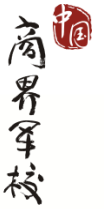


Working Paper No. 2014005



北京大学
汇丰商学院

Peking University HSBC Business School



PHBS

Income Tax Evasion and Optimal Income Taxation

Insook Lee

Copyright © 2014 by Insook Lee. All rights reserved.

PHBS working papers are distributed for discussion and comment purposes only. Any additional reproduction for other purposes requires the consent of the copyright holder.

Income Tax Evasion and Optimal Income Taxation

Insook Lee *

May 17, 2014

Abstract

This paper characterizes an optimal nonlinear income tax that incorporates individuals' decisions regarding tax evasion and labor supply and an optimal rate of tax enforcement when implementing tax code is costly. This enables us to resolve two lingering theoretical ambiguities on how tax rate affects tax evasion and on how probability of detecting tax evasion affects tax evasion. First, an increase in the tax rate is proven to foster tax evasion. Second, we prove that tax evasion responds negatively to an increase in the probability of detection, which invalidates the puzzling case in which enhanced tax enforcement promotes tax evasion. Moreover, we show that the non-asymptotic optimal marginal income tax rate on the richest individuals can be strictly positive.

Key Words: Tax evasion, Optimal nonlinear income tax, Tax evasion, Effect of tax rate on tax evasion, Effect of enforcement rate on tax evasion

JEL Classification: H21, H24, H26

*Peking University, HSBC Business School, Nanshan District, Shenzhen, Guangdong, China
islee@phbs.pku.edu.cn

1 Introduction

Tax evasion is a significant economic problem. For example, in the US, income tax revenue leakage was estimated at \$197 million in 2001 (Slemrod 2007); moreover, in Italy, unpaid taxes were estimated at €285 billion in 2012 (18% of GDP).¹ Difficulties in collecting taxes are even greater in developing countries (IMF 2011). Despite its significance, tax evasion has not been well studied in economics. None of the previous studies of optimal nonlinear income tax have explicitly incorporated tax evasion. However, no matter how optimal an income tax schedule is, it may not be meaningful when taxes are not collected according to the tax schedule. To date, economic theory has only shown that how tax evasion is affected by tax rate and by probability of detecting tax evasion are ambiguous. In particular, the latter result means that tax evasion may rise when governments improve tax enforcement, which is puzzling. The present paper aims to characterize an optimal nonlinear income taxation that incorporates tax evasion, based on which we try to resolve the two theoretical ambiguities regarding the impact on tax evasion of tax rate and that of detecting probability.

Tax evasion has been ignored in the models for optimal nonlinear income taxation, as implementation of any income tax schedule has been guaranteed by the unrealistic assumption of perfect enforcement of the government in the optimal nonlinear income tax literature (e.g., Mirrlees 1971; Jacquet et al. 2013). Obviously, substantive tax evasion would have not existed if perfect enforcement of tax rule were given for free. Rather, implementing tax rules is costly. Thus, the government needs to elaborate on tax enforcement. Tax enforcement variable has been isolated from the optimal income taxation studies, as it is simply assumed. As a consequence of the isolation, optimal tax enforcement is defined as maximizer of net tax revenue, rather than social welfare (Reinganum and Wilde 1985; Mookherjee and Png 1989; Sánchez and Sobel 1993). However, even when an income tax schedule is designed optimally for a society, social optimum is not effectively realized if tax evasion issue is ignored. That is, tax enforcement is an indispensable part of taxation.

The first theoretical framework used to analyze tax evasion behavior was proposed by Allingham and Sandmo (1972). Their model is limited because it does not allow individuals to determine their own working hours by as-

¹Tax evasion in Italy: Big government meets big data (January 8th 2013) *Economist*. Retrieved from <http://www.economist.com/node/21569195>

suming that labor earnings are exogenously given. The subsequent studies of Pencavel (1979), Baldry (1979), and Horowitz and Horowitz (2000) have improved on this limitation by allowing individuals to jointly decide their labor supply and degree of tax evasion; however, they assume that income tax rate is exogenously given. On the other hand, Sandmo (1981) derived an optimal linear income tax rate in the presence of tax evasion; however, Sandmo (1981) does not completely separate tax evasion from labor supply variable. Most importantly, all of the previous studies fail to resolve the ambiguities regarding the effect of tax rate on tax evasion and the effect of a rise in the probability of detecting tax evasion on tax evasion. Furthermore, none of the previous studies elaborate on any further conditions under which the impact of the detecting rate on tax evasion may be positive, which means that governments' efforts to strengthen tax enforcement only invite more tax evasion. Above all, the theoretical indeterminateness regarding tax evasion responses to these two basic policy variables have been left unresolved and almost neglected for decades.

The present paper improves on previous studies by incorporating tax evasion into characterizing an optimal nonlinear income tax schedule. We relax the assumption of perfect tax enforcement by allowing the tax enforcement to be costly. Furthermore, in deriving an optimal income tax schedule that maximizes social welfare, we treat the optimal income tax rate as endogenous to individuals' decisions with respect to both labor supply and degree of tax evasion. In addition to income tax schedule, the present paper derives an optimal probability of detecting income tax evasion that maximizes social welfare.

The paper makes the following contributions. First, it presents formulae for both an optimal nonlinear income tax that incorporates tax evasion and an optimal rate of tax enforcement when tax enforcement is costly. This enables resolution of the lingering theoretical ambiguities regarding income tax evasion. That is, a proof that a rise in the tax rate increases tax evasion is provided. In addition, this paper shows that enhancing tax enforcement clearly does not promote tax evasion. Finally, this study proves that a non-asymptotic optimal marginal income tax rate on the richest can be strictly positive, instead of zero.

This paper is organized as follows. Section 2 provides a theoretical model in which an optimal nonlinear income tax that incorporates responses of tax evasion and labor supply of heterogeneous taxpayers and an optimal rate of tax enforcement can be derived. Section 3 simplifies the optimization

problem stated in Section 2. Section 4 characterizes optimal income taxation (tax rate and tax enforcement rate) and re-visits our main problem of how tax evasion is affected by the tax rate and by the probability of detecting tax evasion. Section 5 concludes the paper.

2 The Model

Consider a society populated by a continuum of individuals whose preference admits a utility representation that satisfies the Von Neumann-Morgenstern axioms. The preference is represented by a function $u(x, l)$, where $x \in \mathbb{R}_+$ is a composite consumption good and $l \in [0, 1)$ is amount of time spent for working. We assume that $u(x, l)$ is strictly concave, continuously differentiable (C^1 function), strictly increasing in x and strictly decreasing in l . Importantly, individuals are differentiated by their innate earning abilities $n \in [\underline{n}, \bar{n}]$ with $\underline{n} > 0$. n is distributed according to a cumulative distribution function (hereafter, CDF) $H(n)$ with full support and a probability density function (hereafter, PDF) $h(n)$. When an individual of ability n spends l amount of time working, his effective labor is $nl \equiv L$. Given the wage rate w , this implies that a given amount of time, l , spent working yields higher income (wL) for an individual with higher n . Thus, the identifier variable n indicates an individual's innate ability for earnings.

When individuals file taxes, the government does not have accurate knowledge of their personal incomes, although it knows whether they earn income from work. Thus, the government assesses taxes based on voluntarily reported income. Then, the government audits on submitted tax reports can detect and penalize tax evasion with probability p , which is less than one. Exploiting this, a rational individual will consider understating his income to the tax authority. Denoting the ratio of reported income to true income as $r \in (0, 1]$, an individual taxpayer can choose a value of r less than one, given a tax schedule T .² One of two possible outcomes will occur. If tax evasion is detected, then disposable income for consumption of an individual of ability n is $x^D = wnl - T(wnl) - \theta\{T(wnl) - T(rwnl)\}$ where θ is penalty rate. If tax evasion is not detected, he can consume as much as $x^{ND} = wnl - T(rwnl)$.

²Since the government tells whether an individual taxpayer earns positive amount of taxable income or not and $\forall n > 0$, none would declare himself as non-taxpayer by reporting zero income (choosing $r = 0$).

In sum, the expected utility of the individual of ability n is stated as

$$pu(wnl - T(wnl) - \theta\{T(wnl) - T(rwnl)\}, l) + (1-p)u(wnl - T(rwnl), l). \quad (1)$$

For short, let this be denoted as $E[u(wnl - T(rwnl), l)]$. Then, given the tax schedule T , the penalty rate θ and enforcement rate p , the individual solves

$$\max_{r, l} E[u(wnl - T(rwnl), l)] \quad (2)$$

obtaining his indirect utility function $v_n \equiv E[u(wnl_n - T(r_n wnl_n), l_n)]$.

To begin with characterizing optimal understatement of income, the first-order condition (henceforth, FOC) for r_n is

$$p\theta u_1^D = (1-p)u_1^{ND} \quad (3)$$

where u_1^D is the marginal utility of consumption if tax evasion is detected and penalized and u_1^{ND} is the marginal utility of consumption if tax evasion is not detected. In other words, the optimal evasion level e_n , which is $1 - r_n$, is chosen to equalize the expected marginal utility of consumption that the taxpayer obtains when tax evasion is detected with the expected marginal utility of consumption that he obtains when tax evasion is not detected. Essentially, tax evasion is an investment in a risky gamble with two possible outcomes (determined by whether tax evasion is detected or not). Thus, the decision regarding degree of tax evasion (e_n) is about the degree to which a taxpayer exposes his disposable income for consumption (which otherwise would certainly be $wnl_n - T(wnl_n)$) to uncertainty (either x^{ND} or x^D). Above all, the income for investing in the tax evasion gamble is procured by taxpayer's labor. Consequently, the uncertainty regarding the final consumption carries over to the marginal value of working. That is, if tax code is enforced with penalizing tax evasion, $-u_2 = u_1^D wn\{1 - (1 + \theta[1 - r_n])T'\}$; otherwise, $-u_2 = u_1^{ND} wn(1 - r_n)T'$. Taking these together, the optimal condition of l_n is derived as

$$pu_1^D wn\{1 - (1 + \theta[1 - r_n])T'\} + (1-p)u_1^{ND} wn(1 - r_n)T' = -u_2. \quad (4)$$

That is, the optimal working hours l_n are chosen to equalize the marginal disutility of working with the expected marginal utility of disposable income for consumption. In turn, from the optimal working hours l_n defined by (4), the effective labor supply of an individual of ability n is derived as $nl_n \equiv L_n$.

Thus, in designing an income tax schedule T , the government can use (4) to know how much labor is supplied by an individual of ability n responding to the income tax schedule T if the tax rates are precisely applied to the intended individual.

In addition, there are profit-maximizing firms that supply the composite good under a perfect competition. Firms' technology is characterized by a production function $F(L)$ which exhibits constant returns to scale and satisfies $F(0) = 0$. Because only relative prices matter, without loss of generality, we normalize the price of the composite good to one. Thus, each firm solves $\max_L F(L) - wL$, where w is the market wage rate paid for the input of effective labor. The FOC of a firm's optimization is $w = F'$, which implies that each firm earns zero profits, based on Euler's homogeneous functions theorem. Hence, ownership of firms does not matter. What matters is the aggregate output of the economy which is stated as $F(\int_{\underline{n}}^{\bar{n}} L_n h(n) dn) = \int_{\underline{n}}^{\bar{n}} w L_n h(n) dn$ because $w = F'$. Utilizing this, we obtain the economy's (ex-ante) resource constraint which is $\int_{\underline{n}}^{\bar{n}} \{p x_n^D + (1 - p) x_n^{ND}\} h(n) dn \leq \int_{\underline{n}}^{\bar{n}} w L_n h(n) dn$.

The government of the society determines the income tax schedule T as well as p and θ . In designing these tax policies, the government respects the utility of each individual and pursues to maximize a social welfare function $W(v_{\underline{n}} \cdots v_n \cdots v_{\bar{n}})$ (henceforth, SWF) which is a weighted sum of individuals' utilities. As the society is populated by a continuum of individuals, the SWF is stated as $W = \int_{\underline{n}}^{\bar{n}} \frac{dG}{dv_n} v_n h(n) dn$, where $\frac{dG}{dv_n}$ is the social marginal value of utility of individuals of ability n . Because the government respects each individual's utility, $\infty > \frac{dG}{dv_n} > 0$ for $\forall n$. For example, if the government assigns equal weight to each individual, $\frac{dG}{dv_n} = 1$ for $\forall n$. To embody more general distributional preferences, we assume that the government does not give greater weight to individuals who enjoy higher level of utility. That is, we let $\frac{d}{dv_n} (\frac{dG}{dv_n}) \leq 0$. This means that the SWF is weakly concave. As a result, imposing taxes is socially desirable. To see why, note that the marginal utility of one unit of the consumption good is higher for individuals with lower values of utility. Thus, the government can increase social welfare through a redistribution of income via taxation as this may yield a social net gain, unless the government puts more weights individuals with higher levels of utility.

Most importantly, the government cannot observe and verify the innate ability of an individual at any cost, due to its nature as an inner characteristic. As a consequence, an interpersonal lump sum transfer is not available as

a tax instrument because this requires the government to observe the identifier variable in order to transfer the correct amounts between designated individuals. Instead, the government can levy income taxes on earnings reported by individuals.³ For tractable analysis, let the income tax schedule T be piece-wise continuously differentiable.

Markedly, announcing an income tax schedule T does not automatically guarantee its perfect execution. Rather, revenue mobilization through tax enforcement involves various phases, each of which is costly. Otherwise, the government would completely eradicate tax evasion by simply choosing $p = 1$. In the first place, an established taxation system and a professional workforce that transparently administers the system are not given for free, although they are an important prerequisite infrastructure for tax enforcement. In addition, maintaining and operating the tax system also incurs administrative costs.⁴ Without such costs for establishing a transparent tax system and for having the system operate well, tax audits alone may not successfully detect and verify fraudulent tax evasion. Auditing is just one part of tax enforcement.

For simplicity, instead of introducing several variables corresponding to various phases of tax administration and enforcement, let us simplify income tax revenue mobilization into a single stage process. Namely, at the site where individuals report their earnings and pay their income tax according to the given tax schedule, the government can verify the extent to which individuals evade their tax liabilities $(1 - r)$ and penalize evasion, with a success rate of p . Thus, p refers to the overall rate of tax enforcement rather than just the audit rate. More importantly, improving the enforcement rate p requires resources that can formally be described by a cost function $\frac{1}{\delta}c(p)$. First, $\delta \in (0, \infty)$ is a given parameter that reflects the efficiency and effectiveness of the government tax enforcement. This parameter captures all factors that affect the performance of the tax system, including institutional and political aspects. Second, $c(p)$ is strictly convex and increasing in p , with $\frac{1}{\delta}c(0) > 0$. This implies that the cost function c is invertible. Thus, the government

³Also, the government can levy excise taxes at the same time. Then, we can innocuously assume that optimal commodity taxes are adopted and already reflected on the consumer price.

⁴For example of US, its tax collection agency Internal Revenue Service (IRS) spends \$10 billion for maintaining individuals' tax return files alone, which must be a small fraction of administrative cost (Guyton et al. 2003). In light of this, the total cost for tax enforcement would be sizable.

decides the level of p by allocating $\frac{1}{\delta}c(p)$ resources to tax enforcement. In practice, however well the tax system may operate, investigative audits on all taxpayers would be too costly for any government to conduct, which can be stated as $\lim_{p \rightarrow 1} \frac{1}{\delta}c(p) = \infty$.

In addition to choosing p , the government decides penalty rate θ that it charges on unpaid tax due when it detects tax evasion.⁵ Presumably, legislating penalty rate of θ is an almost costless measure to curb tax evasion. Nonetheless, the government cannot legislate $\theta = \infty$ to achieve complete tax compliance, as legal penalties must accord with the overall hierarchy of law in the society. That is, in harmony with penalties on greater crimes that seriously threaten the social order and security, the harshness of the penalty for tax evasion cannot be extreme but must be fairly moderate. In this regard, let $\bar{\theta}$ be the maximum penalty on tax evaders under a given legal order of society. After all, the government will select $\theta = \bar{\theta}$ to deter revenue leakage to the extent possible.

All in all, the main problem that the government seeks to solve can be written

$$\max_{T(\cdot) \text{ \& } p} \int_{\underline{n}}^{\bar{n}} \frac{dG}{dv_n} v_n h(n) dn \text{ s.t. } \int_{\underline{n}}^{\bar{n}} \{p x_n^D + (1-p)x_n^{ND}\} h(n) dn \leq \int_{\underline{n}}^{\bar{n}} w n l_n h(n) dn. \quad (5)$$

Notice that the government's budget constraint is not added as it is redundant according to the Walras' law. Moreover, the time line is as follows: (i) income tax schedule T as well as p and $\bar{\theta}$ are announced by the government; and, (ii) individuals make decisions regarding their working hours (l_n) and degree of tax evasion ($e_n = 1 - r_n$).

Having explicitly stated the government's optimization problem, let us begin by finding an optimal income tax schedule $T(\cdot)$ which is a *function* and thus apparently more complicated than finding a single value of an optimal p .

At the outset, all elements in the domain of T should be identified by the source of inequality, namely, individuals' differing innate earning abilities. As noted above, social net gain arises from reduction of the income gap between individuals of high earning abilities and individuals of low earning abilities through the introduction of taxation. However, this equity gain comes with

⁵This is closer to reality than charging constant amount of fine; for instance of the US, the IRS charges a penalty of 20% of unpaid tax liabilities for the case of a substantial understatement of income tax.

an efficiency cost: a decrease in the labor supply. To see why, compare (4) with $u_1 w n = -u_2$ that determines labor supply under non-taxation. Fewer hours of work are provided when taxation is introduced because marginal income tax rate reduces marginal value of working and income tax leaves less income disposable for consumption. On the other hand, as shown in (3), tax rate itself does not directly determine tax evasion decision. Rather, tax evasion depends on disposable income for consumption (x^{ND} and x^D), which is derived from the labor supply once the income tax schedule T and p are determined.

Thus, the first task that the government must undertake for maximizing the SWF is to assess, for each n in the domain of T , how much individuals of ability n will work in response to an income tax schedule T , p and $\bar{\theta}$. The assessment task is not feasible unless the tax rate for each n takes effect precisely on the intended individuals of ability n , which is not automatically satisfied because the government cannot observe n (the innate earning ability of an individual). Individuals have an incentive to mislead the government into applying to them an erroneous tax rate (one intended for individuals of lower ability). In particular, as higher tax liabilities are imposed for individuals of higher earnings abilities, an individual of ability n can be better off with pretending to be of lower ability n' (where $n' < n$) by reducing his working hours so that his effective labor supply is $L_{n'}$ (instead of L_n that is greater than $L_{n'}$). Accordingly, his degree of tax evasion is set at $e_{n'}$ instead of e_n . By doing so, the ability n individual enjoys more consumption with less disutility of working because $\exists l_{n''}$ such that $L_{n'} = n' l_{n''} = n l_{n''}$ and $l_n > l_{n''}$. Taking this into account, the government should provide incentives for individuals to voluntarily reveal their own abilities. Otherwise, the government is stranded ignorant of where it stands in the domain of the income tax function T when designing the tax function T .

In the next section, the conditions necessary for the government to regain the lost link between tax rates and taxpayers' earning abilities will be elaborated further. Once such conditions are met, the government can know labor supply and degree of tax evasion of taxpayers of ability n , using (3) and (4). Based on this, the government can find the optimal tax rate for each n maximizing the SWF. Combining all the optimal tax rates, the optimal tax function can be defined. In other words, once the above-noted conditions have been incorporated into problem (5), the originally complex task of finding the optimal tax function is simplified into a point-wise maximization problem allowing for the use of simple calculus.

3 Simplification of the Optimization Problem

Let us derive the conditions under which the tax rate intended for individuals of ability n is correctly applied to individuals of ability n . Put in another way, we shall elaborate on how the government can provide incentives for individuals to virtually reveal their true ability voluntarily so that, using (3) and (4), it can correctly assess v_n (from l_n and r_n) and maximize the SWF.

First of all, the way an individual disguises himself as another individual of lower ability is to alter (reduce) his working hours. When an individual of ability n mimics the effective labor supply of an individual of ability n' ($L_{n'}$), he earns the same true income ($wL_{n'}$) as an individual of ability n' . However, the government cannot distinguish an individual who is truly of ability n' from one who pretends to be of the ability n' because it only observes earnings after verifying tax evasion. In particular, the individual of ability n who pretends to be of ability n' by altering his working hour obtains utility (payoff) $E[u(wL_{n'} - T(r_{n'}wL_{n'}), \frac{L_{n'}}{n})]$, which differs from $v_n = E[u(wL_n - T(r_nwL_n), \frac{L_n}{n})]$. To neatly express the disguise, denote the payoff to an individual of ability n when he pretends to be of ability n' as $V(n' | n)$. That is, $V(n' | n) = E[u(wL_{n'} - T(r_{n'}wL_{n'}), \frac{L_{n'}}{n})]$. Using this notation, the condition that gives ability n taxpayers the incentive not to pretend those of other ability n' can be stated as $V(n | n) \geq V(n' | n)$. We refer to this condition as an incentive compatibility constraint (hereafter, IC constraint).

In addition to misleading the government into levying an erroneous tax rate which is aimed for other individuals, the ability n individual has another option for responding to a given income tax schedule: to quit being a taxpayer by choosing zero working hours. Because the government knows whether an individual earns some income from work (albeit not the exact amount), the only way for the individual to be exempted from the entire tax liability (i.e., quit being a taxpayer) is to stop working. To take an extreme case, if all earnings were taxed away, no one would work. As another example, when the government provides a very generous aid or subsidies to non-taxpayers who do not earn income, individuals of the lowest ability $\underline{n} > 0$ would benefit from ceasing to work and receiving free aid. Therefore, to induce all the taxpayers to remain taxpayers, $V(n | n)$ must not be less than $u(-T(0), 0)$ (the utility of non-taxpayers). We refer to this requirement as individual-rationality constraint (henceforth, IR constraint) adopting the terminology from the contract theory literature.

Thus, the conditions for regaining the missing link between the tax rate

and the targeted taxpayer are concisely stated as follows:

For $\forall n \in [\underline{n}, \bar{n}]$,

(IC) $V(n | n) \geq V(n' | n)$ for $\forall n' \neq n$

(IR) $V(n | n) \geq u(-T(0), 0)$

Once these two sets of requirements are met, each tax rate in an income tax schedule T is correctly applied to the individuals whom the tax rate targets. Moreover, the government can now locate itself in the domain of an income tax function T as if it knows the earning abilities of individuals. We can thus derive the optimal income tax function T from combining the optimal tax rate for each n . Specifically, with given fixed n , we can obtain the optimal tax rate that maximizes the SWF, as the SWF is a weighted sum of individuals' utilities. As such, the government's task of finding the optimal income tax function is simplified into a point-wise maximization.

In implementing such a simplification, however, one issue remains: for each given ability n , the IC and IR constraint actually involve infinite numbers of inequalities. This difficulty necessitates few more steps of reducing these constraints to more tractable forms.

Lemma 1. For any given $n \in [\underline{n}, \bar{n}]$, the IC constraint is met if and only if (i) L_n is non-decreasing in n ; and, (ii) $V_1(n | n) = 0$ where $V_1(n | n)$ is a partial derivative with respect to the first argument of $V(n | n)$.

Proof. [step 1] (\implies) Suppose that the IC constraint is met for an arbitrarily given $n \in [\underline{n}, \bar{n}]$; namely, $V(n | n) \geq V(n' | n)$ for $\forall n' \in [\underline{n}, \bar{n}]$. Put another way, this means that $n = \arg \max_t V(t | n)$. First of all, this immediately implies that $V_1(n | n) = 0$ because it is the necessary condition for n to be the maximizer.

Next, by way of contradiction, suppose that $L_n = L(n)$ is decreasing in n ; that is, $L' = \frac{dL_n}{dn} < 0$. Pick any arbitrary n' such that $n > n' > \underline{n}$. Since the IC constraint is met, $V(n | n) \geq V(n' | n)$ which implies that $\int_{n'}^n V_1(t | n) dt \geq 0$. Moreover, $\int_{n'}^n V_1(t | n) dt = \int_{n'}^n V_1(t | n) - V_1(t | t) dt$ because we know that $V_1(t | t) = 0$. Since $V_1(t | n) = [pu_1^D w \{1 - (1 + \bar{\theta} - r_t \bar{\theta}) T'\} + (1 - p) u_1^{ND} w (1 - r_t T')]$ and $V_1(t | t) = [pu_1^D w \{1 - (1 + \bar{\theta} - r_t \bar{\theta}) T'\} + (1 - p) u_1^{ND} w (1 - r_t T')]$, $V_1(t | n) - V_1(t | t) = u_2 \frac{1}{n} L' - u_2 \frac{1}{t} L' = u_2 L' [\frac{1}{n} - \frac{1}{t}]$. As a result, $\int_{n'}^n V_1(t | n) dt = \int_{n'}^n u_2 L' [\frac{1}{n} - \frac{1}{t}] dt < 0$ because $L' < 0$, $u_2 < 0$, and $\frac{1}{n} - \frac{1}{t} < 0$ for $\forall t \in (n', n)$. A contradiction to the IC constraint. This proves that $L' \geq 0$.

[step 2] (\impliedby) Conversely, assume that (i) L_n is non-decreasing in n (i.e.,

$L' = \frac{dL_n}{dn} \geq 0$) and (ii) $V_1(n | n) = 0$ this time. Now suppose that the IC constraint is not satisfied. Then, there exists some $n' \in [\underline{n}, \bar{n}]$ such that $V(n | n) < V(n' | n)$. Hence, $n \neq \arg \max_t V(t | n)$ which implies that $V_1(n | n)$ cannot be zero: A contradiction to the assumption (ii). Therefore, this proves that the IC constraint is met. ■

Presumably, we can regard $V_1(n | n) = 0$ as a local IC constraint as it is a necessary condition for $n = \arg \max_t V(t | n)$. As this is met for each $n \in [\underline{n}, \bar{n}]$, satisfying such local IC constraints of different n 's all the way from the highest to the lowest would result in meeting the IC constraints globally. To this end, non-decreasing L_n brings overall coherency because it ensures that individuals with higher innate earning abilities end up with choosing their working hours to earn higher incomes under the given tax schedule. If non-decreasing L_n is violated, some individuals choose to get lower labor earnings than other individuals who are of lower ability than they are, which means that they must have lied about their innate ability, failing the IC constraints.

Now, let us move on the other pair of the requirements: meeting IR constraint.

Lemma 2. For any given $n \in [\underline{n}, \bar{n}]$, if the IC constraint is met, then the IR constraint is met.

Proof. By a way of contradiction, given that the IC constraint is met for an arbitrarily given $n \in [\underline{n}, \bar{n}]$, suppose that its IR constraint is not met. Then, $V(n | n) < u(-T(0), 0)$. Moreover, when a taxpayer of ability n pretends to be a non-taxpayer, his payoff $u(-T(0), 0)$ can be restated as $V(0 | n)$. Thus, $V(n | n) < V(0 | n)$. Since $n \neq 0$, this contradicts to the IC constraint that $V(n | n) \geq V(n' | n)$. ■

Notice that we reduced infinite numbers of inequalities into one inequality (i) and one equation (ii) in **Lemma 1**. Now, we bring these two conditions into conformity with (3) and (4) which the government eventually needs to use for maximizing the SWF.

Lemma 3. For any given $n \in [\underline{n}, \bar{n}]$, the IC constraint and IR constraint are met if and only if (i) $v_n = v_{\underline{n}} + \int_{\underline{n}}^n -\frac{t}{t} u_2 dt$ and (ii) $\int_{\underline{n}}^n -\frac{t}{t} u_2 [p \frac{1}{u_1^p} + (1 -$

$p) \frac{1}{u_1^{ND}}] dt \geq 0.$

Proof. [step 0] As a stepping stone to prove the above statement, we first need to show that for an income tax schedule T , v_n is increasing in n . To show this, pick any n and n' such that $n > n'$. Because $u_2 < 0$, $E[u(w n' l_{n'} - T(r_{n'} w n' l_{n'}), l_{n'})] < E[u(w n \frac{n'}{n} l_{n'} - T(r_{n'} w n \frac{n'}{n} l_{n'}), \frac{n'}{n} l_{n'})] \leq E[u(w n l_n - T(r_n w n l_n), l_n)]$. The first term $v_{n'} = E[u(w n' l_{n'} - T(r_{n'} w n' l_{n'}), l_{n'})]$ is the maximized utility of an ability n' individual. On the other hand, the second term $E[u(w n \frac{n'}{n} l_{n'} - T(r_{n'} w n \frac{n'}{n} l_{n'}), \frac{n'}{n} l_{n'})]$ is a value of utility of an ability n individual when he chooses his working hours as much as $\frac{n'}{n} l_{n'}$ and his report ratio as $r_{n'}$. Clearly, $E[u(w n \frac{n'}{n} l_{n'} - T(r_{n'} w n \frac{n'}{n} l_{n'}), \frac{n'}{n} l_{n'})]$ is always smaller or equal to the maximized utility of the ability n individual $v_n = E[u(w n l_n - T(r_n w n l_n), l_n)]$. Taken together, this results in $v_n > v_{n'}$. Therefore, v_n is increasing in n as $n > n' \implies v_n > v_{n'}$.

[step 1] (\implies) Now let us return to our main statement. For an arbitrarily given $n \in [\underline{n}, \bar{n}]$, suppose that the IC constraint and IR constraint are met. From **Lemma 1**, this implies that $V_1(n | n) = 0$. Based on the Fundamental theorem of calculus, it follows that $V(n | n) = V(\underline{n} | \underline{n}) + \int_{\underline{n}}^n V_1(t | t) + V_2(t | t) dt = V(\underline{n} | \underline{n}) + \int_{\underline{n}}^n -\frac{t}{t} u_2 dt$ because $V_1(t | t) = 0$ and $V_2(t | t) = -\frac{t}{t} u_2$. This means that $v_n = v_{\underline{n}} + \int_{\underline{n}}^n -\frac{t}{t} u_2 dt$ as $V(n | n) = E[u(w L_n - T(r_n w L_n))] = v_n$ and $V(\underline{n} | \underline{n}) = E[u(w L_{\underline{n}} - T(r_{\underline{n}} w L_{\underline{n}}))] = v_{\underline{n}}$.

Since $n \geq \underline{n}$, $v_n \geq v_{\underline{n}}$ based on [step 0]. Because $v_n = v_{\underline{n}} + \int_{\underline{n}}^n -\frac{t}{t} u_2 dt$, this means that $\int_{\underline{n}}^n -\frac{t}{t} u_2 dt \geq 0$. Finally, this implies that $\int_{\underline{n}}^n -\frac{t}{t} u_2 [p \frac{1}{u_1^D} + (1-p) \frac{1}{u_1^{ND}}] dt \geq 0$ because (i) both u_1^D and u_1^{ND} are always strictly positive and (ii) $p \in (0, 1)$.

[step 2] (\impliedby) Conversely, suppose that (i) $v_n = v_{\underline{n}} + \int_{\underline{n}}^n -\frac{t}{t} u_2 dt$, and that (ii) $\int_{\underline{n}}^n -\frac{t}{t} u_2 [p \frac{1}{u_1^D} + (1-p) \frac{1}{u_1^{ND}}] dt \geq 0$, for an arbitrarily given $n \in [\underline{n}, \bar{n}]$. By way of contradiction, assume that the IC constraint is not met. Then, there exists some $n' < n$ such that $V(n | n) < V(n' | n)$. This implies that $V(n' | n) - V(n | n) = \int_n^{n'} V_1(t | n) dt = - \int_n^{n'} V_1(t | n) dt > 0$.

On the other hand, we can restate $v_n = v_{\underline{n}} + \int_{\underline{n}}^n -\frac{t}{t} u_2 dt$ as $V(n | n) = V(\underline{n} | \underline{n}) + \int_{\underline{n}}^n V_2(t | t) dt$ since $V(n | n) = v_n$ and $V_2(t | t) = -\frac{t}{t} u_2$. In turn, $V(n | n) = V(\underline{n} | \underline{n}) + \int_{\underline{n}}^n V_1(t | t) + V_2(t | t) dt$, based on the Fundamental theorem of calculus. This implies that $V_1(n | n) = 0$.

Thus, $\int_n^{n'} V_1(t | n) dt = \int_n^{n'} V_1(t | n) - V_1(t | t) dt$ since $V_1(t | t) = 0$. Since $V_1(t | n) = [p u_1^D w \{1 - (1 + \theta - r_t \theta) T'\} + (1-p) u_1^{ND} w (1 - r_t T')] L' + u_2 \frac{1}{n} L'$

and $V_1(t | t) = [pu_1^D w \{1 - (1 + \bar{\theta} - r_t \bar{\theta}) T'\} + (1 - p)u_1^{ND} w (1 - r_t T')] L' + u_2 \frac{1}{t} L'$,
 $V_1(t | n) - V_1(t | t) = u_2 \frac{1}{n} L' - u_2 \frac{1}{t} L' = u_2 L' [\frac{1}{n} - \frac{1}{t}]$. As a result, $-\int_{n'}^n V_1(t | n) dt = -\int_{n'}^n u_2 L' [\frac{1}{n} - \frac{1}{t}] dt < 0$. This implies that $L' = \frac{dL_n}{dn} > 0$ because $u_2 < 0$ and $[\frac{1}{n} - \frac{1}{t}] < 0$ for $\forall t \in (n', n)$. This is a contradiction to $-\int_{n'}^n V_1(t | n) dt > 0$. Therefore, the above (i) and (ii) implies that the IC constraint is satisfied. Due to **Lemma 2**, this in turn implies that the IR constraint is also met if the above (i) and (ii) are met. ■

It is worth illustrating the economic intuitions underlying the mathematical statements (i) and (ii) of **Lemma 3**. First, note that $v_n = v_n + \int_n^n -\frac{t}{t} u_2 dt$ is the payoff to an individual of ability n for being a taxpayer without misrepresenting himself to the government. The first term on right-hand side is the reservation payoff such that the income tax schedule retains the ability n individual as a taxpayer. The second term on right-hand side is to compensate the individual for not-pretending to be of lower ability by decreasing his working hours.⁶ The government cannot distinguish an individual of ability n who pretends to be of slightly lower ability $n - \varepsilon$ (where $\varepsilon > 0$) from an individual who truly is of ability $n - \varepsilon$. The payoff to the former is $V(n - \varepsilon | n)$, while the payoff to the latter is $V(n - \varepsilon | n - \varepsilon)$. Hence, the surplus that the individual of ability n can exploit as a result of the government's inability to see innate earning ability is $V(n - \varepsilon | n) - V(n - \varepsilon | n - \varepsilon)$. Unless the individual of ability n is compensated by this amount, he would rather mislead the government into believing him to be an ability $n - \varepsilon$ individual. By the same reasoning, the compensation required to deter the ability n individual from pretending to be an individual of the closest lower ability is $\lim_{\varepsilon \rightarrow 0} \frac{V(n - \varepsilon | n) - V(n - \varepsilon | n - \varepsilon)}{\varepsilon} = V_2(n | n)$. Aggregating the surplus of pretending to be of all the other lower abilities results in $\int_n^n V_2(t | t) dt$. This is the second term of the payoff v_n , because $V_2(t | t) = -\frac{t}{t} u_2$. In contract theory terminology, this compensation is known as the 'informational rent.' Presumably, this is a theoretical device for capturing the price that the government pays for limiting individual's freedom. Hypothetically, if the government had absolute power to extract private information regarding innate earning ability, it could deny individuals' freedom to keep their private information and then force them to pay taxes and work as much as they would under non-taxation

⁶Pretending to be of higher ability does not need to be considered since it is strictly dominated strategy.

(without paying them the informational rents). However, such an intrusion on freedom would hardly be politically justifiable or infeasible for the government to conduct or unjustifiable. Although taxation benefits society, it is not necessarily desired by individuals. Individuals have the freedom to keep their private information in pursuit of their own personal interests, which can undermine taxation. Hence, the government needs to pay some costs to induce individuals' cooperation of revealing their private information for implementing socially desirable income taxation. The government should take such costs into account in designing an optimal tax schedule. As shown, the second term contains its own specific notion, so let us introduce the notation $\pi_t \equiv -\frac{t}{u_1} u_2 = V_2(t | t)$. Thus, we now rewrite $\int_{\underline{n}}^n -\frac{t}{u_1} u_2 dt = \int_{\underline{n}}^n \pi_t dt$.

Because the informational rent is neither tax revenue nor produced output, the resources needed for the rent impose an additional constraint in the government's problem. Because the government measures the marginal value of the resources in units of expected consumption goods — not working hours — it converts the resources of the rent into appropriate units by multiplying π_t by $[p\frac{1}{u_1^D} + (1-p)\frac{1}{u_1^{ND}}]$. Thus, the rent paid to a taxpayer of ability n for his private information regarding his earning ability is $\int_{\underline{n}}^n \pi_t [p\frac{1}{u_1^D} + (1-p)\frac{1}{u_1^{ND}}] dt$, which is positive (see (ii) in **Lemma 3**). The government aggregates this term over all the taxpayers using the population weights $h(n)$ to obtain $\int_{\underline{n}}^{\bar{n}} \int_{\underline{n}}^n \pi_t [p\frac{1}{u_1^D} + (1-p)\frac{1}{u_1^{ND}}] dt h(n) dn \geq 0$ in the form which is consistent with other resource constraints that the government faces.

In addition, for any given $n \in [\underline{n}, \bar{n}]$, once both before-tax labor earnings and after-tax (ex-ante) disposable income for consumption of individuals of ability n are determined, the income tax rate for them is automatically determined. Therefore, defining l_n and expected (after-tax) consumption $E[x_n] \equiv px_n^D + (1-p)x_n^{ND}$ is equivalent to defining the income tax rate for ability n individuals. In this light, by defining l_n and $E[x_n]$ that maximize the SWF, the government obtains the optimal income tax rate for ability n individuals. Then, by applying this result to individuals of each ability, the government eventually obtains the optimal income tax schedule T .

In sum, by incorporating the derived conditions for regaining the missing link between the tax rate and the targeted taxpayer into (5), we can rewrite the government's problem as follows: For $\forall n \in [\underline{n}, \bar{n}]$ given,

$$\begin{aligned} & \max_{p \ \& \ \{l_n, E[x_n]\}} \int_{\underline{n}}^{\bar{n}} \frac{dG}{dv_n} (v_{\underline{n}} + \int_{\underline{n}}^n \pi_t dt) h(n) dn \text{ subject to} \\ & (i) \int_{\underline{n}}^{\bar{n}} \{px_n^D + (1-p)x_n^{ND}\} h(n) dn \leq \int_{\underline{n}}^{\bar{n}} w n l_n h(n) dn \end{aligned}$$

$$\text{and (ii) } \int_{\underline{n}}^{\bar{n}} \int_{\underline{n}}^n \pi_t [p \frac{1}{u_1^D} + (1-p) \frac{1}{u_1^{ND}}] dt h(n) dn \geq 0 \quad (6)$$

This is then stated in the Lagrangian form as

$$\mathcal{L} = \{ \int_{\underline{n}}^{\bar{n}} \int_{\underline{n}}^n \{ \frac{d^2 G}{dv_t^2} [\frac{v_n}{n-\underline{n}} + \pi_t] - \lambda \pi_t [p \frac{1}{u_1^D} + (1-p) \frac{1}{u_1^{ND}}] \} h(t) dt dn + \int_{\underline{n}}^{\bar{n}} \lambda \{ w n l_n - [p x_n^D + (1-p) x_n^{ND}] \} h(n) dn \} \quad (7)$$

where λ is the marginal social value on resources for public fund.⁷

4 Optimal Income Tax, Tax Enforcement and Tax Evasion

This section begins with characterizing the optimal tax function, based on (7) and proceeds to define optimal p (rate of tax enforcement). Capitalizing upon these, we investigate how tax evasion is affected by increases in the tax rate and in the tax enforcement rate.

4.1 Optimal Income Tax and Optimal Rate of Tax Enforcement

Because (7) embeds the requirements that enable tax rates to be correctly applied to their target groups of taxpayers, the government uses (7) for obtaining the optimal tax rate for each $n \in [\underline{n}, \bar{n}]$ by finding l_n and $E[x_n]$ that maximize the SWF. Moreover, as the SWF is weakly concave (i.e., $\frac{d^2 G}{dv_n^2} \leq 0$), the FOCs of (7) are sufficient for defining such l_n and $E[x_n]$, which will eventually show characteristics of the optimal income tax schedule. To this end, combining the two FOCs ($\frac{d\mathcal{L}}{dl_n} = 0$ and $\frac{d\mathcal{L}}{dE[x_n]} = 0$)⁸ yields optimality condi-

⁷To be consistent, since the informational rent is forgone resource that could have been used for public fund, the Lagrange multiplier of $\int_{\underline{n}}^n \pi_t [p \frac{1}{u_1^D} + (1-p) \frac{1}{u_1^{ND}}] dt$ is $-\lambda$. At the same time, note that the double integral in $\int_{\underline{n}}^{\bar{n}} \{ \frac{dG}{dv_n} [v_n + \int_{\underline{n}}^n \pi_t dt] \} h(n) dn$ and the constraint (ii) in (6) are simplified following that

$$\begin{aligned} & \int_{\underline{n}}^{\bar{n}} \{ \frac{dG}{dv_n} [v_n + \int_{\underline{n}}^n \pi_t dt] - \lambda \int_{\underline{n}}^n \pi_t [p \frac{1}{u_1^D} + (1-p) \frac{1}{u_1^{ND}}] dt \} h(n) dn \\ &= \int_{\underline{n}}^{\bar{n}} \int_{\underline{n}}^n \{ \frac{d^2 G}{dv_t^2} (\frac{v_n}{n-\underline{n}} + \pi_t) - \lambda \pi_t [p \frac{1}{u_1^D} + (1-p) \frac{1}{u_1^{ND}}] \} dt h(n) dn \\ &= \int_{\underline{n}}^{\bar{n}} \int_{\underline{n}}^n \{ \frac{d^2 G}{dv_t^2} (\frac{v_n}{n-\underline{n}} + \pi_t) - \lambda \pi_t [p \frac{1}{u_1^D} + (1-p) \frac{1}{u_1^{ND}}] \} h(t) dt dn. \end{aligned}$$

⁸ $\frac{d\mathcal{L}}{dl_n} = \frac{d\pi_n}{dl_n} \int_{\underline{n}}^{\bar{n}} [\frac{d^2 G}{dv_t^2} - \lambda \{ p \frac{1}{u_1^D} + (1-p) \frac{1}{u_1^{ND}} \}] h(t) dt + \lambda [w n - \frac{d[p x_n^D + (1-p) x_n^{ND}]}{dl_n}] h(n) = 0$
and $\frac{d\mathcal{L}}{dE[x_n]} = \frac{dl_n}{dE[x_n]} \frac{d\pi_n}{dl_n} \int_{\underline{n}}^{\bar{n}} [\frac{d^2 G}{dv_t^2} - \lambda \{ p \frac{1}{u_1^D} + (1-p) \frac{1}{u_1^{ND}} \}] h(t) dt + \lambda [w n \frac{dl_n}{dE[x_n]} - 1] h(n) = 0$.
Moreover, $\frac{d[p x_n^D + (1-p) x_n^{ND}]}{dl_n} = p \frac{dx_n^D}{dl_n} + (1-p) \frac{dx_n^{ND}}{dl_n} = p w n (1 - T' - \bar{\theta} T' + \bar{\theta} r_n T') + (1 -$

tions of income tax rates which are stated as follows: For each $n \in [\underline{n}, \bar{n}]$,

$$\frac{d\pi_n}{dl_n} \int_n^{\bar{n}} \left[\lambda \left\{ p \frac{1}{u_1^D} + (1-p) \frac{1}{u_1^{ND}} \right\} - \frac{d^2G}{dv_t^2} \right] h(t) dt = wn\lambda [p\{1 + \bar{\theta}(1 - r_n)\} + (1-p)r_n] T' h(n). \quad (8)$$

Note that r_n , u_1^D and u_1^{ND} above are defined by (3) and (4). Tax liability is first assessed on the basis of reported income. However, actual payments to the government may differ from the assessed taxes if tax evasion is detected and penalized. Thus, when the fine for tax evasion paid to the government is perceived as part of the tax payment, $[p\{1 + \bar{\theta}(1 - r_n)\} + (1 - p)r_n]T'$ is regarded as an ex-ante marginal income tax rate.

To illustrate the economic principle underlying the above mathematical formula for the optimal income tax rate, let us consider the following reference case. Suppose, hypothetically, that the government has the clairvoyant power to see earning ability. Then, interpersonal lump sum transfers will be adopted instead of income taxes because the former incurs no efficiency loss, according to the Second Fundamental Welfare Theorem. Moreover, let the amount of transfers taken from an ability n individual mimic his tax payment under the income tax schedule given by (8). In this no-efficiency-loss case, the ability n individual determines his working hours according to $pu_1^D wn + (1 - p)u_1^{ND} wn = -u_2$. Comparing this with (4), the expected value of working ($-u_2$) differs by $wn[p\{1 + \bar{\theta}(1 - r_n)\} + (1 - p)r_n]T'$. Due to **Lemma 4** below, $w > 0$ and $n > 0$, which implies that the individual of ability n works *less* than he would in the hypothetical case proposed above. Therefore, $wn\lambda[p\{1 + \bar{\theta}(1 - r_n)\} + (1 - p)r_n]T'h(n)$ means an efficiency loss to the government as $wn[p\{1 + \bar{\theta}(1 - r_n)\} + (1 - p)r_n]T'$ is weighted by $h(n)$ (the proportion of the population of ability n individuals) and translated in terms of public funds whose marginal social value is λ .

Lemma 4. If T is an optimal income tax schedule, then for $\forall n \in [\underline{n}, \bar{n})$, $[p\{1 + \bar{\theta}(1 - r_n)\} + (1 - p)r_n]T' > 0$; and for $n = \bar{n}$, $T' = 0$.

Proof. First of all, the left-hand side of (8) is greater than zero for $\forall n \in [\underline{n}, \bar{n})$ and zero when $n = \bar{n}$. To see this, first, $\frac{d\pi_n}{dl_n} = \frac{d}{dl_n} \left(\frac{-l_n}{n} u_2 \right) = \frac{-1}{n} u_2 - \frac{l_n}{n} u_{22} > 0$ since u is decreasing in working hours ($u_2 < 0$) and concave ($u_{22} < 0$). Second, $\lambda \left\{ p \frac{1}{u_1^D} + (1-p) \frac{1}{u_1^{ND}} \right\} - \frac{d^2G}{dv_t^2} h(t) > 0$ due to (i) $u_1^D > 0$ and $u_1^{ND} > 0$, (ii) $\frac{d^2G}{dv_t^2} \leq 0$ (weakly concave SWF) and (iii) $h(t) > 0$ (full support).

$p)wn[1 - r_n T']$.

Thus, $\frac{d\pi_n}{dn} \int_n^{\bar{n}} [\lambda \{p \frac{1}{u_1^D} + (1-p) \frac{1}{u_1^{ND}}\} - \frac{d^2G}{dv_t^2}] h(t) dt > 0$ when $n < \bar{n}$. Moreover, it is equal to zero when $n = \bar{n}$ as $\int_{\bar{n}}^{\bar{n}} [\lambda \{p \frac{1}{u_1^D} + (1-p) \frac{1}{u_1^{ND}}\} - \frac{d^2G}{dv_t^2}] h(t) dt = 0$. Second of all, in the right-hand side of (8), $wn\lambda[p\{1 + \bar{\theta}(1 - r_n)\} + (1 - p)r_n] > 0$ for $\forall n \in [\underline{n}, \bar{n}]$ since $r_n \in (0, 1]$, $\bar{\theta} > 0$ and $p \in (0, 1)$ for $\forall n \in [\underline{n}, \bar{n}]$. Therefore, for $n = \bar{n}$, $T' = 0$ following the sign of the left-hand side. By the similar token, as all of w , n , and λ are greater than zero for $\forall n \in [\underline{n}, \bar{n}]$, $[p\{1 + \bar{\theta}(1 - r_n)\} + (1 - p)r_n]T' > 0$ as the left-hand side is strictly positive. ■

To understand what the efficiency loss (left-hand side of (8)) is for, let us run a counterfactual thought experiment. Suppose that, deviating from (8), the government suddenly drops the marginal income tax rate to zero only for ability n taxpayers ($n < \bar{n}$). Then, these individuals will now work more following $pu_1^D wn + (1 - p)u_1^{ND} wn + u_2 = 0$. This means that, only for this specific group of the ability n taxpayers, the efficiency loss falls to zero. However, in the end, this deviation causes much more harm than good because the hypothetical reform entices all taxpayers of abilities greater than n to reduce their working hours to disguise themselves as a taxpayer of ability n for enjoying the zero marginal income tax rate. To prevent this greater loss, the government can compensate all of these higher-ability individuals with their potential surpluses from pretending (informational rents). In particular, for each ability t ($t > n$), the net value of this cost to the government is appraised at $\lambda \{p \frac{1}{u_1^D} + (1 - p) \frac{1}{u_1^{ND}}\} - \frac{d^2G}{dv_t^2}$. The first term refers to the marginal social value of resource forgone to pay the informational rent which otherwise could be used for consumption. The second term is the marginal social value of an increase in the utility of ability t individuals from receiving the rent. Aggregating over all individuals of higher abilities weighted by their population proportions $h(t)$, the cost to the government is $\frac{d\pi_n}{dn} \int_n^{\bar{n}} [\lambda \{p \frac{1}{u_1^D} + (1 - p) \frac{1}{u_1^{ND}}\} - \frac{d^2G}{dv_t^2}] h(t) dt$ obtained by multiplying the aggregated term by $\frac{d\pi_n}{dn}$ to express the cost in the same unit as the left-hand side of (8).

In sum, the optimal income tax rule allows efficiency loss *only* for preventing the cascading decreases in the labor supply of all taxpayers with higher earning abilities, so that more effective labor is supplied by more able individuals (**Lemma 1**). In this way, the government can minimize efficiency losses in maximizing the SWF, as the government can secure resources for public funds which otherwise could not have been produced if taxpayers of

higher abilities would provide less effective labor.

Next, let us move on to finding the optimal rate of tax enforcement. Once the optimal income tax schedule T meets the conditions for regaining the missing link between the tax rate and the target taxpayer, it is redundant to impose such conditions in seeking an optimal rate of tax enforcement p . Moreover, the same p is applied to all taxpayers of different abilities and governs revenue collection of the entire society. In this light, we can rewrite (5) to effectively obtain the optimality condition of p in a more straightforward manner. That is, given the optimal tax schedule T from (8), we replace the market clearing condition in (5) with the government's budget condition for characterizing an optimal rate of tax enforcement p .

$$\max_p \int_{\underline{n}}^{\bar{n}} \frac{dG}{dv_n} v_n h(n) dn \text{ s.t. } R + \frac{1}{\delta} c(p) \leq \int_{\underline{n}}^{\bar{n}} [p\{T(wnl_n) + \bar{\theta}[T(wnl_n) - T(r_n wnl_n)]\} + (1-p)T(r_n wnl_n)] h(n) dn. \quad (9)$$

where R is the revenue requirement for public expenditures. From the FOC of the Lagrangian expression (9), the optimality condition of p is:

$$\int_{\underline{n}}^{\bar{n}} \frac{dG}{dv_n} (u_n^{ND} - u_n^D) h(n) dn = \lambda \left\{ \int_{\underline{n}}^{\bar{n}} (1 + \bar{\theta}) [T(wnl_n) - T(r_n wnl_n)] h(n) dn - \frac{1}{\delta} c'(p) \right\}. \quad (10)$$

Notice that r_n is defined by (3) and identical with the r_n in (8). To understand the formula for the optimal rate of tax enforcement, note that the right-hand side of (10) represents the marginal *net* social benefit from raising p . The first term $\int_{\underline{n}}^{\bar{n}} (1 + \bar{\theta}) [T(wnl_n) - T(r_n wnl_n)] h(n) dn$ represents the increase in expected tax revenue resulting from a small increase in p . We then subtract the second term $\frac{1}{\delta} c'(p)$ that is the increment in enforcement costs needed for the increment in p . Lastly, this net resource for public funds (gain) is appraised at λ per unit. On the other hand, the left-hand side of (10) represents the marginal social loss from the small increase in p as $-\frac{dv_n}{dp} = u_n^{ND} - u_n^D$ is the ensuing increase in the gap between the payoff under being detected (u_n^D) and the payoff under not being detected (u_n^{ND}), which reduces the utility of risk-averse taxpayers of ability n . This is a loss to the government who is a benevolent social planner. Moreover, this loss is aggregated over all abilities weighted by their population proportions $h(n)$ and social values $\frac{dG}{dv_n}$ to obtain the left-hand side of (10). In sum, the optimal p is set to equalize the marginal social value of the *net* increase in expected revenue from the increment in p with the marginal social loss from the increased risk of tax

evasion.⁹

4.2 Tax Evasion Response to Tax Rate and Tax Enforcement Rate

Thus far, we have obtained an optimal income tax schedule that explicitly incorporates tax evasion concerns and an optimal enforcement rate. In contrast to previous studies, the optimal design of nonlinear income taxation presented above allows for responses of both labor supply and tax evasion. The analysis may thus be applicable to various questions related to income tax evasion.

First, we can apply above results to examine how tax evasion is affected by tax rate and by tax enforcement rate. As mentioned at the outset, previous studies discussed how tax evasion is affected by these basic tax policy variables (Allingham and Sandmo 1972; Pencavel 1979; Baldry 1979; Sandmo 1981; Horowitz and Horowitz 2000) but they only showed that the relationships of tax evasion to tax rate and to tax enforcement rate are ambiguous. Neither a clear conclusion nor any clues that may indicate when these relationships are negative (or positive) is provided. Because the optimization in the present study offers a more general environment than previous studies, we are revisiting the pending questions regarding the effect on tax evasion of tax rate and tax enforcement rate, seeking a meaningful clarification of these ambiguities.

We first address the issue of how income tax evasion responds to the marginal income tax rate.

Theorem 1. For any given $n \in [\underline{n}, \bar{n})$, $\frac{\partial e_n}{\partial T'} > 0$. That is, income tax evasion is positively affected by the marginal income tax rate.

Proof. [step 0] As a stepping stone to prove the above statement, we first need to show that $T' > 0$ for $\forall n \in [\underline{n}, \bar{n})$. First, $p\{1 + \bar{\theta}(1 - r_n)\} + (1 - p)r_n > 0$

⁹Notice that this result is more general than previous studies on the optimal tax enforcement (Reinganum and Wilde 1985; Mookherjee and Png 1989; Sánchez and Sobel 1993). First, the current analysis characterizes p that maximizes social welfare whereas they found p that maximizes net tax revenue. More importantly, the present study derives both optimal nonlinear income tax schedule and optimal tax enforcement at the same time while previous studies separated these by assuming a single income tax rate exogenously given.

since $r_n \in (0, 1]$, $\bar{\theta} > 0$ and $p \in (0, 1)$ for $\forall n \in [\underline{n}, \bar{n}]$. From **Lemma 4**, this means that $T' > 0$ for any $n \in [\underline{n}, \bar{n}]$.

[step 1] Also, we need to explicitly state the condition for r_n to be defined by (3) for $\forall n \in [\underline{n}, \bar{n}]$. That is, individuals have interior solution of r_n . In the corner solution of no tax evasion ($r_n = 1$), there is no point to examine how tax evasion is affected by an increase in tax rate. Such a condition is $\frac{dE[u]}{dr}\Big|_{r=1} = \{p\bar{\theta}u_1^D - (1-p)u_1^{ND}\}T' < 0$. Due to the [step 0], this implies that $p\bar{\theta}u_1^D - (1-p)u_1^{ND} < 0$ at $r = 1$. Because $u_1^D = u_1^{ND}$ at $r = 1$, this means that $p\bar{\theta} - (1-p) < 0$.

[step 2] Based on the Implicit Function Theorem, we can obtain $\frac{\partial r_n}{\partial T'} = \frac{-wn\lambda[p\{1+\bar{\theta}(1-r_n)\}+(1-p)r_n]h(n)}{wn\lambda[-p\bar{\theta}+(1-p)]T'h(n)} = \frac{-[p\{1+\bar{\theta}(1-r_n)\}+(1-p)r_n]}{[-p\bar{\theta}+(1-p)]T'}$ using (8) for $\forall n \in [\underline{n}, \bar{n}]$. As shown in the [step 0], $p\{1 + \bar{\theta}(1 - r_n)\} + (1 - p)r_n > 0$. In addition to this, from [step 0] and [step 1], $[-p\bar{\theta} + (1 - p)]T' > 0$. Therefore, $\frac{\partial r_n}{\partial T'} < 0$ for $\forall n \in [\underline{n}, \bar{n}]$. Since $e_n = 1 - r_n$, $\frac{\partial e_n}{\partial r_n} < 0$. Taken together, this finally implies that $\frac{\partial e_n}{\partial T'} > 0$ for $\forall n \in [\underline{n}, \bar{n}]$ since $\frac{\partial e_n}{\partial T'} = \frac{\partial e_n}{\partial r_n} \frac{\partial r_n}{\partial T'}$. ■

As appears in (3) the decision rule for tax evasion, tax rate does not directly affect degree of tax evasion, while it directly affects marginal value of working. So, the intuition underlying **Theorem 1** can be effectively demonstrated under the view that the income for taking tax evasion gamble is purchased (procured) with taxpayer's labor. In this perspective, given p and $\bar{\theta}$, an increment in the tax rate lowers the value of working to taxpayers ($-u_2$), according to (4) and **Lemma 4**. As a result, the price to be paid for investing labor earnings in tax evasion gamble (with a given amount of working) goes down, which entails more gamble of tax evasion. In other words, an increase in the tax rates causes a given amount of working to be less worthy as it yields less ex-ante consumption, due to the ensuing increase in tax liabilities. Responding to this, taxpayers under-report a greater portion of their income to maintain their expected consumption, following (3), because an increase in e_n (a decrease in r_n) entails an increase in expected consumption unless $T' = 0$.

In addition, regarding top earners (individuals of the highest earning ability \bar{n}), $\frac{\partial e_n}{\partial T'}$ cannot be defined for a technical reason. In particular, the tax rate should be able to continuously vary in *both* directions (above and below a given tax rate) in its neighborhood, which is a prerequisite for taking derivatives. By contrast to all other taxpayers (all the individuals of abilities

n below \bar{n}), this is not feasible for the top earners, as variation below a given tax rate (zero) is not allowed due to **Lemma 4**. Putting aside this technical issue, the logic remains the same. That is, taxpayers of the highest earning ability \bar{n} would also want to maintain their expected consumption through an increase in $e_{\bar{n}}$ in response to a change from $T' = 0$ to $T' > 0$.

Importantly, **Theorem 1** resolves the theoretical ambiguity regarding the relationship between tax evasion and the tax rate for all of the population except a handful of top earners. As a matter of fact, on the empirical front of this subject, a number of studies have already found a positive correlation between the marginal tax rate and tax evasion, using various data (e.g., Clotfelter 1983; Baldry 1987; Dubin et al. 1990; Alm et al. 1992).

Let us now re-examine the effect on tax evasion behavior of an increase in the probability that tax evasion is detected and penalized, which has also been left ambiguous by prior literature. As noted above, previous studies (Baldry 1979; Pencavel 1979; Sandmo 1981; Horowitz and Horowitz 2000) have shown that an increase in p can either deter or promote tax evasion. The latter is a particularly perplexing result, as it implies that enhancing tax enforcement promotes tax evasion.

Theorem 2. For any given $n \in [\underline{n}, \bar{n})$, $\frac{\partial e_n}{\partial p} < 0$. That is, income tax evasion is negatively affected by the tax enforcement rate.

Proof. [step 0] First, we need to show that $T' > 0$ for $\forall n \in [\underline{n}, \bar{n})$. Because $r_n \in (0, 1]$, $\bar{\theta} > 0$ and $p \in (0, 1)$ results in $p\{1 + \bar{\theta}(1 - r_n)\} + (1 - p)r_n > 0$, $T' > 0$ for any $n \in [\underline{n}, \bar{n})$ from **Lemma 4**.

[step 1] Based on the Implicit Function Theorem, we can obtain $\frac{\partial r_n}{\partial p} = \frac{-(\bar{\theta}u_1^D + u_1^{ND})}{[p\bar{\theta}^2 u_{11}^D T' wnl_n + (1-p)u_{11}^{ND} T' wnl_n]}$ using (3) for $\forall n \in [\underline{n}, \bar{n})$. Firstly, $\bar{\theta}u_1^D + u_1^{ND} > 0$ because u is increasing in consumption and $\bar{\theta} > 0$. Secondly, because, for $\forall n \in [\underline{n}, \bar{n})$, $\bar{\theta}^2 > 0$, $p \in (0, 1)$, $wnl_n > 0$, $T' > 0$ due to the [step0], and u is concave, $p\bar{\theta}^2 u_{11}^D T' wnl_n + (1-p)u_{11}^{ND} T' wnl_n < 0$ for $\forall n \in [\underline{n}, \bar{n})$. Therefore, $\frac{\partial r_n}{\partial p} > 0$ for $\forall n \in [\underline{n}, \bar{n})$. As $e_n = 1 - r_n$, $\frac{\partial e_n}{\partial r_n} < 0$. All in all, this implies that $\frac{\partial e_n}{\partial p} = \frac{\partial e_n}{\partial r_n} \frac{\partial r_n}{\partial p} < 0$ for $\forall n \in [\underline{n}, \bar{n})$. ■

The rationale for **Theorem 2** is simple. As the likelihood of a bad outcome from investing in the tax evasion gamble increases, given the income tax schedule T and $\bar{\theta}$, the expected return from tax evasion diminishes, making tax evasion less attractive. Thus, responding to an increase in the tax

enforcement rate, taxpayers reduce tax evasion by increasing r_n to secure more x^D (and less x^{ND}). In contrast, an increase in $r_{\bar{n}}$ (a decrease in $e_{\bar{n}}$) of the top earners (individuals of the highest earning ability \bar{n}) does not create the intended effect in their expected consumption because their tax liability does not change due to $T' = 0$. This implies not only that $\frac{\partial e_{\bar{n}}}{\partial p}$ is not defined but that an increase in p would not make a difference in the amount of taxes collected from the ability \bar{n} taxpayers. Therefore, enhanced tax enforcement will clearly decrease overall revenue leakage. Put another way, with **Theorem 2**, we now can invalidate the perplexing case where the government's improvement of tax enforcement causes tax evasion to increase.

As you may notice, the nuisance that the singleton \bar{n} brings to the above theoretical clarifications of **Theorem 1** and **Theorem 2** stems from the fact that the optimal marginal income tax rate on the richest individuals is zero (**Lemma 4**). We discuss this peculiar point in the following subsection.

4.3 Extension: Strictly Positive Marginal Tax Rate on the Richest

In line with the intuition underlying (8), the reason why $T' = 0$ is optimal for taxpayers of the highest earning ability (\bar{n}) is that no other individuals have an incentive to mimic the individuals of the highest earning ability, as all the others are of lower ability. Thus, there is no efficiency loss caused by the ability \bar{n} taxpayers. Moreover, with the IC constraints met, the highest ability taxpayers earn the highest income ($wL_{\bar{n}}$) because their effective labor supply ($L_{\bar{n}}$) is the largest due to (i) in **Lemma 1**. As a matter of fact, the zero marginal income tax rate on top earners (individuals of ability \bar{n}) is one of the few famous and established results in the optimal taxation literature since Mirrlees (1971).¹⁰ This result, however, has been criticized as "a mere theoretical curiosity" (Mankiw et al. 2009), having little practical relevance. For example, all of the OECD countries have statutory income tax schedules that levy strictly positive marginal income tax rates on their richest citizens.

In contrast to the previous studies of optimal income taxation, the present

¹⁰Some tried to get around this with *asymptotic* marginal tax rate by assuming infinity of ability n (thus infinity of the income) (e.g, Diamond 1998; Jacquet et al. 2013) First, such non-existence of highest income lacks relevance to real world. Second, instead of directly obtaining optimal income tax rate, only approaching to it (i.e., not getting at it) would have neither clear policy implications nor solid theoretical foundation for stable equilibrium tax rate.

study allows tax enforcement to be imperfect and costly. Thus, in addition to the efficiency loss of decreases in labor supply, collecting taxes itself can be a source of efficiency loss, which eventually affects social welfare. Incorporating this additional factor of tax enforcement into the analysis of optimal income taxation, we can briefly extend the present analysis to provide theoretical grounds for a strictly positive marginal income tax rate on the top earners by examining a case where the zero marginal income tax rate on the richest in the society causes a decrease in δ .

Before announcing its tax policies (step (i) in the time line displayed in Section 2), the government obtains optimal p and optimal tax function T which are defined by (8) and (10). From this, the government can obtain the level of optimum SWF as well. Let us consider the following case. The government's announcement that the marginal income tax rate on the richest is zero — in stark contrast to the strictly positive tax rates on poorer taxpayers — triggers a decrement in δ . In reality, the tax treatment favorable to the richest, who are typically under the spotlight, often draws more public controversy than the tax treatment favorable to other income groups. Thus, applying the lowest marginal income tax rate to the richest individuals *per se* works as a critical tipping point, inducing nearly all taxpayers to withdraw support and cooperation from the government because they believe that the government lets the most privileged members of society do less part than the parts they do and unduly dumps the tax burden on them. Such a public reaction to the zero marginal income tax rate on the top earners would negatively affect tax compliance and make tax collection more difficult. Consequently, the same amount of money spent on administrating and operating the tax system would yield weaker tax enforcement (and less tax revenue collected), which translates into a marginal decrease in δ . To avoid this fallout, the government would amend the tax schedule of (8) as follows.

Theorem 3. If the zero marginal tax rate on taxpayers of the highest earning ability \bar{n} causes a decrement in δ , then an optimal marginal tax rate for the taxpayers of the highest earning ability \bar{n} (denoted by $T'_{\bar{n}}$) is strictly positive and satisfies $\frac{1}{\delta^2}c(p) \geq w\bar{n}[p\{1 + \bar{\theta}(1 - r_{\bar{n}})\} + (1 - p)r_{\bar{n}}]T'_{\bar{n}}h(\bar{n}) > 0$.

Proof. Suppose that $T' = 0$ on the richest (individuals of the highest earning ability \bar{n}) triggers a decrement in δ . This causes a marginal social loss as much as $-\lambda\frac{1}{\delta^2}c(p)$ which is obtained by taking a derivative with respect to δ of the Lagrangian expression of (9) at its optimal level, based on the

Envelope Theorem. Thus, the social gain of preventing δ from decreasing (keeping the previous level) so that the government can maintain the optimal level of social welfare is $\lambda \frac{1}{\delta^2} c(p)$. Because $\lambda \frac{1}{\delta^2} c(p) > 0$, the government is better off with the amendment of levying a strictly positive marginal income tax rate on the richest as long as a social loss from this alteration is smaller than a social gain from deterring the decrement in δ .

When a new optimal marginal income tax rate on the richest, denoted by $T'_n > 0$, is strictly greater than zero, the government can get the benefit as much as $\lambda \frac{1}{\delta^2} c(p)$ by deterring the decrement in δ . On the other hand, this amendment incurs the marginal social loss as individuals of the highest earning ability \bar{n} supply less labor than before. To see the loss, before the amendment, the richest taxpayers decide hours of working based on $pu_1^D w\bar{n} + (1-p)u_1^{ND} w\bar{n} = -u_2$ because $T' = 0$ for $n = \bar{n}$ due to **Lemma 4**. By contrast, after the amendment, they make their labor supply decision on the basis of $pu_1^D w\bar{n}\{1 - (1 + \bar{\theta}[1 - r_{\bar{n}}])T'_n\} + (1-p)u_1^{ND} w\bar{n}(1 - r_{\bar{n}}T'_n) = -u_2$. Thus, the increase in the marginal income tax rate on the individuals of the highest earning ability \bar{n} reduces their working hours as the marginal disutility of working is decreased by $w\bar{n}\lambda[p\{1 + \bar{\theta}(1 - r_{\bar{n}})\} + (1-p)r_{\bar{n}}]T'_n h(\bar{n})$. Taken together, this implies that a new optimal T'_n should satisfy $\frac{1}{\delta^2} c(p) \geq w\bar{n}[p\{1 + \bar{\theta}(1 - r_{\bar{n}})\} + (1-p)r_{\bar{n}}]T'_n h(\bar{n})$. Otherwise, the amendment of raising the marginal tax rate on the richest to keep δ as the previous level at the social optimum is self-defeating by causing more harm than good.

Moreover, for $\forall n \in [\underline{n}, \bar{n}]$, $r_n \in (0, 1]$, $\bar{\theta} > 0$, $p \in (0, 1)$, and that all of w , n , and λ are greater than zero; in addition, $h(\bar{n}) > 0$ because we have full support. This means that $w\bar{n}[p\{1 + \bar{\theta}(1 - r_{\bar{n}})\} + (1-p)r_{\bar{n}}]T'_n h(\bar{n}) > 0$. ■

In short, the amendment of (8) with $T'_n > 0$ induces the richest individuals to work slightly less; however, this loss is smaller than the gain of saving tax enforcement costs which would otherwise be used for public expenditure. Moreover, *without* relying on asymptotic methods, **Theorem 3** proves that the optimal marginal income tax rate on the top earners can be strictly greater than zero. The condition in **Theorem 3** that the zero marginal income tax rate on the richest in itself negatively affects citizens' cooperation of tax compliance is relevant to the actual practices of taxation in democratic countries.

5 Concluding Remarks

In sum, this paper presents an optimal nonlinear income tax that incorporates individuals' decisions regarding tax evasion and labor supply as well as an optimal rate of tax enforcement that maximizes social welfare when improving rate of tax enforcement is costly. Capitalizing upon this, theoretical ambiguities regarding the effects on tax evasion of the tax rate and the tax enforcement rate (ambiguities that have lingered in the literature for decades) have been resolved. We find that an increase in the tax rate leads to an increase in tax evasion. In addition, we prove that enhancing tax enforcement by increasing the probability of detecting tax evasion does not increase tax evasion.

Finally, this paper shows that the non-asymptotic marginal income tax rate optimal for the richest individuals can be strictly positive, instead of zero, if taxpayers' support for the government (which affects tax compliance) responds negatively when the lowest marginal income tax rate is levied on the richest.

References

- [1] Allingham, Michael G. and Agnar Sandmo: Income tax evasion: a theoretical analysis. *Journal of Public Economics* **1**(3), 323-338 (1972)
- [2] Alm, James, Betty R. Jackson, and Michael McKee: Estimating the determinants of taxpayer compliance with experimental data. *National Tax Journal* **45**(1), 107-114 (1992)
- [3] Baldry, Jonathan C.: Tax evasion and labour supply. *Economics Letters* **3**(1), 53-56 (1979)
- [4] Baldry, Jonathan C.: Income tax evasion and the tax schedule: some experimental results. *Public Finance* **42**(3), 357-83 (1987)
- [5] Clotfelter Charles T.: Tax evasion and tax rates: an analysis of individual returns. *Review of Economics and Statistics* **65**(3), 363-373 (1983)
- [6] Diamond, Peter A.: Optimal income taxation: an example with a u-shaped pattern of optimal marginal tax rates. *American Economic Review* **88**(1), 83-95 (1998)
- [7] Dubin, J.A., M.J. Graetz and L.L. Wilde: The effect of audit rates on the federal individual income tax, 1977-1986. *National Tax Journal* **43**(4), 395-409 (1990)
- [8] Guyton J, O'Hare J, Stavrianos M and Toder E.: Estimating the compliance cost of the U.S. individual income tax. *National Tax Journal* **56**(2), 673-688 (2003)
- [9] Jacquet, Laurence, Etienne Lehmann and Bruno Van der Linden: Optimal redistributive taxation with both extensive and intensive responses. *Journal of Economic Theory* **148**(5), 1770-1805 (2013)
- [10] Horowitz, Ira and Ann R. Horowitz: Tax audit uncertainty and the work-versus-leisure decision. *Public Finance Review* **28**(6), 491-510 (2000)
- [11] IMF Fiscal Affairs Department: Revenue Mobilization in Developing Countries, International Monetary Fund Policy Paper (2011)

- [12] Mirrlees, James A.: An exploration in the theory of optimum income taxation. *Review of Economic Studies* **38**(2), 175-208 (1971)
- [13] Mankiw, N. Gregory, Matthew Weinzierl and Danny Yagan: Optimal taxation in theory and practice. *Journal of Economic Perspectives* **23**(4), 147-74 (2009)
- [14] Mookherjee, Dilip and Ivan Png: Optimal auditing, insurance, and redistribution. *Quarterly Journal of Economics* **104**(2), 399-415 (1989)
- [15] Pencavel, John H.: A note on income tax evasion, labor supply, and nonlinear tax schedules. *Journal of Public Economics* **12**(1), 115-124 (1979)
- [16] Reinganum, Jennifer F. and Louis L. Wilde: Income tax compliance in a principal-agent framework. *Journal of Public Economics* **26**(1), 1-18 (1985)
- [17] Sánchez, Isabel and Joel Sobel: Hierarchical design and enforcement of income tax policies. *Journal of Public Economics* **50**(3), 345-369 (1993)
- [18] Sandmo, Agnar: Income tax evasion, labour supply, and the equity-efficiency tradeoff. *Journal of Public Economics* **16**(3), 265-288 (1981)
- [19] Slemrod, Joel: Cheating ourselves: the economics of tax evasion. *Journal of Economic Perspectives* **21**(1), 25-48 (2007)