Tax Auditing Without Commitment, With an Application to Labor Tax Evasion in Italy

Edoardo Di Porto, Universite' de Lille
Nicola Persico, New York University
Nicolas Sahuguet, HEC Montréal

March 2010.
PRELIMINARY: DO NOT CITE OR DISTRIBUTE

Abstract

We provide an identification result for a general class of auditing games. This result allows us to falsify the joint hypotheses that: (a) the auditor has no commitment powers, and (b) that the auditor has a specific objective function. We apply this identification result to a unique tax auditing dataset on the universe of labor-tax audits in Italy. These audits are carried out by the Italian social security agency INPS. We find that we cannot reject the joint hypotheses that: (a) INPS has a no-commitment strategy; and (b) INPS maximizes the success rate of audits, that is, the probability that the audited is found underreporting by any amount. We then develop a new strategic auditing model in which INPS auditors cannot commit to an auditing strategy, and they maximize the success rate of audits. The Nash equilibrium strategies for firms and auditors are solved for in closed form. Under standard assumptions about the distribution of firm sizes, the equilibrium can reproduce an additional finding from INPS data: namely, that the monetary value of tax evaded is increasing in the true tax base of the firm. Next we impose a parametric restriction on the distribution of (unobserved) tax bases of firms, and we compute the Nash equilibrium of our model in this parametric case. We calibrate the parameters of the Nash equilibrium outcomes to certain moments coming from the INPS dataset. In this way we are able to recover the unobserved parameters governing the distribution of true tax bases of firms, using the information coming from the (non-randomly) audited firms. Finally we use the parameters recovered via this methodology to perform the following counterfactual investigation: how much more revenue could INPS raise if it committed to a (carefully chosen) auditing strategy?

1 Introduction

This paper is concerned with auditing games. These are games in which an auditor allocates scarce auditing resources across a population of firms. The firms are required to report their taxable income, which the auditor can use to target its audits. Firms pay taxes based on their report. A firm which is audited and found cheating pays all the taxes it owes plus a penalty. An equilibrium in an auditing game is: a reporting strategy specifying how much each firm will report given their true tax base; and an auditing strategy specifying the probability that a firm with a certain report is audited. There is a large theoretical literature dealing with auditing games. The challenge in this literature is finding the Bayesian Nash equilibrium of the game.

But, before an equilibrium can be found, a key preliminary question that must be resolved is whether the auditor has commitment power. In other words, the question is whether the auditor behaves as a Stackelberg leader announcing his auditing strategy before firms choose how to report, or whether the auditor moves simultaneously with the firms. The overwhelming majority of the theoretical literature assumes commitment. The ability to commit makes a huge difference both to the equilibrium that prevails, and also to the tax revenue raised in equilibrium. Obviously, if the auditor can commit he can make at least as much revenue as in the no-commitment case. This is because the auditor can always commit to follow the no-commitment equilibrium strategy. In general, we should expect the auditor to make strictly more under commitment.

The specific empirical question addressed in this paper is whether a particular auditor, the Italian agency INPS, makes use of its power of commitment. In order to answer this question, the paper makes several contributions of a more theoretical nature. These contributions are listed below.

1. We provide an identification result for a general class of auditing games. This result allows us to falsify the joint hypotheses that (a) the auditor has no commitment powers, and (b) that the auditor has a specific objective function. For example, we provide a test which can reject the hypothesis that the auditor pursues the objective of reducing the amount of taxes evaded, without being able to publicly commit to any auditing strategy.

2. We apply this identification result to the INPS data set. We find that we cannot reject the joint hypotheses that: (a) INPS follows a no-commitment strategy; and (b) INPS maximizes the success rate of audits, that is, the probability that the audited is found underreporting by any amount.

3. We then develop a new strategic model of tax compliance and enforcement, the novelty being that the auditor maximizes the probability of a successful audit in the absence of commitment. These are, of course, exactly the joint hypotheses about INPS's behavior

that we could not reject. The equilibrium strategies for firms and auditor are solved for in closed form. Under standard assumptions about the distribution of firm size, the equilibrium can reproduce an additional finding from INPS data: namely, that the monetary value of tax evaded is increasing in the size of the firm.

4. We impose a parametric form (Power distribution) on the unobserved distribution of tax bases of firms, and we work out the equilibrium strategies for this parametric case. Under the assumption that the INPS data are generated by this equilibrium, we use the empirical distribution of (observed) reported tax bases of audited firms to back out the unobservable deep parameters of the model, such as the distribution of true tax bases of firms, the amount of cheating firms engage into, etc. In this way, we use the model to correct for the various selection biases that generate our sample.

5. Then, based on the recovered deep parameters, we engage in the counterfactual exercise of asking how much more revenue could INPS collect if, rather than following the equilibrium strategy, it kept the auditing budget fixed but allocated it in a way that made use of commitment and was mindful of deterrence. Specifically, we consider a "commitment strategy" in which only those who report below a threshold are audited, and they are audited with such a high probability that nobody who reports below a threshold cheats. We find that switching to a "commitment strategy" does increase tax revenue relative to the equilibrium of the no-commitment game. Quantitatively, however, the gain is small (in the order of 5%). This is partly because, in the equilibrium of the no-commitment game, our calibration suggests that INPS is already capturing more than 80% of the theoretical maximum revenue attainable, and so there is little room for improvement.

FAQ: How can we know that INPS's problem is really as described is part 2 above?

Why do we believe that the objective of INPS is detection of underreporters? We find that the success rate of INPS—whether an audit finds underreporting by *any* amount—is constant across audit classes. This observation is consistent with a no-commitment model in which the auditor maximizes successful audits. We develop such a model and characterize its equilibrium. We show that in this model the auditor "arbitrages away" any differences in success probability across audit classes, leading to the observed equalization in success rates, just as observed in the INPS data.

FAQ: OK, I believe that INPS's problem is as described is part 2. But is the resulting equilibrium behavior conceptually very different from the commitment case?

Our analysis suggests that the auditing strategy of INPS is not consistent with optimal *deterrence* of tax evasion. Rather, the strategy appears to be geared towards *detection* *of underreporters* (which, incidentally, is not an objective function which has received any attention in the optimal auditing literature). To see the conceptual difference between these two objective functions, consider an auditor facing a mass 1 of firm, each of which has a high income H with probability $1/2$, and a low income L with complementary probability. The auditor has enough resources to audit half the firms. Assume each firm will report truthfully if the probability of being audited exceeds $1/3$, and otherwise it reports L. Now, if the auditor commits to auditing all the firms who report L then it is an equilibrium for all firms to report truthfully. (Indeed, a firm with true income H would face probability 1 of being audited if it misreports.) This is therefore an optimal strategy for an auditor who aims to minimize tax evasion. Notice, however, that the success rate on audits is zero. This cannot please an auditor whose aim is to detect underreporters. Such an auditor should decrease the probability of auditing low reports to just under $1/3$, so as to engineer some underreporting in equilibrium. The general point is that an auditor who aims to detect underreporting may potentially engage in some pretty inefficient strategies, at least from the point of view of minimizing tax evasion.

FAQ: Why does INPS not have, or make use of, the ability to commit?

There are two plausible reasons, we think, which are not mutually exclusive. The first reason is that moving to a "commitment strategy" may not improve tax revenue all that much; indeed, this is what our counterfactual findings indicate. This finding might explain why the apparently inefficient behavior of INPS has survived: because maybe it is not so inefficient. The second reason is basically an agency problem inside INPS. If INPS is worried about incentivizing the inspectors in the field to exert effort, i.e., to work hard at inspecting, the sensible way to do this is to reward them for successful inspections. Let us suppose that INPS does that. The problem is that the commitment strategy requires that some reports with low probability of success are audited, whereas other reports which are more likely to result in a success are not audited. An auditor who follows such a strategy will be hurting his personal success rate,and thus his probability of promotion. This suggest that the dual goals of incentivizing effort and implementing the commitment strategy are sometimes in conflict.

1.1 Related literature

Measuring tax evasion however is very difficult. First, tax evasion is by nature concealed, and the data about compliance is not widely available. Second, compliance goes hand in hand with enforcement, and the observed compliance is the result of a game between tax-payers and enforcement agencies.

There are many theoretical models of auditing game.[1] The first model is Border and Sobel

[1] We define an auditing game as a game of incomplete in which firms choose how much of their income to report, and an auditor decides which firms to audit based partly on the reports. If the firm is not audited then the report determines the taxes paid. The report also determines a penalty, which is levied only if the

(1978), and many variants have followed. See Andreoni and Feinstein (1998) for a survey of work in this area. The main divide between these models goes through the assumption of commitment of the tax agency. Some models assume (in the contract theory tradition) that the tax agency can announce and commit to an audit policy that is known to taxpayers before they file their returns. Other models assume that the tax agency cannot commit to its audit policy and decides after the returns are file which taxpayer to audit. These two assumptions are reasonable but yield very different predictions about tax evasion and auditing results.[2] To our knowledge, our paper is the first to bring any of these theoretical models to data.

2 Tax Evasion and Auditing: A General Framework

We now present the framework we will use. The goal is to have a framework general enough to encompass virtually all of the different assumptions made in existing auditing models. This requires allowing for a very general objective function for the auditor and the firms, for different assumptions about commitment, and for other frictions. The "identification agenda" is the search for a methodology that would allow us to select among these different assumptions, within the context of specific empirical applications.

The players are an auditor and a mass of firms with measure 1.[3] The auditor classifies firms into different audit classes according to any number of firm characteristics which are observable to the auditor (sector, geographic location, etc.). For expositional convenience we start by describing the game with only one audit class, then extend it to the case of several audit classes.

Denote by x the true tax base, which is unobserved by the auditor until an audit is made. The auditor thinks the true tax base is distributed with density f. The function $\pi(x,r,p)$ denotes the auditor's expected payoff from auditing with probability p a firm which reports r and has true tax base x.

Example 1 *If the auditor maximizes the total returns (taxes paid plus revenue from audits) then* $\pi(x,r,p) = tr + p\,(\theta + t)\max(x - r, 0)$.

Maximization of total returns is the conventional assumption in the existing theoretical literature on strategic auditing. Much less common is the following alternative assumption.

Example 2 *If the auditor maximizes the success rate from audits then* $\pi(x,r,p) = pI_{(x-r)>0}$.

firm is audited and did not report honestly.

[2] Other modelling assumptions have not been settled in the litterature and vary from models to models. In particular, the exact objectives of the taxpayers and the audting agencies remain an open question.

[3] Our analysis also applies to the case of auditors and individuals filing income taxes.

Turning to the firms, the function $\kappa(x, r, p)$ represents the expected payoff of a firms with true tax base x who reports r and is audited with probability p. In practice, it can happen that the firm is subject to auditing in different arenas which may, to an extent, indirectly bring to light the same irregularities that our auditor is concerned with. For example, an Italian firm is not only subject to INPS audits, but also to income tax audits carried out by a different auditor.[4] Similarly, it could be that some of the auditor makes are not generated by the game we model here—for example, they could be compelled by law and thus not functional to maximizing π. In these cases we interpret the function κ as expressing the firm's incentives to misrepresent its income after taking into account all the other "extraneous" audits.

We denote by $r(x)$ the report of a firm with true tax base x, and by $p(r)$ the auditing probability chosen for a firm which reports r. B represents the budget constraint on audits. Formally, the equilibrium of the auditing game is defined by the following constrained maximization problem.

$$
\begin{aligned}
p^*(\cdot) \in \arg\max_{p(\cdot)} & \int_a^b \pi(x, r(x), p(r(x)))\, f(x)\, dx \\
\text{subject to: } & \int_a^b p(r(x))\, f(x)\, dx \leq B \\
\text{and, either } & r(x) \in \arg\max_r \kappa(x, r, p^*(r)) \qquad \text{(NOCOMM)} \\
\text{or } & r(x) \in \arg\max_r \kappa(x, r, p(r)). \qquad \text{(COMM)}
\end{aligned}
$$

The first equation represents the auditor's payoff, the second equation the budget constraint. The third and fourth equations are mutually exclusive; they represent the firm's problem under two alternative formulations. Equation (NOCOMM) represents the case in which the auditor cannot commit to an auditing schedule, whereas equation (COMM) represents the commitment case. The majority of the theoretical literature on strategic auditing proceeds under assumption (COMM), but Erard and Feinstein (1994) use (NOCOMM).

Introducing several audit classes The model above is sufficiently general to embed most of the theoretical models of strategic auditing, which generally abstract from the presence of auditing classes. For empirical purposes, however, it is important to allow for the presence of several audit classes in which the auditor classifies firms according to observable (to the auditor) characteristics. We will also have to worry that we, the researchers, may not be able fully to distinguish these audit classes. (More on this later.) Scotchmer (1994) is the first to point out the concerns raised by the presence of latent audit classes.

An audit class is simply a distinct group of firms with some distinguishable characteristics.

[4] In this case the Guardia di Finanza.

Let k index the set of all audit classes that are distinguishable by the auditor. Their relative frequency in the population is given by $G(k)$, with $\sum_k G(k) = 1$. Firms that belong to a given audit class are like all other firms in that they have a privately known taxable income and seek to minimize their tax bill. Also, they face the same penalties if found cheating. What makes them different in the eye of the auditor is the distribution of their taxable income, which the auditor uses to make inference. Conditional on being in class k, the tax base of firms is distributed according to the probability density $f_k(x)$. A firm from audit class k faces a class-specific audit schedule $p_k(\cdot)$.

Introducing inaccurate audits We allow for audits to produce imperfect signals of a firm's true tax base. Formally, we assume that the auditor does not observe a firm's x, but rather a number ξ which is correlated with x and that we call *detected income*. We assume that the auditor maximizes

$$\pi(\xi, r, p).$$

This allows for the possibility that the auditor might not detect underreporting (in which case $\xi \leq r$ even though $x > \xi$) or that the auditor may in fact mistakenly "overdetect" (and in this case $\xi > x$). We model ξ as the realization of a random variable Ξ_k with distribution $\upsilon_k(\xi|x, r)$. Note that we allow the distribution of ξ to depend on the audit class k. This dependence allows for the possibility that it might be more difficult to detect fraud in certain occupations (for example, industries that use part-time labor such as the restaurant industry, construction, agriculture).

In the presence of inaccurate audits, the firm's payoff is potentially a function of ξ, so we will write

$$\kappa(\xi, x, r, p).$$

The assumption of class-specific inaccuracies in audits is made by Macho-Stadler and Perez-Castrillo (1997).

Introducing residual heterogeneity of firms In the base model the only unobserved heterogeneity of firms is x, their true income. A (somewhat unrealistic) implication is that all firms with the same true income report in the same way. We can relax this assumption by simply assuming that the firm's unobserved characteristics are expressed by an N-dimensional vector $\mathbf{x} = (x_1, ..., x_N)$ with density $f_k(\mathbf{x})$. For ease of interpretation we assume that x_1, the first dimension of the vector, represents the tax base, while the other dimensions capture additional heterogeneity which impacts the firm's choice of reporting. The firm's payoff will then be given by $\kappa(\xi, \mathbf{x}, r, p)$, and the firm's equilibrium strategy by $r(\mathbf{x})$. The increased dimensionality permitted in this general formulation allows us to capture, as a special case, the case in which a fraction of the firms is honest and never underreports, while the rest is "normal" and behaves as in the baseline model. A model with these features is analyzed in Section 5.

The framework After introducing all these extensions, the equilibrium of the auditing game is defined by the following constrained maximization problem.

$$\{p_k^*(\cdot)\}_k \in \arg\max_{\{p_k(\cdot)\}_k} \sum_k G(k) \int E\left[\pi(\Xi_k, r_k(\mathbf{x}), p_k(r_k(\mathbf{x})))|\mathbf{x}, r_k(\mathbf{x})\right] \; f_k(\mathbf{x})\, d\mathbf{x}$$

$$\text{st: } \sum_k G(k) \int p_k(r_k(\mathbf{x}))\, f_k(\mathbf{x})\, d\mathbf{x} \leq B$$

$$r_k(\mathbf{x}) \in \arg\max_r E\left[\kappa(\Xi_k, \mathbf{x}, r, p_k^*(r))|\mathbf{x}, r\right] \text{ for each } k, \mathbf{x},$$

where the integrals are understood to be of multiple variables, over the N dimensions of $\mathbf{x}$. The last line captures the no-commitment case. To capture the commitment case it suffices to replace $p_k^*(r)$ with $p_k(r)$.

Assumptions We henceforth maintain the following two assumptions. Taken together, these assumptions characterize this hitherto abstract setting as an auditing game.

Assumption 1 ***(Deterrability)*** *For all k, if $p_k(r) = 1$ then no firm in class k with $x_1 > r$ reports r.*

Assumption 1 says that, if r is audited with sufficient frequency, then the firm's payoff function is such that no type will underreport r. This assumption means that every type can be deterred from misreporting, if the probability of auditing is sufficiently large. The next assumption says that the auditor's expected payoff from auditing someone who reports correctly (or even overreports) cannot be positive.

Assumption 2 ***(Unprofitability of auditing firms who report correctly)*** *For all $k, \mathbf{x}$ we have $E[\pi(\Xi_k, r, p)|\mathbf{x}, r] \leq 0$ when $r \geq x_1$.*

Even if ξ systematically "exagerates" relative to x_1, Assumption 2 can hold if there is a cost of auditing.

3 Identification

We would like to use the data to, inasmuch as possible, learn the nature of the game that auditor and firms play. We would like to know whether the auditor has the ability to commit to an auditing strategy (i.e., whether the (NOCOMM) or the (COMM) version of the game is being played). We would also like to know as much as possible about the objective functions that auditor and firms are maximizing, i.e., the functional forms π and κ. Our job as econometricians is to use the available data to answer these questions. We call this the *identification problem*.

Ideally, we would like the identification strategy to not depend on fine details of the problem (i.e., knowledge of, or assumptions about the distributions $G(k)$ and f_k, for example). We would, also, prefer not to have to solve for the full equilibrium of the game between auditor and firms, because that is often complicated. Most importantly, we would like our methods to be robust to unobservables, that is, we want to allow for the possibility that we, the researcher, may not know as much as the auditor and firms know when they set their strategies. This is an important robustness property, because we often lack access to the full data that the auditor can see. In this spirit of "informational parsimony," we proceed to lay out our assumptions as to what features of the data we can and cannot observe.

In this section we proceed, for expositional ease, as if r can only take integer values (dollars, or cents, in our auditing setup) and ξ also can only take integer values. [5]

What we cannot observe: latent audit classes We assume that we, the researcher, are only able to observe coarse partitions encompassing several auditing classes. We will denote these partitions by K_i. For example, the set of auditing classes observed by the auditor may be $k_1, ..., k_5$, but we, the researcher, are only able to ascertain whether a particular observation belongs to $K_1 = \{k_1, k_2, k_3\}$ or $K_2 = \{k_4, k_5\}$.

What we can observe: empirical averages We assume that we, the researcher, observe individual data on each audit. Audits are indexed by d. For each audit d we observe the reported income r_d, the detected income ξ_d, and what partition $K(d)$ the audited firm belongs to.

Take any function $h(\xi, r)$. Think of it provisionally as the return from auditing a firm who reports r and is found to have a tax base ξ. For each r and each K_i, we want to form the sample average of $h(\xi, r)$ conditional on r and on K_i, which is defined as follows. Let the set of all audits of firms who report r and belong to partition K_i be denoted by

$$D(r, K_i) = \{d : r_d = r, K(d) = K_i\}.$$

Then the average h conditional on r and on K_i is the statistic defined as

$$\overline{h}(r, K_i) = \frac{\sum_{d \in D(r,K_i)} h(\xi_d, r_d)}{\sum_{d \in D(r,K_i)} 1}, \quad (1)$$

and we set $\overline{h}(r, K_i) = 0$ when its denominator is zero. The quantity $\overline{h}(r, K_i)$ is to be

[5] Thus the probability $f_k(\cdot)$ must be understood as having a support that is countable, rather than a continuum.

interpreted as the average return, as computed from the data, from auditing a firm in partition K_i who reports income r. Our identification strategy will be based on studying the properties of $\overline{h}(r, K_i)$.

Of note, $\overline{h}(r, K_i)$ can be computed using solely informations about audits. It is not necessary to have information about the distributions $G(k)$ and f_k, nor even about the probability of being audited p_k. This parsimony is convenient because in order to form p_k, for example, it would be necessary to have information on the universe of all the firms, including those which are not audited. Such information is not necessary for our analysis. Nevertheless, the expected value of $\overline{h}(r, K_i)$ does depend on all these quantities in a way which we describe next.

We think of each point in our data as an i.i.d. realization of a random vector generated by the equilibrium behavior of firms and auditor. Thus, the probability that a random element of our sample $(\xi_d, r_d, K(d))$ is equal to (ξ, r, K_i) is given by

$$\sum_{\substack{k \in K_i \\ \mathbf{x} \in X_k^*(r)}} G(k) f_k(\mathbf{x}) p_k^*(r) v_k(\xi|\mathbf{x}, r), \tag{2}$$

where $p_k^*(r)$ represents the equilibrium probability that a firm in audit class k who reports r is audited, and $X_k^*(r)$ represents the set of $\mathbf{x}$'s which in equilibrium lead a firm in audit class k to report r. The term $G(k) f_k(\mathbf{x}) p_k^*(r)$ represents the probability that a firm belongs to audit class k and has a true tax base $\mathbf{x}$ which in equilibrium leads the firm to report r, and is audited. Using formula (2), the expected value of $\overline{h}(r; K_i)$ is given by

$$\frac{\sum_{\substack{k \in K_i \\ \mathbf{x} \in X_k^*(r)}} E\left[h(\Xi_k, r)|\mathbf{x}, r\right] \; G(k) f_k(\mathbf{x}) p_k^*(r)}{\sum_{\substack{k \in K_i \\ \mathbf{x} \in X_k^*(r)}} G(k) f_k(\mathbf{x}) p_k^*(r)} \tag{3}$$

This formula contains all the functions $G(k), f_k(\cdot), p_k^*(\cdot)$ about which we, the parsimonious researcher, prefer not to make assumptions. Expression (3) is the limit in probability of $\overline{h}(r, K_i)$ as the sample size grows large.

3.1 Identification without commitment

We are now ready to present our identification result. In this section we deal with the case in which the auditor can/does not publicly commit to an auditing schedule. The main result in this section deals with the case in which the auditor's payoff function is linear in the audit probabilities.

Assumption 3 *$\pi(\xi, r, p)$ is a linear affine function of p, that is,*

$$\pi(\xi, r, p) = A(\xi, r) + pC(\xi, r),$$

with $A(\xi, r) \geq 0$.

Assumption 3 holds in Examples 1 and 2 above, and therefore in most of the theoretical literature on strategic auditing. The term $C(\xi, r)$ represents the perceived return from audits, including any costs of auditing, whereas the term $A(\xi, r)$ can be interpreted as the contribution to the auditor's payoff of a firm who reports r and is not audited; typically, this would be the tax paid before the audit, so it makes sense to assume that it is nonnegative. We view Assumption 3 as not overly restrictive. In any case, this assumption can be tested with data, as discussed in the next Proposition.

The next proposition is our identification result. It says, roughly, that if we find some statistic of the data that is equalized across audit classes, then this statistic could well be part of what the auditor is maximizing, provided the auditor has no commitment. Intuitively, an auditor with no commitment will arbitrage his audits across audit classes, i.e., will direct his audits on the classes that promise the highest return from the audit—whatever that return might be. This arbitraging behavior leads, in an equilibrium where firms respond to auditing, to an equalization of the auditor's margins across all audited classes.

Proposition 1 *If one can reject the hypothesis that* $E\left[\overline{h}\left(r, K_i\right)\right]$ *is independent of* r *and* K_i*, then one can reject the joint hypotheses that (a) the auditor can/does not commit to an auditing schedule, and (b) Assumption 3 holds with* $C(\xi, r) = h(\xi, r)$*. Conversely, if a function* $h\left(\xi, r\right)$ *can be found such that one cannot reject the hypothesis that* $E\left[\overline{h}\left(r, K_i\right)\right]$ *is independent of* r *and* K_i*, then one cannot reject the hypotheses that (a) the auditor can/does not commit to an auditing schedule, and (b) Assumption 3 holds with* $C(\xi, r) = h(\xi, r)$*.*

Proof. The proof is made by showing that, if assumption (a) and (b) hold then $E\left[\overline{h}\left(r, K_i\right)\right]$ is independent of r and K_i.

By assumption (a) the auditor cannot commit to an auditing schedule p and so the equilibrium is characterized by the following conditions.

$$\begin{aligned} &\left\{p_k^*\left(\cdot\right)\right\}_k \in \arg\max_{\left\{p_k\left(\cdot\right)\right\}_k} \sum_k G\left(k\right) \int E\left[\pi(\Xi_k, r_k\left(\mathbf{x}\right), p_k\left(r_k\left(\mathbf{x}\right)\right))|\mathbf{x}, r_k\left(\mathbf{x}\right)\right] \; f_k\left(\mathbf{x}\right) d\mathbf{x} \\ &\text{st: } \sum_k G\left(k\right) \int_a^b p_k\left(r_k\left(\mathbf{x}\right)\right) f_k\left(\mathbf{x}\right) d\mathbf{x} \leq B \\ &\qquad r_k\left(\mathbf{x}\right) \in \arg\max_r E\left[\kappa(\Xi_k, \mathbf{x}, r, p_k^*\left(r\right))|\mathbf{x}, r\right] \text{ for each } k, \mathbf{x}. \end{aligned} \tag{4}$$

Let $r_k^*\left(\mathbf{x}; p_k^*\left(r\right)\right)$ denote the reporting strategy that solves (4). Since condition (4) involves $p^*\left(r\right)$, not $p\left(r\right)$, the behavior of firms is a function of the auditor's expected equilibrium strategy, not of the actual strategy employed by the auditor. We shall therefore write, for ease of notation, $r_k^*(\mathbf{x}; p_k^*(r)) = r_k^*(\mathbf{x})$. Form the Lagrangean for the auditor's problem:

$$\mathcal{L}(\{p_k(\cdot)\}_k; \lambda) = \sum_k G(k) \int E[\pi(\Xi_k, r_k^*(\mathbf{x}), p_k(r_k^*(\mathbf{x})))|\mathbf{x}, r_k^*(x)] \;\; f_k(\mathbf{x})\, d\mathbf{x}$$
$$- \lambda \left[\sum_k G(k) \int p_k(r_k^*(\mathbf{x}))\, f_k(\mathbf{x})\, d\mathbf{x} - B \right].$$

Use assumption 3 to substitute into the Lagrangean, which upon rearrangement reads

$$\sum_k G(k) \int \{E[C(\Xi_k, r_k^*(\mathbf{x}))|\mathbf{x}, r_k^*(\mathbf{x})] - \lambda\}\, p_k(r_k^*(\mathbf{x})) \;\; f_k(\mathbf{x})\, d\mathbf{x}$$
$$+ \sum_k G(k) \int E[A(\Xi_k, r_k^*(\mathbf{x}))|\mathbf{x}, r_k^*(\mathbf{x})]\, f_k(\mathbf{x})\, d\mathbf{x} + \lambda B.$$

The first term of the Lagrangean can be written as

$$\sum_k G(k) \sum_r \int_{X_k^*(r)} \{E[C(\Xi_k, r)|\mathbf{x}, r] - \lambda\}\, p_k(r) \;\; f_k(\mathbf{x})\, d\mathbf{x}$$
$$\sum_k G(k) \sum_r p_k(r) \left[\int_{X_k^*(r)} \{E[C(\Xi_k, r)|\mathbf{x}, r] - \lambda\} \;\; f_k(\mathbf{x})\, d\mathbf{x} \right]$$

As the Lagrangean is linear in each $p_k(\cdot)$, the necessary conditions for optimality of the auditor's strategy are that, if $p_k^*(r) > 0$ then $E[C(\Xi_k, r)|\mathbf{x} \in X_k^*(r), r] \geq \lambda$.

Now, suppose by contradiction that the strict inequality $E[C(\Xi_k, r)|\mathbf{x} \in X_k^*(r), r] > \lambda$ holds for some r. Then at the optimum it must be $p_k^*(r) = 1$. Because $p_k^*(r) = 1$ Assumptions 1 and 2 together imply that

$$E[A(\Xi_k, r)|\mathbf{x} \in X_k^*(r), r] + E[C(\Xi_k, r)|\mathbf{x} \in X_k^*(r), r] \leq 0 \tag{5}$$

for that r. But, since $E[C(\Xi_k, r)|\mathbf{x} \in X_k^*(r), r] > \lambda \geq 0$, and $E[A(\Xi_k, r)|\mathbf{x} \in X_k^*(r), r] \geq 0$ by Assumption 3, inequality (5) cannot hold. This contradiction proves that at the optimum it must be $E[C(\Xi_k, r)|\mathbf{x} \in X_k^*(r), r] = \lambda$ for all r such that $p_k^*(r) > 0$. We may rewrite this condition as

$$E[C(\Xi_k, r)|\mathbf{x} \in X_k^*(r), r] = \lambda \text{ for all } r \text{ such that } p_k^*(r) > 0. \tag{6}$$

Now, remember that from (3) we had

$$
\begin{aligned}
E\left[\overline{h}\left(r,K_i\right)\right] &= \frac{\sum_{k\in K_i} G\left(k\right) p_k^*\left(r\right) \sum_{\mathbf{x}\in X_k^*(r)} E\left[h(\Xi_k,r)|\mathbf{x},r\right]\ f_k\left(\mathbf{x}\right)}{\sum_{k\in K_i} G\left(k\right) p_k^*\left(r\right) \sum_{\mathbf{x}\in X_k^*(r)} f_k\left(\mathbf{x}\right)} \\
&= \frac{\sum_{k\in K_i} G\left(k\right) p_k^*\left(r\right) \left(\sum_{\mathbf{x}\in X_k^*(r)} f_k\left(\mathbf{x}\right)\right)\ E\left[h(\Xi_k,r)|\mathbf{x}\in X_k^*\left(r\right),r\right]}{\sum_{k\in K_i} G\left(k\right) p_k^*\left(r\right) \sum_{\mathbf{x}\in X_k^*(r)} f_k\left(\mathbf{x}\right)}
\end{aligned}
$$

From assumption (b) we know that $h(\xi,r)=C(\xi,r)$, and substituting into $E\left[\overline{h}\left(r,K_i\right)\right]$ we get

$$
E\left[\overline{h}\left(r,K_i\right)\right] = \frac{\sum_{k\in K_i} G\left(k\right) p_k^*\left(r\right) \left(\sum_{\mathbf{x}\in X_k^*(r)} f_k\left(\mathbf{x}\right)\right)\ E\left[C(\Xi_k,r)|\mathbf{x}\in X_k^*\left(r\right),r\right]}{\sum_{k\in K_i} G\left(k\right) p_k^*\left(r\right) \sum_{\mathbf{x}\in X_k^*(r)} f_k\left(\mathbf{x}\right)} = \lambda
$$

where the last equality makes use of (6). We have shown that, if hypotheses (a) and (b) hold then $E\left[\overline{h}\left(r,K_i\right)\right]$ is equal to a constant independent of r and K_i. ■

This proposition provides a straightforward identification strategy: if we suspect that we are in the no-commitment case, we can try out various "economically reasonable" functions $h(\xi,r)$ and check which, if any, has the property that it is equalized across all reports that are audited. If such a function is found, then this is identified as $C(\xi,r)$, and we cannot reject the hypothesis of lack of commitment.

This identification strategy is robust to details, in the sense that it is robust to the many frictions we have built into our model, and it is informationally parsimonious—it does not require us to know $G\left(k\right)$, $f_k\left(\cdot\right)$, or $p_k^*\left(\cdot\right)$, or even solve for the equilibrium behavior of firms.

The identification strategy can be extended to the case in which Assumption 3 need not hold—and so the payoff function is allowed to be nonlinear in the probability of auditing—but only at a heavy price in terms of additional assumptions. First, it is necessary to form $p\left(r\right)$, and therefore data on the universe of non-audited firms are needed. Second, it is no longer possible to accommodate the presence of latent auditing classes; we need to assume that we can observe auditing classes just as well as the auditor. Finally, and perhaps most onerous, we will require the existence of exogenous variation in the auditor's budget, not observed by the firms but observed by us. (Of course, it is perfectly all right if the firms are aware of the existence of this variation, provided they cannot observe the realization of it). Under these stringent conditions is is possible to test whether "economically reasonable" functions $h(\xi,r,p)$ can generate the data we observe. This result is presented in Appendix F.

3.2 Identification with commitment

When the auditor can/does publicly commit to an auditing schedule it is more difficult to give a "detail-free," characterization, even a partial one, of the equilibrium of the game between auditor and firms.

An important special case, which is well-studied case in the literature, is the case in which auditor and audited play a constant-sum game, so that $\pi(x, r, p(r)) = -\kappa(x, r, p(r))$, and in addition $\pi(x, r, p(r))$ is as in Example 1. In this case the solution of the auditing game has been characterized, see e.g. Sanchez and Sobel (1993). Generically, the optimal auditing policy is to set $p_k(r_k(x)) = \frac{t}{t+\theta}$ for those types x such that $\frac{1-F_k(x)}{f_k(x)}$ exceeds a threshold, and zero for all other types. Under this auditing policy, the probability of auditing any report r is so high (when it is positive) that no-one wants to cheat by reporting r. Instead, the cheaters in group k will choose to underreport at those levels r' which are not audited at all under p_k. In terms of testable identification, this model has some very sharp predictions: the testable implications are that all audits must be unsuccessful. Clearly, this prediction is preserved even in the presence of latent audit classes. So one might be inclined to think that, at least within this specific set of primitives, we have "robust" identification. However, Macho-Stadler and Perez-Castrillo (1997) show that this result is not robust to introducing class-specific inaccuracies in audits, of the kind that we have introduced through Ξ_k. In addition, it is clear from the proofs in Sanchez and Sobel (1993) that the assumption that $\pi(x, r, p(r)) = -\kappa(x, r, p(r))$ is essential to get the "zero success" result to hold in equilibrium. We read this body of literature as suggesting that a robust and informationally parsimonious identification strategy does not exist.

3.3 Summary of identification

The message from this section is that a robust, informationally parsimonious identification strategy is available which jointly tests for the assumption of no commitment and any specification of the auditor's objective function. It is based on checking what it is that the auditor is equalizing across audit classes. If we find an economically resonable objective h which is being equalized across audit classes then we cannot reject the hypothesis that the auditor has no commitment and that the return from an audit is given by h. It is noteworthy that this identification strategy is basically agnostic about the objective function of firms. This is convenient in that it is not necessary to make specific assumptions about the nature of the firms' decision problem, in order to get identification. It is a drawback, however, in that the identification analysis per se does not give us any information about what the firms might be maximizing.

If we are unsuccessful and do not find a reasonable h that is equalized across audit classes, then we need to consider the hypothesis that the auditor has commitment power. In this case we believe that a robust and informationally parsimonious identification strategy does not exist. This is because the whole body of theoretical literature on strategic auditing highlights the sensitivity of the equilibrium to different assumptions about the objective function of the firms, about the accuracy of audits, etc.

4 Empirical Analysis of INPS Data

In this section we apply this strategy to the INPS data. We check whether there is any objective h which is being equalized across audit classes. We will argue that the data do not allow us to reject the joint assumptions of no commitment *and* maximization of detection of underreporters. In a later section we will return to the question of whether these joint assumptions about INPS's motives and behavior are reasonable.

4.1 The environment

Our data comes from labor-tax auditing of Italian firms. In Italy it is the employers' responsibility to pay labor taxes on its employees. These taxes are analogous to Social Security contributions in the US, but they are higher (they hover around 40% of the worker's gross compensation).[6] Every year the Italian Social Security Institute (INPS) inspects a number of firms in order to verify that they paid their labor taxes. This is done by auditors who visit the firms' locations and check for violations. The auditor can interview the workers he finds and check administrative and accounting records. An employer found underreporting is assessed a fine equal to the money underreported plus 33% of it.[7]

Our dataset is composed of the universe of INPS audits in 2000-2005, except for two sectors: agriculture and self-employed workers.[8] This unique dataset was created in order to get some insight into labor tax evasion and undocumented work.[9] Each observation is an audit. For each audit, the data consists in some firm characteristics (number of declared workers, production sector, regional location) and some characteristic of the audit and its outcome (lenght of the time window that is the object of audit,[10] the amount of underreported taxes, the number of undeclared workers detected). In all, we have 474,645 inspections developed on 396,065 different firms, an average of around 80,000 per year.[11] Most of these firms (90%,

[6] For most workers these taxes amount to 40-42% of gross wages, but they are 38% for workers classidied as "artisans," and only 23% for specific types of workers who are not permanently employed. Our data does not distinguish among these various types of workers.

[7] To be more precise is 33% on annual base. WHAT DOES THAT MEAN?

[8] These two sectors are subjected to a separate auditing process on which we have no data.

[9] Describe what is in Di Porto, 2009 and add to references.

[10] Every audit examines only a specific time window, say, the two most recent years of activity. If a firm is audited twice, the window of the second audit cannot by law overlap with the first audit's.

[11] Since there are around 1,660,000 Italian firms, this means that INPS audits almost 5% of them every year.

or about 430,000) report 10 or fewer workers, reflecting the well-known prevalence of small firms in Italy.

To match the model to the data we need measures of what income the firms reported and of what evasion was found, if any. For evasion detected we will use two related variables. The variable *evasioni* is the amount of money that INPS assesses it is owed, if any. The variable *risultato* is a dummy created based on the previous variable, and it equals 1 if the audit resulted in an assessed fine in any amount. The dataset does not contain the reported income, but it contains the number of employees the firm reported having. We will use this as a proxy for reported income. The variable *settori* codes the ATECO industry sector codes to which the audited firm belongs.[12]

4.2 Selecting the sample, and summary statistics

In order to be consistent with the theoretical framework of optimal auditing, our sample should only contain audits which are discretionarily intiated by INPS with the goal of uncovering underreports. However, the administrative process that generates our data is multifaceted, and thus we need to decide what to do with "anomalous" audits that are not discretionarily intiated by INPS with the goal of uncovering underreporters. Our strategy will be to exclude them from the sample. It is important to note that we are not only erring on the side of caution; this strategy also has a theoretical justification, because eliminating these "anomalous" audits does not invalidate the analysis we intend to carry out. As mentioned on page 6 when we discussed the interpretation of κ, these excluded audits may influence the behavior of the firms, but that will not matter for our analysis: the impact of the extraneous audits folds into the definition of κ, and Proposition 1 holds regardless of their presence.

Our sample is determined as follows. First, we drop the roughly 171,000 observations in which firms are audited in a month in which they declare zero workers. These are not audits of self-employed workers, which as we mentioned do not appear in our data. Rather, these are firms which closed down (or went bankrupt) before the month in which they are audited, and who therefore report zero workers in the month in which they are audited. Unfortunately we do not know what number of workers they did report before they closed down, so even if we wanted to correlate the audit with their true report we could not do that. But, in fact, in many cases a post-bankruptcy audit is not an audit aimed at uncovering underreports of taxes, but rather part of a procedure aimed at recovering unpaid taxes (about which there is no uncertainy in INPS's records) out of the bankruptcy process. For both these reasons we eliminate observations where *dipendenti* equals zero.

Next we use the variable *origine* to screen out several types of interactions between INPS

[12] There are nine such sectors, with the numbers from 1 to 9 corresponding to, respectively: Energy, Water, Gas; Mining and Chemical; Manufacturing and Mechanical; Food and Textiles; Constructions; Commerce; Transportation; Credit and Insurance; and, finally, Services.

and the public which are not audits in the sense of our models. We keep in the sample only the roughly 175,000 audits which are coded as *controlli incrociati* and *mirate*. These are the audits that are discretionarily intiated by INPS with the goal of uncovering underreporters.

What is left out is, first, about 5,000 audits coded *fallimenti* which are initiated in connection with bankruptcy and which we eliminate for the same reason mentioned above—these are part of bankruptcy process and not true audits. Next, we have 27,000 interactions coded *scoperture* which are triggered when INPS detects a mismatch between the number of workers declared by the firm and the amount of taxes paid. This mismatch is not cheating in the sense that our models intend it: a firm who wanted to cheat would underreport both the number of workers and the taxes paid. Moreover, these audits are triggered automatically and they are not discretionary. So we eliminate them from the sample. A third type of anomalous audit is the almost 79,000 *segnalazioni*, "whistleblower audits" initiated following a complaint, typically by an alleged employee who claims that they were not declared to the tax authority—in other words, that the firm underreported its employee count. These audits are (a) not discretionary, because INPS is required by law to follow up; and (b) they are based on a piece of information (the whistleblower) which is not contemplated in auditing models, including ours. Therefore, we eliminate whistleblower audits for our sample.

After eliminating the audits mentioned above we are left with 176,230 discretionary audits which are initiated by INPS based only on documentary information about the firm. These audits are allocated following a strategy devised by the top management at INPS. The strategic guidelines, which are updated throughout the year, direct auditors in a given region to focus on specific types of businesses, such as truckers, or ice-cream parlors, etc. These discretionary audits correspond to the auditing activity contemplated in the auditing models. Therefore, we will restrict attention to these audits. To the extent that the auditing strategy is centrally designed, our model with a single auditor fits well the institutional environment.[13] There is no explicit statement, however, about INPS's objective function. Therefore, it is left to us to infer it empirically from the data.

Table 1 reports the summary statistics. We divide the 176, 230 audits into audits of small firms, which we define as firms which declare 10 or fewer employees, and audits of large firms. Small firms represent a very large fraction of all Italian firms (and roughly 90% of our entire sample). For 175,991 of these firms we know their sector (the remaining 239 are missing the sector variable). Among these firms we have 151,806 small firms and 24,185 large firms. Audits of small and large firms differ in the industry composition, as one would expect, with small firms being a larger fraction of the audited population in certain sectors.

[13]While on paper the auditing strategy is decided centrally, the reader may wonder about the incentives of individual auditors to potentially subvert the centrally-decided strategy. Individual auditors are compensated on a fixed wage plus a "productivity premium" based on the amount of unpaid taxes recovered in their region. We view these incentives as rather low-powered for two reasons. First, an individual auditor has a negligible effect on how much is recovered in his entire region. Second, as a practical matter the auditor's union has always resisted the notion that the productivity premium might be withheld. Perhaps as a result, the productivity premium has historically never been denied to any region.

	Small Firms	Large Firms
EnGasAcq	453 (0.3%)	83 (0.3%)
IndEstrChim	1,600 (1%)	656 (2,7%)
ManMetmecc	8,563 (5.7%)	3,303 (13.7%)
IndAliTess	16,807 (11.1%)	4,419 (18.3%)
Const	35,427 (23.3%)	6,910 (28.6%)
ComPubbEs	69,385 (45.7%)	5,087 (21%)
Trasp	1,587 (1%)	775 (3.2%)
CredAss	5,927 (3.9%)	1,684 (6.7%)
Serv	12,057 (7.9%)	1,268 (5.2%)
Total	151,806 (100%)	24,185 (100%)
Risultato (std. dev.)	0.40 (0.49)	0.54 (0.50)
Evasioni (std. dev.)	6,953 (35,950)	57,601 (306,704)
Evasioni conditional on evasioni > 0 (std. dev.)	17,683 (55,650)	108,297 (413,972)
Dipendenti (std. dev.)	3.09 (2.35)	51.13 (346.32)

Table 1: Summary Statistics. Large firms are those with more than 10 reported employees.

The probability of a successful audit is smaller for small firms: among all 176,230 audits, 40% of the small firms audit result in a fine being paid, and 54% of large firms audits. When evasion is measured by the amount of the fine, there is more evasion detected in the large firms sample. Moreover, when an evasion is detected, the fine paid averages 17,683 euros for small firm audits and 108,297 euros for large firms audits.

The distribution of reported sizes is given in Figure 1.

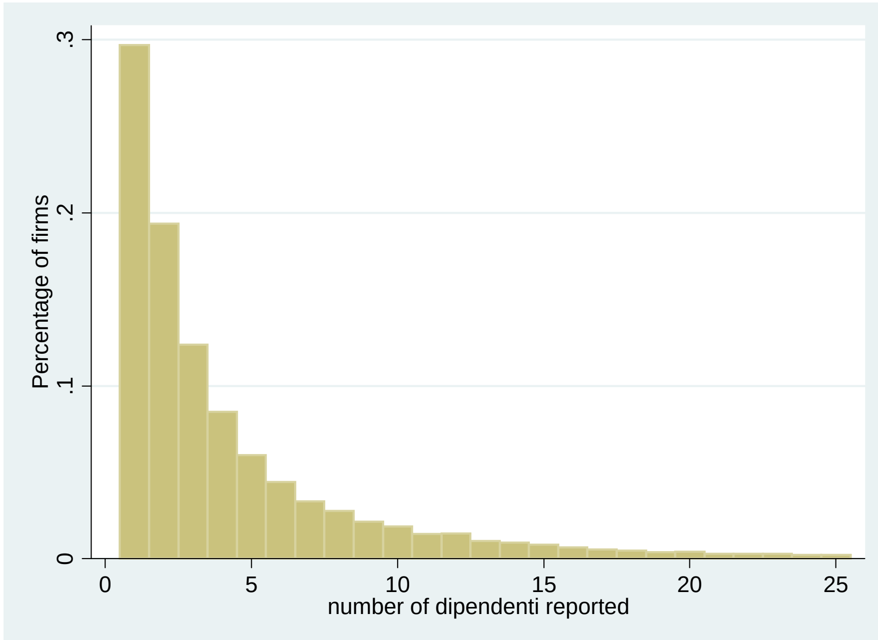

Figure 1: Distribution of reported number of employees.

4.3 Results for small firms

The dependent variable in the regression of Table 2, *risultato,* is a dummy which is 1 if the audit resulted in a fine of any amount. According to our test, if the firm maximizes the probability of detecting evasion, then the probability of detection should be the same for any reported number of *dipendenti* and for any sector (the variables *settori#*). In our empirical specification we allow for additional flexibility by controlling for the interaction between dipendenti and industry codes (the variables *sett#dip*). This allows for the probability of detection to vary with dipendenti in a sector-specific way. Despite this flexibility, few of the coefficients (4 out of 20) are significantly different from zero at the 10% level. We interpret this widespread lack of explanatory power as evidence that none of the independent variables in our regression help improve the probability of detection. One must not overemphasize this interpretation, because the F statistic indicates joint significance of the independent variables. Nevertheless, Table 2 does point to the difficulty for an auditor of predicting the probability of the success of an audit based on the variables we have, so that any audit is perceived as “equally likely to succeed” by INPS. This is consistent with the assumption that INPS maximizes the probability of a successful audit, and that INPS does not have, or does not make strategic use of, the power to commit.

The next regression, Table 3, is identical to the first except that the dependent variable is *evasioni*, the amount of money that INPS assesses it is owed. According to our theory, if F/f is increasing, then there is a positive correlation between number of reported employees

```
Linear regression                                      Number of obs =  151806
                                                       F( 17,151788) =  125.86
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.0141
                                                       Root MSE      =  .48565
```

risultato	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
dipendenti	.0113146	.0107671	1.05	0.293	-.0097886	.0324178
settori2	.0323488	.0405487	0.80	0.425	-.0471259	.1118235
settori3	.0297063	.0357099	0.83	0.405	-.0402844	.0996969
settori4	.0706489	.0351828	2.01	0.045	.0016913	.1396065
settori5	.0412871	.03489	1.18	0.237	-.0270966	.1096708
settori6	.0248906	.0347275	0.72	0.474	-.0431745	.0929557
settori7	.263766	.040259	6.55	0.000	.1848593	.3426728
settori8	.0852163	.0360221	2.37	0.018	.0146138	.1558189
settori9	-.0192634	.0352476	-0.55	0.585	-.0883479	.0498211
sett2dip	.0106818	.0117732	0.91	0.364	-.0123935	.033757
sett3dip	.0076794	.0109501	0.70	0.483	-.0137825	.0291413
sett4dip	.0085667	.0108717	0.79	0.431	-.0127415	.0298749
sett5dip	.0059091	.0108192	0.55	0.585	-.0152962	.0271145
sett6dip	.0125863	.0108006	1.17	0.244	-.0085826	.0337551
sett7dip	-.0026538	.011755	-0.23	0.821	-.0256934	.0203857
sett8dip	.0085434	.0111466	0.77	0.443	-.0133036	.0303905
sett9dip	.0258065	.0110239	2.34	0.019	.0041999	.0474132
_cons	.2935722	.0345996	8.48	0.000	.2257578	.3613866

Table 2: Predicting successful audits in small firms (robust standard errors).

and amount of the misreport, i.e., firms who report more workers also misreport by more. The coefficient on dipendenti is positive and close to significant at the 10 percent level, and moreover coefficients on the interaction terms are all positive and most are robustly significant. This evidence supports the finding that firms who report more employees also underreport by more. This, as we will see in the next section, is a prediction of our theoreitcal model under the assumption that F/f is increasing.

4.4 Results for large firms

For large firms, our identification strategy delivers mixed results. From the summary statistics we know that, on average, the probability of a successful audit is larger for large than for small firms (54% versus 40%). In other words, an inspector could substitute a search of a small firms with a search of a large firm and increase his probability of success.

One might attribute this difference to the (unobserved) cost of carrying out an audit in a larger firm. This interpretation not consistent with the evidence presented in Table 4, however. The regression in Table 4 is identical to that of Table 2, except that it is performed on the sample of firms which report more than 10 employees. First, we notice that the coefficients of the variable *dipendenti*, as well as those of the variables interacted with *dipendenti*, are very small. This means that size per se does not appreciably raise the probability of a successful audit above the 41% level (the constant in the regression which, incidentally, is about equal to the average success rate among small firms). Rather, the "excess return"

Linear regression

Number of obs =	151806
F(17,151788) =	86.86
Prob > F =	0.0000
R-squared =	0.0173
Root MSE =	35662

evasioni	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
dipendenti	808.5379	500.0176	1.62	0.106	-171.4863	1788.562
settori2	-876.3222	2157.041	-0.41	0.685	-5104.078	3351.434
settori3	-17.18535	1608.122	-0.01	0.991	-3169.072	3134.702
settori4	-788.2723	1420.655	-0.55	0.579	-3572.727	1996.182
settori5	-1072.676	1371.121	-0.78	0.434	-3760.045	1614.694
settori6	-1552.089	1345.561	-1.15	0.249	-4189.362	1085.183
settori7	4871.017	2353.191	2.07	0.038	258.8114	9483.223
settori8	-2936.275	2402.829	-1.22	0.222	-7645.771	1773.221
settori9	-3803.104	1390.893	-2.73	0.006	-6529.226	-1076.982
sett2dip	1471.74	759.6196	1.94	0.053	-17.09889	2960.579
sett3dip	1443.75	580.073	2.49	0.013	306.8186	2580.681
sett4dip	1162.011	530.0129	2.19	0.028	123.1968	2200.826
sett5dip	603.4178	509.9039	1.18	0.237	-395.9835	1602.819
sett6dip	325.6284	504.9198	0.64	0.519	-664.0041	1315.261
sett7dip	1958.261	788.0929	2.48	0.013	413.6145	3502.907
sett8dip	3355.358	1088.62	3.08	0.002	1221.685	5489.031
sett9dip	1086.043	549.4998	1.98	0.048	9.035142	2163.052
_cons	3525.649	1334.122	2.64	0.008	910.7978	6140.501

Table 3: Predicting returns from audits in small firms (robust standard errors).

from audits comes through some of the sector dummies. This observation casts doubt on the interpretation that the difference in success rates of audits is due to the larger cost of performing audits on bigger firms. Adding to the puzzle, the sector with the biggest "excess return" to an audit is *settori8*, corresponding to credit and insurance.

It is possible that the cross-sector differences in returns to an audit that are present in the large firm sample reflect an unobserved "complexity of audit" cost that varies across large firms in different sectors. It is difficult to rule out the existence of such unobservable differences; however, we saw no evidence of them in the small firms sample. In our own reading, the evidence in the large firm sample provides weak support for the no-commitment, success-maximizing model.

4.5 Discussion of findings

For the "small firms" sample, which comprises about 90% of the firms in Italy, our results suggest that INPS behaves, in the aggregate, as an agency which maximizes the probability of finding tax cheaters and, moreover, has no ability to publicly commit to an auditing strategy. Is such a conclusion plausible? We think it is, and that the observed auditing behavior reflects an implicit incentive scheme whereby individual inspectors are rewarded (with promotions, for example, or by other means) as a function of the number of times they find a firm which is not reporting correctly. Furthermore, we believe that this implicit compensation scheme deals with an agency problem between INPS and its inspectors. We will now outline the

Linear regression

Number of obs =	24185
F(17, 24167) =	27.93
Prob > F =	0.0000
R-squared =	0.0129
Root MSE =	.49553

risultato	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
dipendenti	.0000339	.0000113	2.99	0.003	.0000117	.0000561
settori2	.0677106	.058369	1.16	0.246	-.0466963	.1821175
settori3	.1732998	.0556894	3.11	0.002	.0641451	.2824544
settori4	.1043166	.0556147	1.88	0.061	-.0046916	.2133247
settori5	.09056	.0553061	1.64	0.102	-.0178434	.1989634
settori6	.1064535	.0554327	1.92	0.055	-.0021981	.215105
settori7	.1833129	.0578605	3.17	0.002	.0699028	.2967231
settori8	.2866815	.056091	5.11	0.000	.1767397	.3966234
settori9	.165802	.0567756	2.92	0.004	.0545183	.2770857
sett2dip	.0000168	.0000122	1.38	0.169	-7.11e-06	.0000407
sett3dip	-.0001349	.0000581	-2.32	0.020	-.0002489	-.000021
sett4dip	-.0002117	.000096	-2.21	0.027	-.0003999	-.0000235
sett5dip	-.0000141	.000056	-0.25	0.801	-.0001239	.0000956
sett6dip	-.000141	.0000736	-1.92	0.055	-.0002853	3.26e-06
sett7dip	-.0000187	.0000412	-0.45	0.650	-.0000994	.000062
sett8dip	-.0000656	.0000177	-3.70	0.000	-.0001003	-.0000309
sett9dip	.0000632	.0000315	2.00	0.045	1.33e-06	.000125
_cons	.4125532	.0549474	7.51	0.000	.3048528	.5202536

Table 4: Predicting successful audits for large firms (robust standard errors).

elements of the agency problem

The first element of the agency relationship is that effort exerted by the inspector is not observable. This means that it is difficult to induce an inspector to monitor very intensely a group of firms which in equilibrium do not to evade taxes: the inspector will simply shirk and report finding no evasion. Notice that this is precisely the behavior required by the optimal strategy with commitment! The general point is that commitment bites exactly when it is necessary to audit well-behaved groups. If, as seems natural, the problem of unobservable effort is solved by rewarding the productivity if audits, then it will not be possible to generate "commitment behavior."

This argument does not explain why the implicit incentive scheme should reward detection and disregard the *amount* of evasion detected. We believe that a possible reason might lie in the peculiar technology of assessing tax evasion. Suppose that the inspector can manipulate the amount of evasion ascertained, but cannot as easily manipulate whether any evasion is ascertained. In this case it would be optimal not to reward the size of the tax evasion ascertained, because this is more easily manipulable; rather, it would be optimal to tie compensation to the less manipulable correlate of effort, which is whether any evasion was ascertained. In support of our argument, we note that there is a very large amount of overascertaining.

5 A New Theoretical Model: No-Commitment Auditing to Maximize Detections

This section develops and analyzes a new theoretical model of tax evasion and enforcement that is consistent with the empirical findings of the previous section. The model is a special case of the general model introduced in Section 2.

We restrict attention to a single audit class. Therefore, there is a continuum of firms with true tax base x distributed according to the density $f(x)$ on the interval $[a, b]$. A firm reporting a type r pays taxes $t \cdot r$. We make the stipulation that in equilibrium no firm can report below a, the lowest possible income; in other words, all firms must pay at least the taxes corresponding to income a.

Following Erard and Feinstein (1994), we assume that there is a proportion λ of honest firms and a proportion $(1 - \lambda)$ of strategic firms. Honest firms always report the true value x and pay taxes $t \cdot x$. A strategic firm chooses which tax base r to report to the tax authority in order to minimize its expected taxes. In doing so, the firm recognizes that it faces an probability schedule $p(r)$ that relates the report r to the probability of being audited by the auditor. Firms are aware that in case of an audit, the true characteristic x will be discovered. The taxes are then assessed on the true level x and a penalty is added that is proportional to the amount of evasion. In case of an audit, a firm x that reported $r \leq x$, pays a total of $t \cdot x + \theta(x - r)$. We will be looking for a separating equilibrium in which strategic firms report their income according to a strategy $\rho(x)$ that is strictly increasing in x.

The auditor observes the report of the firms and chooses an audit schedule $p(r)$. We assume that the auditor does not have the power to commit to an audit schedule, and that it maximizes the number of successful audits. The first assumption situates this model as a special case of the model analyzed in Section 3.1; the second assumption ensures that the auditor's objective function satisfies Assumption 3.

The auditor chooses how many firms to audit by equalizing the expected probability of a successful audit to a (constant) marginal cost of an audit. This marginal cost is denoted by $(1 - q)$, and it can be any number between zero and one. We choose this formulation for ease of exposition. This formulation is seen to be equivalent to the more common formulation in which the auditor has a budget constraint on the number of firms it can audit, once we reinterpret the Lagrange multiplier on the budget constraint as the marginal cost of an audit.

5.1 The firm's problem

A firm with type x chooses which $r \leq x$ to report so as to maximize

$$p(r)(x - tx - \theta(x - r)) + (1 - p(r))(x - tr). \tag{7}$$

We will construct an equilibrium in which all strategic firms will misreport. In that case the constraint $r \leq x$ is never binding and the first-order conditions associated with (7) are necessary conditions for a maximum. They are

$$p'(r)(r-x)(t+\theta) + p(r)(t+\theta) - t = 0, \tag{8}$$

which can be rewritten as

$$x^*(r) = r + \frac{p(r) - \frac{t}{\theta+t}}{p'(r)}, \tag{9}$$

where $x^*(r)$ denotes the true type of a strategic firm which in equilibrium reports $r < x^*(r)$. We note for future reference that if $p(r) \geq \frac{t}{\theta+t}$ then it is optimal for the firm to report its true tax base r.

Concavity of the objective function with respect to r is a sufficient condition for the first order conditions to identify a global maximum. Concavity means that, for all x and $r < x$, the second derivative of the objective function with respect to r must be negative:

$$(t+\theta)[p''(r)(r-x) + 2p'(r)] \leq 0. \tag{10}$$

5.2 The auditor's problem

Observing a report r, the auditor realizes that it can come from an honest firm with true type $x = r$, or from a strategic firm that underreported taxes with true type $x^*(r) > r$. Since the auditor seeks to maximize the probability of a successful audit and does not have the power to commit to an audit schedule, the auditor's best response is to only audits reports which have the highest probability of being made by cheating firms. This implies that, in equilibrium, all reports audited with positive probability need to lead to the same probability of success.

The auditor uses Bayes' Rule to assess the probability that a firm reporting r underreported its taxes. Among the honest types, the mass who report in the interval of length Δ centered around r are approximately $f(r) \cdot \Delta$. Among the strategic types, the mass who report in that same interval are approximately $f(x^*(r)) \cdot x^*(\Delta)$, where we denote $x^*(\Delta) = x^*(r + \Delta/2) - x^*(r - \Delta/2)$. Therefore, the probability of an honest type conditional on reporting in the interval is

$$\frac{\lambda f(r)\Delta}{\lambda f(r)\Delta + (1-\lambda) f(x^*(r)) x^*(\Delta)}.$$

Dividing by Δ and letting $\Delta \to 0$ yields

$$\frac{\lambda f(r)}{\lambda f(r) + (1-\lambda) f(x^*(r)) \frac{dx^*(r)}{dr}}. \tag{11}$$

A constant success of audits means that on the range of reports audited, the probability of honest types must be constant and equal to $q \in (0,1)$. Indeed, if this probability were larger (respectively, smaller) than q then the expected success rate on every audit would be lower (resp., larger) than the marginal cost of an audit, which cannot be the case in equilibrium. Therefore, in equilibrium it must be

$$\frac{\lambda f(r)}{\lambda f(r) + (1-\lambda) f(x^*(r)) \frac{dx^*(r)}{dr}} = q \tag{12}$$

Denoting

$$\frac{\lambda (1-q)}{(1-\lambda) q} = \gamma,$$

we can rewrite (12) as

$$f(x^*(r)) \frac{dx^*}{dr} = \frac{\lambda (1-q)}{(1-\lambda) q} f(r) = \gamma f(r). \tag{13}$$

5.3 Equilibrium

Let us start by establishing the support of the reporting strategies must be of the form $[a, \rho^*(b)]$. Remember that in equilibrium no firm can report below a, the lowest possible income. Since a firm with income a will not report above its true income, it must be $\rho^*(a) = a$. Further, the range of the reporting strategy is an interval. To see this, observe that if a report r is not used by any strategic firm, then $p^*(r)$ must be zero since the audits at that report would be only of honest firms. But a zero auditing probability would lead all firms that report more than r to want to deviate to that report. This means that if tax report r is made in equilibrium, then all reports below r are also used by some firm.

Let us further observe that in any equilibrium with some evasion it must be

$$p^*(\rho^*(b)) = 0. \tag{14}$$

This boundary condition comes from the following observation. If $p^*(\rho^*(b)) > 0$ and $\rho^*(b) < b$, then we would have an immediate contradiction since a firm with type b would rather report a bit higher than $\rho^*(b)$ avoiding all audits.

Next comes a formal definition of the equilibrium in this game.

Definition 1 *For any $q \in (0,1)$, an equilibrium of the auditing game is a reporting strategy $\rho^*(\cdot)$ with associated inverse strategy $\rho^{*-1}(\cdot) = x^*(\cdot)$, and an auditing schedule $p^*(\cdot)$ with support $[a, \rho^*(b)]$ that solve the firm's first and second order conditions (9) and (10), the auditors' indifference condition (13), and the boundary condition (14).*

The next proposition shows that an equilibrium exists and characterizes it.

Proposition 2 ***(Equilibrium of the auditing game)*** *For any $q \in (0,1)$ there exists an equilibrium of the auditing game. It is given by a reporting strategy $\rho^*(\cdot)$ and an auditing schedule $p^*(\cdot)$ that solve:*

$$\rho^*(x) = F^{-1}(F(x)/\gamma)$$
$$p^*(r) = \max\left\{\frac{t}{\theta+t}(1-\exp\left(-\int_r^{F^{-1}(1/\gamma)} \frac{1}{\rho^{*-1}(y)-y}\,dy\right),0\right\}$$
$$\gamma = \frac{\lambda(1-q)}{(1-\lambda)q}.$$

Proof. Let us first characterize the equilibrium reporting strategies. Integrating both sides of (13) yields

$$\gamma F(r) = F(x(r)) + k. \tag{15}$$

Since $x^*(a) = a$, it follows from (15) evaluated at $r = a$ that $k = 0$.Therefore, for a generic $r > a$ we have

$$x^*(r) = F^{-1}(\gamma F(r)), \text{ or equivalently} \tag{16}$$
$$\rho^*(x) = F^{-1}\left(\frac{F(x)}{\gamma}\right). \tag{17}$$

We now characterize the equilibrium auditing schedule. Using (16) to substitute into (9) we get

$$\frac{p^{*\prime}(r)}{p^*(r)-\frac{t}{\theta+t}} = \frac{1}{F^{-1}(\gamma F(r))-r}. \tag{18}$$

Integrating both sides yields:

$$\ln\left(\frac{t}{\theta+t}-p^*(r)\right) = -\int_r^{\rho^*(b)} \frac{1}{F^{-1}(\gamma F(y))-y}dy + k$$
$$p^*(r) = \frac{t}{\theta+t} - K\exp\left(-\int_r^{\rho^*(b)} \frac{1}{F^{-1}(\gamma F(y))-y}dy\right)$$

K is set equal to $\frac{t}{\theta+t}$, to ensure that $p^*(\rho^*(b)) = 0$ as per the boundary condition (14).

Finally, Lemma 3 in Appendix B verifies that, given the audit schedule $p^*(\cdot)$, the firm's reporting strategy $\rho^*(\cdot)$ satisfies the second order conditions (10). ■

Equation (17) shows that a larger value of γ correponds to a lower value of $\rho^*(x)$, that is, greater underreporting in equilibrium. This makes sense: a large γ correponds by definition to a low value of q, which means a high marginal cost of an audit. Equation (17) also implies that, in any equilibrium in which there is some underreporting, γ must be greater than 1.

Finally, equation (17) pins down the report of the highest income firm,

$$\rho^*(b) = F^{-1}(1/\gamma).$$

This expression shows that, as the marginal cost of funds increases, the interval of the reports being audited shrinks. All strategic firms report in that interval, and thus as that interval shrinks, strategic firms become a greater percentage of the set of firms who report in that interval. Only honest firms report above that interval.

If F is log-concave then we can further characterize the equilibrium strategies. Log-concavity means that $f(x)/F(x)$ is decreasing in x.The assumption of log-concavity is relatively mild, in that many common cumulative distribution functions are log-concave, including: the Uniform, the Power distribution, the Normal, the Gamma, the extreme value, the exponential, the Pareto, and many others. (See Bagnoli and Bergstrom 1989). If F is log-concave then we can show that the amount of underreporting is increasing in the true income.

Lemma 1 ***(increasing cheating)*** *If F is log-concave then $x - \rho^*(x)$ is increasing in x, that is, strategic firms with higher true tax base underreport by more.*

Proof. See Appendix B. ■

6 Calibration of the Theoretical Model

The power distribution on $[a, b]$ is given by $F(x) = \left(\frac{x-a}{b-a}\right)^{\beta}$ with $\beta > 0$. We now apply Proposition 2 to derive closed forms for the equilibrium of the auditing game when the firms' tax base is distributed according to a power distribution.

Proposition 3 ***(Equilibrium of the auditing game with Power distribution)*** *Suppose the tax base of firms is distributed on $[a, b]$ according to a power distribution $F(x) = \left(\frac{x-a}{b-a}\right)^{\beta}$. For any $q \in (0, 1)$ there exists an equilibrium of the auditing game. It is given by a reporting strategy $\rho^*(\cdot)$ and an auditing schedule $p^*(\cdot)$ that solve:*

$$\begin{aligned}
\rho^*(x) &= a + \frac{\alpha}{\alpha+1}(x-a) \\
p^*(r) &= \max\left\{\frac{t}{\theta+t}\left[1 - \left(\frac{\alpha+1}{\alpha}\frac{r-a}{b-a}\right)^{\alpha}\right], 0\right\} \\
\alpha &= \frac{1}{\gamma^{(1/\beta)} - 1} \\
\gamma &= \frac{\lambda(1-q)}{(1-\lambda)q}.
\end{aligned}$$

Proof. From (18) we get

$$\begin{aligned}\frac{p^{*\prime}(r)}{p^{*}(r)-\frac{t}{\theta+t}} &= \frac{1}{\gamma^{(1/\beta)}(r-a)+a-r} \\ &= \frac{1}{(r-a)}\frac{1}{\gamma^{(1/\beta)}-1},\end{aligned}$$

and integrating both sides yields

$$\ln\left(\frac{t}{\theta+t}-p^{*}(r)\right)=\ln(r-a)\frac{1}{\gamma^{(1/\beta)}-1}+\kappa,$$

where κ denotes a constant of integration. Taking exponentials on both sides leads to

$$p^{*}(r)=\frac{t}{\theta+t}-K(r-a)^{\alpha},$$

where $K=(\exp\kappa)$ is a constant that will be computed momentarily and we denote

$$\alpha=\frac{1}{\gamma^{(1/\beta)}-1}.$$

Note that $\alpha>0$ because $\gamma>1$. Finally, from (17) we have

$$\rho^{*}(x)=a+\frac{1}{\gamma^{(1/\beta)}}(x-a)=a+\frac{\alpha}{\alpha+1}(x-a).$$

The constant K is computed using the fact that $p^{*}(\rho^{*}(b))=0$. Rewrite this condition as

$$\frac{t}{\theta+t}-K\left(\frac{\alpha}{\alpha+1}(b-a)\right)^{\alpha}=0,$$

whence $K=\frac{t}{\theta+t}\left(\frac{\alpha+1}{\alpha(b-a)}\right)^{\alpha}$. Substituting back into the probability of auditing yields

$$p^{*}(r)=\frac{t}{\theta+t}\left[1-\left(\frac{\alpha+1}{\alpha}\frac{r-a}{b-a}\right)^{\alpha}\right].$$

■

We want to calibrate the parameters of the model on the INPS data. For realism's sake, we don't want the model to allow a huge firm (FIAT, say) reporting just one employee and, in so doing, being indistinguishable from a mom-and-pop store. Therefore we need to incorporate the possibility that firms are observably different to the auditor. We do this by allowing firms (audited and not) to be partitioned into several audit classes. An audit class is made up of firms which share some characteristic observable to the auditor (legal structure, location, productive sector, energy consumption etc.), different from the firm's report, that is correlated with their true tax base. Formally, an audit class k is defined by three parameters known to the auditor: a_k and b_k the lowest and highest possible true types of firms in the class, and the parameter β_k which characterizes the Power distribution within that class. No firm in audit class k can report below a_k, (implicitly, because the auditor audits such underreports with probability 1) but firms are free to report anywhere within (a_k, b_k). In general, the size of the interval (a_k, b_k) implicitly measures the auditor's residual uncertainty about a firm's true tax base, after all available (non-report) information has been evaluated to assign the firm to an audit class. FIAT is presumably in an audit class where a_k equals thousands of employees, and so this formulation avoids the possiblity of FIAT reporting very few employees.

We do not observe the audit classes that the auditor sees. Therefore, we choose to incorporate audit classes in a simple way, specifically, in a way that all firms in a given audit class report in a given interval. To illustrate our methodology, consider for example the audit class formed of all firms which are *unincorporated* legal entities. In practice, the auditor reasonably expects that any unincorporated firm is "small." In the model, this knowledge is introduced by lumping all these firms within an audit class with true tax base within $(1, b)$, for some b. While we do not observe the true tax base in our data, we observe the reports of the firms which are audited. In equilibrium, firms in this audit class are audited only if they report M or fewer employees, where $M = \rho^* (b) < b$. So we fix M arbitrarily, say $M = 10$, and then we use Proposition 3 to recover the value of the unobserved b that would cause the strategic firms in this class to report in the interval $(1, M)$. With this procedure, we capture the presence of unobservable audit classes which, in our data, show up as the group of firms which are audited when they report within the interval $(1, 10)$. We can then repeat the procedure by partitioning the set of reports into adjacent intervals $(a_k, M_k), (a_{k+1}, M_{k+1}), ...$ thus partitioning the entire set of reports as coming from several distinct audit classes. We can make the audit classes small and sector-specific if we want, thus allowing for fine-grained information on the part of the auditor.

Now we detail the methodology which, after we fix a and M, is used to infer all relevant parameters of an audit class. The procedure is based on matching moments. The first moment we match is the fraction of firms who report in the interval (a, M) and are found *not* to have underreported. This fraction should be close to, but smaller than λ, the *unconditional* fraction of honest firms in the model. This is because a firm is in our sample *only conditional* on being audited. In the model, the fraction of firms that are honest conditional on being audited is smaller than λ, because the honest firms with a high tax base report (truthfully) a large number and are not audited. Therefore, strategic firms are disproportionally present among the firms being audited. According to the model, the ratio of honest to strategic firms *among those audited* is given by

$$\frac{\lambda \int_a^M p^* (r) \, dF (r)}{(1-\lambda) \int_a^M p^* (r) \, dF (\rho^{*-1} (r))} \tag{19}$$

This ratio should equal the ratio of honest to strategic firms in our data, which we denote by C_1.

The second moment is the average number of employees reported by firms who report in (a, M). Call this C_2. In the theory, this quantity is given by

$$(1-\lambda)\int_a^b \rho^*(x)\,dF(x) + \lambda \int_a^M r dF(r). \tag{20}$$

The third moment we match is the total amount of euros evaded conditional on evading a positive amount. This amount, divided by 26,500 (the gross average salary in Italy in 2007[14]) translates these monetary amounts into "employee-equivalent underreports." We call this quantity C_3. The theoretical expression that corresponds to these amounts is

$$\int_a^b [x-\rho^*(x)]\,dF(x). \tag{21}$$

Equations (19)$= C_1$, (20)$= C_2$, and (21)$= C_3$, together with the condition $\rho^*(b) = M$, form a system of four equations in four unknowns. After substituting for $p^*(\cdot)$ and $\rho^*(\cdot)$ from Proposition 3 and after much algebra (detailed in Appendix C), the system of equations is reduced to the following.

$$\begin{aligned} \frac{\lambda}{(1-\lambda)\left(\frac{\alpha+1}{\alpha}\right)^{\beta}} &= C_1 \\ (1-\lambda)(1+C_1)(a+\alpha C_3) &= C_2 \\ \frac{(M-a)}{\alpha}\frac{\beta}{\beta+1} &= C_3 \end{aligned} \tag{22}$$

We think of a, M, C_1, C_2, C_3 as empirically observable parameters, and we want to know which values of the unknown parameters λ, α, β are compatible with any given constellation of observable parameters. Consider for example the audit class $(a = 0, M = 10)$, which is composed of all the firms who have true tax base between 1 and $b > 10$, and which are audited only when they report below 10. In this case we can compute that C_1, the ratio of honest to strategic firms among the firms which report less than 11 employees in our data, equals 0.6/0.4. We have C_2, the average number of employees reported by firms who report in $(0, 10)$, is given by 3.09 employees. Finally C_3, the total amount evaded conditional on evading a positive amount is 17,683 euros, which translated into employee-equivalents yields 0.67 employees. Given these parameters, the system of equations (22) can be solved numerically to yield $[\lambda = 0.621\,51, \alpha = 4.\,874, \beta = 0.484\,91]$. The computed value of λ means that the fraction of honest firms is 62%, slightly higher than the 60% of firms which are found not to cheat when audited. The difference is given by the fact that strategic auditing

[14] "Domanda di lavoro e retribuzioni nelle imprese italiane", Rapporto Unioncamere 2008, page 23.

selectively hits precisely those firms which are more likely to misreport. The parameter α is best understood in terms of the fraction $\alpha/(\alpha+1)$, which represents the fraction of the true tax base reported by strategic firms. In this case this fraction equals 0.83, which means that strategic firms report 83 percent of their true tax base. The highest true tax base in the audit class (unobserved because firms who report this much are not audited) equals $10\cdot(\alpha+1)/\alpha = 12$. Finally, the fact that $\beta < 1$ implies that the density of true tax bases is decreasing.

For the audit class identified by audited reports in $(a = 11, M = 25)$, which represent about 9 percent of our sample, we have that C_1, the ratio of honest to strategic firms equals 0.47/0.53. We have C_2, the average number of employees reported, is given by 15.39 employees. Finally C_3, the total amount evaded conditional on evading a positive amount is 29,950 euros, which translated into employee-equivalents yields 1.13 employees. Solving the system of equations (22) yields $[\lambda = 0.49937, \beta = 0.60788, \alpha = 4.6840]$. The highest true tax base in the audit class (unobserved because firms who report this much are not audited) equals $11 + (25-11)(\alpha+1)/\alpha = 28$.

For the audit class identified by audited reports in $(a = 26, M = 50)$, which represent about 3 percent of our sample of audits, we have that C_1, the ratio of honest to strategic firms equals 0.43/0.57. We have C_2, the average number of employees reported, is 36 employees. Finally C_3, the total amount evaded conditional on evading a positive amount is 58,000 euros, which translated into employee-equivalents yields 2.19 employees. Solving the system of equations (22) yields $[\lambda = 0.55145, \alpha = 9.0171, \beta = 4.6438]$. The parameter β being greater than 1 will lead to a poor fit of (23) to the data. What we learn from this is that sometimes, matching the moments in (22) *perfectly* has a significant cost in terms of fit. One way around this problem is to relax the matching of the moments in (22). If we accept the second equation in (22) to equal 35.235 rather than 36, which seems like a small cost, then (22) yields a solution $[\lambda = 0.47148, \alpha = 5.4795, \beta = 1]$ which produces a better fit of (23) to the data. The implication is that, for this audit class, our model underpredicts slightly the reported number of employees.

We do not pursue the analysis of audit classes with reports larger than 50 employees since these reports are only about 2% of our sample.

6.1 Fit of the calibrated model

Given these parameter configurations we can plot the predicted distribution of reports by audited firms, and compare it to the distribution of reports in our data. The results of this comparison are not a foregone conclusion, despite the fact that our parameters have been calibrated on the empirical distribution (histogram) of reports. Indeed, our calibration procedure was not designed to match the *shape* of the histogram but rather its average..Therefore, the forthcoming comparison can give us a sense of how well our calibrated model fits the data.

The equilibrium probability of observing an audit of a firm which reports r is given by

$$p^*(r) \cdot \left[(1-\lambda) f\left(\rho^{*-1}(r)\right) \frac{\partial \rho^{*-1}(r)}{\partial r} + \lambda f(r) \right], \tag{23}$$

The first multiplicative term represents the probability of being audited conditional on reporting r. The term in brackets represents the density of firms which report r, which is a mixture of strategic and honest firms. Figure 2 plots this function as the green line against the histogram of the empirical distribution of reported sizes. For small firms and for medium firms the fit is rather good, particularly considering that our calibration procedure was not designed to match the shape of this histogram but rather just one moment, its average.. For large firms (reports between 26 and 50) the fit is not good when we apply the parameter configuration with $\beta > 1$. The fit becomes better when we apply the parameter configuration obtained by forcing $\beta = 1$ and relaxing the matching of C_2 in (22), which in the figure is referred to as "relaxed identification." Even closer fits could be achieved by further relaxing the matching of C_2, which would give rise to values of β smaller than 1.

Figure 2 also plots the predicted distribution of true tax bases (the orange curve). As expected, this curve extends further to the right than the histogram, up to $b > M$. This is because the model predicts that some firm have a tax base greater between M and b. Among these firms, the strategic ones will report between a and M and they will be audited with positive probability, so their audits will show up in our histogram. The honest among these firm will report above M, and their audits will not show up in the histogram.

7 Counterfactual: Commitment Strategy

In this section we carry out a counterfactual experiment. We start from the no-commitment equilibrium and use the model to calculate: (a) the fraction of firms audited, which is a measure of the resources devoted to this audit class; (b) the amount of money raised in equilibrium, which is the sum of taxes raised and penalties assessed. We then turn to the strategy in which the same number of firms is audited, but the probability of auditing is such that only firms who report below a threshold are audited and no audited firm cheats. We call this the "commitment strategy" because it is the strategy that requires the most commitment. Indeed, under this strategy *all* audits are unsuccessful and, moreover, a simple deviation in the type of reports audited could deliver a very high success rate of audits. We then compare the amount of money raised under the two strategies.

We develop the analysis in the first two subsections working only on the audit class $[0, 10]$. In the last subsection we extend the same analysis to the other audit classes.

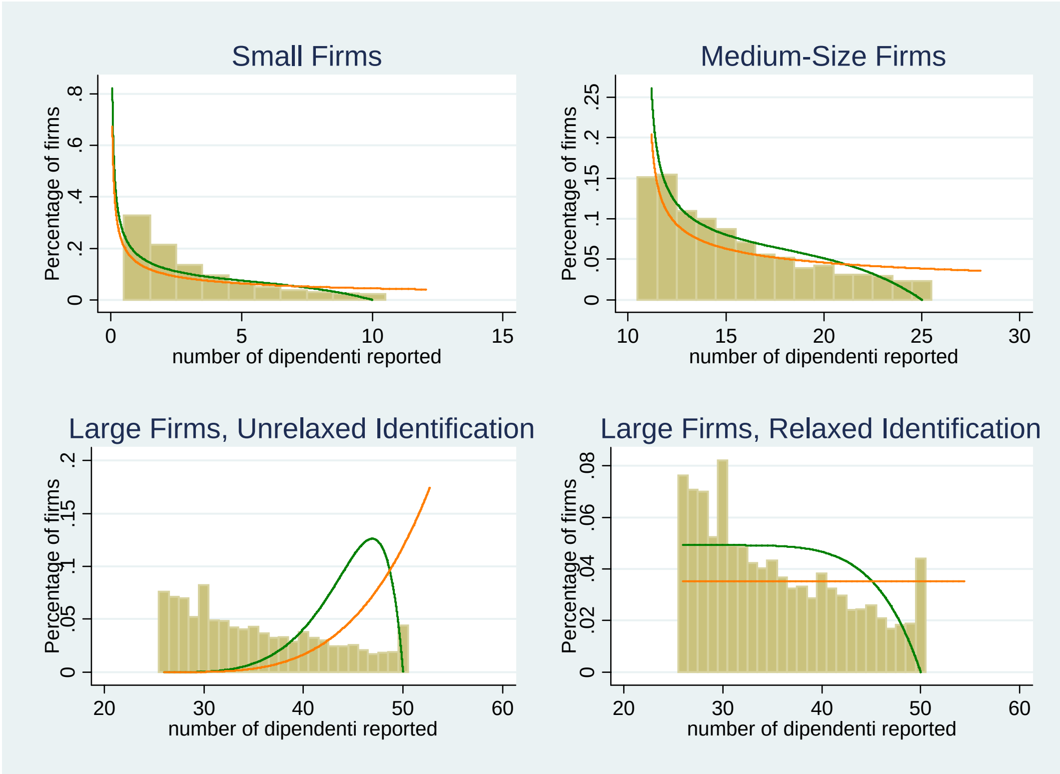

Figure 2: The histogram represents the empirical percentage of audited firms who report a certain size; the green line the predicted percentage of audited firms by reported size; the orange line the predicted distribution of true firm size.

7.1 No-commitment equilibrium: fraction of firms audited and money raised

We want the fraction of firms audited to match aggregate statistics available from INPS. Crude statistics that are publicly available suggest that every year about 2-3 percent of firms who declare less than 10 reported employees are audited. When audited, we know from our data that the average firm's books are checked going back somewhat longer than 2 years. This "backward looking" span of the audit increases the deterrence power of auditing; we factor in this effect very crudely by jacking up the "effective" probability of auditing to 5 percent. In our model, the fraction of firms audited among those that report between a and M is constructed starting from (23) and is given by

$$\frac{\int_a^M p^*(r) \cdot \left[(1-\lambda) f\left(\rho^{*-1}(r)\right) \frac{\partial \rho^{*-1}(r)}{\partial r} + \lambda f(r)\right] \, dr}{\int_a^M \left[(1-\lambda) f\left(\rho^{*-1}(r)\right) \frac{\partial \rho^{*-1}(r)}{\partial r} + \lambda f(r)\right] \, dr} \tag{24}$$

We want this term to equal 5 percent. We substitute out for $p^*(r)$ and write this condition as

$$\frac{\int_a^M \left[1-\left(\frac{\alpha+1}{\alpha}\frac{r-a}{b-a}\right)^\alpha\right]\cdot\left[(1-\lambda) f\left(\rho^{*-1}(r)\right)\frac{\partial\rho^{*-1}(r)}{\partial r}+\lambda f(r)\right]\,dr}{\int_a^M \left[(1-\lambda) f\left(\rho^{*-1}(r)\right)\frac{\partial\rho^{*-1}(r)}{\partial r}+\lambda f(r)\right]\,dr}=\frac{\theta+t}{t}0.05 \qquad (25)$$

The parameters t and θ are not pinned down by the analysis in Section 6. We set $t = 0.4$ to capture an approximately 40% tax rate on gross wages, and we let θ range freely to achieve the desired equality. The parameter $\widehat{\theta}$ so obtained will capture the 33% penalty on the amount underreported, plus additional costs (psychological, legal, etc.) involved in being found in violation of the tax code. We expect therefore that $\widehat{\theta} \geq 0.33$. Solving equation (25) numerically based on the parameters calibrated in Section 6, we get $\widehat{\theta} = 6.87$. This means that the *perceived* cost for being found cheating is estimated to be 6.87 times the amount underreported. This is a very large number compared to the monetary fine which is 0.33. We attribute this discrepancy partly to the psychological, legal, etc. costs of being found in violation of the tax code, and partly to the fact that we have thrown out some audits from our data (whistleblower, etc). These audits will provide additional deterrence which is not incorporated in our empirical analysis.

The money raised, expressed in proportion to the salary of a worker, is given by

$$\begin{aligned}&\lambda\int_a^b tr\cdot f(r)\,dr+(1-\lambda)\int_a^M tr\cdot f\left(\rho^{*-1}(r)\right)\frac{\partial\rho^{*-1}(r)}{\partial r}\,dr\\&+(1-\lambda)\int_a^M p^*(r)\left[(t+0.33)\left(\rho^{*-1}(r)-r\right)\right]\cdot f\left(\rho^{*-1}(r)\right)\frac{\partial\rho^{*-1}(r)}{\partial r}\,dr. \qquad (26)\end{aligned}$$

The first addend is the amount of taxes raised from honest firms; using Lemma 2, it equals $\lambda t\left[a+\frac{\beta}{\beta+1}(b-a)\right] = \lambda t\cdot 3.9187 = \lambda\cdot 1.5675$. The second term is the amount of taxes paid by strategic firms, and it equals $(1-\lambda)t\left[a+\frac{\alpha}{\alpha+1}\frac{\beta}{\beta+1}(b-a)\right] = (1-\lambda)t\cdot 3.2516$ (see Appendix D.1 for details on all computations in this Section). Finally, the the third term is the money raised from audited cheaters and is computed numerically to be equal to $(1-\lambda)(t+0.33)(0.02813)$. The total money raised from strategic firms is

$$\begin{aligned}&(1-\lambda)\left[t\cdot 3.2516+(t+0.33)(0.02813)\right]\\=\;&(1-\lambda)\left[(0.4)\cdot 3.2516+(0.4+0.33)(0.02813)\right]\\=\;&(1-\lambda)\,1.3212\end{aligned}$$

7.2 Commitment strategy: money raised for the same amount of firms audited

Consider now the counterfactual policy of auditing with maximal probability all firms which report less than some threshold T (yet to be determined), and of not auditing any firm who reports T or more. As explained in Section 5.1, if $p(r) \geq t/\left(t+\widehat{\theta}\right)$ then a firm with true tax base r will report truthfully. Therefore, the policy we propose is to audit with probability $t/\left(t+\widehat{\theta}\right)$ all firms who report less than T. Under this policy, there is no cheating among firms who report below T, and the success rate of audits is zero. The threshold T is determined by the budget constraint. According to equation (24), a fraction equal to

$$0.05 \int_a^M \left[(1-\lambda) f\left(\rho^{*-1}(r)\right) \frac{\partial \rho^{*-1}(r)}{\partial r} + \lambda f(r) \right] dr \tag{27}$$

of all firms in the audit class is audited. This expression equals 0.047 (see Appendix D.2 for details). This number needs to equal to the fraction of firms audited under the commitment strategy. Under the commitment strategy, all firms (strategic or not) with tax base below T will report truthfully and be audited with probability $\frac{t}{t+\widehat{\theta}}$. No other firm will report in that range. Therefore, the fraction of firms audited under the commitment strategy is given by the equation

$$\begin{aligned} \int_a^T \frac{t}{t+\widehat{\theta}} f(r)\ dr &= \frac{t}{t+\widehat{\theta}} F(T) \\ &= \frac{t}{t+\widehat{\theta}} \left(\frac{T-a}{b-a} \right)^{\beta}. \end{aligned}$$

Equating this to 0.047 and solving for T yields

$$T = a + (b-a) \left(\frac{t+\widehat{\theta}}{t} 0.047 \right)^{\frac{1}{\beta}} = 12 \left(0.047 \frac{0.4+6.87}{0.4} \right)^{\frac{1}{\beta}} = 8.6710.$$

The money raised from strategic firms under the commitment strategy equals the amount declared by firms with tax base below T, plus the amount declared by firms with tax base

	Class 0-10	**Class 11-25**	**Class 26-50**
λ	0.62	0.5	0.47
β	0.48	0.61	1
α	4. 88	4.68	5.48
b	12	28	54.38
$\widehat{\theta}$	6.87	8.45	6.95
Amount raised from honest firms	$\lambda \cdot 1.57$	$\lambda \cdot 6.97$	$\lambda \cdot 16.08$
Amount raised from strategic firms in equilibrium	$(1-\lambda)\,1.32$	$(1-\lambda)\cdot 6.55$	$(1-\lambda)\cdot 15.26$
Amount raised from strategic firms under the commitment strategy	$(1-\lambda)\,1.47$	$(1-\lambda)\,6.80$	$(1-\lambda)\,15.78$

Table 5: Calibrated parameters, and counterfactuals.

greater than T, which is exactly T. Formally,

$$
\begin{aligned}
& (1-\lambda)\left[\int_a^T trf(r)\,dr + tT(1-F(T))\right] \\
= \; & (1-\lambda)\left[t\int_a^T r\left(\frac{1}{b-a}\right)^{\beta}\beta(r-a)^{\beta-1}\,dr + tT\left(1-\left(\frac{T-a}{b-a}\right)^{\beta}\right)\right] \\
= \; & (1-\lambda)\left[t\left(\frac{T-a}{b-a}\right)^{\beta}\beta\left[\frac{T-a}{\beta+1}+a\frac{1}{\beta}\right]+tT\left(1-\left(\frac{T}{12}\right)^{\beta}\right)\right] \\
= \; & (1-\lambda)\,t\,[2.4188+1.2640] \\
= \; & (1-\lambda)\,0.4\,(2.4188+1.2640) = (1-\lambda)\,1.4731
\end{aligned}
$$

7.3 Other audit classes

We can repeat the same procedure for the other two audit classes. This is done in Appendix D.3 and D.4. For the audit class 26-50, we focus on the parameter constellation corresponding to $\beta = 1$. The results for all audit classes are presented in the table below.

We see that the commitment strategy, which is to audit all reports below a threshold with probability $\frac{t}{t+\widehat{\theta}}$, always delivers an improvement in the amounts of money raised relative to the (no-commitment) equilibrium strategy. When we focus on the amount raised from strategic firms, which are the only ones whose report is affected by auditing anyway, the improvement from the commitment strategy is slightly above 10% for the 0-10 audit class (1.32 to 1.47), and smaller for the other audit classes.

Also, the table indicates that the amount of money raised from strategic firms in equilibrium is relatively close to the one obtained from honest firms (85% in the 0-10 audit class, and better in the other classes), which represents the theoretical maximum revenue obtainable. In this sense, despite the fact that in equilibrium there is no commitment and the wrong objective function is maximized, our calibration suggests that the amount of money left on the table is not too large. Therefore, there is not much room for improvement from moving to a commitment strategy. Of course, this conclusion is true only at the level of auditing budget. It is possible that, were INPS to reduce the fraction of firms audited and simultaneously move to a commitment strategy, that the gains from commitment might be significantly larger.

8 Conclusion

We have provided an identification theorem (a statistical test) that allows us to reject the assumption that an auditor maximizes the number of successful audits in the absence of commitment. We have applied this test to audit data from the Italian labor-tax auditing agency INPS. The results suggest that the behavior of INPS is reasonably close to the paradigm of no-commitment and maximization of successful audits. We have presented a new theoretical model of strategic auditing, in which an auditor maximizes the number of successful audits in the absence of commitment. We have solved for the Bayesian Nash equilibrium of this model and, under the assumption that the INPS data are generated by the equilibrium, we have used observables (auditing data) to back out the unobservable deep parameters of the model, such as the distribution of true tax bases of firms, the amount of cheating firms engage into, etc. Then, based on the recovered deep parameters, we have engaged in the counterfactual exercise of asking how much more revenue could INPS collect if, rather than following the equilibrium strategy, it kept the auditing budget the same but allocated it in a way that made use of commitment and was mindful of deterrence. Specifically, we consider the "commitment strategy" in which only those who report below a threshold are audited, and they are audited with such a high probability that nobody who reports below a threshold cheats. We find that switching to this "commitment strategy" does increase tax revenue relative to the equilibrium of the no-commitment game However, the gain is small (in the order of 5%). This is partly because, in the equilibrium of the no-commitment game, INPS is already capturing more than 80% of the theoretical maximum revenue attainable, and so there is little room for improvement.

Our theoretical auditing model can accommodate "honest" firms, that is, firms which not willing to cheat on their taxes under any circumstances. The model can also accommodate a difference between the monetary cost of being found in violation of tax law, and a firm's perception of it. In the calibration we find that more than 50% of firms are unwilling to ever cheat on their taxes, and that the perceived cost of being found in violation is much higher than the actual monetary cost. These high estimate of the "degree of honesty" of firms are in line with empirical work and lore in the taxation literature suggesting that taxpayers behave as if they were overestimating the expected cost of being caught cheating. These estimates may, however, be out of line with the stereotype of Italians being lax in their tax-paying habits.

A caveat to our counterfactual analysis (but not to the rest of the paper) is that, when we change INPS's auditing strategy, we keep fixed the distribution of true tax bases of firms (which we obtain from the calibration exercise). Realistically we should expect that, if a permanent change in auditing behavior was implemented, firms would respond by changing the number of employees. Since we do not take this channel into account, our counterfactual results are should be interpreted as "short run" results. Even so, the counterfactual findings are interesting because their main thrust is that moving to a "commitment strategy" does not much improve tax revenue. To the extent that bureucrats evaluate policy with a focus on the short term, this finding might explain why the apparently inefficient behavior of INPS has survived.

References

[1] Andreoni J., Erard B. and Feinstien J., "Tax Compliance", *Journal of Economic Literature*, 1998, p. 818-860.

[2] Bagnoli, Mark and Ted Bergstrom (1989). "Log-Concave Probability and Its Applications." Manuscript, University of Michigan.

[3] Border K. and Sobel J., "Samurai Accountant: A Theory of Auditing and Plunder", *Review of Economic Studies*, 1987, **54** ,p. 525-540.

[4] Chander P. and Wilde L., "A General Characterisation of Optimal Income Tax Enforcement", *Review of Economic Studies*, 1998, **65**, p 165-183.

[5] Erard B. and Feinstein J., (1994) "Honesty and evasion in the tax compliance game", *RAND Journal of Economics*, 25(1), 1994, p. 1-19.

[6] Macho-Stadler I. and Perez-Castrillo J., "Optimal Auditing with Heterogeneous Income Sources", *International Economic Review*, 38(4), 1997, p. 951–968.

[7] Reinganum J. and Wilde, L., "Equilibrium Verification and Reporting Policies in a Model of Tax Compliance", *International Economic Review*, 27(3), 1986, p. 739-760.

[8] Sanchez, I. and Sobel, J. "Hierarchical Design and Enforcement of Income Tax Policies," Journal of Public Economics 50 (1993), 345-369.)

[9] Scotchmer, S., "Audit Classes and Tax Enforcement Policy", *The American Economic Review*, 77(2), 1987, p. 229-233.