j} \bar{V}_{C_i}^{j} \frac{\partial B(Y_{j}, E)}{\partial E} N_{j}} \frac{\partial B(Y_{j}, E)}{\partial Y_{j}} \geq 0, \]

for \( j = 1, ..., n - 1 \).

**Proof:** see the Appendix.

The right side of the equation (29) is interpretable as the marginal income tax rate implemented for any type \( j \) when the optimal tax structure and the public expenditure to avoid corruption in tax administration are chosen simultaneously. Intuitively, the optimal public expenditure on anti-corruption efforts influences the minimum bribe required by the taxman for all levels of the taxpayers’ reported income. Therefore, the marginal income tax that neutralizes the incentives by all types with abilities higher than \( j \) to mimic type \( j \) through tax evasion takes into account the marginal effects of the government’s anti-corruption expenditure on the taxpayers’ cost of tax evasion, i.e., \( B(Y_{j}, E) \). As explained above, since each potential mimicker would choose the mimicking option giving the highest utility, all these self-selection constraints cannot bind at the same time. It is the best mimicking strategy that the binding self-selection constraint should prevent which, in turn, constitutes the realized version of the expression \( \sum_{j=1}^{n-1} \sum_{i=j+1}^{n} \lambda_{i,j} \bar{V}_{C_i}^{j} \frac{\partial B(Y_{j}, E)}{\partial E} \).

By analogy to Corollary 1, we consider two special cases directly following from Proposition 3 in the following Corollary 3.

**Corollary 3.** If each individual would potentially mimic the adjacent type with lower productivity, the optimal marginal tax structure in Proposition 3 simplifies to \( T'(Y_n, E) = 0 \) and

\[ T'(Y_j, E) = -\frac{\lambda_{j+1,j} \bar{V}_{C_{j+1}}^{j+1}/\sum_{i=1}^{n} \bar{V}_{C_i}^{j+1} \frac{\partial B(Y_{i}, E)}{\partial E} N_{i}}{1 + \lambda_{i,j} \bar{V}_{C_i}^{j}/\sum_{i=1}^{n} \bar{V}_{C_i}^{j} \frac{\partial B(Y_{i}, E)}{\partial E} N_{i}} \frac{\partial B(Y_{j}, E)}{\partial Y_{j}} > 0, \]

for \( j = 1, ..., n - 1 \). Instead, if each individual of type \( j > 1 \) would potentially mimic type 1, then \( T'(Y_j, E) = 0 \) for \( j = 2, ..., n \) and

\[ T'(Y_1, E) = -\frac{\partial B(Y_{1}, E)/\partial Y_{1}}{1 + N_{1} \partial B(Y_{1}, E)/\partial E} > 0. \]
The proof of Corollary 3 immediately follow from Proposition 3 by making the same assumptions about binding self-selection constraints as the proof of Corollary 1 (see the Appendix).

Policy rule (30) is the realization of policy rule (29) if individuals would potentially mimic the adjacent type with lower productivity. As such, it also coincides with the policy rule that would follow in a two-type model, i.e., it is the analogue to (17) in an economy where public policy includes direct expenditures to incentivize the taxman. Policy rule (31) shows the optimal marginal income tax rate implemented for type 1 in the case where all individuals with abilities higher than type 1 would potentially mimic type 1. Since the optimal tax and expenditure policy prevents all individuals with higher productivity than type 1 from mimicking this type, the marginal tax rate is determined by the number of taxpayers of type 1, and how the reservation bribe reacts to changes in the before-tax income of type 1 and the anti-corruption expenditure, respectively.

Let \( \Delta_j = \beta_j V^j_C / \gamma \geq 1 \) denote the distributional weight the government attaches to individuals of type \( j \) at the social optimum, where \( \gamma = \sum_{j=1}^{n-1} \sum_{i=j+1}^{n} \lambda_{ij} V^i_C \frac{\partial B(Y_j, E)}{\partial E} \) from the social first-order condition for \( E \). Corollary 4 below is an immediate consequence of the calculations behind Proposition 3.

**Corollary 4.** The policy rules in Proposition 3 can be reformulated to read

\[
T'(Y_j, E) = \frac{1 - \Delta_j \frac{\partial B(Y_j, E)}{\partial Y_j}}{\Delta_j} \geq 0
\]

for \( j = 1, ..., n \).

**Proof:** see the Appendix.

Corollary 4 is analogous to Corollary 2, with the exception that the marginal cost of public funds now equals the marginal welfare benefit of \( E \). Therefore, since policy rule (32) is an alternative formulation of policy rule (29), it is also consistent with special cases (30) and (31) in Corollary 3. Here, where the optimal tax structure and the salary of the taxman are chosen simultaneously, we can see that the distributional weight the government attaches to any type \( j \) depends on the effects of \( E \) on the reservation bribe. More specifically, the higher the overall marginal effects of the government’s anti-corruption expenditure on the reservation bribe, the lower the distributional weight attached to any type
This paper develops a Mirrleesian model of optimal redistributive taxation in an economy with potential corruption in the tax administration. In our framework, there are two potential channels for individuals to mimic people with lower income: by reducing their labor supply (as in conventional models of optimal taxation) or by tax evasion accomplished through bribing the tax collector. The latter mechanism is novel in the literature on optimal taxation and also the focus of the present paper. Thus, public sector concerns for incentive compatibility lead the policy maker to design the marginal tax schedule such that tax evasion becomes undesirable for each individual. The contribution of the paper is to characterize the marginal tax schedule and show how this schedule can be designed to avoid tax evasion. In turn, this tax schedule reflects how the discrepancy between the taxpayers’ reported and actual income affects the reservation bribe required by the tax collector to misreport individual income to the policy maker, and how public expenditures targeting the incentives of the tax collector affect this bribe. Our characterization is general in the sense of applying regardless of which (out of several possible) self-selection constraints that binds for a given ability-type. The policy rules for marginal income taxation can be expressed in terms of two key variables: the distributional weights attached to individuals, and how sensitive the reservation bribe is to changes in the reported income.

Future research may take several directions. One would be to integrate corruption in the tax administration (which is the focus of our paper) with other possible mechanisms underlying tax evasion such as imperfect monitoring. This would enable us to examine the simultaneous design of marginal income tax schedules and penalty schedules, which is clearly an interesting topic. In addition, to the extent that tax evasion is more common among the wealthy, another relevant extension would be to analyze the policy implications of tax evasion in a dynamic framework with taxes on capital income and/or wealth. We hope to address both these issues in future research.
7. Appendix

Brief explanation of Footnote 7:

Suppose \( T(Y_1) < T(Y_2) \), i.e., \( Y_1 - C_1 < Y_2 - C_2 \). Let

\[
\bar{V}^1(C_2, Y_2) = \max U^1(C, Y), \text{ s.t., } C \leq Y - (Y_2 - C_2) - B(Y_2),
\]

and

\[
V^1(C_1, Y_1) = \max U^1(C, Y), \text{ s.t., } C \leq Y - (Y_1 - C_1).
\]

For \( B(Y_2) > 0 \), since

\[
Y - (Y_1 - C_1) > Y - (Y_2 - C_2) > Y - (Y_2 - C_2) - B(Y_2),
\]

then

\[
C \leq Y - (Y_2 - C_2) - B(Y_2) \implies C < Y - (Y_1 - C_1).
\]

Therefore, \( V^1(C_1, Y_1) \geq \bar{V}^1(C_2, Y_2) \).

Proof of Proposition 1:

Solving the government’s maximization problem (15) facing the endogenous bribe \( B(Y_1; I_0) \), we obtain

\[
\frac{-V^2_{Y_2}}{V^2_{C_2}} = 1. \tag{33}
\]

\[
T'(Y_2) = 0 \tag{34}
\]

The social first-order conditions for \( Y_1 \) and \( C_1 \) read

\[
\beta_1 N_1 V^1_{Y_1} - \lambda \bar{V}^2_{Y_1} + \gamma N_1 = 0, \tag{35}
\]

\[
\beta_1 N_1 V^1_{C_1} - \lambda \bar{V}^2_{C_1} - \gamma N_1 = 0. \tag{36}
\]

Combine these equations to derive

\[
\gamma N_1 \left( \frac{V^1_{Y_1}}{V^1_{C_1}} + 1 \right) = \lambda \bar{V}^2_{C_1} \left( \frac{V^2_{Y_1}}{V^2_{C_1}} - \frac{V^1_{Y_1}}{V^1_{C_1}} \right). \tag{37}
\]

By using the private first-order condition \( V^i_{Y_1}/V^i_{C_1} + 1 = T'(Y_i) \), we have

\[
T'(Y_1) = \frac{\lambda \bar{V}^2_{C_1}}{\gamma N_1} \left( \frac{V^2_{Y_1}}{V^2_{C_1}} - \frac{V^1_{Y_1}}{V^1_{C_1}} \right). \tag{38}
\]
Now, use

\( \bar{V}_{C_1}^2 = V_{C_2}^2 \).

\( \bar{V}_{Y_1}^2 = -(1 + \partial B_1 / \partial Y_1). \)

Substitute into equation (38) and rearrange

\( T'(Y_1) = -\frac{\lambda \bar{V}_{C_1}^2 / \gamma N_1}{1 + \lambda \bar{V}_{C_1}^2 / \gamma N_1} \frac{\partial B_1}{\partial Y_1} > 0. \)

**Proof of Proposition 2:**

For type 1, the social first-order conditions for \( Y_1 \) and \( C_1 \) read

\( \beta_1 N_1 V_{Y_1}^1 - \sum_{i=2}^{n} \lambda_i,1 \bar{V}_{Y_1}^i + \gamma N_1 = 0, \)

\( \beta_1 N_1 V_{C_1}^1 - \sum_{i=2}^{n} \lambda_i,1 \bar{V}_{C_1}^i - \gamma N_1 = 0. \)

Following the same procedure of calculation of two types model, we obtain

\( T'(Y_1) = -\frac{\sum_{i=2}^{n} \lambda_i,1 \bar{V}_{C_1}^i / \gamma N_1}{1 + \sum_{i=2}^{n} \lambda_i,1 \bar{V}_{C_1}^i / \gamma N_1} \frac{\partial B_1}{\partial Y_1} > 0. \)

For type 2, since all the type with higher productivity than type 2 would evade tax and potentially mimicking type 2, the social first-order conditions for \( Y_2 \) and \( C_2 \) read

\( \beta_2 N_2 V_{Y_2}^2 - \sum_{i=3}^{n} \lambda_i,2 \bar{V}_{Y_2}^i + \gamma N_2 = 0, \)

\( \beta_2 N_2 V_{C_2}^2 - \sum_{i=3}^{n} \lambda_i,2 \bar{V}_{C_2}^i - \gamma N_2 = 0. \)

we obtain

\( T'(Y_2) = -\frac{\sum_{i=3}^{n} \lambda_i,2 \bar{V}_{C_2}^i / \gamma N_2}{1 + \sum_{i=3}^{n} \lambda_i,2 \bar{V}_{C_2}^i / \gamma N_2} \frac{\partial B_2}{\partial Y_2} > 0. \)

And so on, for type \( n-1 \), we obtain

\( T'(Y_{n-1}) = -\frac{\lambda_{n,(n-1)} \bar{V}_{C_{n-1}}^n / \gamma N_{n-1}}{1 + \lambda_{n,(n-1)} \bar{V}_{C_{n-1}}^n / \gamma N_{n-1}} \frac{\partial B_{n-1}}{\partial Y_{n-1}} > 0. \)
There is no other types mimicking type \( n \), we obtain

\[
T'(Y_n) = 0.
\]

We conclude that for type \( j = 1, \ldots, n-1 \),

\[
T'(Y_j) = -\sum_{i=j+1}^{n} \frac{\lambda_{i,j} V_{i,C_j}^i/\gamma N_j}{1 + \sum_{i=j+1}^{n} \lambda_{i,j} V_{i,C_j}^i/\gamma N_j} \frac{\partial B_j}{\partial Y_j} > 0.
\]

**Proof of Corollary 1:**

The proof of Corollary 1 immediately follow from Proposition 2 by making the appropriate assumptions about binding self-selection constraints. Specifically, policy rule (22) is obtained by assuming that the best evasion strategy is to mimic the tax payment of the adjacent type with lower productivity, which the binding self-selection constraint serves to prevent. That is, for \( i = 2, 3, \ldots n \), if \( j = i - 1 \), \( \lambda_{i,j} > 0 \), otherwise, \( \lambda_{i,j} = 0 \).

Instead, if each individual of type \( j > 1 \) would potentially mimic type 1, then we solve the Lagrangean function (20) assuming that for \( i = 2, 3, \ldots n \), \( \lambda_{i,1} > 0 \), while for any \( j \neq 1 \), \( \lambda_{i,j} = 0 \). Hence, the proof of policy rule (23) is obtained.

**Proof of Corollary 2:**

The social first-order condition of Lagrangean function (20) for \( C_j \) can be written as

\[
\beta_j N_j V_{C_j}^j - \gamma N_1 = \sum_{i=j+1}^{n} \lambda_{i,j} V_{i,C_j}^i.
\]

Substitute equation (51) into equation (21),

\[
T'(Y_j) = -\frac{(\beta_j N_j V_{C_j}^j - \gamma N_j)/\gamma N_j}{1 + (\beta_j N_j V_{C_j}^j - \gamma N_j)/\gamma N_j} \frac{\partial B_j}{\partial Y_j}.
\]

Rearrangement gives

\[
T'(Y_j) = -\frac{\beta_j N_j V_{C_j}^j/\gamma N_j - 1}{\beta_j N_j V_{C_j}^j/\gamma N_j} \frac{\partial B_j}{\partial Y_j}.
\]

This can also be written as

\[
T'(Y_j) = \left( \frac{1}{\beta_j V_{C_j}^j/\gamma} - 1 \right) \frac{\partial B_j}{\partial Y_j}.
\]
Define the social welfare weight attached to type \( j \), \( \delta_j = \beta_j V C_j / \gamma \), we have

\( T'(Y_j) = \frac{1 - \delta_j \partial B_j}{\delta_j \partial Y_j} \geq 0. \)

**Calculation of Equation (27):**

The optimal consumption of the type \( i \) satisfies

\( \tilde{C}_i(C_j, Y_j, E) \equiv \tilde{Y}_i(C_j, Y_j, E) - (Y_j - C_j) - B(Y_j, E). \)

Meanwhile, at the optimal point, the first order condition (26) gives

\( \frac{\partial V^i(\tilde{C}_i(C_j, Y_j, E), \tilde{Y}_i(C_j, Y_j, E)) + \partial V^i(\tilde{C}_i(C_j, Y_j, E), \tilde{Y}_i(C_j, Y_j, E))}{\partial Y_i} \equiv 0. \)

Then we obtain

\( \frac{\partial \tilde{V}^i(C_j, Y_j, E)}{\partial E} = \frac{\partial V^i(\tilde{C}_i, \tilde{Y}_i)}{\partial C_i} \frac{\partial \tilde{C}_i}{\partial E} + \frac{\partial V^i(\tilde{C}_i, \tilde{Y}_i)}{\partial Y_i} \frac{\partial \tilde{Y}_i}{\partial E}. \)

Since

\( \frac{\partial \tilde{C}_i(C_j, Y_j, E)}{\partial E} = \frac{\partial \tilde{Y}_i(C_j, Y_j, E)}{\partial E} - \frac{\partial B(Y_j, E)}{\partial E}, \)

by envelop theorem, we obtain

\( \frac{\partial \tilde{V}^i(C_j, Y_j, E)}{\partial E} = -\frac{\partial V^i(\tilde{C}_i, \tilde{Y}_i)}{\partial C_i} \cdot \frac{\partial B(Y_j, E)}{\partial E} \leq 0. \)

Substitute equation (13) into equation (60) could be rewritten as

\( \frac{\partial \tilde{V}^i(C_j, Y_j, E)}{\partial E} = -\frac{\partial V^i(C_j, Y_j, E)}{\partial C_j} \cdot \frac{\partial B(Y_j, E)}{\partial E} \leq 0. \)

**Proof of Proposition 3:**

The social first-order conditions for \( Y_1, C_1 \), and \( E \) read

\( \beta_1 N_1 V_{Y_1}^1 - \sum_{i=2}^{n} \lambda_{i1} \tilde{V}_{Y_1}^i + \gamma N_1 = 0, \)

\( \beta_1 N_1 V_{C_1}^1 - \sum_{i=2}^{n} \lambda_{i1} \tilde{V}_{C_1}^i + \gamma N_1 = 0, \)

\( \frac{\partial \mathcal{L}}{\partial E} = 0, \)
From equation (62) and (63), we obtain

\[ T'(Y_1, E) = -\sum_{i=2}^{n} \frac{\lambda_i \bar{V}_i}{\gamma N_1} \frac{\partial B(Y_1, E)}{\partial Y_1} > 0. \]

From equation (64), we obtain

\[ -\sum_{j=1}^{n-1} \lambda_{n,j} \frac{\partial \bar{V}^n(j, Y_j, E)}{\partial E} - \sum_{j=1}^{n-2} \lambda_{(n-1),j} \frac{\partial \bar{V}^{n-1}(j, Y_j, E)}{\partial E} - \ldots - \lambda_{21} \frac{\partial \bar{V}^2(C_1, Y_1, E)}{\partial E} = \gamma. \]

which could be rewritten as

\[ -\sum_{j=1}^{n-1} \sum_{i=j+1}^{n} \lambda_{ij} \frac{\partial \bar{V}^i(j, Y_j, E)}{\partial E} = \gamma. \]

Substitute equation (61) into (67), we obtain

\[ \sum_{j=1}^{n-1} \sum_{i=j+1}^{n} \lambda_{ij} \bar{V}_i^j \frac{\partial B(Y_j, E)}{\partial E} = \gamma \]

which can be rewritten as

\[ \sum_{j=1}^{n-1} \sum_{i=j+1}^{n} \lambda_{ij} \frac{\partial \bar{V}^i(j, Y_j, E)}{\partial E} = \gamma. \]

Substitute equation (68) into (65), we obtain

\[ T'(Y_1, E) = \frac{\sum_{i=2}^{n} \lambda_i \bar{V}_i / \sum_{j=1}^{n-1} \sum_{i=j+1}^{n} \lambda_{ij} \bar{V}_i^j \frac{\partial B(Y_j, E)}{\partial E} N_1}{1 + \sum_{i=2}^{n} \lambda_i \bar{V}_i / \sum_{j=1}^{n-1} \sum_{i=j+1}^{n} \lambda_{ij} \bar{V}_i^j \frac{\partial B(Y_j, E)}{\partial E} N_1} \frac{\partial B(Y_1, E)}{\partial Y_1} > 0. \]

Following the same procedure of calculation of Proposition 2, we obtain for type \( j = 1, \ldots, n-1 \),

\[ T'(Y_j, E) = \frac{\sum_{i=j+1}^{n} \lambda_{ij} \bar{V}_i^j / \sum_{j=1}^{n-1} \sum_{i=j+1}^{n} \lambda_{ij} \bar{V}_i^j \frac{\partial B(Y_j, E)}{\partial E} N_j}{1 + \sum_{i=j+1}^{n} \lambda_{ij} \bar{V}_i^j / \sum_{j=1}^{n-1} \sum_{i=j+1}^{n} \lambda_{ij} \bar{V}_i^j \frac{\partial B(Y_j, E)}{\partial E} N_j} \frac{\partial B(Y_j, E)}{\partial Y_j} > 0 \]

Since there is no other types mimicking type \( n \), we obtain

\[ T'(Y_n) = 0 \]

**Proof of Corollary 4:**

The social first-order condition of Lagrangean function (28) for \( C_j \) can be written as

\[ \beta_j N_j V_{C_j}^j - \gamma N_1 = \sum_{i=j+1}^{n} \lambda_{ij} \bar{V}_i^j. \]
Substitute equation (68) and (72) into policy rule in Proposition 3, i.e., equation (29),

\[ (73) \quad T'(Y_j) = -\frac{(\beta_j N_j V^j_{C_j} - \gamma N_j) / \gamma N_j}{1 + (\beta_j N_j V^j_{C_j} - \gamma N_j) / \gamma} \frac{\partial B_j}{\partial Y_j}. \]

Rearrangement gives

\[ (74) \quad T'(Y_j) = -\frac{\beta_j N_j V^j_{C_j} / \gamma N_j - 1}{\beta_j N_j V^j_{C_j} / \gamma} \frac{\partial B_j}{\partial Y_j}. \]

This can also be written as

\[ (75) \quad T'(Y_j) = \left( \frac{1}{\beta_j V^j_{C_j} / \gamma} - 1 \right) \frac{\partial B_j}{\partial Y_j}. \]

Define \( \Delta_j = \beta_j V^j_{C_j} / \gamma \geq 1 \), where \( \gamma = \sum_{j=1}^{n-1} \sum_{i=j+1}^{n} \lambda_{ij} \dot{V}^i_{C_j} \frac{\partial B(Y_j, E)}{\partial E} \). We have

\[ (76) \quad T'(Y_j) = \frac{1 - \Delta_j \frac{\partial B(Y_j, E)}{\partial Y_j}}{\Delta_j}. \]

REFERENCES


