# Optimal Self-Employment Income Tax Enforcement[*]

Saki Bigio

Department of Economics

New York University

Eduardo Zilberman

Department of Economics

New York University

February 2010

## Abstract

Most models of optimal income tax enforcement assume that income is either random or solely remunerates labor, neglecting that auditing strategies may depend on observable inputs. This paper outlines a model to optimally monitor self-employed entrepreneurs when, in addition to reported profits, the tax collection agency also observes the number of workers employed (or any other input variable) at each firm. We show that, by conditioning the monitoring strategy only on labor input, it is optimal for the IRS to audit firms in a way that generates some empirical regularities, like the *missing middle*. We also show that the optimal direct mechanism can be implemented by an indirect monitoring strategy that is consistent with actual IRS practices. In particular, the IRS calculates inputted income as function of labor. Whenever an entrepreneur reports profits that are lower than inputted income, she is randomly monitored. Finally, we formalize a model of optimal presumption taxation, in which inputted income is the tax base, to compare revenue collection across tax systems.

Keywords: optimal auditing, tax evasion, informal sector, missing middle, entrepreneurship, presumptive taxation.

JEL Classification: D21, H26.

1

# 1  Introduction

The optimal income tax enforcement literature[1] mostly focuses on individual taxpayers whose income is either exogenous or solely remunerates labor/effort supply decisions.[2] However, Slemrod [2007] reports that in the U.S., only 1 percent of wages and salaries are underreported to the tax collection agency (henceforth the IRS), whereas 43 percent of self-employment business income is not.[3] This evidence suggests that the degree by which the IRS can enforce taxes varies across occupations. In particular, taxes on wages and salaries are more easily enforced than on self-employment business income.[4]

In practice, wage taxes are virtually always enforced. Thus, the number of workers at each firm seems to be costless information available to the IRS. This paper shifts the focus from individuals to firms, and asks the following: given that employment per each firm is observable, what is the optimal scheme to monitor heterogeneous entrepreneurs?[5]

This paper makes three contributions. First, we show that, when the monitoring strategy is conditioned only on labor input, it is optimal for the IRS to audit firms in a way that generates some empirical regularities. In particular, the model generates a *missing middle*, that is the fact that medium-scale firms are scarce. Second, when auditing probabilities depends both on labor input and reported income, we show that optimal auditing theory is consistent with actual IRS practices. Finally, by allowing taxpayers to self-report their income subject to audits and penalties, we show that the IRS collects more revenue than by imposing a tax base presumed from labor input. On the technical side, we solve a mechanism design problem that goes beyond quasi-linearity and that features a type-dependent individual rationality constraint.

In Section 2, we introduce a general model, in which heterogenous entrepreneurs (or self-employed business owners) can underreport their individual income, that is, the profits generated by their own firms. The source of heterogeneity is an unobserved random managerial ability that enhances productivity in a plant exhibiting decreasing returns to scale.

---

[1] The seminal paper is Reinganum and Wilde [1985], which is inspired by the costly state verification model of Townsend [1979]. Notable contributions are Border and Sobel [1987], Mookherjee and Png [1989], Sánchez and Sobel [1993], Cremer and Gahvari [1996], Chander and Wilde [1998], Bassetto and Phelan [2008]. The first theoretical work on tax noncompliance is Allingham and Sandmo [1972]. Recent surveys are Andreoni et al. [1998], Slemrod and Yitzhaki [2002], and Sandmo [2005].

[2] An exception is Parker [2010], who introduces entrepreneurship and dynamic incentives in a two-period model.

[3] These figures account for 10 and 109 billion dollars, respectively.

[4] Kleven et al. [2010] define tax evasion rate as the share of reported income that is underreported. Using data from Denmark, they estimate it as being 8.1% for self-employment income, 0.4% for earnings, 0.3% for third-party reported income, and 37% for self-reported income.

[5] In our model, firms and entrepreneurs make up a single unit. Henceforth, we use both terms interchangeably.

Entrepreneurs face an audit probability that depends on labor input and reported income, which are assumed to be costlessly observed by the IRS. If monitored, the firm must pay a linear penalty on the underreported amount.

The model is a two-stage game. In the first stage, given the managerial ability distribution, a budget-constrained IRS commits to a monitoring strategy that depends on employment and reported income. In the second stage, firms take into account this monitoring strategy and choose labor inputs and reported income. To solve this model, we invoke the revelation principle, and cast it into a direct mechanism design problem.

A novelty of this paper is that it allows the monitoring strategy to depend on an observable input choice, interpreted as labor. In the absence of strategic interactions, entrepreneurs with higher ability would hire more labor. However, employment also acts as a signal of the ability, thus entrepreneurs may distort their labor input strategically in order to signal a different ability.

It turns out that the most general version of the problem is hard to solve. More precisely, the problem's formulation implies that, on top of the multidimensional aspect, the agent's objective is not quasi-linear and the individual rationality constraint is type dependent. Nonetheless, we characterize the solution for two subproblems, one in which the IRS conditions the monitoring strategy only on reported income, and another in which it only conditions on employment.

The first problem has been studied previously (e.g., Sánchez and Sobel [1993]). In this context, the optimal monitoring strategy is to randomly audit firms which report below some threshold level of profits. In equilibrium, only low-ability entrepreneurs report honestly. This result is quickly reviewed in Section 3.1. In this mechanism, larger firms are the set of evaders, and every firm produces efficiently.

In Section 3.2, we study the opposite case in which the auditing probability depends solely on labor input. In contrast, the optimal strategy is to randomly monitor firms that employ labor above some threshold, while the rest are not monitored at all. Intuitively, labor input is a signal about the true managerial ability. Since profits increase with ability, by monitoring larger firms, the IRS collects taxes from the most profitable ones. As a result, smaller firms are the set of evaders. Moreover, in order to avoid detection, some of the firms distort labor input.

Hence, in contrast with the optimal tax enforcement literature, this mechanism is consistent with three empirical regularities: (1) a negative relationship between firm size, measured by employment, and amount evaded (Dabla-Norris et al. [2008]); (2) some of the evaders, if interpreted as "informals," producing inefficiently (la Porta and Shleifer [2008]); and (3) the existence of a missing middle (Tybout [2000]).

Not surprisingly, some papers in the development literature explain these facts in an adapted version of the Lucas [1978] model of entrepreneurial choice that allows an informal sector.[6] Our model is closely related to this literature, except that auditing policies are optimal here which, in principle, could revert these theoretical results. Consequently, it is optimal for the IRS to audit firms in a way that corroborates these facts.

In Section 4, we provide a partial characterization of the general solution that explains auditing practices in many countries. In particular, we restrict the monitoring strategy to be bang-bang, as the solutions to the previous two particular cases suggest. We show that any solution to the direct mechanism can be implemented by an indirect monitoring strategy as follows: The IRS calculates inputted income as a function of labor. Whenever an entrepreneur reports profits that are higher than inputted income, she is not monitored. If reported income is lower than inputted income, the IRS randomly monitors her. This auditing scheme is consistent with actual IRS practices, as we illustrate in Section 4.1, with anecdotal evidence.

The use of presumptions of income to tax or audit (as in this result) is the central principle of presumptive taxation.[7] Although it has been widely used in many countries, there is little theory formalizing optimal presumptive methods. This paper is a step towards this direction.[8]

A natural question arises in this framework. What if inputted income, as a function of labor, is designed to tax instead of audit entrepreneurs? In Section 5, we formalize and solve a model of presumptive taxation, in which the inputted income is the tax base. The optimal solution is characterized for tax systems in which presumptives of income are, and are not, rebuttable.[9] In particular, whenever presumptives of income are irrebuttable, the problem fits a standard mechanism design problem. Hence, labor is not distorted only at the top-type, and the informational rent is fully appropriated at the bottom-type. In contrast, in the rebuttable tax system, the reservation value is type-dependent, and standard tricks in the literature are not readily applicable. As before, employment is distorted everywhere

---

[6]Rauch [1991] is the seminal contribution. See also Fortin et al. [1997], Choi and Tum [2005], Amaral and Quintin [2006], Antunes and Cavalcanti [2007], Dabla-Norris et al. [2008], de Paula and Scheinkman [2009].

[7]There is a large practitioner literature on presumptive taxation (e.g., Tanzi and de Jantscher [1989] and Thuronyi [2004]).

[8]Scotchmer [1987] and Macho-Stadler and Pérez-Castrillo [2002] also consider a framework in which the IRS observes a signal of taxpayers' true income. In contrast, income is exogenous in both papers. In Scotchmer [1987], taxpayers are grouped into classes, and auditing probabilities depend on reported income and the class to which the individual belongs. Macho-Stadler and Pérez-Castrillo [2002] assume taxpayers do not know the realization of the signal.

[9]Under a rebuttable presumptive tax system, a taxpayer can claim and prove that her presumptive income is higher than her actual income. In contrast, under an irrebuttable presumptive tax system, taxpayers can not claim any revision. See Tanzi and de Jantscher [1989] for more details.

except at the top-type. However, the IRS potentially appropriates the full informational rent not only at the bottom-type, but also among a continuum of low-types.

The mathematical structure of the problem allows us to compare revenue collection across tax systems, even under the unsolved general case, in which the monitoring strategy depends both on labor and reported income. In Section 6, we show that the revenue collected in the rebuttable presumptive tax system is a lower bound on the revenue collected in the general case. To derive this result, we assume that there is no cost to claim and prove that someone's income was presumed wrongly. If this assumption is violated, presumptive tax methods might generate more revenue. This comparison suggests some explanations on why modern tax systems are the rule, but presumptive taxation is still observed in many countries (see Bird and Wallace [2004] for a list).

## 2   General model

We consider a two-stage game in which entrepreneurs remit taxes to the IRS. Taxes may be potentially evaded. In the first stage, the IRS commits to a monitoring scheme that depends on observable labor input and reported income. In the second stage, firms take into account the monitoring scheme, and choose labor input and reported income.

There is a continuum of firms of measure one. Each firm is owned and managed by a single entrepreneur, who experiences a random managerial ability $z$, which is her privately observed type. There is a single good produced with a single input, labor $n$. The production technology is $zf(n)$ common to all firms. Wages are the numeraire, and $p$ is the price of the good. Therefore, pre-tax profits are given by $\pi(n, z) = pzf(n) - n$.

**Assumption 1.** *The function $f : N \to \mathbb{R}_+$ is strictly increasing, strictly concave, and twice continuously differentiable, with $f(0) = 0$ and $N \subseteq \mathbb{R}_+$.*

Assumption 1 implies that the efficient labor employment is $n^*(z) = f'^{-1}\left(\frac{1}{pz}\right)$.

**Assumption 2.** *$z$ is a random variable, independently and identically distributed according to $G : Z \to [0, 1]$, with density $g(z) = G'(z) > 0$ for all $z \in Z$, where $Z \equiv [\underline{z}, \overline{z}]$, and $\underline{z} > 0$.*

A profit tax rate, $\tau$, is imposed exogenously by the government. After observing her own type $z$, the entrepreneur decides how much income to report to the IRS. We denote reported profits by $x$, chosen from the space $X \subseteq \mathbb{R}_+$. Thus, $\tau(\pi(n, z) - x)$ is the amount evaded by the entrepreneur. The IRS (the principal) costlessly observes labor $n$ and reported income $x$. However, it is able to observe ability $z$, and hence actual income, only if it audits the firm at a cost. If an entrepreneur is not audited, she transfers $\tau x$ to the tax authority, otherwise

5

the firm is assessed by $\max\{\mu\tau(\pi(n,z) - x), 0\}$. $\mu > 1$ is a linear penalty on the amount evaded. Notice we assume the IRS does not reward overreporting.

Implicitly, we assume all penalties are enforced even if $\mu\tau(\pi(n,z) - x) > \pi(n,z) - \tau x$, that is, penalties are higher than post-tax profits. Alternatively, if limited liability is a concern, we could assume that $\mu \in (1, \frac{1}{\tau}]$, which is enough to guarantee penalties are payable only with post-tax profits.

The IRS knows the distribution of firms, $G(\cdot)$. In the first stage, given $G(\cdot)$, it commits to a monitoring strategy, which is an auditing probability function $\varphi : N \times X \rightarrow [0, 1]$, dependent on employment and reported income.

In the second stage, given $\varphi$, the entrepreneur's problem is to maximize her expected profits:

$$\max_{n \geq 0, x \geq 0} \pi(n,z) - \tau x - \varphi(n,x)\max\{\mu\tau(\pi(n,z) - x), 0\} \tag{1}$$

Since the IRS does not reward overreporting, optimality implies that declared profits never exceed actual profits, that is $x \leq \pi(n,z)$. Hence, without loss of generality, we can set $\max\{\mu\tau(\pi(n,z) - x), 0\} = \mu\tau(\pi(n,z) - x)$, and restrict the set of reported income to be $[0, \pi(n,z)]$.

Moreover, when $x < \pi(n,z)$, it is optimal for a firm to declare its actual profits rather than $x$ if $x$ induces an auditing probability of $\varphi(n,x) \geq \frac{1}{\mu}$. Indeed, by comparing expected profits, we obtain $(1 - \tau)\pi(n,z) \geq \pi(n,z) - \tau x - \varphi(n,x)\mu\tau(\pi(n,z) - x)$ if and only if $\varphi(n,x) \geq \frac{1}{\mu}$. Consequently, without loss of generality, we can restrict auditing probabilities to the set $[0, \frac{1}{\mu}]$ instead of $[0, 1]$.

In this paper, the IRS is an agency responsible only for auditing and collecting taxes. Choosing tax rates, penalties and its budget is beyond its scope. These variables are usually chosen by other government spheres like the Treasury or Congress.[10]

The IRS is assigned a budget, $C$, and then chooses a monitoring strategy in order to maximize the government's revenue. When it audits a firm, it pays a cost dependent on both labor input and managerial ability, $c : N \times Z \rightarrow \mathbb{R}_+$, such that total auditing costs must be less than the budget.

**Assumption 3.** *The IRS has limited resources, $C < \frac{1}{\mu}\int_Z c(n^*(z), z)dG(z)$.*

This assumption states that the IRS does not have enough budget to audit all firms with intensity higher than $\frac{1}{\mu}$. If this assumption does not hold, by setting $\varphi(n,x) = \frac{1}{\mu}$ for all $(n,x) \in N \times X$, the IRS could fully enforce taxes making the posed question uninteresting.

---

[10]For example, in August 2007, The U.S. Government Accountability Office (U.S. GAO) published a report (http://www.gao.gov/new.items/d071062.pdf) suggesting the Congress to require IRS to periodically adjust penalties for inflation.

Before proceeding with the analysis, we quickly solve for the *full-information* case, that is, when $z$ is observable. When $x < \pi(n, z)$, the IRS prefers entrepreneurs to report their true profits instead of $x$ whenever $\varphi(n, x) \leq \frac{1}{\mu}$. Indeed, $\tau\pi(n, z) \geq \tau x + \varphi(n, x)\mu\tau(\pi(n, z) - x)$ if and only if $\varphi(n, x) \leq \frac{1}{\mu}$. Hence, the best the IRS can do is to induce a given entrepreneur $z$ to produce efficiently, $n^*(z)$, and report her true profits, $x^*(z) = \pi(n^*(z), z)$. This is achieved by the following monitoring strategy:

$$\varphi^*(n, x) = \begin{cases} \frac{1}{\mu} & , \text{ if } x \neq x^*(z) \text{ and } n \neq n^*(z) \text{ for some } z \in Z \\ 0 & , \text{ otherwise} \end{cases}.$$

With some abuse of notation, we use the superscript $*$ to denote the full-information solution in order to highlight that it induces the efficient employment per firm. This mechanism works only through off-equilibrium threats. As long as the IRS commits to this monitoring technology, it enforces all taxes without spending any resources.

When $z$ is not observable for the IRS, an adverse selection problem arises. In order to increase her expected profits, an entrepreneur may distort her labor and report less income, such that she is monitored with a potential smaller auditing probability.

To solve this problem, we follow a mechanism design approach, in which the IRS offers menus $(n, x, \varphi(n, x))$ for all $(n, x) \in N \times X$, and firms pick a particular one in order to maximize expected profits. By the revelation principle, it is enough to restrict attention within the class of direct mechanisms. Let $\phi(\cdot)$ be the monitoring strategy on the set $Z$.

**Assumption 4.** *The function $\phi : Z \to [0, 1]$ is piecewise continuously differentiable.*

The idea is to solve for the optimal direct mechanism $(n(z), x(z), \phi(z))$ for all $z \in Z$, and then construct a map between $\phi$ and $\varphi$ in the following way:

$$\varphi(n, x) = \begin{cases} \phi(z) & , \text{ if there exists } z \text{ s.t. } (n, x) = (n(z), x(z)) \\ \bar{\varphi}(n, x) & , \text{ otherwise} \end{cases},$$

where $\bar{\varphi}(n, x)$ is a high enough off-equilibrium threat to deter any deviation to such $(n, x)$.[11]

In addition to the budget constraint (BC), it is sufficient to add two more constraints in the principal's problem, the incentive compatibility constraints (IC) and the individual rationality constraints (IR).

Let a type $z$ entrepreneur's expected profits be

$$\Pi(n, x, \phi, z) = \pi(n, z) - \tau x - \phi\mu\tau(\pi(n, z) - x).$$

---

[11]Notice we are implicitly requiring that if $z_1 \neq z_2$ and $(n(z_1), x(z_1)) = (n(z_2), x(z_2))$, then $\phi(z_1) = \phi(z_2)$. This requirement is satisfied under standard sorting conditions.

Given $G(\cdot)^{12}$, the IRS problem is to

$$\max_{n(\cdot),x(\cdot),\phi(\cdot)} \int_Z \left[\tau x(z) + \phi(z)\mu\tau(\pi(n(z),z) - x(z))\right] dG(z)$$

$$s.t.$$

$(F) \quad \phi(z) \in [0, 1/\mu], x(z) \in [0, \pi(n(z), z)], n(z) \in N, \forall z \in Z$

$(IR) \quad \Pi(n(z), x(z), \phi(z), z) \geq (1 - \tau)\pi(n^*(z), z), \forall z \in Z$

$(IC) \quad \Pi(n(z), x(z), \phi(z), z) \geq \Pi(n(\tilde{z}), x(\tilde{z}), \phi(\tilde{z}), z), \forall\{z, \tilde{z}\} \in Z \times Z$

$(BC) \quad \int_Z \phi(z)c(n(z), z)dG(z) \leq C.$

The feasibility condition (F) requires that the set of offered menus correspond to feasible probabilities, declarations and labor input.

If a type $z$ entrepreneur declares her true profits, $x(z) = \pi(n(z), z)$, the firm's post-tax profits are $(1 - \tau)\pi(n(z), z)$. Hence, any mechanism must assign expected profits of at least $(1-\tau)\pi(n^*(z), z)$ to the entrepreneur.[13] This feature is captured by the individual rationality (IR) constraints.

Notice that for a given $z$, if $\phi(z) = \frac{1}{\mu}$, expected profits are $(1 - \tau)\pi(n(z), z)$. Therefore, (IR) implies that it is optimal for firms to produce efficiently and report true profits, i.e. $n(z) = n^*(z)$ and $x(z) = x^*(z)$. Hence, if Assumption 3 were violated, the IRS could recover the full-information revenue collection by setting $\phi(z) = \frac{1}{\mu}$ for all $z \in Z$. However, it would spend part of its resources $C$.

Incentive compatibility (IC) requires that no firm has an incentive to choose a different labor input or reported income than the amount induced by the mechanism.[14] Finally, the IRS is required to respect its budget constraint (BC).[15]

It turns out that the general problem is hard to solve analytically. First, agent's objective is not quasi-linear in $\phi$. Second, it features multidimensional screening (on $n$ and $x$). Finally, reservation values are type dependent. We are not aware of any paper that solves a mechanism design problem that displays these three properties.

Henceforth, we characterize the solution to two polar cases, one in which the IRS restricts

---

[12]Throughout the paper, we assume $n(\cdot)$ and $x(\cdot)$ are also piecewise continuously differentiable.

[13]Recall that $n^*(z)$ is the efficient labor, and it maximizes $(1 - \tau)\pi(n(z), z)$.

[14]As opposed to traditional mechanism design applications, like monopoly screening, the "supply" of agent's choice variables ($n$ and $x$, in this case) are not controlled by the principal. Thus an agent can also deviate to an off-scheduled decision $(n, x)$. Given (IR) implies that any mechanism must assign at least $(1-\tau)\pi(n^*(z), z)$ to each entrepreneur, this problem is trivially circumvented by setting $\bar{\varphi}(n, x) = \frac{1}{\mu}$ for all off-scheduled $(n, x)$.

[15]The budget constraint might not be satisfied off-equilibrium.

the design of menus on the space of reported income, as in the previous literature, and another in which the design of menus is restricted only on the labor space. We show that conditioning only on employment fits some empirical regularities better. Finally, we provide a partial characterization to the general problem, which is enough to explain IRS actual practices.

## 2.1 Discussion

### 2.1.1 Universe of firms

The relevance of our theory hinges on the plausibility of the assumption that profits may be hidden. However, the degree to which firms can hide profits depends on the particular context.

First, firms might be caught in a web of business-to-business transactions that facilitate enforcement for tax reasons. For example, whenever a downstream firm buys from an upstream firm, value-added taxes along the production chain generate tax credits to be used against future tax liabilities. Thus, this transaction is observable by the IRS, and compliant firms have an incentive to deal among themselves (Kopczuk and Slemrod [2006] and de Paula and Scheinkman [forthcoming]).

Second, access to the financial sector generates information that can be used by the IRS to enforce taxes (Gordon and Li [2009]). Indeed, the IRS can easily track transactions with credit cards, transfers between bank accounts, formal loans, and so on.

Finally, when the use of accounting books, necessary to run complex business operations, are known to many employees, the firm is less likely to hide them successfully from the IRS. Reward schemes can be designed such that workers have an incentive to denounce, and prove through the true books, any eventual irregularity (Kleven et al. [2009]).

Consequently, our theory is more readily applicable for a self-employed entrepreneur, whose operations require a low degree of complexity. In particular, we have in mind a firm that sells exclusively to final consumers or transacts mostly with similar firms. These transactions are usually carried on through cash or any hidable asset. Moreover, in order to operate, these firms do not rely as much on formal credit and accounting books (or can easily hide them).

These characteristics are common across firms in developing countries, particularly, but not exclusively, in the informal sector. More specifically, in these countries, there is a large number of labor-intensive firms, such as small agricultural units, family businesses (restaurants, small shops, etc.), independent contractors (builders, carpenters, etc.), and so on, that fit the description above. Nevertheless, to a lesser extent, our theory is also applica-

ble to more developed countries. Slemrod [2007], for example, reports that 43 percent of self-employment business income in the U.S. should have been reported to the IRS.

Moreover, we can refine the universe of firms based on fixed observable characteristics, such as the industry, location, and so on. Thus, we can focus on an appropriate subset of firms that display the characteristics stated above. For example, street-sellers in New York City or small manufacturers in Rio de Janeiro.

### 2.1.2 Observability of labor

Crucial to our analysis is the assumption that labor input can be observed completely without cost. In developed countries, this seems realistic. Wallis [2000], for instance, argues that the administration of the Social Security payroll tax provides an enormous amount of information to the government on wages and salaries. Moreover, business managers are usually responsible to withhold labor taxes. It is unlikely that managers and workers would collude to evade wage taxes and hide labor. Indeed, only 1 percent of wages and salaries are not reported to the IRS in the U.S. (Slemrod [2007]). In Denmark, earning taxes are also virtually enforced (Kleven et al. [2010]). Furthermore, even in an environment in which wage tax evasion is frequent, as long as wages are partially declared, the IRS still has information about the employee and the firm for which he works.

Consequently, the IRS can track the number of workers at each firm. In particular, a computer program that matches reports from both workers and firms is enough. Hence, if there is any cost, this should be a fixed one of data processing.[16] Paying a fixed cost to observe employment at each firm does not change the optimal mechanism. However, the IRS might not choose to observe labor input, if the extra revenue from conditioning the monitoring strategy on labor is not enough to pay this fixed cost.

On the other hand, in some sectors, especially in developing countries, the assumption that labor input can be observed at no cost is extreme. In particular, legal and illegal workers are highly substitutable in the informal sector. Therefore, an extension that allows firms to report labor, subject to a penalty if caught misreporting, is desirable. In other words, some of the workers are illegal (i.e., hidden from the IRS), but labor can be observed if the IRS pays a cost. Adding another choice variable to the model further complicates its tractability. Hence, we follow an alternative approach to deal with this criticism.

First, $n$ can be interpreted differently, as long as it is an observable choice variable and $\pi(\cdot, z)$ is strictly concave. For example, it can be the number of seats in a restaurant, the

---

[16]In the U.S., firms issue W-2 forms, one for each worker, detailing her identity and the amount of wages paid. Each form is sent to the IRS and the relevant employee. The latter uses the W-2 form to file her income tax return. See Logue and Slemrod [2009] for more details.

size of a building, the surface area of warehouses, or any other costlessly observable proxy for the firm size.

Second, assume production is carried out with two inputs, one that is costlessly observable, $n_o$, and another that is unobservable, $n_u$. With a slight abuse of notation, let the production technology be $f(n_o, n_u) = n_o^{\alpha_o} n_u^{\alpha_u}$, where $\alpha_o \in (0, 1)$, $\alpha_u \in (0, 1)$ and $\alpha_o + \alpha_u < 1$. Furthermore, assume the monitoring strategy can be conditioned only on the observable input $n_o$. Therefore, the model is general enough to accommodate this environment.

Indeed, normalizing the price of the observable input, entrepreneurs' expected profits are

$$pzn_o^{\alpha_o} n_u^{\alpha_u} - n_o - wn_u - \tau x - \varphi(n_o, x)\mu\tau(pzn_o^{\alpha_o} n_u^{\alpha_u} - n_o - wn_u - x), \qquad (2)$$

where $w$ is the price of the unobservable input. Taking the first order condition with respect to $n_u$, which is chosen in the last of stage of the game, one obtains

$$n_u(n_o, z) = \left(\frac{\alpha_u pz n_o^{\alpha_0}}{w}\right)^{\frac{1}{1-\alpha_u}}.$$

By plugging $n_u(n_o, z)$ back into (2), expected profits become

$$p_o z_o n_o^{\frac{\alpha_o}{1-\alpha_u}} - n_o - \tau x - \varphi(n_o, x)\mu\tau(p_o z_o n_o^{\frac{\alpha_o}{1-\alpha_u}} - n_o - x), \qquad (3)$$

where $z_o = z^{\frac{1}{1-\alpha_u}}$ and $p_o = \left(\frac{p}{w^{\alpha_u}}\right)^{\frac{1}{1-\alpha_u}} \left[\alpha_u^{\frac{\alpha_u}{1-\alpha_u}} - \alpha_u^{\frac{1}{1-\alpha_u}}\right]$.

If $f(n) = n^{\alpha}$ with $\alpha \in (0, 1)$, which satisfies Assumption 1, then (3) is mathematically equivalent to the original expected profits, although the variables have different meanings.

We provide two interpretations to this extension. First, $n_u$ and $n_o$ can be unskilled and skilled labor, respectively. Assume skilled workers face a high cost of being illegal, and there is a friction to replace them. Hence, their employer might not have enough bargain power to hide them, even if she is willing to. Second, $n_u$ can be labor input, while $n_o$ is any input bought from an intermediary. As long as this business-to-business transaction is observable by the IRS, say because there is a value-added tax along the production chain, information about $n_o$ is also readily available.

### 2.1.3 Presumptive taxation

Similar problems have been addressed in the literature. The novelty here is that the model allows the monitoring strategy to be dependent on a single characteristic of the taxpayer, $n$. The possibility to condition taxes or audits on observable variables is the central principle of presumptive taxation, which has been widely practiced around the globe. Tanzi and

de Jantscher [1989] and Thuronyi [2004] describe a variety of presumptive taxation methods. Some of them are currently used, even in developed countries.

Curiously, there has been little optimal tax (and auditing) theory formalizing this idea (exceptions are Scotchmer [1987] and Macho-Stadler and Pérez-Castrillo [2002], see footnote 8). If $n$ is observable, taxes could potentially depend on it, but in modern tax systems, an individual is required to self-assess and report her income. Since reported income is the tax base, we impose $\tau(n, x) = \tau(x)$.[17] Interestingly, even in modern tax systems, presumptive methods play an important role to the extent that they determine which reports should be audited. This possibility is incorporated in our model.

In Section 5, we formalize and solve a model of optimal presumptive taxation, in which the tax base (or presumptive income) depends on $n$.

### 2.1.4 Institutional environment

We assume a fixed budget is assigned to the IRS, so the marginal benefit and cost of monitoring need not be equalized. Thus, from a net revenue perspective, taxes are not necessarily collected efficiently. However, taxes would be efficiently enforced if either the choice of penalties or the IRS budget is internalized.[18]

Many authors argue that penalties on tax evasions are determined by non-economic factors,[19] such as a common ethical norm that respects a proportionality principle. In fact, although prohibitively high penalties would induce full enforcement at no cost, this would be disproportionate to the punishment faced for other crimes.

The fact that the IRS budget is fixed makes the framework well suited to analyze more general problems, in which the IRS choice of a monitoring strategy is a subproblem. For example, it could be cast in a political economy setup in which the Congress assigns a budget to the IRS (see Sánchez and Sobel [1993], for an example). In addition, the Congress periodically assigns budgets to different governmental institutions. Hence, in a static environment, fixing the IRS budget is a realistic assumption.

### 2.1.5 Sales taxes

By changing the interpretation of the variables in the model, the analysis carried out throughout the paper fits an environment in which the entrepreneur faces a sales tax rate, $\tau_s$, instead

---

[17]In addition, for tractability, we assume a linear tax rate $\tau(x) = \tau x$.

[18]Since Becker [1968], the optimality of penalties and the amount of resources to be assigned to enforcement activities has been widely studied by economists. See Polinsky and Shavell [2000] for a recent survey.

[19]Rosen [2005], for example, argues that "existing penalty systems try to incorporate *just retribution*. Contrary to the assumptions of the utilitarian framework, society cares not only about the end result (getting rid of the cheaters) but also the processes by which the result is achieved."

of a profits one.

Suppose the entrepreneur, who uses a single input $n$ to produce, must self-assess and report her sales proceeds, rather than profits, to the IRS. Let $s$ be reported sales proceeds. A type $z$ entrepreneur's problem is to maximize expected profits.

$$\max_{n \geq 0, s \geq 0} p_s z f(n) - w_s n - \tau_s s - \varphi_s(n, s) \mu_s \tau_s \max\{p_s z f(n) - s, 0\} \tag{4}$$

Here, $w_s$ is the price of the input, and the monitoring strategy $\varphi_s : N \times S \to [0, 1]$ is defined on the product space of labor and reported sales proceeds. A linear penalty over misreported sales proceeds is assessed when a firm is audited.

Redefine $w_s = 1 - \tau_s$, and let $x = s - n$. Straightforward algebra shows that, by setting $\varphi_s(n, s) = \varphi(n, s - n) = \varphi(n, x)$, (1) and (4) are equivalent, but taxes, penalties, prices, and $x$ are interpreted differently.

One way to interpret this model is as follows: $n$ is an intermediate good used in the production of a final good. If there is a value-added tax $\tau_s$ along the production chain, upon the purchase of $n$ units at a numeraire, firms earn a tax credit $\tau_s$ per unit of input. We assume this transaction is observed by the IRS, and so is $n$. On the other hand, a firm must pay an *ad valorem* tax $\tau_s$ over sales proceeds, which does not generate tax credits for final consumers, thus sales are not observable and may be hidden.

### 2.1.6 Tax evasion and firms

The definition of firm used in this paper is not very sophisticated. It consists of a risk-neutral self-employed entrepreneur, who manages a plant and leads a team of workers. In particular, we study the optimal tax enforcement under asymmetric information and perfect competition. However, a few recent studies stresses that business income tax evasion is affected by the internal organization of the firm and the nature of the market. Namely, Crocker and Slemrod [2005] and Chen and Chu [2005] develop shareholder-manager agency models, while Bayer and Cowell [2009] study tax evasion in an oligopolistic setting.[20] In all of these papers, auditing probabilities are exogenous. Hence, it would be interesting to

---

[20]Crocker and Slemrod [2005] study the contractual relationship between a shareholder and a tax specialist who has private information regarding the extent of legally permissible reduction in taxable income. The authors show that penalties imposed on the tax manager are more effective in reducing evasion. Chen and Chu [2005], for instance, develop a moral hazard model, in which a risk-averse manager exerts costly effort that affects outcomes. If the risk-neutral shareholder decides to engage in evasion, contracts cannot be legally contingent on the outcome of an audit. This incompleteness creates a distortion in the manager's effort and reduces the efficiency of the contract. Finally, Bayer and Cowell [2009] study tax evasion in a Cournot environment. If the monitoring strategy is conditioned on every firm's observable behavior, then it manipulates both market decisions and compliance behavior. This potentially generates less tax evasion and an efficiency improvement.

extend these models and study the implications of optimal tax enforcement in these different setups. It would also be interesting to study the optimal tax enforcement in a version of our model that displays agency costs within the firm.

# 3 Specific cases

## 3.1 Screening over $x$

In this section we describe the optimal mechanism when the IRS restricts the design of menus only on the space of reported income $X$. This particular case has been studied in the literature (see footnote 1 for references). Optimal strategies under this mechanism are denoted by the superscript $x$.

Since the monitoring strategy does not induce distortions in the labor choice, firms employ efficiently, $n^x(z) = n^*(z)$ for all $z \in Z$. It is easy to check that (IR) is immediately satisfied.[21]

Rewriting pre-tax profits as $y(z) = \pi(n^*(z), z)$, the IRS problem is to

$$\max_{\phi(\cdot),x(\cdot)} \int_Z [\tau x(z) + \phi(z)\mu\tau(y(z) - x(z))]dG(z)$$

s.t.

$(F)$ $\quad \phi(z) \in [0, 1/\mu]\,, x(z) \in [0, y(z)], \forall z \in Z$

$(IC)$ $\quad \tau x(z) + \phi(z)\mu\tau(y(z) - x(z)) \le \tau x(\tilde{z}) + \phi(\tilde{z})\mu\tau(y(z) - x(\tilde{z})), \forall z, \tilde{z} \in Z \times Z$

$(BC)$ $\quad \int_Z \phi(z)c(n^*(z), z)dG(z) = C.$

This problem is a variant of the one analyzed by Sánchez and Sobel [1993]. The main differences are that they directly consider an endowment economy, and that the cost function in their paper is constant. Indeed, by setting $c(n(z), z) = c$, and with an appropriate change from the space of managerial ability $Z$ to the space of pre-tax income, $Y \equiv \{y(z)|z \in Z\}$, our problem is easily converted into theirs.[22]

**Proposition 1.** *(Sánchez and Sobel [1993], adapted) If* $\frac{1-G(z)}{g(z)} \frac{f(n^*(z))}{c(n^*(z),z)}$ *is strictly decreasing,*

---

[21]Recall that $x(z) \le \pi(n^*(z), z)$. (IR) is violated if $\pi(n^*(z), z) - \tau x(z) - \phi(z)\mu\tau(\pi(n^*(z), z) - x(z)) < (1 - \tau)\pi(n^*(z), z)$, which holds if and only if $\phi(z) > \frac{1}{\mu}$ and $x(z) < \pi(n^*(z), z)$, but this contradicts entrepreneurs' optimality.

[22] Notice that $Y = [y(\underline{z}), y(\overline{z})]$, and since $y(\cdot)$ is an increasing function, we can define a distribution function by $F(y) = G(z^{-1}(y))$ for all $y \in Y$. With some abuse of notation, let $x(y) = x(z^{-1}(y))$, $x(\tilde{y}) = x(z^{-1}(\tilde{y}))$, $\phi(x(y)) = \phi(z^{-1}(y))$, and $\phi(x(\tilde{y})) = \phi(z^{-1}(\tilde{y}))$. This setup is equivalent to Sánchez and Sobel [1993], as long as the image of $x(\cdot)$ is $Y$. Even if there does not exist $y \in Y$, such that $x(y) = i$, for some $i \in Y$, the IRS can simply offer $(\frac{1}{\mu}, i)$.

*then the optimal and unique IRS monitoring scheme is*

$$\phi^x(z) = \begin{cases} \frac{1}{\mu} & if \quad z < z^x \\ 0 & if \quad z \geq z^x \end{cases},$$

*where $z^x$ solves $\int_{\underline{z}}^{z^x} c(n^*(z), z) dG(z) = \mu C$.*

Intuitively, any strictly increasing monitoring strategy is not incentive compatible, since every entrepreneur would report less and be monitored with less intensity. The IRS is budget-constrained, so the best it can do is to monitor low-ability entrepreneurs, and let the high ones report the profits of a particular threshold type, $z^x$, that clears the budget constraint.

The result in this section relies on the assumption that $\frac{1-G(z)}{g(z)} \frac{f(n^*(z))}{c(n^*(z),z)}$ is strictly decreasing, which is a variation of the standard monotone hazard-rate requirement that $\frac{1-G(z)}{g(z)}$ is strictly decreasing. This assumption is satisfied if the ratio $\frac{f(n^*(z))}{c(n^*(z),z)}$ does not increase in $z$ at a higher rate than $\frac{1-G(z)}{g(z)}$ decreases in $z$. However, even if this assumption is violated, there always exists a solution to this problem in which taxpayers are divided into, at most, three groups according to their income. The lowest income group is audited with intensity $\frac{1}{\mu}$, while the highest is not monitored at all. All individuals in the middle group are audited with some probability in $(0, \frac{1}{\mu})$. See Sánchez and Sobel [1993] for more details.

Finally, in the reported income space, $X$, the monitoring is strategy is

$$\varphi(x) = \begin{cases} \frac{1}{\mu} & \text{, if } x < y(z^x) \\ 0 & \text{, if } x \geq y(z^x) \end{cases}.$$

## 3.2 Screening over $n$

In this section, we study optimal auditing when the IRS only exploits information about employment, that is, when auditing probabilities depends only on $n$. Let the superscript $n$ denote the optimal strategies in this section.

Since the monitoring strategy does not induce distortions in income reports, for all $z$:

$$x^n(z) = \begin{cases} \pi(n^n(z), z) & \text{if} \quad \phi^n(z) \geq 1/\mu \\ 0 & \text{if} \quad \phi^n(z) < 1/\mu \end{cases}.$$

This trigger strategy arises from the fact that expected profits are linear in $x(z)$.[23] This result would not hold without linear penalties on the amount evaded or in an environment

---

[23]We are implicitly assuming throughout the paper that, whenever indifferent between evading or not, taxpayers do not evade. Similarly, whenever indifferent between monitoring intensities, the IRS chooses to audit with probability $1/\mu$.

with risk aversion. This implies that the type set, $Z$, is partitioned into two sets: firms that hide all profits (full evaders), and firms that comply with taxes (compliers).

The IRS can audit firms with intensity of, at most, $\frac{1}{\mu}$. Hence, with some abuse of notation, we can redefine expected profits as $\Pi(n(z), \phi(z), z) = (1 - \mu\tau\phi(z))\pi(n(z), z)$. We rewrite the IRS problem as to

$$\max_{n(\cdot), \phi(\cdot)} \int_Z \mu\tau\phi(z)\pi(n(z), z)dG(z)$$

s.t.

$(F)$    $\phi(z) \in [0, 1/\mu]$, $n(z) \geq 0, \forall z \in Z$

$(IR)$    $\Pi(n(z), \phi(z), z) \geq (1 - \tau)\pi(n^*(z), z), \forall z \in Z$

$(IC)$    $\Pi(n(z), \phi(z), z) \geq \Pi(n(\tilde{z}), \phi(\tilde{z}), z), \forall\{z, \tilde{z}\} \in Z \times Z$

$(BC)$    $\int_Z \phi(z)c(n(z), z)dG(z) \leq C.$

The following lemma is standard.

**Lemma 1.** *(IC) is equivalent to*

$(LIC)$    $(1 - \tau\mu\phi(z))\pi_n(n(z), z)\dfrac{dn}{dz}(z) - \mu\tau\pi(n(z), z)\dfrac{d\phi}{dz}(z) = 0$

$(M)$    $\dfrac{dn}{dz} \geq 0$

*for all $z \in Z$ such that $\frac{d\phi}{dz}(z)$ and $\frac{dn}{dz}(z)$ exist, and*

$(IC')$    $\lim_{z\uparrow z'}\Pi(n(z), \phi(z), z') = \lim_{z\downarrow z'}\Pi(n(z), \phi(z), z')$

*for any $z' \in Z$ such that either $\frac{d\phi}{dz}(z')$ or $\frac{dn}{dz}(z')$ does not exist.*

The previous lemma states that $n(\cdot)$ is implementable by some auditing probability $\phi(\cdot)$ if and only if (LIC), (M), and (IC') hold. However, expected profits are not a quasi-linear function, so we cannot use standard methods to solve for the optimal mechanism. Our approach involves two steps. First, we ignore (IC) and solve the problem assuming $z$ is observable. We call this solution the *n-full-information*.[24] Then, we modify $n(\cdot)$ such that the IRS still achieves the n-full-information revenue and satisfies the (IC).

This method works here because the IRS always chooses to monitor the low-types with

---

[24]We label n-full-information to distinguish from the full-information solution in Section 2.

zero probability. Hence, any distortion on their labor decision does not affect the IRS n-full-information revenue collection.

**Assumption 5.** $\frac{\pi(n^*(z),z)}{c(n^*(z),z)}$ *is strictly increasing in* $z$.

Once $\pi(n^*(z), z)$ is strictly increasing, this assumption is satisfied for a non-increasing cost function in $z$. If the cost is increasing, all we require is that it increases at a lower rate than the profit function. Also, this assumption is consistent with a constant cost function that has been widely used in the literature.[25]

**Lemma 2.** *Assume* $z$ *is observable. There exists a threshold* $z^n$ *such that it is optimal to set* $\phi^n(z) = \frac{1}{\mu}$ *for all* $z \geq z^n$, *and* $\phi^n(z) = 0$ *for all* $z < z^n$. $z^n$ *is uniquely determined by* $\int_{z^n}^{\bar{z}} c(n^*(z), z) dG(z) = \mu C$. *Finally,* $(n^*(z), \phi^n(z))$ *for* $z \in Z$ *is the optimal mechanism.*

*Proof.* Since $z$ is observable by assumption, we can solve this problem by backward induction. In the second period, given $\phi(\cdot)$, a firm $z$ chooses to report $x^n(z)$, and employ $n^*(z)$. In the first period, the IRS sets $\phi(z)$ in order to maximize $\int_Z \mu\tau\phi(z)\pi(n^*(z), z)dG(z)$, subject to (F) and (BC). Let $\lambda$ be the lagrange multiplier associated with $(BC)$, which binds by Assumption 2. Then,

$$\phi^n(z) = \begin{cases} 0 & \text{if} \quad \mu\tau\pi(n^*(z), z) - \lambda c(n^*(z), z) < 0 \\ \xi \in \left[0, \frac{1}{\mu}\right] & \text{if} \quad \mu\tau\pi(n^*(z), z) - \lambda c(n^*(z), z) = 0 \\ \frac{1}{\mu} & \text{if} \quad \mu\tau\pi(n^*(z), z) - \lambda c(n^*(z), z) > 0 \end{cases}.$$

Since $\frac{\pi(n^*(z),z)}{c(n^*(z),z)}$ is strictly increasing, there is a unique $z^n$, such that $\frac{\pi(n^*(z^n),z^n)}{c(n^*(z^n),z^n)} = \frac{\lambda}{\mu\tau}$. □

Here, we use the superscript $n$ anticipating the next proposition, which states that the monitoring strategy is the same whether $z$ is observable or not.

Lemma 2 relies on Assumption 5, which induces the IRS to more intensively monitor high-ability entrepreneurs. In contrast, if this assumption is violated, the optimal mechanism could induce the larger firms to become the set of evaders,[26] or even a disconnected set of evaders.[27]

---

[25]It can be argued that large firms take longer to monitor than smaller firms, so $c(\cdot, z)$ might be increasing. Moreover, for a high-ability entrepreneur, it could be easier to circumvent the law, and hide her income, making $c(n, \cdot)$ increasing. On the other hand, there may be a visibility effect that reduces the informational cost associated with monitoring larger firms, hence $c(\cdot, z)$ could be decreasing. Thus, the validity of Assumption 5 is an empirical question.

[26]If $\frac{\pi(n^*(z),z)}{c(n^*(z),z)}$ is strictly decreasing, for example.

[27]Notice $\pi(n^*(z), z)$ is strictly increasing and convex, so $c(n^*(z), z)$ needs to increase at a higher rate in order to violate the conditions required by Lemma 2. Moreover, Assumption 5 can be relaxed as follows: The set $\left\{ z : \frac{\pi(n^*(z),z)}{c(n^*(z),z)} = \frac{\lambda}{\mu\tau} \right\}$ is connected, and $\frac{\pi(n^*(\bar{z}),\bar{z})}{c(n^*(\bar{z}),\bar{z})} \geq \frac{\lambda}{\mu\tau}$. Since this relaxed assumption is stated in terms of the endogenous variable $\lambda$, we prefer to state Assumption 5 the way it is.

Intuitively, an auditing probability $\frac{1}{\mu}$ suffices to guarantee that an entrepreneur report her true income. If the IRS observes $z$, as long as the cost of auditing increases at a lower rate than profits, it exhausts its budget over high-ability, and hence profitable, entrepreneurs. This strategy maximizes revenue.

Once $z$ is privately observable, an adverse selection problem arises. In this case, a subset of compliant firms could potentially distort their own labor input, become evaders, and avoid being monitored at all. Under the presence of adverse selection, the IRS must account for the (IC) upon its maximization problem.

**Proposition 2.** *Let $z^n$ solve $\int_{z^n}^{\bar{z}} c(n^*(z), z)dG(z) = \mu C$. Define $z_0$ such that $(1-\tau)\pi(n^*(z^n), z^n) = \pi(n^*(z_0), z^n)$. Let*

$$
n^n(z) = \begin{cases} n^*(z) & \text{if } \underline{z} \le z < z_0 \text{ and } \underline{z} < z_0 \\ n^*(z_0) & \text{if } \max\{\underline{z}, z_0\} \le z < z^n \\ n^*(z) & \text{if } z^n \le z \end{cases},
$$

*and $\phi^n(z)$ be the one obtained when $z$ is observable. Then, $n^n(\cdot)$ and $\phi^n(\cdot)$ solve the principal problem.*

*Proof.* Notice that this mechanism generates the n-full-information revenue collection. Moreover, (BC), (F), (M), (LIC), and (IC') are trivially satisfied by construction. Therefore, all we need is to check that (IR) holds for all $z \in (z_0, z^n)$.

Define $h : [z_0, z^n] \to \mathbb{R}_{++}$, such that

$$
h(z) = \frac{(1-\tau)\pi(n^*(z), z)}{\pi(n^*(z_0), z)},
$$

and notice that $h(z^n) = 1$. If $h'(\cdot) \ge 0$, $\pi(n^*(z_0), z) \ge (1-\tau)\pi(n^*(z), z)$ for all $z \in (z_0, z^n)$. Indeed,

$$
h'(z) = \frac{(1-\tau)p}{\pi(n^*(z_0), z)^2}[f(n^*(z_0))n^*(z) - f(n^*(z))n^*(z_0)] \ge 0,
$$

by Assumption 1 and $n^*(z) \ge n^*(z_0)$ for all $z \in [z_0, z^*]$. $\qquad\square$

Figure 1 illustrates how Proposition 2 works. The left axis plots the optimal auditing probability, whereas the right one plots the optimal employment. The x-axis is the type space $Z$. After finding the n-full-information solution (top panel), we adjust labor, $n(\cdot)$, such that (IC), (IR), (BC) and (F) are satisfied, and government's revenue does not change (bottom panel). Notice this result does not hinge on a monotone hazard-rate assumption.

This solution is not unique. For example, the IRS could distort the labor choice for all $z < z^n$, making it flat. We rule out this possibility based on efficiency grounds. In other

**n-first-best mechanism (*z* is observable)**

$1/\mu$

$\phi$

$n$

$n^*(z)$

$x(z) = 0$

$x(z) = \pi(n^*(z), z)$

$\underline{z}$      $z^n$      $\overline{z}$

**optimal mechanism (*z* is privately observable)**

$1/\mu$

$\phi$

$n$

$n^n(z)$

$x(z) = 0$

$x(z) = \pi(n^*(z), z)$

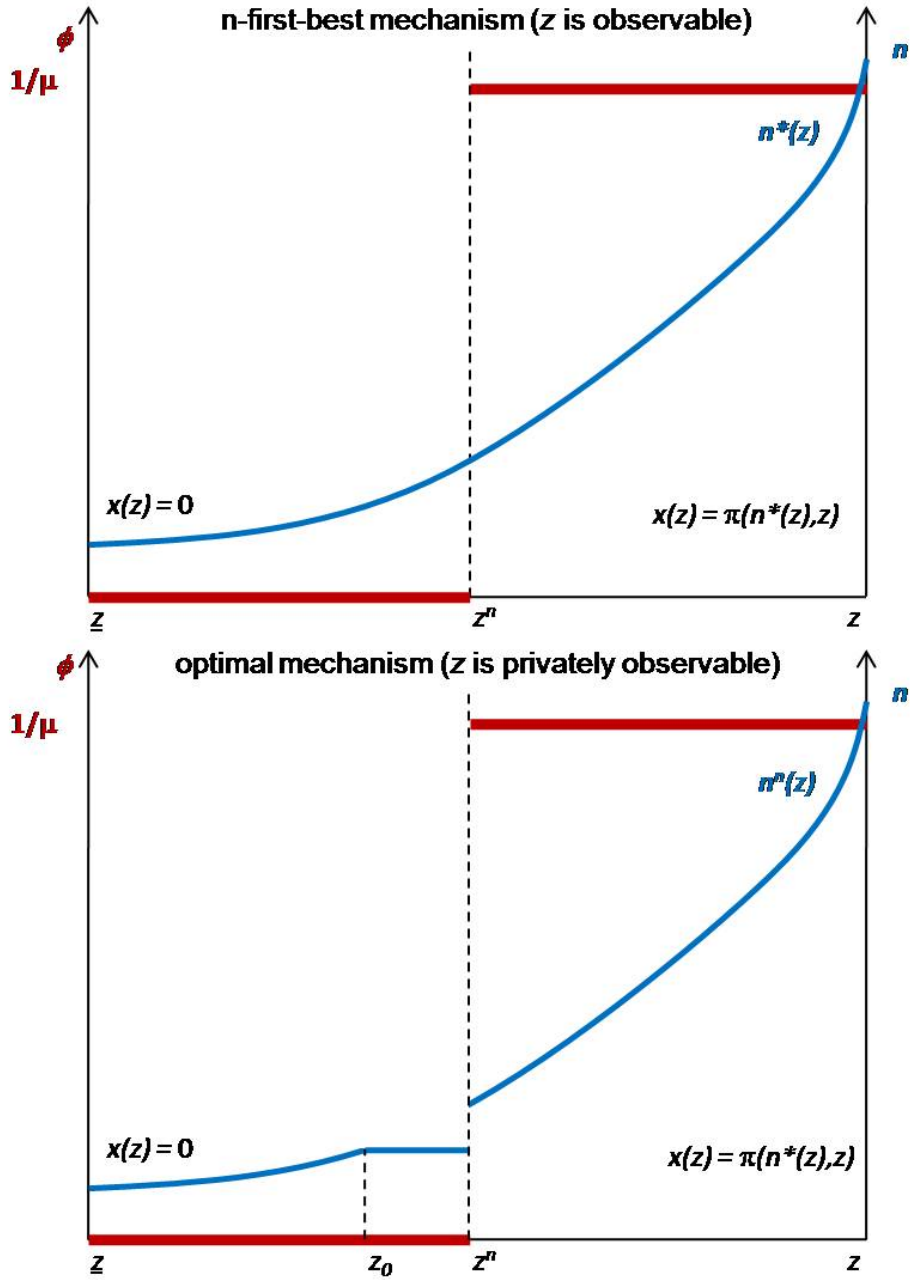$\underline{z}$      $z_0$      $z^n$      $\overline{z}$

Figure 1: optimal mechanism

19

words, this solution generates the highest revenue, by distorting labor decisions as less as possible.

In the labor space, $N$, the monitoring is strategy is

$$\varphi(n) = \begin{cases} 0 & \text{, if } n \leq n^*(z_0) \\ \frac{1}{\mu} & \text{, if } n > n^*(z_0) \end{cases}.$$

Since the direct mechanism does not assign auditing probabilities for any $n \in (n^*(z_0), n^*(z^n))$, we set $\varphi(n) = \frac{1}{\mu}$ when such labor decisions are observed. (IR) guarantees these off-equilibrium threats are enough to deter deviation.

## 3.3 Discussion

When the IRS screens only over reported income, an optimal mechanism induces cross-sectional characteristics that are inconsistent with the empirical evidence. It induces a positive relationship between firm size, measured by employment, and amount evaded, while Dabla-Norris et al. [2008] estimate a negative correlation.

Also, under this mechanism, every firm produces efficiently. If we interpret hidden profits as a measure of informality,[28] this result is also inconsistent with the view that informal firms lack efficiency. See, for example, la Porta and Shleifer [2008] and the references therein.

By conditioning the monitoring strategy on employment, the model is able to match these empirical regularities.[29] Interestingly, since potentially larger firms remain small to avoid detection, this mechanism generates a *missing middle*, that is, the fact that medium-scale firms are scarce (see Tybout [2000]). Notice this result follows from the game-theoretical approach to the problem, and not from a primitive assumption on the distribution of managerial ability.

In the development literature, there is a class of models that is closely related to the ideas explored in this section. In models of entrepreneurship, such as Lucas [1978], an informal sector arises whenever there is an exogenous cost to operate in the formal sector (e.g., bureaucracy cost, bribing, taxes, or minimum wages), and some exogenous feature that induces larger firms to operate legally (e.g., an auditing probability that increases with firm size, penalties that are proportional to profits, or credit constrains for small firms). In equilibrium, for some threshold type, the expected value of operating in both sectors must

---

[28]Of course, informality encompasses a wider range of activities than only tax noncompliance. See Schneider and Enste [2000] for a taxonomy of the underground economy.

[29]Scotchmer [1987] and Macho-Stadler and Pérez-Castrillo [2002] also show that the use of signals can help select audits in a way that reduces the regressive bias, i.e., richer taxpayers evade proportionally more taxes.

be equal. Hence, high-ability entrepreneurs self-select into the formal sector. Rauch [1991] was the first to formalize this idea (see footnote 6 for more references).

These models usually predict (1) a negative relationship between firm size and amount evaded, (2) the fact that informal firms produce inefficiently, and (3) a missing middle. This paper also contributes to this literature, since it shares the same flavor, although the monitoring strategy is endogenous.[30] Consequently, it is not surprising that we are able to deliver these empirical regularities. However, by endogeneizing the monitoring strategy, we claim that the emergence of an informal sector with these characteristics is also consistent with optimal policy. In principle, the optimal monitoring strategy could not be increasing in the firm size, and one of the driving forces to match these facts would be missing.

Moreover, as Dharmapala et al. [2009] point out, neither the development nor the public finance literature address "under what conditions it is optimal for policy to treat differentially firms of different size, perhaps in a way that generates a missing middle." As in this paper, Dharmapala et al. [2009] explain this phenomenon in an environment with heterogenous firms and asymmetric information. However, in their framework there is no evasion since output is observable, but it is costly to implement a tax rate that varies across firms. As a result, the principal exempts low-output firms from paying taxes. In our model employment, rather than output, is observable, the tax rate is fixed, and it is costly to audit potential evaders. As a result, the principal does not audit low-ability entrepreneurs. Hence, in both models, the smallest firms face the lowest effective tax burden, and the missing middle is generated by optimal policy.

# 4  Screening over $x$ and $n$: a partial characterization

As stated above, the general problem, in which the IRS screens over reported income and labor, may be hard to solve. Nonetheless, in this section, we provide a partial characterization based on the intuition built by studying the mechanism that conditions solely on $n$ or $x$. In particular, we restrict the set of auditing probabilities to be bang-bang, that is $\phi(z) \in \{0, \frac{1}{\mu}\}$ for all $z \in Z$. For this particular case, we show that for any optimal direct mechanism, there is an optimal indirect mechanism, in which the monitoring strategy resembles IRS actual practices.

We say a decision $(n(\cdot), x(\cdot))$ is implementable if there exists an auditing probability $\phi(\cdot)$, such that $(x(z), n(z), \phi(z))$ for $z \in Z$ satisfies (IC). If $(n(z), x(z), \phi(z))$ for $z \in Z$ also satisfies (IR), we say $(n(\cdot), x(\cdot))$ is rational-implementable by $\phi(\cdot)$, and $(n(z), x(z), \phi(z))$ for $z \in Z$ is

---

[30]Actually, it is straightforward to incorporate this mechanism into this class of models. Hence, we also provide a micro-foundation for an auditing probability usually used in these models.

a rational-implementable allocation.

The next lemma sates necessary and sufficient conditions for implementability.

**Lemma 3.** *If $\phi(z) \in \{0, \frac{1}{\mu}\}$ for all $z \in Z$, a decision $(n(\cdot), x(\cdot))$ is implementable by some auditing probability $\phi(\cdot)$ if and only if*

$$\pi_n(n(z), z)(1 - \phi(z)\mu\tau)\frac{dn}{dz}(z) - \tau(1 - \phi(z)\mu)\frac{dx}{dz}(z) = 0$$

$$\frac{dn}{dz}(z) \geq 0$$

*for all $z \in Z$ such that $\frac{d\phi}{dz}(z)$, $\frac{dx}{dz}(z)$ and $\frac{dn}{dz}(z)$ exist, and*

$$\lim_{z \uparrow z'} \Pi(n(z), x(z), \phi(z), z') = \lim_{z \downarrow z'} \Pi(n(z), x(z), \phi(z), z')$$

*for any $z' \in Z$ such that $\frac{d\phi}{dz}(z')$, $\frac{dx}{dz}(z')$ or $\frac{dn}{dz}(z')$ do not exist.*

The proof of this lemma is standard. If $n(z) \leq n^*(z)$ for all $z \in Z$, the first two conditions imply that $\frac{dx}{dz}(\cdot) \geq 0$ and $\frac{dn}{dx}(\cdot) \geq 0$ almost everywhere.

The next proposition presents an indirect monitoring strategy that implements any rational-implementable allocation that has the bang-bang property. It constructs an "inputted income" function of labor, $I : N \to \mathbb{R}_+$, such that the IRS monitors with the highest possible intensity every entrepreneur that reports profits below inputted income. As we argue later, this strategy resembles IRS actual practices in many countries.

**Proposition 3.** *If $\phi(\cdot) \in \{0, \frac{1}{\mu}\}$, any rational-implementable decision $(n(\cdot), x(\cdot))$ by $\phi(\cdot)$ is also implementable by the following indirect monitoring strategy:*

$$\varphi(n, x) = \begin{cases} 0 & \text{if } x > I(n) \\ \phi(z) & \text{if } x = I(n), n = n(z) \text{ and } x = x(z), \text{ for some } z \in Z \\ \frac{1}{\mu} & \text{otherwise} \end{cases},$$

*where $I : N \to X$ is defined by*

$$I(n) = \begin{cases} x(z) & \text{if } \exists z \text{ such that } n(z) = n \text{ and } \phi(z) = 0 \\ \frac{1}{\tau}\left[(1 - \tau)f^{-1}\left(\frac{1}{1-\tau}f(n)\right) - n\right] & \text{otherwise} \end{cases}.$$

*Proof.* For a given mechanism, a type $z$ can either deviate to a $(n, x)$ such that $\varphi(n, x) = 0$, or another $(n, x)$ such that $\varphi(n, x) = \frac{1}{\mu}$. We show below that none of these deviations are profitable.

22

Suppose a given type $z$ deviates to $(n, x)$ such that $\varphi(n, x) = \frac{1}{\mu}$. Hence, her expected profits can not be greater than $(1 - \tau)\pi(n, z)$. By (IR), this deviation is not profitable.

Suppose $z$ deviates to $(n, x)$ such that $\varphi(n, x) = 0$. Assume there exists $z'$ such that $n(z') = n$. $x \geq I(n)$ implies $x \geq x(z')$. If $\phi(z') = 0$, by (IC), $\pi(n(z), x(z), \phi(z), z) \geq \pi(n, x(z'), 0, z) \geq \pi(n, x, 0, z)$, hence this deviation is not profitable.

If $\phi(z') = \frac{1}{\mu}$ or there does not exist such $z'$, by (IR), the deviation does not occur whenever $(1 - \tau)\pi(n^*(z), z) \geq \pi(n, z) - \tau x$. Rearranging, $\tau x \geq \pi(n, z) - (1 - \tau)\pi(n^*(z), z)$. Maximizing the RHS with respect to $z$ yields $f(n) = (1 - \tau)f(n^*(z))$. Substituting it in the RHS of the inequality, and rearranging terms back again, one gets that whenever $x \geq \frac{1}{\tau}\left[(1 - \tau)f^{-1}\left(\frac{1}{1-\tau}f(n)\right) - n\right]$, the deviation is not profitable.

<div style="text-align: right">□</div>

This proposition shows one possible way to construct a map between the direct and the indirect mechanism. It sets off-equilibrium auditing intensities high enough to deter deviation. This map is not unique. If $(n(\cdot), x(\cdot))$ is indirectly implementable by $\varphi$, it is also implementable by any indirect monitoring strategy that sets higher off-equilibrium auditing probabilities. Notice the optimal $I(\cdot)$ is fully characterized by the optimal rational-implementable allocation $(n(z), x(z), \phi(z))$ for $z \in Z$.

Figure 2 summarizes Proposition 3. For the sake of exposition, we impose that the IRS monitor high-ability entrepreneurs with intensity $\frac{1}{\mu}$. This problem fits an IRS that is concerned with the efficiency of big firms, or is restricted by social norms to spend all its budget auditing the richest entrepreneurs. In Appendix B, under additional assumptions on the production technology and the type distribution, we fully characterize the solution to this problem.[31]

The solid line corresponds to the set of direct menus designed by the IRS. If the entrepreneur picks one of the menus in the right and steeper curve, she is monitored with intensity $\frac{1}{\mu}$. If she picks one in the left curve, she is not monitored at all. Notice the IRS does not design direct menus for all possible pairs $(n, x)$. The dashed line is the function $I(\cdot)$, which overlaps with the left set of menus.

## 4.1 IRS actual practices

If we interpret $I(\cdot)$ as "inputted" income, calculated based on $n$, an observable characteristic of the firm, Proposition 3 states that one way to optimally audit firms is as follows: Whenever

---

[31] Figure 2 displays some properties of the solution of a calibrated version of this problem. See Appendix B for more details.
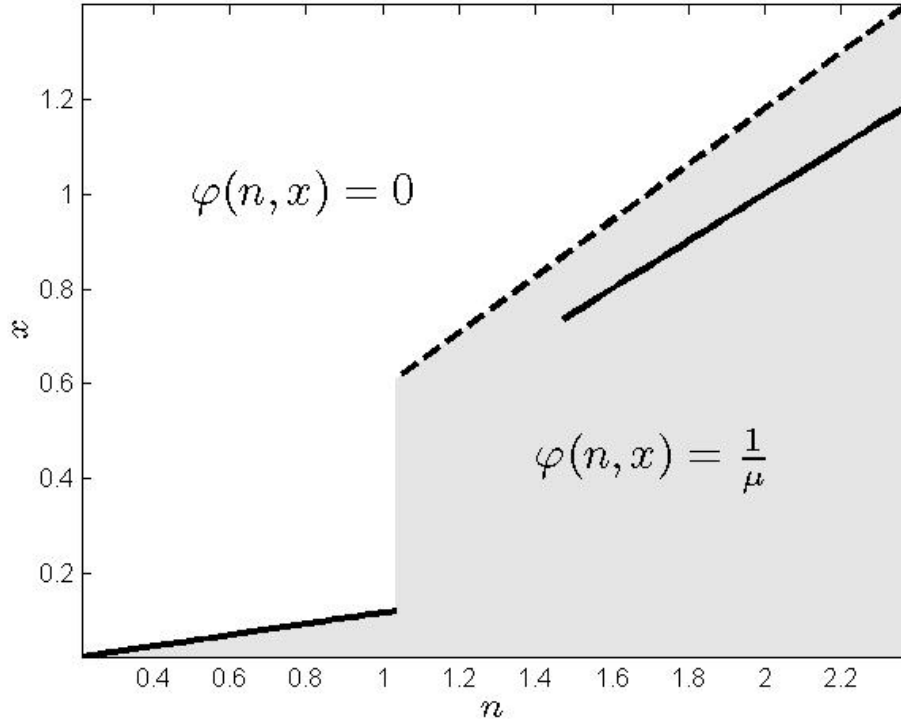
Figure 2: implementation

an entrepreneur declares less than her inputted income, the IRS monitors her with intensity $\frac{1}{\mu}$. This strategy fits IRS actual practices in several countries.

In Italy, for example, a small or medium sized firm can be audited if it reports sales proceeds that are lower than a presumed level. Presumed sales proceeds are statistically inferred from easily observable variables, like surface area of offices and warehouses, number of employees, type of customers, and so on. See Arachi and Santoro [2007] and Santoro [2008] for more details.

Similarly, in Israel, the *tachshiv* system was developed to assess self-employed and small business income. Tax liability was determined by assessed income, which was estimated on a wide range of easily observable characteristics, like location, business type, posted prices, and so on. In 1975, the *tachshiv* was replaced by a book-keeping system. However, anecdotal evidence suggests it continues to be used unofficially as an auditing strategy. See Wilkenfeld [1973] and Yitzhaki [2007] for more details.

To the extent that the rules used to input income are known, as in Italy, declared income tends to converge to presumed income. Hence, these systems are also interpreted as a method of presumptive taxation (Arachi and Santoro [2007]).

Andreoni et al. [1998] describe how audit policy is conducted in the U.S. for individual

income tax returns. In a first stage, intensive audits are conducted on a stratified random sample. Then, these results are used to assess the likelihood that a report contains evasion. Slightly over one-half of all audit selections are based at least partly on this method. Therefore, the IRS also relies on observable information to perform audits, and follows a similar strategy to the one predicted by the model. However, the rules used to assign each report a likelihood that it contains irregularities are strictly guarded.

Finally, inputted income, $I(\cdot)$, should be designed not only to presume true income, but also to deter deviation, since entrepreneurs might be willing to change $n$ in order to be inputted a lower income. It is not clear whether the IRS accounts for these possible deviations when designing $I(\cdot)$.

# 5   Optimal presumptive taxation

In this section, we adapt our model and shift the discussion towards presumptive tax systems. This is important because it allows us to compare revenue collection, even with the unsolved general model outlined in Section 2. In particular, Section 6 provides an insight on why presumptive tax methods are scarce in developed countries, but still widely present in undeveloped and developing ones.

We have in mind a particular system, in which presumed income $I(\cdot)$ is a function of labor input, as in the previous section, but instead of using it to allocate auditing probabilities, the IRS use it as the "presumptive" tax base. In other words, an entrepreneur who employs $n$ must transfer $\tau I(n)$ to the IRS.

In this tax system, entrepreneurs are not required to self-assess and report their income, thus audits and punishments are ruled out from the model. Furthermore, since labor is costlessly observable, the IRS has no expenses and entrepreneurs fully comply with taxes.

In this setup, entrepreneurial post-tax profits are

$$\pi(n, z) - \tau I(n),$$

so, invoking the revelation principle, the IRS must solve the following mechanism design problem in order to maximize revenue. Let $P : Z \to \mathbb{R}_+$ be the direct presumptive tax base, i.e., presumed income as a function of the managerial ability, which is assumed to be piecewise continuously differentiable.

$$\max_{P(\cdot),n(\cdot)} \int_Z \tau P(z)dG(z)$$

$s.t.$

$(F) \quad n(z) \geq 0, \forall z \in Z$

$(IC) \quad \pi(n(z),z) - \tau P(z) \geq \pi(n(\tilde{z}),z) - \tau P(\tilde{z}) \; \forall z, \tilde{z} \in Z \times Z$

$(IR) \quad \pi(n(z),z) - \tau P(z) \geq O(z)$

The individual rationality constraint requires that for an entrepreneur with ability $z$, her post-tax income must be greater than, or equal to, her reservation utility, which is equal to $O(z)$.

Following Tanzi and de Jantscher [1989], presumptions of income are either rebuttable or not. Under a rebuttable presumptive tax system (RPTS), income may be rebutted when taxpayers can prove their presumptive income is higher than their actual income. In contrast, under an irrebuttable presumptive tax system (IPTS), taxpayers are not allowed to claim any revision.

Specifying functional forms to the entrepreneurs' reservation utility is a useful way to model these institutions. Under the RPTS, if $m(z)$ is the cost a given type $z$ faces to claim, and prove that her presumptive income is too high, her outside value is $O(z) = (1 - \tau)\pi(n(z),z) - m(z)$.

On the other hand, in the IPTS, the reservation utility for all $z \in Z$ is set to be a nonnegative constant, $O(z) = \underline{\Pi}$. In particular, if everyone's outside option is to be employed in a firm that pays the market wage, then $\underline{\Pi} = w$. However, one might think of $\underline{\Pi}$ as being the value of home production, or even some political cost that prevents the government from extracting someone's entire income.[32]

In Sections 5.1 and 5.2, we derive the optimal presumptive tax base for irrebuttable and rebuttable presumptions of income, respectively.

## 5.1 Irrebuttable presumptive tax system

When $O(z) = \underline{\Pi}$, the mechanism design problem is standard, given that $\tau P(z)$ plays the role of transfers from agent $z$ to the principal. Assuming from now on that $\frac{G(z)-1}{g(z)}$ is strictly increasing, one can follow the steps in Fudenberg and Tirole [1991] to prove the following

---

[32]Of course, one can argue that, even in the IPTS, the outside option depends on the entrepreneur's managerial ability. If this is the case, by choosing an appropriate functional form to $m(\cdot)$, the optimal mechanism would be analogous to the one that solves the RPTS problem.

proposition. Let the superscript $i$ denote the optimal solution to the mechanism design problem with irrebuttable presumptions of income.

**Proposition 4.** *If* $\underline{z} + \frac{G(\underline{z})-1}{g(\underline{z})} \geq 0$, *then*

$$n^i(z) = f'^{-1}\left(\frac{1}{p}\left[z + \frac{G(z)-1}{g(z)}\right]^{-1}\right)$$

$$P^i(z) = \frac{1}{\tau}\left[\pi(n^i(z), z) - \underline{\Pi} - \int_{\underline{z}}^{z} pf(n^i(s))ds\right]$$

*for all* $z \in Z$.

Obviously, standard results present in the mechanism design literature hold here. First, the top type $\bar{z}$ is the only one employing labor efficiently. Second, except for the lowest type $\underline{z}$, the principal is not able to fully appropriate the agents' informational rent. Third, transfers do not depend on the tax rate $\tau$. Notice we do not restrict optimal transfers to be non-negative. Indeed, if $\underline{\Pi}$ is large enough, transfers are negative for some of the lower types.

## 5.2   Rebuttable presumptive tax system

In the RPTS, $O(z) = (1 - \tau)\pi(n(z), z) - m(z)$. For simplicity, we assume that an entrepreneur faces a cost, proportional to post-tax profits, to claim and prove that her profits were presumed wrongly. That is, $m(z) = \gamma(1 - \tau)\pi(n(z), z)$ for all $z \in Z$, with $\gamma \in [0, 1]$. Hence, her reservation utility is $O(z) = (1 - \tau)(1 - \gamma)\pi(n(z), z)$. One can think of $\gamma$ as the opportunity cost to forgo production, or even a policy variable like the degree of beaurocracy needed to claim a revision. Since the decision whether or not to claim a revision is taken in the last stage of the game, any mechanism must assign at least $(1-\tau)(1-\gamma)\pi(n^*(z), z)$ to the entrepreneur. In equilibrium, no one claims her actual income is lower than her presumptive income, so we can ignore any cost the IRS would face to uncover true income.

To proceed with the analysis, we make an additional assumption on the production technology that eases the proof of Proposition 5.

**Assumption 6.** $f(n) = n^\alpha$, $0 < \alpha < 1$.

Under this assumption, $n^*(z) = (\alpha p z)^{\frac{1}{1-\alpha}}$, and $\pi(n^*(z), z) = [\alpha^{\frac{\alpha}{1-\alpha}} - \alpha^{\frac{1}{1-\alpha}}](pz)^{\frac{1}{1-\alpha}}$. An additional restriction on the distribution function enables us to characterize a closed-form solution.

**Assumption 7.** $g(z) + \left[1 - [(1-\tau)(1-\gamma)]^{\frac{1-\alpha}{\alpha}}\right](g(z) + zg'(z)) \geq 0$, *for all* $z \in [\underline{z}, \max\{\underline{z}, z^r\})$,

where $z^r$ is defined in Proposition 5.

This assumption is not readily interpreted. Some tedious algebra shows that if $\frac{G(z)-1}{g(z)}$ is strictly increasing (which is assumed for the rest of the paper), Assumption 7 is satisfied for a high enough lower bound $\underline{z}$ of the type space $Z$. However, some distributions exhibit this property without any restrictions in the support (e.g., a uniform).[33]

In contrast with the previous section, the firm reservation value is type-dependent. Using Lewis and Sappington [1989] terminology, a type-dependent reservation value generates countervailing incentives. On the one hand, a firm has incentives to hire less labor and pay less taxes. On the other hand, since the reservation value, $(1-\tau)(1-\gamma)\pi(n^*(z),z)$, is increasing in $z$, an entrepreneur is also tempted to overstate $z$ by hiring more labor. Therefore, the solution is not necessarily characterized by the full informational rent extraction of the lowest type, and standard tricks in the literature are not readily applicable here.[34]

Let $U(z) = \pi(n(z),z) - \tau P(z) - (1-\tau)(1-\gamma)\pi(n^*(z),z)$ be the entrepreneurs' informational rent. Here, standard sorting conditions hold, so the (IC) is equivalent to the local incentive compatibility constraint (LIC) $\pi_n(n(z),z)\frac{dn}{dz}(z) - \tau\frac{dP}{dz}(z) = 0$ (or $\frac{dU}{dz}(z) = pf(n(z)) - (1-\tau)(1-\gamma)pf(n^*(z))$, applying the envelope theorem[35]), and a monotonicity (M) condition $\frac{dn}{dz}(z) \geq 0$, for all $z \in Z$. The condition (M) is innocuous, as we show ignoring this constraint from the problem, and verifying it is not violated.

We cast this problem into an optimal control problem, in which expected rents $U$ are the state variable and labor $n$ is the control. Hence, the IRS problem is to

$$\max_{n(\cdot),U(\cdot)} \int_{\underline{z}}^{z^n} [\pi(n(z),z) - U(z)]dG(z) + \Theta$$

$$s.t.$$

$$(IR) \quad U(z) \geq 0, \forall z \in Z$$

$$(LIC) \quad \frac{dU}{dz}(z) = pf(n(z)) - (1-\tau)(1-\gamma)pf(n^*(z)), \text{ a.e.}$$

where $\Theta = \int_{\underline{z}}^{\overline{z}}(1-\tau)(1-\gamma)\pi(n^*(z),z)dG(z)$.

---

[33]If $g'(\cdot)$ is bounded below, by continuity, this assumption holds for sufficiently small combination of $\tau$ and $\gamma$. The same argument holds for a sufficiently high $\alpha$.

[34]See Maggi and Rodríguez-Clare [1995] and Jullien [2000] for a comprehensive treatment of type-dependent individual rationally constraints. In Jullien [2000], the principal can also exclude some of the agents from the game, however in our environment, this is not a possibility since the IRS does not control the labor supply.

[35] This equation provides a clear interpretation of countervailing incentives. It specifies how the slope of the information rents, $U(z)$, should change with $z$, in order to not violate incentive compatibility. The first term captures the incentives to understate $z$, and declare less profits. The second term captures the temptation to overstate $z$, and increase the reservation value.

This problem is similar to the one analyzed by Maggi and Rodríguez-Clare [1995] in another context. In the Appendix A, we follow the approach in their paper to prove the following proposition. Let the superscript $r$ denote the optimal strategies.

**Proposition 5.** *If* $\max\{\underline{z}, z^r\} + \frac{G(\max\{\underline{z}, z^r\})-1}{g(\max\{\underline{z}, z^r\})} \geq 0$ *then*

$$
n^r(z) = \begin{cases} \left\{ p\alpha \left[ z + \frac{G(z)-1}{g(z)} \right] \right\}^{\frac{1}{1-\alpha}} & \text{if } z \geq \max\{\underline{z}, z^r\} \\ [(1-\tau)(1-\gamma)]^{\frac{1}{\alpha}} (\alpha p z)^{\frac{1}{1-\alpha}} & \text{if } z < \max\{\underline{z}, z^r\} \end{cases}
$$

$$
U^r(z) = \begin{cases} p \int_{\underline{z}}^{z} [f(n^r(s)) - (1-\tau)(1-\gamma)f(n^*(s))]ds & \text{if } z \geq \max\{\underline{z}, z^r\} \\ 0 & \text{if } z < \max\{\underline{z}, z^r\} \end{cases}
$$

$$
P^r(z) = \frac{1}{\tau} \left[ \pi(n^r(z), z) - (1-\tau)(1-\gamma)\pi(n^*(z), z) - U^r(z) \right],
$$

*where* $z^r \in [0, \bar{z}]$ *is the unique root that solves the following equation in* $s$

$$
\frac{1 - G(s)}{g(s)} - \left\{ 1 - [(1-\tau)(1-\gamma)]^{\frac{1-\alpha}{\alpha}} \right\} s = 0
$$

*Moreover, this solution is continuous. Finally, (M) is satisfied.*

*Proof.* See Appendix A ☐

Intuitively, to overstate her reservation utility, which is strictly increasing and convex, the agent has an incentive to report a higher $z$. On the other hand, she also has an incentive to understate $z$ in order to pay less taxes. For high types, $z \geq \max\{\underline{z}, z^r\}$, the second effect dominates and we have the "no distortion at the top" result with the level of production away from the full-information solution. These types get a positive surplus. However, for $z < \max\{\underline{z}, z^r\}$, the second effect is not so strong, and the set of menus is designed in a way that respects (IC) and (IR). The principal appropriates the informational rent of these types. Notice the threshold $z = z^r$ is the highest type for which $\frac{dU^r}{dz}(z) = 0$, thus $n^r(z)$ is adjusted downward to set $\frac{dU^r}{dz}(z) = 0$ for all $z < \max\{\underline{z}, z^r\}$, otherwise lower types would overstate $z$. Finally, a little algebra shows that transfers are positive for all types.

# 6   Presumptive vs. modern tax methods

As Tanzi and de Jantscher [1989] point out, in general, rebuttable presumptions are the rule. Hence, in this section, we compare the RPTS with the general model described in Section

2, which we refer to as the modern tax system (MTS).[36]

As a benchmark, we assume that presumptions of income can be rebutted at no cost, that is, $\gamma = 0$. If, for example, $\gamma$ is a policy variable, then it can be determined by non-economic factors. In particular, the government can follow a principle of justice that protects individuals from the burden caused by a particular authority's mistake.

If the budget $C$ is assigned optimally to the IRS, then the net revenue collected in the RPTS is a lower bound on the net revenue collection in the MTS. Indeed, by imposing $C = 0$ and $\phi(z) = 0$ for all $z \in Z$ into the general model outlined in Section 2, the resulting optimization problem is equivalent to the one solved in Section 5.2, which deals with the RPTS. Since assigning $C = 0$ and $\phi(z) = 0$ for all $z \in Z$ are feasible choices, the MTS must generate at least the same net revenue as the RPTS with $\gamma = 0$.

Intuitively, if $\phi(z) = 0$ for all $z \in Z$ in the MTS, the IRS can always shape reported income $x$ to mimic the presumptive tax base $P$, obtained in the RPTS. This is achieved through high enough off-equilibrium auditing intensities that deter off-schedule deviations. In the MTS, since the IRS can additionally audit entrepreneurs to cope with taxes, it may furthermore increase revenue collection.

Governments are also concerned with the overall efficiency of the chosen tax system. In the RPTS, agents distort labor to pay less taxes, so efficient labor employment can be recovered only through lump sum taxes. Similarly, in the MTS, an entrepreneur distorts labor in order to be monitored with less intensity, so auditing probabilities that depend only on reported income is enough to recover efficiency (see Section 3.1). If policy instruments do not depend on labor and $\gamma = 0$, then a similar argument to the one in the previous paragraph proves that the net revenue collected in the RPTS is again a lower bound on the net revenue collection in the MTS. Furthermore, even if the policy instrument is conditioned on labor, the MTS can achieve the efficient level of employment once a large enough budget $C$ is assigned to the IRS.[37] Hence, the MTS might also be preferred for efficiency.[38]

Consequently, if $\gamma$ is small enough, as the technological cost needed to implement the MTS becomes relatively cheaper, the MTS tends to be adopted. In the 20th century, for instance, the technology that allows the IRS to accurately track and monitor the taxpayers in a timely fashion became available and affordable. It is not surprising that, during this period, most countries adopted a variant of the MTS as a source of revenue collection.[39]

---

[36]Recall that in the general model, the monitoring strategy depends on both reported income and labor input. Also, taxpayers are subject to audits and penalties, if caught underreporting income.

[37]Obviously, a high enough budget might not maximize the net revenue collection.

[38]These results should be read with caution. General equilibrium effects, through prices or the distribution of firms, can change the conclusions. An entrepreneurial choice model à la Lucas [1978] seems the appropriate framework to further investigate these questions.

[39]See Wallis [2000] for a discussion on the transition from a property-tax based system to an income-based

However, although the cost associated with this technology plummeted over the last century, presumptive tax methods are still widely in use, especially to collect revenue from a specific subset of taxpayers.[40] Thuronyi [2004] highlights a few reasons. First, in modern tax systems, taxpayers with low income might face a high compliance burden, as the IRS might face a high administrative burden to monitor such taxpayers. Second, as long as it is easier to hide income than the indicators used to presume income, tax evasion is less pervasive in presumptive tax systems. Third, presumptive methods may lead to a more equitable taxation when noncompliance and administrative corruption are widely spread. Fourth, rebuttable presumptions encourage a taxpayer to carry proper accounts, once it facilitates the proof of an eventual claim that her income was wrongly presumed. Finally, some presumptive methods may lead to a more efficient outcome.[41]

The mechanism design approach presented here provides another insight on the reason presumptive methods are still around, especially in the underdeveloped and developing world. Assume that $\gamma$, the share of post-tax profits lost due to revising someone's presumptive income, is high enough. Then, it is possible to construct examples showing that the RPTS generates more net revenue than the MTS. This is trivially true in a society that cannot allocate budget to the IRS. Notice $\gamma$ tend to be prohibitively high in countries plagued with corruption, unsecured property rights, bad governance, overwhelming bureaucracy, unequal access to courts, and so on. From a net revenue collection perspective, presumptive methods may be optimal in these countries.

A recent literature has emphasized the role of technological and political factors in explaining fiscal capacity. Acemoglu [2005] and Besley and Persson [2009], for example, model fiscal capacity simply as a cap on the tax rate.[42] Under our analysis, one can think of fiscal capacity as the maximum revenue the IRS is able to raise, given a fixed tax rate. Hence, if $\gamma$ is low enough, the transition from the RPTS to the MTS is also a way of building fiscal

---

tax system in the U.S.

[40]See Bird and Wallace [2004] for a list of countries that rely on a wide range of presumptive methods.

[41]Since we do not solve the general model outlined in Section 2, it is not clear whether the presumptive tax system we analyze generates less output than the MTS, although this is trivially true if the budget is high enough.

[42]Acemoglu [2005] develops a model, in which a self-interested ruler can divert taxes or invest in public goods, an action that increases future tax revenues. If the state is too weak and the ruler is not able to collect future revenues, he underinvests in public goods. On the other hand, if she rules a strong state, high taxes will be imposed, discouraging private investment. Therefore, power must be balanced in order to encourage all parties to undertake investments. The optimal cap on the tax rate arises as way to balance power between the ruler and the citizens. Besley and Persson [2009], for instance, model fiscal capacity as *ex-ante* costly investments under uncertainty. In their two-period model with two groups facing conflicting interests, common interest public goods, such as the threat of war, are conductive to building fiscal capacity. See also Acemoglu et al. [forthcoming] and Besley and Persson [2010], as though the political science view on these issues, e.g., Levi [1988], Migdal [1988] or Tilly [1990].

capacity. If the resources needed to implement the MTS cannot be raised in the RPTS, the society might find itself trapped in an equilibrium with low fiscal capacity, even if this transition is politically feasible. In the RPTS, net revenue collection depends positively on $\tau$ and $\gamma$. In particular, it is maximized at $\tau = 1$, or $\gamma = 1$. Hence, to escape the trap and move towards a more modern tax system, bad institutions, like those mentioned above, might be designed.[43]

# 7    Conclusion

This paper develops a framework to study optimal tax enforcement when employment size (or any other choice variable) acts as a signal of productivity. In addition to reported income, tax authorities condition auditing probabilities on this variable, incorporating optimal behavior of firms into their design. In turn, taking as given the monitoring strategy, firms employ workers and evade taxes optimally.
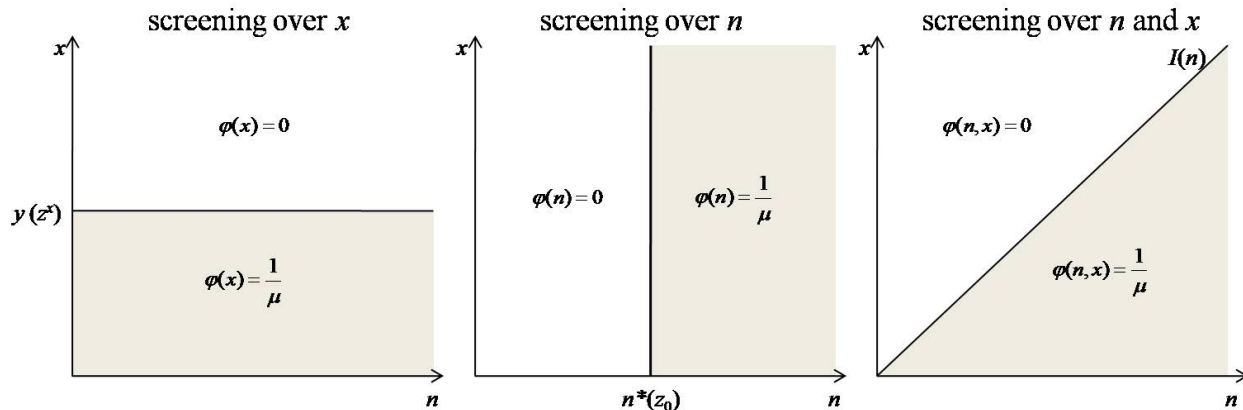


Figure 3: summary

Figure 3 summarizes how optimal allocations can be implemented. The first and second panels describe the cases studied in Section 3, when the IRS designs monitoring strategies conditioned only on reported income $x$ or employment $n$, respectively. While the first case has been previously studied in the literature, the second is a contribution of this paper. We find that some facts about the informal sector are only consistent with optimal auditing for the latter case.

In Section 4, we show that any optimal allocation can be implemented indirectly in the way it is summarized in the third panel of Figure 3. Note it combines features from the two extreme cases in a way that resembles IRS actual practices. However, we provide only a

---

[43]See Robinson [1998] for more theories of "bad policy."

partial characterization for the curve depicted in the third panel. The combination of a non quasi-linear objective, multidimensional choice, and a type-dependent outside option make the general problem hard to solve analytically.

If the IRS has enough budget to monitor every entrepreneur with intensity $\frac{1}{\mu}$, the full-information revenue is recovered. Recall that in the full-information mechanism, every entrepreneur produces efficiently and reports her true income. If a type $z$ is monitored with intensity $\phi(z) = \frac{1}{\mu}$, expected profits are $(1 - \tau)\pi(n, z)$. (IR) implies that the optimal labor input is the efficient one, $n^*(z)$, and then income is reported truthfully.

Conversely, assume the IRS has no budget, $C = 0$, and then is restricted to monitor every firm with intensity zero in equilibrium. Under Assumptions 6 and 7, the optimal mechanism is the one described in Proposition 5 with $\gamma = 0$. It induces the "no distortion at the top" result, labor input is distorted away from its efficient level, and the principal appropriates the informational rent of the smallest types. To see this notice that if $\phi(z) = 0$ for a given type $z$, her expected profits are $\pi(n, z) - \tau x$. Thus is enough to let $x$ play the role of $P$ and set $\gamma = 0$ in Section 5.2.

When the budget is greater than zero, but not high enough, the solution might combine the use of both random auditing and labor distortions. In the Appendix, we solve a particular problem, in which the IRS exhausts its budget monitoring high-ability entrepreneurs. A closer inspection of this solution suggests that it might be particularly hard to compute how auditing intensities should be distributed across types in a way that respects incentive compatibility and individual rationality. In particular, the optimal monitoring strategy might be discontinuous (as in both polar cases), inducing jumps in the optimal mechanism. We leave the solution to the general model described in Section 2, as an open question for further research.

# Appendix

## A - Proof of Proposition 5

Let $\lambda$ be the costate variable associated with the state $U$, then the Hamiltonian is

$$H(U, n, \lambda, z) = [\pi(n, z) - U]g(z) + \lambda[pf(n) - (1 - \tau)(1 - \gamma)pf(n^*(z))],$$

and the Lagrangean is

$$L = H + \theta U,$$

where $\theta$ is the multiplier of the (IR) constraint.

Like in Maggi and Rodríguez-Clare [1995], the idea is to follow a guess-and-verify method. We look for a solution with a continuous costate variable, which exists. Following Seierstad and Sydsæter [1987], since $H(U, n, \lambda, z)$ is concave in $U$, below are the set of sufficient conditions for an optimum.

1. $\left[z + \frac{\lambda(z)}{g(z)}\right] pf'(n^r(z)) = 1$;

2. $\left[z + \frac{\lambda(z)}{g(z)}\right] pf''(n) < 0$, for all $n \geq 0$ and $z \in Z$;

3. $\frac{d\lambda(z)}{dz} = -\frac{\partial L}{\partial U} = g(z) - \theta(z)$;

4. $\frac{dU^r(z)}{dz} = pf(n^r(z)) - (1 - \tau)(1 - \gamma)pf(n^*(z))$;

5. $\theta(z)U^r(z) = 0$; $\theta(z) \geq 0$; $U^r(z) \geq 0$;

6. $\lambda(\underline{z})U^r(\underline{z}) = 0$; $\lambda(\underline{z}) \leq 0$; $\lambda(\overline{z})U^r(\overline{z}) = 0$; $\lambda(\overline{z}) \geq 0$.

1 and 2 guarantee that the Hamiltonian is globally maximized in $n^r(z)$. 3 and 4 are the costate and state equations, respectively. 5 accounts for the (IR). Finally, 6 states the transversality conditions.

The key step is to set an appropriate value for $\lambda$, and then verify that our solution satisfies 1 through 6. We set

$$\lambda(z) = \begin{cases} G(z) - 1 & \text{if } z \geq \max\{\underline{z}, z^r\} \\ \left[[(1 - \tau)(1 - \lambda)]^{\frac{1-\alpha}{\alpha}} - 1\right] zg(z) & \text{if } z < \max\{\underline{z}, z^r\} \end{cases}.$$

Condition 1 implies that

$$n^r(z) = \begin{cases} \left\{p\alpha\left[z + \frac{G(z)-1}{g(z)}\right]\right\}^{\frac{1}{1-\alpha}} & \text{if } z \geq \max\{\underline{z}, z^r\} \\ [(1 - \tau)(1 - \gamma)]^{\frac{1}{\alpha}}(\alpha pz)^{\frac{1}{1-\alpha}} & \text{if } z < \max\{\underline{z}, z^r\} \end{cases}.$$

Notice that once 1 is satisfied, also 2 is, so the Hamiltonian is globally maximized at $n^r(\cdot)$. From 3,

$$\theta(z) = \begin{cases} 0 & \text{if } z \geq \max\{\underline{z}, z^r\} \\ g(z) + \left[1 - (1-\tau)^{\frac{1-\alpha}{\alpha}}\right](g(z) + zg'(z)) & \text{if } z < \max\{\underline{z}, z^r\} \end{cases}.$$

Hence, Assumption 7 implies $\theta(\cdot) \geq 0$.

From 4,

$$\frac{dU^r}{dz}(z) = \begin{cases} \frac{1-G(z)}{g(z)} - \left\{1 - [(1-\tau)(1-\gamma)]^{\frac{1-\alpha}{\alpha}}\right\}z & \text{if } z \geq \max\{\underline{z}, z^r\} \\ 0 & \text{if } z < \max\{\underline{z}, z^r\} \end{cases}.$$

From 6, $U(\underline{z}) = 0$ and $U(\overline{z}) \geq 0$. Then, all we need is to verify 5. Indeed, solving the differential equation above,

$$U^r(z) = \begin{cases} p \int_{\underline{z}}^{z} [f(n^r(s)) - (1-\tau)(1-\gamma)f(n^*(s))]ds & \text{if } z \geq \max\{\underline{z}, z^r\} \\ 0 & \text{if } z < \max\{\underline{z}, z^r\} \end{cases},$$

which is greater than, or equal to, zero by construction. Also by construction, this solution is continuous.

Finally, given that the monotone hazard rate holds, (M) is trivially satisfied.

## B - The IRS monitors high-ability entrepreneurs

In this Appendix, we present the solution for a mechanism in which the IRS monitors high-ability entrepreneurs with probability $\frac{1}{\mu}$. Here, we rely on Assumptions 6 and 7, with $\gamma = 0$, and $z^n$ playing the role of $z^r$. Recall that $z^n$ is the threshold defined in Lemma 2.

Let the superscript $o$ denote the optimal strategies. Since the IRS monitors high-ability types with intensity $\frac{1}{\mu}$, their expected profits are $(1-\tau)\pi(n(z), z)$. Thus by (IR), $n^o(z) = n^*(z)$ and $x^o(z) = \pi(n^*(z), z)$ for all $z \geq z^n$.

Notice we can set $\phi(z) = 0$ for all $z < z^n$, ignore the budget constraint, and solve this problem within the low-ability type subspace. The IRS problem is to

$$\max_{n(\cdot),x(\cdot)} \int_{\underline{z}}^{z^n} \tau x(z) dG(z) + \int_{z^n}^{\bar{z}} \tau \pi(n^*(z), z) dG(z)$$

s.t.

$$(F) \quad 0 \le x(z); 0 \le n(z), \forall z < z^n$$

$$(IR) \quad \pi(n(z), z) - \tau x(z) \ge (1 - \tau)\pi(n^*(z), z), \forall z < z^n$$

$$(IC) \quad \pi(n(z), z) - \tau x(z) \ge \pi(n(\tilde{z}), z) - \tau x(\tilde{z}), \forall z, \tilde{z} < z^n$$

$$(IC') \quad \lim_{z \uparrow z^n} \pi(n(z), z) - \tau x(z) = \lim_{z \downarrow z^n} (1 - \tau)\pi(n^*(z), z).$$

This problem is similar to the one outlined in Section 5.2, except that transfers are bounded below by zero (entrepreneurs cannot declare negative profits), and there is a terminal condition (IC') that guarantees incentive compatibility at the threshold $z^n$.

The requirement that $x(z) \ge 0$, is innocuous. All firms declare a positive amount in equilibrium, as as it can be verified by solving this problem ignoring (F).

Let $U(z) = \pi(n(z), z) - \tau x(z) - (1 - \tau)\pi(n^*(z), z)$ be the entrepreneur's expected informational rent. Here, standard sorting conditions hold, so the (IC) is equivalent to a local incentive compatibility (LIC) $\frac{dU}{dz}(z) = pf(n(z)) - (1 - \tau)pf(n^*(z))$, and a monotonicity (M) condition $\frac{dn}{dz}(z) \ge 0$, for all $z \le z^n$. As (F), (M) is also innocuous, as it can be verified from the solution that ignores this constraint.

As in Section 5.2, we cast this problem into an optimal control problem, in which expected rents $U$ are the state variable and labor $n$ is the control.

$$\max_{n(z),U(z)} \int_{\underline{z}}^{z^n} [\pi(n(z), z) - U(z)] dG(z) + \Omega$$

s.t.

$$(IR) \quad U(z) \ge 0, \forall z < z^n$$

$$(LIC) \quad \frac{dU}{dz}(z) = pf(n(z)) - (1 - \tau)pf(n^*(z)), \text{ a.e.}$$

$$(IC') \quad U(z^n) = 0,$$

where $\Omega = \int_{\underline{z}}^{z^n} \pi(n^*(z), z) dG(z) + \tau \int_{\underline{z}}^{\bar{z}} \pi(n^*(z), z) dG(z)$.

Define $\hat{n}(z) = f^{-1}((1 - \tau)f(n^*(z))) = (1 - \tau)^{\frac{1}{\alpha}}(\alpha p z)^{\frac{1}{1-\alpha}}$.

**Proposition 6.** *An optimal mechanism when the IRS is constrained to monitor high-ability*

*entrepreneurs with intensity $\frac{1}{\mu}$ is*

$$(\phi^o(z), n^o(z), U^o(z)) = \begin{cases} (0, \hat{n}(z), 0) & \textit{if } \underline{z} \leq z < z^n \\ (\frac{1}{\mu}, n^*(z), 0) & \textit{if } \overline{z} \geq z \geq z^n \end{cases}.$$

*Also, (M) and (F) are satisfied.*

The proof follows the same steps used to prove Proposition 5, and is omitted. Since the final state, $U(z^n) = 0$, is fixed, the only change needed from the previous set of sufficient conditions for an optimum (in Appendix A) is the transversality condition. It is straightforward to verify that this solution satisfies these conditions.

In contrast with the results from Section 5.2, expected informational rent is zero for all types. The reason is that the extra constraint, (IC'), imposes a fixed terminal condition to the differential equation (LIC).

We use this solution as an example to illustrate the result in Section 4. We plot Figure 2 setting $z \sim U[1, 2]$, $\alpha = 2/3$, $p = 1$, $\tau = 0.2$, $\mu = 2$, and $C$ is enough to monitor 30% of the firms.

# References

D. Acemoglu. Politics and economics in weak and strong states. *Journal of Monetary Economics*, 52(7):1199–1226, October 2005. 31

D. Acemoglu, D. Ticchi, and A. Vindigni. Emergence and persistence of inefficient states states. *Journal of the European Economic Association*, forthcoming. 31

M. G. Allingham and A. Sandmo. Income tax evasion: A theoretical analysis. *Journal of Public Economics*, 1(3-4):323–338, November 1972. 2

P. S. Amaral and E. Quintin. A competitive model of the informal sector. *Journal of Monetary Economics*, 53(7):1541–1553, October 2006. 4

J. Andreoni, B. Erard, and J. Feinstein. Tax compliance. *Journal of Economic Literature*, 36(2):818–860, June 1998. 2, 24

A. R. Antunes and T. Cavalcanti. Start up costs, limited enforcement, and the hidden economy. *European Economic Review*, 51(1):203–224, January 2007. 4

G. Arachi and A. Santoro. Tax enforcement for smes: Lessons from the italian experience? *eJournal of Tax Research*, 5(2):225–243, 2007. 24

M. Bassetto and C. Phelan. Tax riots. *Review of Economic Studies*, 75(3):649–669, July 2008. 2

R. Bayer and F. Cowell. Tax compliance and firms' strategic interdependence. *Journal of Public Economics*, 93(11-12):1131–1143, December 2009. 13

G. S. Becker. Crime and punishment: An economic approach. *Journal of Political Economy*, 76(2):169–217, March/April 1968. 12

T. Besley and T. Persson. The origins of state capacity: Property rights, taxation and politics. *American Economic Review*, 99(4):1218–1244, September 2009. 31

T. Besley and T. Persson. State capacity, conflict and development. *Econometrica*, 78(1): 1–34, January 2010. 31

R. M. Bird and S. Wallace. Is it really so hard to tax the hard-to-tax? the context and role of presumptive taxes. In J. Alm, J. Marinez-Vazquez, and S. Wallace, editors, *Taxing the Hard-to-Tax: Lessons from Theory and Practice*, chapter 5, pages 121–158. Elsevier B.V., 2004. 5, 31

K. C. Border and J. Sobel. Samurai accountant: A theory of auditing and plunder. *Review of Economic Studies*, 54(4):525–540, October 1987. 2

P. Chander and L. L. Wilde. General characterization of optimal income tax enforcement. *Review of Economic Studies*, 65(1):165–183, January 1998. 2

K.-P. Chen and C. Y. C. Chu. Internal control vs. external manipulation: A model of corporate income tax evasion. *Rand Journal of Economics*, 36(1):151–164, Spring 2005. 13

J. P. Choi and M. Tum. Corruption and the shadow economy. *International Economic Review*, 46(3):817–836, August 2005. 4

H. Cremer and F. Gahvari. Tax evasion and the general income tax. *Journal of Public Economics*, 60(2):235–249, May 1996. 2

K. J. Crocker and J. Slemrod. Corporate tax evasion with agency costs. *Journal of Public Economics*, 89(9-10):1593–1610, September 2005. 13

E. Dabla-Norris, M. Gradstein, and G. Inchauste. What causes firms to hide output? the determinants of informality. *Journal of Development Economics*, 85(1-2):1–27, February 2008. 3, 4, 20

A. de Paula and J. A. Scheinkman. The informal sector: An equilibrium model and some empirical evidence from brazil. PIER Working Paper 09-044, 2009. 4

A. de Paula and J. A. Scheinkman. Value added taxes, chain effects and informality. *American Economic Journal: Macroeconomics*, forthcoming. 9

D. Dharmapala, J. Slemrod, and J. D. Wilson. Tax policy and the missing middle: Optimal tax remittance with firm-level administrative costs. Working Paper, 2009. 21

B. Fortin, N. Marceau, and L. Savard. Taxation, wage controls and the informal sector. *Journal of Public Economics*, 66(2):293–312, November 1997. 4

D. Fudenberg and J. Tirole. *Game Theory*. MIT Press, 1991. 26

R. Gordon and W. Li. Tax structures in developing countries: Many puzzles and a possible explanation. *Journal of Public Economics*, 93(7-8):855–866, August 2009. 9

B. Jullien. Participation constraints in adverse selection models. *Journal of Economic Theory*, 93(1):1–47, July 2000. 28

H. J. Kleven, C. T. Kreiner, and E. Saez. Why can modern governments tax so much? an agency model of firms as fiscal intermediaries. NBER Working Paper 15218, 2009. 9

H. J. Kleven, M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez. Unwilling or unable to cheat? evidence from a randomized tax audit experiment in denmark. NBER Working Paper 15769, 2010. 2, 10

W. Kopczuk and J. Slemrod. Putting firms into optimal tax theory. *American Economic Review Papers and Proceedings*, 96(2):130–134, May 2006. 9

R. la Porta and A. Shleifer. The unofficial economy and economic development. *Brookings Papers on Economic Activity*, pages 275–363, Fall 2008. 3, 20

M. Levi. *Of Rule and Revenue.* University of California Press, 1988. 31

T. R. Lewis and D. E. M. Sappington. Countervailing incentives in agency problems. *Journal of Economic Theory*, 49(2):294–313, December 1989. 28

K. D. Logue and J. Slemrod. Of coase, calabresi, and optimal tax liability. John M. Olin Center for Law & Economics Working Paper 96, 2009. 10

R. E. Lucas. On the size distribution of business firms. *Bell Journal of Economics*, 9(2): 508–523, Autumn 1978. 4, 20, 30

I. Macho-Stadler and J. D. Pérez-Castrillo. Auditing with signals. *Economica*, 69(273):1–20, February 2002. 4, 12, 20

G. Maggi and A. Rodríguez-Clare. On countervailing incentives. *Journal of Economic Theory*, 66(1):238–263, June 1995. 28, 29, 34

J. S. Migdal. *Strong Societies and Weak States: State-Society Relations and State Capabilities in the Third World.* Princeton University Press, 1988. 31

D. Mookherjee and I. Png. Optimal auditing, insurance, and redistribution. *Quarterly Journal of Economics*, 104(2):399–415, May 1989. 2

S. C. Parker. Entrepreneurial growth and dynamic tax evasion incentives. Working Paper, 2010. 2

A. M. Polinsky and S. Shavell. The economic theory of public enforcement of law. *Journal of Economic Literature*, 38(1):45–76, March 2000. 12

J. E. Rauch. Modeling the informal sector formally. *Journal of Development Economics*, 35 (1):33–48, January 1991. 4, 21

J. F. Reinganum and L. L. Wilde. Income tax compliance in a principal-agent framework. *Journal of Public Economics*, 26(1):1–18, February 1985. 2

J. A. Robinson. Theories of "bad policy". *Journal of Economic Policy Reform*, 2(1):1–46, February 1998. 32

H. S. Rosen. *Public Finance*. McGraw-Hill, 2005. 12

I. Sánchez and J. Sobel. Hierarchical design and enforcement income tax policies. *Journal of Public Economics*, 50(3):345–369, March 1993. 2, 3, 12, 14, 15

A. Sandmo. The theory of tax evasion: A retrospective view. *National Tax Journal*, 58(4): 643–663, December 2005. 2

A. Santoro. Taxpayer's choices under studi di settore: What do we know and how we can interpret it? *Giornale degli Economisti e Annali di Economia*, 67(2):161–184, July 2008. 24

F. Schneider and D. H. Enste. Shadow economies: Sizes, causes and consequences. *Journal of Economic Literature*, 38(1):77–114, March 2000. 20

S. Scotchmer. Audit classes and tax enforcement policy. *American Economic Review Papers and Proceedings*, 77(2):229–233, May 1987. 4, 12, 20

A. Seierstad and K. Sydsæter. *Optimal Control Theory with Economic Applications*. North-Holland, 1987. 34

J. Slemrod. Cheating ourselves: The economics of tax evasion. *Journal of Economic Perspectives*, 21(1):25–48, Winter 2007. 2, 10

J. Slemrod and S. Yitzhaki. Tax avoidance, evasion and administration. In A. J. Auerbach and M. Feldstein, editors, *Handbook of Public Economics*, volume 3, chapter 22, pages 1423–1470. Elsevier B.V., 2002. 2

V. Tanzi and M. C. de Jantscher. The use of imputed income in modern tax systems. In A. Chiancone and K. Messere, editors, *Changes in Revenue Structures*, pages 37–51. Wayne State University Press, 1989. 4, 11, 26, 29

V. Thuronyi. Presumptive taxation of the hard-to-tax. In J. Alm, J. Marinez-Vazquez, and S. Wallace, editors, *Taxing the Hard-to-Tax: Lessons from Theory and Practice*, chapter 4, pages 101–120. Elsevier B.V., 2004. 4, 12, 31

C. Tilly. *Coercion, Capital and European States, AD 990-1992*. Blackwell, 1990. 31

R. M. Townsend. Optimal contracts and competitive markets with costly state verification. *Journal of Economic Theory*, 21(2):265–293, October 1979. 2

J. R. Tybout. Manufacturing firms in developing countries: How well do they do and why? *Journal of Economic Literature*, 38(1):11–44, March 2000. 3, 20

J. J. Wallis. American government finace in the long run: 1790 to 1990. *Journal of Economic Perspectives*, 14(1):61–82, Winter 2000. 10, 30

H. C. Wilkenfeld. *Taxes and People in Israel*. Harvard University Press, 1973. 24

S. Yitzhaki. Cost-benefit analysis of presumptive taxation. *FinanzArchiv: Public Finance Analysis*, 63(3):311–326, September 2007. 24