Decentralizing Deterrence, with an Application to Labor Tax Auditing.*

Edoardo Di Porto
Universita La Sapienza, Roma

Nicola Persico
New York University

Nicolas Sahuguet
HEC Montréal

April 2011

Abstract

Deterrence of illegal activities is frequently carried out by many atomistic auditors (tax auditors, law enforcement agents, etc.). Not much is known either normatively about the best way to incentivize the auditors, nor positively about what these incentives look like in real world organizations. This paper focuses on the positive question. It proposes a game-theoretic model of decentralized deterrence and an empirical test, based on the equilibrium of the model, to identify the incentives of individual auditors. In the special (but important) case of tax enforcement, the paper fully characterizes the equilibrium of a strategic auditing game *and* provides a method to calibrate its parameters based on audit data.

Applying the model and method to Italian auditing data provides "proof of concept:" the methods are practical and tractable. We are able to provide an estimate of tax evasion based on (non-random) audit data alone. Counterfactual simulation of the model quantifies the costs and benefits of alternative auditing policies. We compare decentralized enforcement with a counterfactual "commitment policy," and compute the loss from the former. Thus we are able to quantify the costs of decentralizing enforcement.

*We thank Dan Silverman and Joel Slemrod for useful comments.

1 Introduction

Economic theory has long studied the problem of optimally (or at least effectively) generating deterrence, going back to the works of Becker (1968) and Ehrlich (1979). Border and Sobel's (1987) seminar article poses this question within an *auditing game*, a game where the deterrence-generating action can be conditioned on a report by the auditee. A large body of theoretical literature follows in their footsteps (Andreoni and Feinstein 1998 provide a good review). This literature proves its relevance, if nothing else because it has many important applications: auditing is a core mission of many important organizations including law enforcement agencies, tax authorities, and various regulators.

The auditing literature has not, to date, been brought to data. One reason, perhaps, is that the relevant data have been scarce and what is available (audit data) represents a highly selected sample of the population subject to audit. Another reason might be that the theoretical auditing literature largely ignores a key aspect: incentivizing auditors.

The issue of incentivizing auditors is moot in most of the auditing literature, due to the assumption that there is a unitary actor (a single auditor or regulator) who can effortlessly commit to any auditing policy. However in many environments, enforcement is not actually carried out by a unitary actor but by a multitude of individual "auditors" (police officers, tax inspectors, etc.) whose individual behavior has negligible impact but whose aggregate behavior generates deterrence. Little is known about how these auditors are incentivized in these organizations. Theory suggests that incentivizing these "atomistic" auditors is tricky.

Example 1 ***(Difficulty of incentivizing auditors)*** *A tax authority faces a distribution of firms with unknown income, each of which makes a report and pays taxes based on it. The tax authority chooses which firms to audit based on the firms' report, and subject to a budget constraint on audits. A well-known result (see Border and Sobel 1987) is that the strategy that maximizes tax returns is the following "extremal" strategy: all firms which report less than some threshold* T *are audited with a probability large enough that no firm wants to underreport below that threshold; and, no firm which reports at or above the threshold is audited. Given this strategy, no firm is ever caught underreporting (all cheating firms make sure to report no less than, and in fact exactly* T*). Suppose now we wanted to decentralize this auditing policy, which means we want individual auditors to carry out the required audits at a cost, say, of* ε *per audit. If we reward these auditors based on the evasion they uncover, then they will want to deviate from the deterrence-maximizing auditing policy and audit some firms which report at or above the threshold. If we don't reward them at all, then they would shirk, and no evasion would be deterred or detected.*

This example illustrates the challenge in decentralizing incentives: whereas in the familiar (single-agent) agency problem "selling a share of the firm" to the agent is an effective incentive scheme, when there are many agents this need not be so. In our case, rewarding auditors based on detection does not automatically promote the agency's mission and may, to a degree, conflict with it. Now, if theory tells us that decentralizing deterrence is tricky, how should we expect real world organizations to decentralize deterrence? The literature is silent on this issue.

This paper takes aim at these two gaps in the literature: both estimating an equilibrium auditing model, and finding out how deterrence is actually decentralized. We take the viewpoint that the incentives given to the population of auditors cannot be entirely observed, in part because these incentives are implicit, that is, given through promotion or through non-fully contractible schemes. With this premise in mind, we do the following:

1. We introduce a general auditing game in which deterrence arises as the result of equilibrium play by many atomistic auditors. The model is general enough to encompass settings such as tax auditing, police searches, and selective prosecution. We provide an empirical test to diagnose the exact form of the incentives given to the individual auditors. The test is based on the properties of the equilibrium.

2. We show theoretically that, in this auditing game, rewarding individual auditors based on their "marginal contribution" to the organizational mission is not the best way of promoting the institutional mission. For example, in the context of a tax auditing game we show that rewarding auditors in proportion to the amount of tax evasion they detect (which is their marginal contribution to the institutional mission of minimizing aggregate amount of taxes evaded) may be inferior to rewarding them every time they detect a cheater. This result suggests that the simple incentive scheme which rewards the detection of cheaters need not be a bad strategy for decentralizing deterrence.

3. We then restrict attention to the special (and important) case of tax auditing. We develop a new game-theoretic model in which auditors are rewarded for detecting cheaters.

4. We provide a method for calibrating this new tax auditing model. The calibration is based on audit data. These data represent a highly selected sample: very few firms are audited, and the decision to audit depends both on the auditor's strategy and on the firm's reported tax base. Our method uses the structure of the equilibrium to correct for these selection biases.

5. We apply the tools developed above to auditing data from INPS, the Italian agency that audits firms to ensure that they paid their labor taxes. First we apply the diagnostic test mentioned in part 1 and find that we cannot reject the hypothesis that INPS auditors maximize the number of cheaters detected. This empirical finding is supported by the theoretical finding mentioned in part 2. Then, under the assumption that the INPS data are generated by the equilibrium of the game mentioned in part 3, we use the calibration method mentioned in part 4 to back out the unobservable deep parameters of the model, such as the distribution of true tax bases of firms.

6. We do the counterfactual exercise of asking how much more revenue INPS could collect if it somehow could solve the decentralization problem and, rather than leaving it to its auditors to choose whom to audit, it could centrally implement a "deterrence strategy" inspired by the literature on optimal auditing. This is the strategy introduced in Example 1: only firms which report below a threshold are audited, and these are audited with high probability. Based on the "deep parameters" recovered in part 5, we find that switching to such a "deterrence 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 the calibration method of part 4 suggests that INPS is already capturing more than 80% of the theoretical maximum revenue attainable, and so there is little room for improvement.

1.1 Contributions of the paper

The main contribution of the paper is to connect auditing theory with data. We develop an equilibrium model, and a method for estimating the key parameters of that model from auditing data alone. The method takes careful account of the multiple sources of equilibrium selection that generate the auditing data.

We believe our paper is the first to bring to data any game-theoretic model of auditing. Our analysis suggests that the ideas developed in the theoretical auditing literature can actually inform empirical research, although too little emphasis has been placed until now on decentralized models, that is, agency models in which auditing is carried out by many agents.

Applying the model and method to Italian auditing data provides "proof of concept:" the methods are practical and tractable. We are able to provide an estimate of tax evasion based on (non-random) audit data alone.

Counterfactual simulation of the model quantifies the costs and benefits of alternative auditing policies. We compare decentralized enforcement with a counterfactual "commitment policy," and compute the loss from the former. Thus we are able to quantify the costs of decentralizing enforcement.

Aside from these "big picture" contributions, the technical results contained in Proposition 1 and Section 5 are novel.

1.2 Related literature

There are many theoretical models of auditing.[1] Almost always these models assume that the tax agency can directly commit to an auditing policy. These models, therefore, do not address the issue of decentralized enforcement. However there are two exceptions: Scotchmer (1986) and Erard and Feinstein (1994) study the equilibrium of a game in which the tax agency cannot commit. The key difference with the model analyzed in Sections 5 and 6 is that our auditors maximize the probability of finding a cheater, whereas in these other works the agency maximizes the expected returns from auditing. This is an important difference: not only are the testable implications different but, furthermore, there are normative reasons to explore our set of assumptions (refer to Example 6).[2] None of the models in this literature explores empirical applications.

There are a few theoretical models of delegated auditing. Melumad and Mookherjee (1986) study the optimal incentive scheme to be used in delegating to a single auditor. In their case the auditor has an impact on the aggregate, and so she can be incentivized by conditioning her compensation on aggregate outcomes. We focus instead on the atomistic auditors case where such incentives would not be effective. Sanchez and Sobel (1993) study a model of delegated tax auditing in which the (single) auditor seeks to maximize tax revenue, whereas the principal also cares about the distributional impact of taxes, i.e., whether the rich or the poor bear the burden. This divergence of objectives leads the principal to underfund the auditor's budget.

The idea that delegated deterrence is imperfect is not new. In the context of police enforcement, Persico (2002) shows that if individual police officers maximize successful interdiction then crime need not be minimized. This result can be interpreted as exposing the limits of decentralized deterrence. In the context of antitrust enforcement, Harrington (2010) makes a related point; he shows that the objective of an antitrust authority to maximize the number of successfully prosecuted cartels can be at odds with the social objective of minimizing the number of cartels that form.

The identification result presented in Proposition 1 is somewhat related to a test for racial bias developed by Knowles et al. (2001). See also Anwar and Fang (2006) for a different but related test for racial bias. The connection is that racial bias in these papers is a parameter in the police objective function, and so these tests address a special case of the question addressed in Proposition 1. Relative to this strand of the literature, the result of Proposition

[1] We define an auditing game as a game of incomplete information 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 firm is audited and did not report honestly. The first model of optimal auditing is Border and Sobel (1978), and many variants have followed. See Andreoni and Feinstein (1998) for a survey of work in this area.

[2] Example 6 shows that it may be better for a principal to make his agents maximize the probability of finding a cheater rather than the expected returns from auditing, even if the principal maximizes the latter!

1 is not parametric (it is about estimating an unknown function h rather than a parameter) and, also, is derived in a much more general environment.

2 A General Auditing Game

We now present a rather general framework that can encompass most auditing games which are present in the literature. For expositional convenience we start by describing a simple version of the game where we assume only one audit class, then extend it to allow for several audit classes.

2.1 The framework, simplified

The players are: a mass of auditors and a mass of auditees, both with measure 1. Each auditee has a true type x and reports a number r. The auditee's true type is unobserved by the auditor unless the auditee is audited. These true types are distributed with density $f(x)$. The function $\pi(x, r, p)$ denotes the auditor's expected payoff from auditing with probability p an auditee who reports r and has type x. The function $\kappa(x, r, p)$ represents the expected payoff of a auditees with type x who reports r and is audited with probability p. Each auditor selects a probability $p(r)$ with which he is going to audit reports r, subject to the constraint that the auditor can make no more than B audits. Thus B captures the relative scarcity of auditing resources. The function $r(x)$ denotes the report of an auditee with true type x.

A Nash equilibrium of this game solves the following program:

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

In words, each auditor selects the function $p(\cdot)$ that maximizes his expected payoff, subject to a budget constraint and subject to the constraint that the auditees are best responding to the aggregate function $p^*(\cdot)$ which aggregates the actions of all auditors. Of note, the auditor regards the function $p^*(\cdot)$ as given when choosing his strategy $p(\cdot)$. This assumption reflects the atomistic size of each auditor.[3]

[3] Since the auditors have mass 1, we are justified in defining $p^*(\cdot)$ as we do in the first line of the programming problem.

The budget constraint says that the individual auditor's total effort cannot exceed B. This constraint will hold with equality in equilibrium, and therefore B can be interpreted as specifying a target level of effort which is determined by a principal. Specifically, consider an environment in which the auditors trade off the payoff function π against a cost of effort $c\left(\int_a^b p\left(r\left(x\right)\right) f\left(x\right) dx\right)$. The principal's problem is to set π in order to induce a desired level of effort e^*. This is a classic agency problem in which the principal elicits effort by promising π. After the payoff function π has been set by the principal, the agent's problem reduces to the one we study with $B = e^*$.

We now present some applications of this general framework.

Example 2 *In the tax auditing context let r represent a tax report, x the true tax base of the taxpayer, t the tax rate and θ the penalty that is applied to those who underreport. We take t and θ to be determined exogenously (statutorily). Then we have $\kappa(x, r, p) = p\left(x - tx - \theta\left(x - r\right)\right) + \left(1 - p\right)\left(x - tr\right)$ for $r \leq x$, where κ represents the expected cost of underreporting. Moreover,*

1. *If auditors maximize the revenue from the audits then $\pi(x, r, p) = p\cdot\left(\theta + t\right)\max\left(x - r, 0\right).$*
2. *If auditors maximize the total returns (taxes paid plus revenue from audits) then $\pi(x, r, p) = tr + p \cdot \left(\theta + t\right)\max\left(x - r, 0\right).$*
3. *If auditors maximize the success rate of audits then $\pi(x, r, p) = pI_{(x-r)>0}$.*

Example 3 *In the context of selective prosecution (a district attorney who selects which cases to prosecute), r represents the case's observable characteristics, x the true underlying facts to be ascertained (including whether the crime has been committed), and p the probability that a case with characteristics r is prosecuted. The function $\kappa(x, r, p^*\left(r\right))$ represents the expected cost of misrepresenting as r the case's true type x (cover-up).*

A distinctive feature of selective prosecution is that the crime has been committed already, so there is no question of deterrence. The next example shows how to introduce deterrence into this framework.

Example 4 *In a policing context in which we care about the deterrence effect of searches, we let $\pi(x, r, p) = p \cdot C(x, r, p)$ where $C(x, r, p)$ represent the probability that a citizen of type x who reports type r is a criminal, given that citizens who report r are policed with probability p. Typically we expect the function C to be decreasing in p, due to the deterrence effect. In this framework, therefore, the function C is a reduced form that embeds the potential criminals' behavior. If we wish to ignore the possibility of misrepresentation we may let $r \equiv x$. The function $\pi(x, r, p)$ represents the payoff of a police officer who maximizes the expected return from his searches.*

2.2 Enriching the framework

The model above is sufficiently general to embed most of the theoretical models of strategic auditing. For empirical purposes, however, it is important to enrich the model by considering several extensions.

Several auditors The auditees may be subject to simultaneous auditing by other auditors, over and above the auditor that is the focus of out interest. For example, an Italian firm is not only subject to INPS audits, but also to income tax audits carried out by a different auditor, the Guardia di Finanza. In these cases we interpret the function κ as expressing the auditee's incentives to misrepresent its income *after taking into account* all the other "extraneous" audits.

Several audit classes It is important to allow for the presence of several audit classes in which the auditor classifies auditees according to characteristics which are observable to the auditor. (We will also have to worry that we, the researchers, may not be able fully to distinguish these audit classes; more on this later.)

We assume that the auditor classifies auditees into different audit classes according to any number of auditee characteristics which are observable to the auditor. An audit class is simply a group of auditees who share a specific combination of observable characteristics. Auditees that belong to a given audit class are distinctive in the eye of the auditor due to the distribution of their type, which the auditor uses to make inference. 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 $\sum_k G(k) = 1$. Conditional on being in class k, the type of auditees is distributed according to the probability density $f_k(x)$. A auditee from audit class k faces a class-specific audit schedule $p_k(\cdot)$.

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

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

Where ξ is the realization of a random variable Ξ_k with distribution $v_k(\xi|x, r)$.

In the tax auditing context, introducing ξ 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 $\xi > x$). 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 auditee'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).

Residual heterogeneity of auditees In the base model the only unobserved heterogeneity of auditees is x, their type. A (somewhat stark) implication is that all auditees with the same type report in the same way. We can relax this assumption by simply assuming that the auditee'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 portion of the type which is payoff-relevant for the auditor (in the tax auditing case, the firm's true tax base). The remaining dimensions capture additional heterogeneity which impacts the auditee's choice of reporting. The auditee's payoff will then be given by $\kappa(\xi, \mathbf{x}, r, p)$, and the auditee'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 auditees 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.

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

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

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

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

where the integrals are understood to be of multiple variables, over the N dimensions of $\mathbf{x}$.

3 Identification of the Auditors' Objective Function

We take the viewpoint that the auditors' incentives are often implicit and anyway cannot be observed directly. Our goal in this section is to use the available data to learn the objective functions that auditor and auditees are maximizing, i.e., the functional forms π and κ. We call this the *identification problem*. Proposition 1 at the end of this section provides an answer to the identification problem.

For ease of exposition, in this section we proceed as if r can only take integer values (dollars, or cents, in the tax auditing framework) and ξ also can only take integer values.[4]

3.1 Assumptions

Ideally, we would like the identification strategy not to depend on knowledge of the fine details of the problem (i.e., our knowledge of, or assumptions about the distributions $G(k)$ and f_k, for example). In the same spirit, 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 auditees 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 some (relatively weak) assumptions about the structure of the game and about what features of the data we can observe.

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

Assumption 1 ***(Deterrability)*** *For all k, if $p_k(r) = 1$ then no auditee in class k with $x_1 > r$ chooses to report r.*

Assumption 1 says that, if r is audited with sufficiently high probability, then the auditee'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. Although the assumption is stated in "behavioral terms," it can easily be restated (though more cumbersomely) in terms of primitives. We chose not to do so to make the statement more transparent.

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 auditees 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 "exaggerates" relative to x_1, Assumption 2 can hold if there is a cost of auditing.

Assumption 3 ***(Auditor's marginal reward to effort is constant)*** *$\pi(\xi, r, p)$ is a linear affine function of p, that is,*

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

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

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

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 auditee who reports r and is not audited; in the tax auditing context, this would be the tax paid before the audit, so it makes sense to assume that it is nonnegative. Assumption 3 embodies the notion of atomistic auditors. If an auditor is atomistic, then the return to his action p must be linear in p. We view Assumption 3 as not very restrictive; this assumption holds in Example 2 above, and in most models featuring decentralized deterrence. In any case, this assumption can be tested with data, as discussed in the next Proposition.

What we cannot observe: latent audit classes We allow for the possibility that we, the researcher, are only able to observe coarse partitions encompassing several audit classes. We will denote these partitions by K_i. For example, the set of audit 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 auditee belongs to.

Take any function $h(\xi, r)$. Think of it provisionally as the return from auditing a auditee 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 auditees 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

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

and we set $\overline{h}(r, K_i) = 0$ when its denominator is zero. The quantity $\overline{h}(r, K_i)$ is to be interpreted as the average return, as computed from the data, from auditing a auditee 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 an attractive feature of the methodology we propose.

We think of each point in our data as an i.i.d. realization of a random vector generated by the equilibrium behavior of auditees 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 auditee 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 auditee in audit class k to report r. The term $G(k) f_k(\mathbf{x}) p_k^*(r)$ represents the probability that a auditee belongs to audit class k and has a true tax base $\mathbf{x}$ which in equilibrium leads the auditee 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, avoid making assumptions. Expression (3) is the limit in probability of $\overline{h}(r, K_i)$ as the sample size grows large.

3.2 Identification result

We are now ready to present our identification result. The next proposition 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. Intuitively, an auditor 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 auditees respond to auditing, to an equalization of the auditor's margins across all audited classes.

Proposition 1 *If one can reject the hypothesis that $E\left[\overline{h}(r, K_i)\right]$ 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 $h(\xi, r)$ can be found such that one cannot reject the hypothesis that $E\left[\overline{h}(r, K_i)\right]$ 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 $E\left[\overline{h}(r, K_i)\right]$ 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.

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

Let $r_k^*(\mathbf{x}; p_k^*(r))$ denote the reporting strategy that solves (4). Since condition (4) involves $p^*(r)$, not $p(r)$, the behavior of auditees 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 Lagrangian for the auditor's problem:

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

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

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

The first term of the Lagrangian can be written as

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

As the Lagrangian 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 $E\left[C(\Xi_k, r)|\mathbf{x} \in X_k^*(r), r\right] \ge \lambda$.

Now, suppose by contradiction that the strict inequality $E\left[C(\Xi_k, r)|\mathbf{x} \in X_k^*(r), r\right] > \lambda$ 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

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

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

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

Now, remember that from (3) we had

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

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

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

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

This proposition provides a straightforward identification strategy: 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 $C(\xi, r)$. 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 $G(k)$, $f_k(\cdot)$, or $p_k^*(\cdot)$, or even solve for the equilibrium behavior of auditees.

Of note, this identification strategy is agnostic about the objective function of auditees. This is convenient in that it is not necessary to make specific assumptions about the nature of the auditees' 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 auditees might be maximizing.

Finally, we acknowledge that Proposition 1 has a slight "data mining" flavor: since little structure is imposed on the function h, there is bound to be some h which satisfies the independence requirement. Therefore the value of the identification strategy rests on how reasonable the resulting h is. There are several requirements that can increase our confidence that the specific h we find is not the outcome of data mining. First, we can do additional "placebo" tests, showing that the expression $E\left[\overline{h}\left(r, K_i\right)\right]$ is *not* independent of some other characteristic other than those which the auditor cannot arbitrage over. (The theory would not predict such independence). Second, we can ask that the h we find be "simple," such that it could actually be used in a real-world compensation scheme. Third, one can ask that it be "plausible:" that there be a theoretical justification for why the atomistic auditors would be endowed with that specific h. In our empirical application we will apply these three criteria.

4 Setting the Auditors' Incentives

Proposition 1 demonstrates how the data can be used to identify the objective function which is actually pursued by auditors. In this section we ask what objective function the principal would actually *want* to give the auditors. We do not seek a full characterization of the optimal objective function, because the characterization is sure to be highly sensitive to modeling assumptions and thus, we feel, of little practical relevance. We ask, instead, a more limited question. Taking the perspective of a principal who is charged with implementing an institutional mission, we ask whether it is optimal for the principal to give individual auditors "a stake" in the institutional mission. We show that, generically, the answer is no. In fact, in the context of tax auditing we show that it may be better to incentivize the auditors based on the *number of cheaters* they have uncovered, even though this is not "a stake" in the institutional mission, which is to minimize the *amount* of taxes evaded. This answer is in contrast with the standard prescription from single-agent agency theory, which is that, in the absence of "frictions" such as differences in risk aversion, etc., the optimal incentive scheme is to "sell the firm" (or at least a stake in it) to the agent. The root of the difference lies in the fact that, when the agents (auditors) are many, giving the auditors "a stake" in the institutional mission does not account for the deterrence effect.

4.1 The centralized problem

Let us start with the *centralized* problem—a planner's or principal's problem, unconstrained by the decentralization incentives. The principal sets an aggregate auditing strategy $p\left(r\right)$ which specifies the probability of being audited for an auditee who reports r, to maximize the average of a "mission function." The mission function $\pi^M(x, r, p)$ embodies the institutional payoff from auditing with probability p an auditee who has type x and reports r. In the tax auditing framework, for example, the conventional assumption about the institutional mission is maximization of total returns so that $\pi^M(x, r, p) = tr + p \cdot (\theta + t) \max\left(x - r, 0\right).$ Formally, the equilibrium of the centralized auditing problem is defined by the following constrained maximization problem.

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

The principal's choice of strategy is subject to a budget constraint, and also subject to the constraint that the auditees are best responding to the aggregate function $p(\cdot)$. Notice a key difference between constraint (COMM) and constraint (NOCOMM) on page 6: whereas the latter involves $p^*(\cdot)$, the one in the present problem involves $p(\cdot)$. In words, equation (NOCOMM) represents the case in which the auditor cannot affect the aggregate auditing schedule, whereas equation (COMM) represents the case in which it can.

The first order conditions of the Lagrangian associated to the centralized problem are given by

$$
\frac{\partial}{\partial p(\widehat{r}(z))} \int_a^b \left[\pi^M(x, \widehat{r}(x), p(\widehat{r}(x))) - \lambda p(\widehat{r}(x))\right]\; f(x)\, dx = 0 \text{ for all } z,
$$

where $\widehat{r}(z)$ denotes the solution to the incentive compatibility constraint (COMM). Problem (COMM) depends on the function $p(\cdot)$, a property which reflects the deterrence effect of choosing a policy p. The first order conditions of the Lagrangian can be rewritten as follows.

$$
\left[\pi_3^M(x, \widehat{r}(z), p(\widehat{r}(z))) - \lambda\right] f(z) + \int_a^b \pi_2^M(x, \widehat{r}(x), p(\widehat{r}(x))) \frac{\partial \widehat{r}(x)}{\partial p(\widehat{r}(z))}\; f(x)\, dx = 0 \text{ for all } z.
\tag{7}
$$

The term $\partial \widehat{r}(z) / \partial p(\widehat{r}(z))$ inside the integral reflects the deterrence effect: changing the audit schedule $p(\cdot)$ at the single point $\widehat{r}(z)$ affects the report of types z. Moreover, all the reports by the other types are also affected through non-local effects on the incentive compatibility constraint (COMM). The integral accounts for the non-local changes.

4.2 The decentralized problem

By comparison, the first order conditions for the decentralized problem on page 6 are given by

$$
[\pi_3(x, r^*(z), p(r^*(z))) - \lambda]\, f(z) = 0 \text{ for all } z,
\tag{8}
$$

where $r^*(z)$ denotes the solution to the incentive compatibility constraint (NOCOMM). Unlike in (7) there is no integral term in this equation because in the decentralized problem no individual auditor can affect $r^*(z)$.

4.3 Why giving the auditor a stake in the agency mission is not the optimal incentive scheme in the decentralized problem

Suppose that the auditor in the decentralized problem is incentivized at the marginal rate of π_3^M, which defines the *auditor's individual contribution to the institutional mission***.** Then in equilibrium equation (8) needs to hold with $\pi_3 = \pi_3^M$. But then generically equation (8) *cannot* hold. This simple observation is recorded next.

Conclusion 5 *Setting up a reward system based on the auditor's individual contribution to the institutional mission, will generally not implement the solution of the centralized problem. The reason is that the marginal contribution of an atomistic agent does not account for the deterrence effect.*

A practical implication of this observation can be seen in the tax auditing framework. In that framework the conventional assumption is that the institutional mission is maximization of total returns from taxation: $\pi^M(x, r, p) = tr + p \cdot (\theta + t) \max(x - r, 0)$. >From this it follows that $\pi_3^M(x, r, p) = (\theta + t) \max(x - r, 0)$; this means that rewarding the auditor's individual contribution to the institutional mission means, in fact, rewarding the auditor in proportion to the amount of tax evasion he uncovers. As we have seen, there is no presumption that this should be an optimal scheme, and indeed, there should be a presumption that the principal should be able to do better than this. We now present an example in which another "simple" incentive scheme, namely rewarding agents based on the number of cheaters they find, regardless of the magnitude evaded, does better in terms of the institutional mission: it yields higher total returns from taxation.

Example 6 *Let's consider the following set-up introduced in Erard and Feinstein (1994). The tax base x is uniformly distributed on $[0, 1]$. The tax rate is $t = 50\%$ and the fines are $\theta = 50\%$ of the underreported amount. The resources of the auditor are such that $B = 10\%$ of the reports can be audited. Suppose that there is a fraction $\lambda = 50\%$ of firms that report honestly their tax base independently of the auditors' strategy. Let's assume first that auditors maximize the amount of evasion he uncovers, which means that an auditor's objective function is given by $\pi_3^M(x, r, p) = (\theta + t) \max(x - r, 0)$. Then in equilibrium tax payers underreport in by a constant amount T, so that $r = x - T$ (they report 0 if $x \leq T$) and the probability of auditing a firm which reports r is given by $p(r) = \frac{1}{2}\left(1 - \exp\left(\frac{r-(1-T)}{T}\right)\right)$. The total tax revenue (taxes +fines) associated with the auditor's objective is $R = 0.174$. See Appendix A for details.*

Let's now consider the alternative objective of maximizing detection, $\left(\pi_3^M(x, r, p) = pI_{(x-r)>0}\right)$. In equilibrium, tax payers underreport by a constant ratio, $r = \frac{\alpha}{\alpha+1}y$. The auditing strategy consists in a probability of auditing reports $p(r) = \frac{1}{2}\left(1 - \left(\frac{\alpha+1}{a}r\right)^{\alpha}\right)$. The revenue associated is $R = 0.179$. See Appendix A for details.

The revenue is 3% larger when the auditor maximizes detection than when he maximizes the revenue generated by the audits. The intuition for why auditors who seek to maximize detections (MD) create more deterrence than those who seek to maximize returns from audits (MR) is as follows. MR leads auditors to audit large firms (more precisely, firms who report large amounts, which in a monotone strategy equilibrium is the same thing), because these firms are more likely to evade more taxes. This "size effect" is absent if auditors MD. Therefore, we should expect that in equilibrium auditors who MD are more willing to audit small firms, compared to auditors who MR. Therefore when auditors MD the aggregate audit strategy places greater probability of auditing on small firms. This means that the aggregate audit strategy under MD is closer to the extremal strategy of the Border-Sobel setup with commitment discussed in Example 1.

Example 1 can be interpreted as showing that if a principal seeks to deter tax evasion, it can be better for that principal to reward his agents for detecting cheaters, rather than for finding large underreports. To draw this conclusion, it is necessary to consider the wage bill that the principal needs to pay. We assume that the principal sets the rewards in both systems so as to exactly compensate the auditors for their effort, but only just. The auditor's effort in Example 1 is represented by the 10% fraction if firms which is audited. Since the effort is kept constant when we compare the two incentive schemes, the wage bill for the principal will also be constant. Therefore Example 1 can be interpreted as showing that, when the principal switches to rewarding detection the wage bill remains the same, and deterrence increases.

5 A Special Auditing Game: Tax Auditing to Maximize Detections

This section develops and analyzes a new theoretical model of tax evasion and enforcement. Relative to the general model introduced in Section 2, the tax auditing model is special in that the auditees (firms) are assumed to pursue a very specific objective, which is to minimize the amount of taxes paid (inclusive of penalties for detected underreporting). We also initially restrict attention to a single audit class.

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 a probability schedule $p(r)$ that relates the report r to the probability of being audited. Firms are aware that in case of an audit, the true characteristic x will be discovered and then taxes will be assessed on the true level x and a penalty added which is proportional to the amount of evasion. Thus 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 construct 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 the auditor maximizes the number of successful audits. The first assumption situates this model as a special case of the decentralized model analyzed in Section 3; 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

$$p(r)(x - tx - \theta(x - r)) + (1 - p(r))(x - tr). \tag{9}$$

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 (9) are necessary conditions for a maximum. They are

$$p'(r)(r - x)(t + \theta) + p(r)(t + \theta) - t = 0, \tag{10}$$

which can be rewritten as

$$x^*(r) = r + \frac{p(r) - \frac{t}{\theta+t}}{p'(r)}, \tag{11}$$

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:

$$(t+\theta)\left[p''(r)(r-x)+2p'(r)\right]\leq 0. \tag{12}$$

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

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

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

$$\frac{\lambda f(r)}{\lambda f(r)+(1-\lambda)f(x^*(r))\frac{dx^*(r)}{dr}}=q \tag{14}$$

Denoting

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

we can rewrite (14) as

$$f(x^*(r))\frac{dx^*}{dr}=\frac{\lambda(1-q)}{(1-\lambda)q}f(r)=\gamma f(r). \tag{15}$$

5.3 Equilibrium

Let us start by establishing the support of the reporting strategies must be an interval 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 must be 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

$$p^*(\rho^*(b)) = 0. \tag{16}$$

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 tiny 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 (11) and (12), the auditors' indifference condition (15), and the boundary condition (16).*

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:*

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

Proof. Let us first characterize the equilibrium reporting strategies. Integrating both sides of (15) yields

$$\gamma F(r) = F(x(r)) + k. \tag{17}$$

Since $x^*(a) = a$, it follows from (17) evaluated at $r = a$ that $k = 0$. Therefore, for a generic $r > a$ we have

$$\begin{aligned} x^*(r) &= F^{-1}(\gamma F(r)), \text{ or equivalently} &(18)\\ \rho^*(x) &= F^{-1}\left(\frac{F(x)}{\gamma}\right). &(19)\end{aligned}$$

We now characterize the equilibrium auditing schedule. Using (18) to substitute into (11) we get

$$\frac{p^{*\prime}(r)}{p^*(r) - \frac{t}{\theta+t}} = \frac{1}{F^{-1}(\gamma F(r)) - r}. \tag{20}$$

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 (16).

Finally, Lemma 3 in Appendix A verifies that, given the audit schedule $p^*(\cdot)$, the firm's reporting strategy $\rho^*(\cdot)$ satisfies the second order conditions (12). ■

Equation (19) shows that a larger value of γ corresponds to a lower value of $\rho^*(x)$, that is, greater underreporting in equilibrium. This makes sense: a large γ corresponds by definition to a low value of q, which means a high marginal cost of an audit. Equation (19) also implies that, in any equilibrium in which there is some underreporting, γ must be greater than 1. Finally, equation (19) 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. (For a collection of results related to log-concave distributions 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 A. ■

6 Calibration Methodology for the Tax Auditing Model

For the purpose of calibrating the model we specialize the analysis to the case in which the firms' tax base is distributed according to a Power distribution. The c.d.f. of a Power distribution on $[a, b]$ is given by $F(x) = \left(\frac{x-a}{b-a}\right)^{\beta}$ with $\beta > 0$. The power distribution is log-concave, and its density can be decreasing or increasing depending on whether β is smaller or greater than 1. The uniform distribution arises when $\beta = 1$. Despite being a one-parameter distribution, the Power distribution will prove sufficiently flexible for the purpose of matching our data.

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. See Appendix A. ■

We want to use the equilibrium strategies in Proposition 3 to calibrate the unknown parameters λ, α, β.

For realism's sake we need to introduce audit classes in our calibration exercise. This is because it is unrealistic that a huge firm (automobile production, say) could report just one employee and fool the auditor into thinking that it is a mom-and-pop store. Therefore we need to incorporate the possibility that firms are observably different to the auditor, so that a report of 1 worker from General Motors triggers an audit for sure, whereas a report of 1 worker from a mom and pop store might not. We therefore introduce 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 otherwise free to report anywhere within (a_k, b_k). In this formulation, 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. General Motors is presumably in an audit class where a_k equals thousands of employees, and so this formulation avoids the possibility of GM reporting very few employees.

Unfortunately, we may often not know what audit classes the auditor assigns firms to. Therefore, in our calibration exercise we choose to be crude in the way we incorporate audit classes. We define an audit class as all firms which would be audited with positive probability if they report within a given interval (a_k, M_k). We can then partition the set of reports into contiguous non-overlapping intervals $(a_k, M_k), (a_{k+1}, M_{k+1}), \ldots$ thus partitioning the entire set of reports into distinct audit classes. In this way, each audit class is associated with an interval of audited reports.[5]

Having identified an audit class with an interval (a, M), we now we detail the methodology used to infer all relevant parameters of the audit class. The procedure is based on matching moments from audit data. The first moment is the fraction of audited firms who report in the interval (a, M) and are found *not* to have underreported.[6] According to the model, the ratio of honest to strategic firms *among those audited* is given by

$$\frac{\lambda \int_a^M p^*(r)\, dF(r)}{(1-\lambda) \int_a^M p^*(r)\, dF\left(\rho^{*-1}(r)\right)} \tag{21}$$

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

[5] This definition is convenient for calibration purposes, but it is somewhat unnatural from the viewpoint of the auditors. For the auditors, an audit class is defined in terms of some observable characteristic, which then gives rise to a specific distribution of true tax bases. From their perspective, our construction amounts to imposing that the distribution of true tax bases in the audit class which we associate with (a_k, M_k) is in fact (a_k, b_k), where $b_k > M_k$ because $M_k = \rho^*(b_k) < b_k$. From the viewpoint of an auditor then, the structure we have given results in separate audit classes each characterized by a Power distribution with support (a_k, b_k) and parameter β_k. Each interval (a_k, b_k) partly overlaps with the next interval (a_{k+1}, b_{k+1}), but in equilibrium no strategic firm in audit class k and true tax base in (a_{k+1}, b_k) reports above $M_k = a_{k+1}$.

[6] 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.

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

$$(1-\lambda)\int_a^b \rho^*(x)\,dF(x) + \lambda \int_a^M r dF(r). \tag{22}$$

The third moment we match is the average underreport of audited firms. We call this quantity C_3. The theoretical expression that corresponds to this amounts is

$$\int_a^b [x-\rho^*(x)]\,dF(x). \tag{23}$$

Equations (21)$=C_1$, (22)$=C_2$, and (23)$=C_3$, 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 A.2), the system of equations is reduced to the following.

$$\begin{aligned} \frac{\lambda}{(1-\lambda)\left(\frac{\alpha+1}{\alpha}\right)^{\beta}} &= C_1 \\ (1-\lambda)(1+C_1)(a+\alpha C_3) &= C_2 \\ \frac{(M-a)}{\alpha}\frac{\beta}{\beta+1} &= C_3 \end{aligned} \tag{24}$$

The parameters a, M, C_1, C_2, C_3 will be empirically observable, and the unknown parameters are λ, α, β. The system of equations (24) represents our calibration tool. If a solution exists to this nonlinear system of equations, then the solution identifies the "deep parameters" of the model without error. If a solution does not exist, then we may look for estimates $\widehat{\lambda}, \widehat{\alpha}, \widehat{\beta}$ which minimize some weighted sum of the distances between expression (21) and C_1, expression (22) and C_2, and expression (23) and C_3.

7 Illustrative Application: The INPS Auditing Data

We view the preceding analysis as developing a replicable, almost mechanical procedure for the structural estimation of decentralized auditing games: measure (the auditors' incentive scheme using Proposition 1), cut (build a model and solve for the equilibrium, as we did in Section 5), and fit (calibrate the model using audit data). In this section we illustrate the procedure by applying it to the INPS data.

7.1 The INPS data

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).[7] Every year the Italian Social Security Institute (INPS) inspects a number of firms in order to verify that they paid their labor taxes. An employer found underreporting is assessed a fine equal to the money underreported plus 33% of it. 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. Audits are carried out by auditors who select the firm out of a large list of firms, visit the firms' location and check for violations. The auditor can interview the workers he finds and check administrative and accounting records. 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 (length 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%, 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 *amount of evasion* (*evasioni)* is the amount of money that INPS assesses it is owed, if any. The variable *success* (*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 *sectors (settori)* codes the ATECO industry sector codes to which the audited firm belongs.[12]

[7] For most workers these taxes amount to 40-42% of gross wages, but they are 38% for workers classified 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.

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

[9] For a more precise description of the data and of the process of building the dataset, see Di Porto (2009).

[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.

[12] There are nine such sectors, with the numbers from 1 to 9 corresponding to, respectively: Energy, Water, Gas; Mining and Chemical Industries; Manufacturing and Mechanical Industries; Food and Textiles; Construction; Wholesale and Retail Trade; Transportation; Finance and Insurance; and, finally, Services. They correspond to the 1981 version of the ATECO codes.

In order to be consistent with the theoretical framework of optimal auditing, our sample should only contain audits which are discretionarily initiated by INPS with the goal of uncovering underreports. However, the administrative process that generates our data is multi-faceted, and thus we need to decide what to do with audits that are not discretionarily initiated by INPS with the goal of uncovering underreporters. Our strategy will be to exclude them from the sample. We detail this process in Appendix B.1. It is important to note that the rationale for eliminating non-discretionary audits is not (only) to err on the side of caution; this strategy also has a theoretical justification, because eliminating these audits does not invalidate the analysis we intend to carry out. As mentioned on page 8 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.

After eliminating non-discretionary audits 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: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. 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.

[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**
Sector 1: Energy, water, gas	453 (0.3%)	83 (0.3%)
Sector 2: Mining and chemical industries	1,600 (1%)	656 (2,7%)
Sector 3: Manufacturing and mechanical industries	8,563 (5.7%)	3,303 (13.7%)
Sector 4: Food and textile	16,807 (11.1%)	4,419 (18.3%)
Sector 5: Construction	35,427 (23.3%)	6,910 (28.6%)
Sector 6: Wholesale and retail trade	69,385 (45.7%)	5,087 (21%)
Sector 7: Transportation	1,587 (1%)	775 (3.2%)
Sector 8: Finance and Insurance	5,927 (3.9%)	1,684 (6.7%)
Sector 9: Services	12,057 (7.9%)	1,268 (5.2%)
Total	151,806 (100%)	24,185 (100%)
Success of audit (Risultato) (std. dev.)	0.40 (0.49)	0.54 (0.50)
Amount of evasion (Evasioni) in euros (std. dev.)	6,953 (35,950)	57,601 (306,704)
Amount of evasion conditional on > 0 (std. dev.)	17,683 (55,650)	108,297 (413,972)
Number of employees (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 distribution of reported sizes is given in Figure **??**.

7.2 Using Proposition 1 to identify auditors' objective function

We check whether there is any objective h which is being equalized across audit classes. Inspired by Example 6, we conjecture that this objective is the detection of a cheater, independent of the amount of cheating.[14] Accordingly, the dependent variable in Table 2 is the fraction of audited firms which are found cheating in a given region/year. The independent variables are the fraction of audited firms in each region/year which belong to each sector. In specification (2) we control for reported firm size, to check whether the return from auditing large firms is different. In both specification we introduce region fixed effects to capture the idea that the game is being played within region, so an inspector from one region cannot arbitrage by auditing a firm from another region.

According to the theory, whatever heterogeneity determines the variation in the dependent variables within a region, observables should not predict the probability of finding a cheater. Table 2 is broadly supportive of this statement since no sectoral variable is significant in either specification (1) or (2). In fact, a Wald test carried out on specification (2) does not reject the hypothesis that the sectoral coefficients are jointly equal to zero. However, the coefficient on the dummy "large firm" is significant in specification (2), indicating that firms which report more than 10 employees are more likely to be found cheating. This is an indication that firm size matters.

The large R^2 in Table 2 is attributed to region fixed effects. This is in line with the theory: institutional considerations suggest that auditors should not be able to arbitrage across regions, and the data bear this out. From a statistical viewpoint, we learn that there is variation to be explained, but sectors don't explain it; in this sense, the significance of region fixed effects serves as a a sort of "placebo test" for the non-significance of sectoral variables.

Albeit at a very exploratory level, Table 2 reveals that firm size matters. To further explore the question of firm size, we split the sample into large and small firms. Moreover, we make each individual audit into a data point, with the result that now we have a very large number of observations. Table 3 reports the results for this regression. The dependent variable in the regression of Table 3 is the *success* variable (*risultato).* 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 employees (*dipendenti)* and for any sector (the variables *sector#*). In our empirical specification we allow for additional flexibility by controlling for the interaction between number of employees and sectors (the variables *sector# * employees*). This allows for the probability of detection to vary with the number of employees in a sector-specific way.

[14] In the language of Section 3 we posit that $h(\xi, r) = 1$ if $r < \xi$, and zero otherwise.

Table 2:
Dependent Var: Fraction of Audited Firms Found Cheating

VARIABLES	(1)	(2)
Fraction of audits in sector 1	2.33	1.39
	(0.121)	(0.489)
Fraction of audits in sector 3	-0.43	1.13
	(0.776)	(0.485)
Fraction of audits in sector 4	-1.14	0.99
	(0.391)	(0.389)
Fraction of audits in sector 5	-0.86	1.00
	(0.512)	(0.459)
Fraction of audits in sector 6	-1.12	1.00
	(0.431)	(0.489)
Fraction of audits in sector 7	2.81	2.50
	(0.242)	(0.310)
Fraction of audits in sector 8	-1.21	0.14
	(0.445)	(0.939)
Fraction of audits in sector 9	-1.23	0.98
	(0.389)	(0.443)
Large firm indicator		1.50***
		(0.000)
Time trend	-0.02*	-0.01
	(0.066)	(0.224)
Constant	1.43	-0.73
	(0.296)	(0.588)
Region Fixed Effect	yes	yes
Observations	120	120
R-squared	0.2411	0.387
Number of id	20	20

Robust pval in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Focusing first on small firms (column 1), we find that, despite the flexibility allowed by our specification, few of the coefficients (4 out of 20) are significantly different from zero at the 10% level. Furthermore, the R^2 coefficient is very small (0.01). We interpret this widespread lack of explanatory power as evidence that most of the independent variables in our regression do not 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 3 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 the auditor. Based on Proposition 1, this is consistent with the assumption that auditors maximizes the probability of a successful audit, and that INPS does not have, or does not make strategic use of, the power to commit so that the game is indeed one of decentralized auditing. In Table 7 of Appendix 7.2, we verify the robustness of this finding by controlling for region, for a time trend, and to restricting the population to firms which report 5 or fewer employees. Of note, the region coefficients are generally significantly different from zero, which again we interpret as a successful "placebo" test.

Column (3) is identical except that the dependent variable is *amount of evasion (evasioni)*, the amount of money that INPS assesses it is owed. According to Proposition 1, if auditors maximize detections, then under a mild regularity condition there should be a positive correlation between number of reported employees and amount of the misreport, i.e., firms who report more workers also misreport by more. The coefficient on the number of employees 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. Also, to the extent that most coefficients on the terms involving the number of employees are not zero, column (3) supports the conclusion that auditors are not maximizing the *amount of evasion* detected.

Summing up, for the "small firms" sample, which comprises about 90% of the firms in Italy, the empirical results are generally in line with the theory. The results suggest that INPS behaves, in the aggregate, as a group of auditors which maximize the probability of finding tax cheaters in a decentralized way.

Large firms are different: 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%). This fact was also picked up in Table 1. This implies that, in principle, an auditor 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 is not consistent with the evidence presented in Table 2, however. The regression in column (2) is identical to column (1), 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 *number of employees*, as well as those of the variables interacted

Table 3:
Dependent Var: Detection Success and Amount of Evasion

VARIABLES	(1) Success Small firms	(2) Success Large firms	(3) Evasion Small firms
Number of employees	0.01	0.00***	808.54
	(0.293)	(0.003)	(0.106)
sector 2	0.03	0.07	-876.32
	(0.425)	(0.246)	(0.685)
sector 3	0.03	0.17***	-17.19
	(0.405)	(0.002)	(0.991)
sector 4	0.07**	0.10*	-788.27
	(0.045)	(0.061)	(0.579)
sector 5	0.04	0.09	-1,072.68
	(0.237)	(0.102)	(0.434)
sector 6	0.02	0.11*	-1,552.09
	(0.474)	(0.055)	(0.249)
sector 7	0.26***	0.18***	4,871.02**
	(0.000)	(0.002)	(0.038)
sector 8	0.09**	0.29***	-2,936.28
	(0.018)	(0.000)	(0.222)
sector 9	-0.02	0.17***	-3,803.10***
	(0.585)	(0.004)	(0.006)
sect2 * employees	0.01	0.00	1,471.74*
	(0.364)	(0.169)	(0.053)
sect3 * employees	0.01	-0.00**	1,443.75**
	(0.483)	(0.020)	(0.013)
sect4 * employees	0.01	-0.00**	1,162.01**
	(0.431)	(0.027)	(0.028)
sect5 * employees	0.01	-0.00	603.42
	(0.585)	(0.801)	(0.237)
sect6 * employees	0.01	-0.00*	325.63
	(0.244)	(0.055)	(0.519)
sect7 * employees	-0.00	-0.00	1,958.26**
	(0.821)	(0.650)	(0.013)
sect8 * employees	0.01	-0.00***	3,355.36***
	(0.443)	(0.000)	(0.002)
sect9 * employees	0.03**	0.00**	1,086.04**
	(0.019)	(0.045)	(0.048)
Constant	0.29***	0.41***	3,525.65***
	(0.000)	(0.000)	(0.008)
Observations	151,806	24,185	151,806
R-squared	0.014	0.013	0.017

Robust pval in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

with *number of employees*, 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" 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 *sector 8*, corresponding to Finance 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 from column (c) provides weak support for the no-commitment, success-maximizing model *within the large firm population*.

Separately from the results in Tables 2 and 3, including the "placebo" tests, is the hypothesis persuasive that auditors maximize the probability of finding a cheater, at least among small firms? We think so. First, this compensation scheme is "simple," such that it could actually be used in a real-world compensation scheme. Second, it is "plausible" in the sense that if a principal seeks to deter tax evasion, it can be better for that principal to reward his agents for detecting cheaters, rather than for finding large underreports (see Example 1).

7.3 Backing out the calibrated parameters

The system of equations (24) is our calibration tool. 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 these observable parameters. If a solution exists to this nonlinear system of equations, then the solution identifies the "deep parameters" of the model without error. If a solution does not exist, then we may look for estimates $\widehat{\lambda}, \widehat{\alpha}, \widehat{\beta}$ which minimize some weighted sum of the distances between expression (21) and C_1, expression (22) and C_2, and expression (23) and C_3.

We choose to divide our sample into three audit classes: firms which report in the intervals (0,10), (11-25), and (26-50). 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.

To illustrate how the calibration exercise is implemented, 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 we need a value for C_3, the average number of underreported employees. This we do not have in the data, but we

	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

Table 4: Calibrated parameters.

approximate it by taking the total amount evaded conditional on evading a positive amount (17,683 euros), and dividing it by 26,500 (the gross average salary in Italy in 2007[15]); we thus obtain a proxy for C_3 which in this case equals 0.67 "employee-equivalent underreports."

Given these parameters, the system of equations (24) can be solved numerically to yield $[\lambda = 0.621\,51, \alpha = 4.874, \beta = 0.484\,91]$. Details about the numerical solution are provided in Appendix C. 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 due to the fact that strategic auditing 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.

The same methodology can be applied to audit classes $(a = 11, M = 25)$ and $(a = 26, M = 50)$. In this last audit class, unlike in the first two, solving the system of equations (24) yields a parameter β which leads to a poor fit of (25) to the data. We explain how we dealt with this issue in Appendix C. Finally, all the parameters recovered through this calibration exercise are presented in Table 4.

7.4 Fit of the calibrated model

Given the parameters collected in Table 4, we plot the *predicted distribution* of reports by audited firms, and compare it to the *distribution of reports in our data.* This is a meaningful test of fit because the calibration procedure was *not designed to match the shape* of the report distributions, but rather certain moments of it. Therefore, the forthcoming comparison can give us a sense of how well our model (including the restriction to the Power distribution) fits the data.

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

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

$$p^*(r) \cdot \left[(1-\lambda) f\left(\rho^{*-1}(r)\right) \frac{\partial \rho^{*-1}(r)}{\partial r} + \lambda f(r)\right], \tag{25}$$

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 **??** 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 relaxing the matching of C_2 in (24), which in the figure is referred to as "relaxed identification." We refer to Appendix C for details about the difference between third and fourth panel.

Figure **??** 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 some $b > M$. 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.

8 Counterfactual Exercises With INPS Data

8.1 Counterfactual I: compute the amount of tax evasion

In this section we compute the revenue raised in equilibrium *as a fraction* of the total revenue theoretically achievable if every firm paid their taxes in full.

The maximal amount of tax revenue theoretically achievable in our model is what all firms would pay if they were honest, which is

$$\int_a^b tr \cdot f(r)\ dr. \tag{26}$$

The money actually raised in equilibrium is given by

$$\begin{aligned} &\lambda \int_a^b tr \cdot f(r)\ dr + (1-\lambda) \int_a^M tr \cdot f\left(\rho^{*-1}(r)\right) \frac{\partial \rho^{*-1}(r)}{\partial r}\ dr \\ &+ (1-\lambda) \int_a^M p^*(r) \left[(t+0.33)\left(\rho^{*-1}(r) - r\right)\right] \cdot f\left(\rho^{*-1}(r)\right) \frac{\partial \rho^{*-1}(r)}{\partial r}\ dr. \end{aligned} \tag{27}$$

	Class 0-10	**Class 11-25**	**Class 26-50**
$\widehat{\theta}$	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	$(1-\lambda)\,1.32$	$(1-\lambda)\cdot 6.55$	$(1-\lambda)\cdot 15.26$

Table 5: Calibrated parameters, and counterfactuals.

The first addend is the amount of taxes raised from honest firms. The second addend is the amount of taxes declared up-front by strategic firms. The the third addend is the money raised from audited cheaters.

At this point we could simply use the calibrated parameters from Table 4, plug them into the strategies derived Proposition 3, and evaluate these expressions numerically. However, this would not be fully satisfactory because the resulting expression for $p^*(r)$ would be too large relative to ballpark estimates. Indeed, while detailed data about $p^*(r)$ are not available, crude statistics that are publicly available suggest that only 2% of firms who declare less than 10 employees are audited every year. According to our model, however, the probability of auditing must be in the order of 30% to generate the comparatively mild underreporting found in the data. This observation raises the question of why Italian firms are so compliant, given the low probability of being audited and the mild penalties for being found in violation. We do not have an answer to this question except by resorting to unmeasured psychological, legal, etc. costs of being found in violation of the tax code. So in what follows we will calibrate a new parameter, $\widehat{\theta}$ which captures the penalties plus the non-monetary costs of being found cheating.

We explain in Section 6 how we calibrate the parameter $\widehat{\theta}$ (essentially, $\widehat{\theta}$ is chosen so as to generate the observed amount of underreporting given an aggregate amount of yearly audits equal to around 2% of the total population of firms). Our calibration yields values for $\widehat{\theta}$ above 6. This means that the *perceived* cost for being found cheating is estimated to be more than 6 times the amount underreported. This is a very large number compared to the monetary fine which is 0.33. We will return to this issue in the conclusions.

Plugging our calibrated parameter $\widehat{\theta}$ into the expression for $p^*(r)$ yields a "realistic" auditing strategy. We then compute values for the money raised in equilibrium (expression 27) and for the tax revenue theoretically achievable if all firms were honest (expression 26). The results for all audit classes are presented in the table below.

Table 5 implies, for the 0-10 class for example, that the ratio of revenue actually raised over maximal revenue achievable (expression 27 over 26) is given by

$$\frac{\lambda \cdot 1.57 + (1-\lambda)\,1.32}{1.57} = 0.94,$$

where we have used $\lambda = 0.62$ from Table 4. This means that, despite the relatively modest auditing budget and despite the decentralization problem, INPS is achieving 94% of the maximal revenue achievable if all firms reported truthfully. Results for the other audit classes are similarly high.

The root of this perhaps surprising result can be traced back to the raw data. For firms in the $(0, 10)$ audit class C_3, the average underreport, equals 0.67, meaning that the average strategic firm in that class reports 0.67 fewer employees than it actually has. Moreover, 62% of firms are actually honest, which means that the average number of underreported employees is actually about $0.67 \cdot 0.62 = 0.42$. Since these firms have on average 3.1 reported employees, and slightly more actual employees, on average these firms underreport less than 14% of their actual workforce. This modest amount of underreporting is consistent with the small margin for improvement we find in our counterfactual exercise.

8.2 Counterfactual II: the costs of decentralization

In this section we relax the constraint that the auditing schedule be incentive compatible for the auditors. We seek, instead, to identify the revenue-maximizing auditing schedule within a class of "simple" strategies which are generally not decentralizable.

The class of strategies we focus on is composed of strategies all of which result in the same number of firms audited in equilibrium. In this sense, all strategies in a class are equally expensive in terms of resources. Moreover, every strategy in the class is "simple" in that it is a step function of the reported tax base.

Definition 2 *Fix B, the number of firms to be audited, and denote $\widehat{\tau} = t/(t + \widehat{\theta})$. The class of* simple strategies $\mathcal{S}(B)$ *denotes the set of auditing strategies in which: (a) all reports below some T are audited with the same probability $p \leq \widehat{\tau}$; (b) no reports above T are audited; (c) exactly B firms are audited. The simple strategy in which $p = \widehat{\tau}$ is referred to as the* extremal *strategy.*

Restricting attention to simple strategies rules out sophisticated strategies in which the audit probability is finely modulated based on the report. Despite this limitation, we think simple strategies are of special policy relevance precisely because of their simplicity.

A noteworthy strategy is the extremal strategy. This strategy audits all reports below a threshold T with such high probability as to deter any cheating among the audited reports. Indeed, the extremal strategy has been shown to be revenue-maximizing in a setup very similar to ours (see Border and Sobel 1981, Macho-Stadler and Perez-Castrillo 1997). However, the extremal strategy is not necessarily revenue-maximizing in our setup: there may be another strategy which raises more revenue using the same amount of audits. This is shown in Example 8 in Appendix A.[16] Therefore, it is an open question whether the extremal is revenue-maximizing in our setup. We show in this section that, given our calibrated parameter values, the extremal strategy is revenue-maximizing among all simple strategies that are as expensive as the equilibrium outcome in terms of auditing resources.

The strategic firms' behavior under a simple strategy is characterized by a threshold $\widehat{x}$. Types with tax base lower than $\widehat{x}$ will report a (the minimum possible report) and pay ta, plus if they are audited they also incur a (subjective) cost $\left(\widehat{\theta}+t\right)(x-a)$; types higher than $\widehat{x}$ will report T (or slightly above that) and pay tT without ever being audited. Therefore the types who choose to report a rather than T are those for whom $ta+p\left(\widehat{\theta}+t\right)(x-a)\leq tT$, hence $\widehat{x}$ solves

$$T-a=\frac{p}{\widehat{\tau}}\left(\widehat{x}-a\right), \tag{28}$$

where we denote $\widehat{\tau}=t/\left(\widehat{\theta}+t\right)$. A simple strategy must meet the budget constraint, which is the sum of audits of strategic and honest firms: the latter are audited when their tax base is below T, and so the budget constraint reads

$$B=p\left[(1-\lambda)F\left(\widehat{x}\right)+\lambda F\left(T\right)\right] \tag{29}$$

Finally, the revenue raised by a simple strategy is given by

$$R=(1-\lambda)\left[ta+p\cdot(t+\theta)\int_a^{\widehat{x}}(x-a)f(x)\,dx+t\,(T-a)\left(1-F\left(\widehat{x}\right)\right)\right]+\lambda t\mathbb{E}\left(X\right). \tag{30}$$

When F corresponds to the Power distribution, we can use expressions (28) and (29) to explicitly solve for T and $\widehat{x}$ as a function of p. We can then express the revenue as a function of p alone, and maximize it as a function of p. Of course, only the portion of revenue raised from strategic firms (the expression in brackets in 30) will vary with p. The corresponding pictures, one for each auditing class, are displayed in Figure **??**; these pictures make use of the calibrated parameters and set B equal to the fraction of firms audited in our data (see Appendix E for details.)

The revenue in each panel of Figure **??** is increasing as p grows towards $\widehat{\tau}$, showing that revenue increases as we approach the extremal strategy. This suggests that, among the set of simple strategies which cost the same as the equilibrium strategy, the extremal strategy should be revenue-maximizing.

However, we note that the revenue exhibits a discontinuity at $p=\widehat{\tau}$.[17] The presence of

[16] The non-standard feature of our setup is the presence of the honest firms. Honest firms are not responsive to auditing (they never cheat anyway), yet they soak up auditing resources. Therefore the presence of honest firms dilutes the deterrence power of the auditing. Moreover, the extent of dilution depends on the distribution of honest firms—it will be stronger around those report levels where honest firms are more concentrated. Therefore the presence of honest firms shapes the optimal auditing schedule.

[17] The existence of this discontinuity is instructive. The discontinuity is caused by two factors. First,

	Class 0-10	**Class 11-25**	**Class 26-50**
$\widehat{\tau}$	0.055	0.045	0.054
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 extremal strategy	$(1-\lambda)\,1.47$	$(1-\lambda)\,6.80$	$(1-\lambda)\,15.78$

Table 6: Calibrated parameters, and counterfactuals.

this discontinuity requires the revenue raised from the extremal strategy to be computed separately. This is done in Appendix E, and the results are reported in Table 6. Comparing the third row in Table 6 with Figure **??** reveals the discontinuity in revenue at $p = \widehat{\tau}$, and proves the next observation.

Conclusion 7 *Given our calibrated parameters and setting an auditing budget equal to the equilibrium one, the extremal strategy raises more revenue than any other simple strategy with the same budget.*

Comparing the second and third rows in Table 6 shows that the extremal strategy always delivers more revenue than the equilibrium strategy for the same amount of auditing resources. Yet the improvement is small. Focusing 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 still for the other audit classes. Moreover, both these amounts are quite close to the one obtained from honest firms, which represents the theoretical maximum revenue obtainable. In this sense, despite the fact that in the decentralized equilibrium the wrong objective function is maximized, our calibration suggests that the amount of money left on the table is not too large. Of course, this conclusion is true only given the equilibrium auditing budget. It is possible that, were INPS to reduce the fraction of firms audited, the gains from moving to an extremal strategy might be larger.

when $p = \widehat{\tau}$ all firms report truthfully, while for any $p \in [0, \widehat{\tau})$ a bunch of firms cheat and report a. Still, this discontinuity in behavior need not *per se* result in revenue discontinuities, since the firms' payoffs is continuous at $p = \widehat{\tau}$. The firm's payoff is continuous because at $p = \widehat{\tau}$ a bunch of firms are indifferent between reporting truthfully or cheating and risking the penalty $\widehat{\theta}$. The revenue is *almost* the opposite of the firms' payoffs, *except* the revenue is computed using penalty $\theta < \widehat{\theta}$; therefore, revenue is not continuous at $p = \widehat{\tau}$. Since $\theta < \widehat{\theta}$, compared to the firms, the auditor strictly prefers to raise revenue through truthful reporting rather than through the penalties imposed on the cheaters. This observation accounts for the discontinuity in payoffs at $p = \widehat{\tau}$. The interest of this observation is that, when firms incur non-economic costs from being audited and found in violation of the law, then the game between auditor and audited is no longer a zero-sum game, and this has implication for the kind of audit strategies that are likely to be optimal in the mechanism-design problem. In particular, relative to the zero-sum version of the mechanism-design problem, there should be a preference for strategies that deter, rather than punish.

9 Conclusion

We view this paper as a first attempt to bring to data the theoretical literature on strategic auditing. The goal is important because having a model that can claim to quantitatively capture the strategic interaction between auditors and audited would allow us to (a) understand how well a given auditing policy is succeeding; and (b) to perform counterfactual experiments of enormous policy relevance (think tax evasion).

We believe this paper makes progress along several fronts. First, the paper suggests shifting the focus from a stylized one-auditor model with commitment to a more realistic many-auditors model with incentives. Second, the paper provides a method for recovering what the auditors' incentives actually are, based on audit (administrative) data. Third, in the special but important case of tax evasion, the paper introduces a variant of existing models *and* a method to calibrate it based on audit data. The method reads the data through the lens of equilibrium behavior and, in so doing, accounts for the huge selection process that generates the data. The parameters recovered through the calibration process allow us to perform counterfactual experiments.

When applied to the Italian data, the calibrated model seems able to capture accurately the behavior of the auditors "playing against" the 90% of firms with 10 employees or less. The model does less well on its ability to explain the game involving the 10% of firms with more than 10 employees, in the sense that it is harder to make precise sense of the auditor's behavior vis a vis these firms. The model we calibrate contains two parameters, λ and $\widehat{\theta}$, which capture respectively a reluctance on the part of firms to underreport taxable income, and an overassessment (relative to the mere economic consequences) of the cost of being caught cheating. In our calibration these parameters are assigned large values, which we interpret as suggestive that other factors, besides mere economic calculus, influence the firms' decision to cheat on their taxes. We believe this finding is not an artifact of our calibration procedure, but rather reflects a real economic phenomenon. The high estimate of the "degree of honesty" of firms is 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. If replicated in other datasets, this phenomenon will deserve further attention.

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 should be interpreted as "short run" results.

On top of the novel results presented in the context of tax compliance and auditing, the paper has the ambition of developing what we hope can become a replicable procedure for calibrating strategic auditing models: measure (the auditors' incentive scheme using Proposition 1), cut (build a model and solve for the equilibrium, as we did in Section 5), and fit (calibrate the model using audit data). We believe that this procedure can be applied to other auditing environments, but clearly this will not always be straightforward. A step of this procedure which we think is likely to require "custom tailoring" is that, given any objective function for the auditors, the procedure calls for solving the equilibrium of the auditing game. Solving for this equilibrium may not be routine.

References

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

[2] Anwar, Shamena and Hanming Fang (2006). "An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence." *American Economic Review* 96(1): 127-151.

[3] Bagnoli, Mark and Ted Bergstrom (2005). "Log-Concave Probability and Its Applications." *Economic Theory* 26, 445–469.

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

[5] Becker, G. 1968. "Crime and punishment: an economic approach." *Journal of Political Economy,* 76, p. 169–217.

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

[7] Di Porto, Edoardo. 2009. "Audit tax compliance and irregular work: The Italian case". SESS Euro PhD diss. Berlin: University of Berlin.

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

[9] Ehrlich, I. (1979). *The economic approach to crime.* In Criminology Review Yearbook, ed. S. Messinger and and E. Bittner. Beverly Hills: Sage Publications.

[10] Harrington, Joseph (2010) "When is an Antitrust Authority not Aggressive Enough in Fighting Cartels?" Forthcoming, *International Journal of Economic Theory.*

[11] Knowles, John, Persico, Nicola, and Todd, Petra (2001). "Racial Bias in Motor Vehicle Searches: Theory and Evidence." *Journal of Political Economy* 109(1): 203–229.

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

[13] Melumad N. and Mookherjee D., "Delegation as Commitment: The Case of Income Tax Audits." *The RAND Journal of Economics* Vol. 20(2),1989, p. 139-163.

[14] Persico Nicola, "Racial Profiling, Fairness, and Effectiveness of Policing." *The American Economic Review* 92(5), December 2002, pp. 1472-97.

[15] 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.

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

[17] Scotchmer, S., "Equilibrium Verifications and Reporting Policies in a Model of Tax Compliance." *International Economic Review*, 27(3), 1986, p. 739-760.

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

APPENDIX (NOT FOR PUBLICATION)

A Omitted Proofs

A.1 Details of Example 6

Maximizing monetary returns of audits:

Reporting Strategy: firms underreport by T if $y \geq T$, report 0 if $y \leq T$.

Auditing strategy:

$$p\left(r\right)=\frac{1}{2}\left(1-\exp\left(\frac{r-(1-T)}{T}\right)\right)$$

The equilibrium is an application of results found in Erard and Feinstein (1994).

We now compute the parameter T so that the number of audits corresponds to 10% of firms audited:

$$\begin{aligned}\int_0^{1-T}\frac{1}{2}\left(1-\exp\left(\frac{x-(1-T)}{T}\right)\right)dx+\frac{T}{2}(\frac{1}{2}\left(1-\exp\left(\frac{T-1}{T}\right)\right)) &= 0.1\\ T &= 0.669607\end{aligned}$$

We now compute the revenue collected when this auditing strategy is used (tax collected + fines):

$$\begin{aligned}R &= \lambda t\int_0^1 xdx\\ &\quad+(1-\lambda)\,T\frac{T}{2}\frac{1}{2}\left(1-\exp\left(\frac{-(1-T)}{T}\right)\right)\\ &\quad+\int_0^{1-T}\frac{1}{2}\left(1-\exp\left(\frac{r-(1-T)}{T}\right)\right)(r/2+T)\,dr\\ &\quad+\int_0^{1-T}r/2\left(1-\frac{1}{2}\left(1-\exp\left(\frac{r-(1-T)}{T}\right)\right)\right)dr\\ &= 0.17424\end{aligned}$$

Maximizing detection

Reporting strategy: firms with income y report $r=\frac{\alpha}{\alpha+1}y$.

Auditing strategy: $p(r) = \frac{1}{2}\left(1 - \left(\frac{\alpha+1}{a} r\right)^{\alpha}\right)$

The equilibrium is an application of results derived in section 5.

We now compute the parameter α so that the number of audits corresponds to 10% of firms audited

$$\begin{aligned}
\int_0^{\alpha/(a+1)} \frac{1}{2}\left(1 - \left(\frac{\alpha+1}{\alpha} r\right)^{\alpha}\right)\left(\lambda + (1-\lambda)\frac{\alpha+1}{\alpha}\right) dr &= 0.1 \\
\alpha &= 0.441\,39
\end{aligned}$$

The revenue is calculated as follows:

$$\begin{aligned}
R &= (1-\lambda)\int_0^{\frac{\alpha}{\alpha+1}} \left[p(r)\left(tr + (t+\theta)\left(\frac{\alpha+1}{\alpha} r - r\right)\right) + ((1-p(r))\,tr)\right] f(r)\,dr + \lambda t \int_0^1 x dx \\
R &= 0.17896.
\end{aligned}$$

Lemma 2 *If X is distributed on $[a,b]$ according to $F(x) = \left(\frac{x-a}{b-a}\right)^{\beta}$, then $\mathbb{E}(X) = a + \frac{\beta}{\beta+1}(b-a)$.*

Proof.

$$\begin{aligned}
\mathbb{E}(X) &= \int_a^b x dF(x) = a + \int_a^b (x-a)\,dF(x) = a + \int_a^b (x-a)\left(\frac{1}{b-a}\right)^{\beta} \beta (x-a)^{\beta-1} dx \\
&= a + \left(\frac{1}{b-a}\right)^{\beta} \beta \int_a^b (x-a)^{\beta} dx = a + \left(\frac{1}{b-a}\right)^{\beta} \beta \int_0^{b-a} y^{\beta} dy \\
&= a + \left(\frac{1}{b-a}\right)^{\beta} \beta \cdot \left(\left.\frac{y^{\beta+1}}{\beta+1}\right|_{y=0}^{b-a}\right) = a + \left(\frac{1}{b-a}\right)^{\beta} \frac{\beta}{\beta+1}(b-a)^{\beta+1} = a + \frac{\beta}{\beta+1}(b-a)
\end{aligned}$$

■

Lemma 3 *Given the audit schedule $p^*(\cdot)$, the firm's reporting strategy $\rho^*(\cdot)$ satisfies the second order conditions (12).*

Proof. From, 12, we need to show that $(t+\theta)\left[p''(r)(r-x) + 2p'(r)\right] \leq 0$. We have

$$p^{*\prime}(r) = -\frac{t}{\theta+t} \exp\left(-\int_r^{\rho(b)} \frac{1}{(x(y)-y)} dy\right) \frac{1}{(x(r)-r)},$$

$$p^{*\prime\prime}(r) = -\frac{t}{\theta+t}\exp\left(-\int_r^{\rho(b)}\frac{1}{(x(y)-y)}dy\right)\left[\frac{1}{(x(r)-r)^2}+\frac{\partial}{\partial r}\frac{1}{((x(r)-r))}\right].$$

We can then compute

$$\begin{aligned}
&p''(r)(r-x(r))+2p'(r) \\
&= \frac{t}{\theta+t}\exp\left(-\int_r^{r(b)}\frac{1}{(x(y)-y)}dy\right)(r-x(r))\left[\frac{1}{(x(r)-r)^2}-\frac{\partial}{\partial r}\frac{1}{(x(r)-r)}\right].
\end{aligned}$$

The sign of $(p''(r)(r-x)+2p'(r))$ is the opposite as the sign of $\left(\frac{1}{(x(r)-r)^2}-\frac{\partial}{\partial r}\frac{1}{(x(r)-r)}\right)$.

Now, we have

$$\begin{aligned}
\frac{1}{(x(r)-r)^2}-\frac{\partial}{\partial r}\frac{1}{(x(r)-r)} &= \frac{1}{(x(r)-r)^2}+\frac{x'(r)-1}{(x(r)-r)^2} \\
&= \frac{x'(r)}{(x(r)-r)^2} > 0.
\end{aligned}$$

Thus reporting $r^*(x)$ is indeed a local maximum for a firm with true tax base x. If any other local maxima exist which do not satisfy the first order conditions, they must be at the corners of the feasible report sets, either $r=0$ or $r=x$. But neither can be a local maximum, for otherwise there would have to be at least one other solution to the first order conditions between $r^*(x)$ and 0, or $r^*(x)$ and x, whereas we know that $r^*(x)$ is the unique solution to the first order conditions. Therefore, reporting $r^*(x)$ is also a *global* maximum for a firm with true tax base x. ■

Proof of Lemma 1.

Proof. Proving that $x-r(x)$ is increasing is equivalent to show that $r'(x)\leq 1$. We have :

$$r'(x) = \frac{f(x)}{\gamma f(r(x))}$$

>From decreasing hazard rate property we have

$$\frac{f(r(x))}{F(r(x))} \geq \frac{f(x)}{F(x)},$$

using the fact that $F(x)=\gamma F(\rho(x))$, we get:

$$\gamma f(r(x)) \geq f(x).$$

whence $r^{'}(x) \leq 1$. ■

Proof of Proposition 3.

Proof. >From (20) 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 exponential 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 (19) 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

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

■

Example 8 *Let F be uniform on $[0,1]$. Set $t=\theta=\widehat{\theta}=0.5$. Assume $B=0.1$, which means that up to 10% of firms can be audited, and a fraction λ of honest firms equal to 0.5.*

The "extremal strategy" is to audit with probability $p=\pi=0.5$ all firms which report less *than $T = 0.2$. Under this strategy the strategic firms in $[0, 0.2]$ will report truthfully, and so 0.1 on average, while all others will report T. Total revenue raised from strategic firms equals $t\,[0.2\,(0.1) + 0.8T] = 0.09$. Total audits, including strategic and honest firms, under this strategy are exactly 0.1.*

Consider now the following simple strategy: audit with probability $p = 0.4$ all firms who report less than $T = 0.2222$. Using (28), it follows that all strategic firms with type less than $\widehat{x} = 0.27778$ report zero and so the average revenue from them is $p\,(t+\theta)\left(\frac{\widehat{x}}{2}\right)$, while the rest report T. Total revenue raised from strategic firms is $\widehat{x}p\,(t+\theta)\left(\frac{\widehat{x}}{2}\right) + (1-\widehat{x})\,tT = 0.097$. Total audits under this strategy are $(p\cdot\widehat{x})\,(1-\lambda) + \lambda pT = 0.1$.

A.2 Deriving the System of Nonlinear Equations (24)

In this appendix we show how we go from the three equations (21)$= C_1$, (22)$= C_2$, and (23)$= C_3$, together with the condition $\rho^*(b) = M$, to the system of equations (24).

Equation (21) reads

$$\begin{aligned} & \frac{\lambda \int_a^M p^*(r)\,dF(r)}{(1-\lambda)\int_a^M p^*(r)\,dF\left(\rho^{*-1}(r)\right)} \\ = & \frac{\lambda \int_a^M p^*(r)\left(\frac{1}{b-a}\right)^\beta \beta\,(r-a)^{\beta-1}\,dr}{(1-\lambda)\int_a^M p^*(r)\left(\frac{1}{b-a}\right)^\beta \beta\left(\frac{\alpha+1}{\alpha}\right)^\beta (r-a)^{\beta-1}\,dr} \\ = & \frac{\lambda}{(1-\lambda)\left(\frac{\alpha+1}{\alpha}\right)^\beta}. \end{aligned} \tag{31}$$

Equation (22) reads

$$\begin{aligned} & (1-\lambda)\int_a^b \rho^*(x)\,dF(x) + \lambda\int_a^M r\,dF(r) \\ = & (1-\lambda)\int_a^b \rho^*(x)\,dF(x) + \lambda\int_a^b \rho^*(x)\left(\frac{\alpha}{\alpha+1}\right)^\beta dF(x) \\ = & \left[(1-\lambda) + \lambda\left(\frac{\alpha}{\alpha+1}\right)^\beta\right]\int_a^b \rho^*(x)\,dF(x) \\ = & \left[(1-\lambda) + \lambda\left(\frac{\alpha}{\alpha+1}\right)^\beta\right]\int_a^b \rho^*(x)\,dF(x) \\ = & \left[(1-\lambda) + \lambda\left(\frac{\alpha}{\alpha+1}\right)^\beta\right]\left[a + \frac{\alpha}{\alpha+1}\left(\int_a^b x\,dF(x) - a\right)\right] \end{aligned} \tag{32}$$

where the second integral in the second line reflects the change of variables $r = \rho^*(x) =$

$a+\frac{\alpha}{\alpha+1}(x-a)$, $dF(r)=dF\left(a+\frac{\alpha}{\alpha+1}(x-a)\right)=d\left(\frac{\alpha}{\alpha+1}\frac{x-a}{b-a}\right)^{\beta}=\left(\frac{\alpha}{\alpha+1}\right)^{\beta}dF(x)$.

Equation (23) reads

$$\begin{aligned}
&\int_a^b [x-\rho^*(x)]\,dF(x)\\
=\;&\int_a^b \left[x-a-\frac{\alpha}{\alpha+1}(x-a)\right]dF(x)\\
=\;&\left[1-\frac{\alpha}{\alpha+1}\right]\int_a^b (x-a)\,dF(x)=\frac{1}{\alpha+1}\left[\int_a^b x\,dF(x)-a\right]. \qquad (33)
\end{aligned}$$

Equations (31), (32), and (33), together with the condition $\rho^*(b)=M$, form a system of four equations in four unknowns. This system is given by

$$\begin{aligned}
a+\frac{\alpha}{\alpha+1}(b-a) &= M\\
\frac{\lambda}{(1-\lambda)\left(\frac{\alpha+1}{\alpha}\right)^{\beta}} &= C_1\\
\left[(1-\lambda)+\lambda\left(\frac{\alpha}{\alpha+1}\right)^{\beta}\right]\left[a+\frac{\alpha}{\alpha+1}\left(\int_a^b x\,dF(x)-a\right)\right] &= C_2\\
\frac{1}{\alpha+1}\left[\int_a^b x\,dF(x)-a\right] &= C_3
\end{aligned}$$

Substitute from the fourth into the third equation to get the equivalent system

$$\begin{aligned}
a+\frac{\alpha}{\alpha+1}(b-a) &= M\\
\frac{\lambda}{(1-\lambda)\left(\frac{\alpha+1}{\alpha}\right)^{\beta}} &= C_1\\
\left[(1-\lambda)+\lambda\left(\frac{\alpha}{\alpha+1}\right)^{\beta}\right][a+\alpha C_3] &= C_2\\
\frac{1}{\alpha+1}\left[\int_a^b x\,dF(x)-a\right] &= C_3
\end{aligned}$$

Substitute from the second into the third equation to get the equivalent system

$$\begin{aligned} a+\frac{\alpha}{\alpha+1}(b-a) &= M \\ \frac{\lambda}{(1-\lambda)\left(\frac{\alpha+1}{\alpha}\right)^{\beta}} &= C_1 \\ (1-\lambda)\left[1+C_1\right]\left[a+\alpha C_3\right] &= C_2 \\ \frac{1}{\alpha+1}\left[\int_a^b x dF(x)-a\right] &= C_3 \end{aligned}$$

Use the formula $\mathbb{E}(X)=a+\frac{\beta}{\beta+1}(b-a)$ to substitute into the fourth equation to get the equivalent system

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

Eliminate the first equation by substituting into the fourth, to get the equivalent system

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

which is the system of equations (24).

B Ancillary Material Related to the Application to INPS Data

B.1 Creating the sample

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 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.

B.2 Robustness checks for Section 7.2

Dependent Var: Detection Success

VARIABLES	(1) Regions	(2) Year	(3) <5 employees
number of employees	0.01	0.01	-0.01
	(0.353)	(0.275)	(0.564)
sector 2	0.05	0.04	-0.01
	(0.229)	(0.278)	(0.862)
sector 3	0.04	0.04	-0.03
	(0.286)	(0.246)	(0.494)
sector 4	0.08**	0.08**	0.00
	(0.025)	(0.020)	(0.972)
sector 5	0.05	0.05	-0.01
	(0.121)	(0.123)	(0.890)
sector 6	0.04	0.04	-0.04
	(0.260)	(0.298)	(0.421)
sector 7	0.27***	0.27***	0.19***
	(0.000)	(0.000)	(0.000)
sector 8	0.10***	0.10***	0.03
	(0.008)	(0.005)	(0.601)
sector 9	-0.00	-0.01	-0.08*
	(0.912)	(0.786)	(0.085)
sector 2 * employees	0.01	0.01	0.04
	(0.402)	(0.372)	(0.130)
sector 3 * employees	0.01	0.01	0.04*
	(0.490)	(0.484)	(0.051)
sector 4 * employees	0.01	0.01	0.05**
	(0.388)	(0.448)	(0.027)
sector 5 * employees	0.01	0.01	0.03
	(0.612)	(0.580)	(0.115)
sector 6 * employees	0.01	0.01	0.05**
	(0.263)	(0.252)	(0.021)
sector 7 * employees	-0.00	-0.00	0.04
	(0.839)	(0.820)	(0.107)
sector 8 * employees	0.01	0.01	0.04**
	(0.467)	(0.484)	(0.039)
sector 9 * employees	0.03**	0.03**	0.06***

Continued on next page

Table 7 -

VARIABLES	(1)	(2)	(3)
	(0.019)	(0.021)	(0.003)
Region 2	0.08***		
	(0.000)		
Region 3	0.19***		
	(0.000)		
Region 4	0.02**		
	(0.029)		
Region 5	0.15***		
	(0.000)		
Region 6	-0.03***		
	(0.004)		
Region 7	0.19***		
	(0.000)		
Region 8	0.04***		
	(0.000)		
Region 9	0.10***		
	(0.000)		
Region 10	0.22***		
	(0.000)		
Region 11	-0.10***		
	(0.000)		
Region 12	0.16***		
	(0.000)		
Region 13	0.05***		
	(0.000)		
Region 14	0.13***		
	(0.000)		
Region 15	0.06***		
	(0.000)		
Region 16	0.20***		
	(0.000)		
Region 17	0.09***		
	(0.000)		
Region 18	0.25***		
	(0.000)		

Continued on next page

Table 7 -

VARIABLES	(1)	(2)	(3)
Region 19	0.23***		
	(0.000)		
Region 20	0.07***		
	(0.000)		
Year		-0.02***	
		(0.000)	
Constant	0.18***	40.85***	0.33***
	(0.000)	(0.000)	(0.000)
Observations	151,806	151,806	127,366
R-squared	0.034	0.018	0.011

Robust pval in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

C Numerical Solutions to the Nonlinear System of Equations (24)

C.0.1 Class below 10

Consider 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 $(1, 10)$, 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.

$$\begin{aligned}
\frac{\lambda}{(1-\lambda)\left(\frac{\alpha+1}{\alpha}\right)^{\beta}} &= \frac{0.6}{0.4} \\
(1-\lambda)\left(1+\frac{0.6}{0.4}\right)\alpha 0.67 &= 3.09 \\
\frac{10}{\alpha}\frac{\beta}{\beta+1} &= 0.67
\end{aligned}$$

Solution is: $\{[\lambda = 0.621\,51, \alpha = 4.\,874, \beta = 0.484\,91]\}$

C.0.2 Class 11-25

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 (24) yields $[\lambda = 0.499\,37, \beta = 0.607\,88, \alpha = 4.\,684\,0]$. The highest true tax base in the audit class (unobserved because firms who report this much are not audited) equals $11 + (25 - 11)(\alpha + 1)/\alpha = 28$.

$$\begin{aligned}
\frac{\lambda}{(1-\lambda)\left(\frac{\alpha+1}{\alpha}\right)^{\beta}} &= \frac{0.47}{0.53} \\
(1-\lambda)\left(1+\frac{0.47}{0.53}\right)(11+\alpha\cdot 1.13) &= 15.39 \\
\frac{(25-11)}{\alpha}\frac{\beta}{\beta+1} &= 1.13
\end{aligned}$$

Rewrite slightly as

$$\begin{aligned}
\frac{\lambda}{(1-\lambda)\left(\frac{\alpha+1}{\alpha}\right)^{\beta}} &= 0.886\,79 \\
(1-\lambda)(1.886\,79)(11+\alpha(1.13)) &= 15.39 \\
14\frac{\beta}{\beta+1} &= \alpha(1.13)
\end{aligned}$$

Solution is: $\{[\lambda = 0.499\,37, \beta = 0.607\,88, \alpha = 4.\,684\,0]\}$

C.0.3 Class 26-50

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 (24) yields $[\lambda = 0.551\,45, \alpha = 9.\,017\,1, \beta = 4.\,643\,8]$.

Solving the system of equations (24) yields

$$\begin{aligned}
\frac{\lambda}{(1-\lambda)\left(\frac{\alpha+1}{\alpha}\right)^{\beta}} &= \frac{0.43}{0.57} \\
(1-\lambda)\left(1+\frac{0.43}{0.57}\right)(26+\alpha 2.19) &= 36 \\
\frac{(50-26)}{\alpha}\frac{\beta}{\beta+1} &= 2.19
\end{aligned}$$

, Solution is: $\{[\lambda = 0.551\,45, \alpha = 9.\,017\,1, \beta = 4.\,643\,8]\}$

The parameter β being greater than 1 will lead to a poor fit of (25) to the data. What we learn from this is that sometimes, matching the moments in (24) *perfectly* has a significant cost in terms of fit. One way around this problem is to relax the matching of the moments in (24). If we accept the second equation in (24) to equal 35. 235 rather than 36, which seems like a small cost, then (24) yields a solution which produces a better fit of (25) to the data. This is done in the next subsection.

C.0.4 Class 26-50 relaxing $C_2 = 36$

Here we follow a different procedure. We fix $\beta = 1$, whereby from the third equation in (24) we get

$$24\frac{1}{1+1} = \alpha(2.19)$$

Solution is: $\alpha = 5.4795$. Substitute this into the first equation

$$\frac{\lambda}{(1-\lambda)\left(\frac{5.4795+1}{5.4795}\right)} = \frac{0.43}{0.57}$$

Solution is: $\lambda = 0.47148$. Substitute into the LHS of the second equation to get

$$(1-0.47148)\left(1+\frac{0.43}{0.57}\right)(26+(5.4795\cdot 2.19)) = 35.235.$$

So if we allow C_2 to equal 35.235 rather than 36, then the parameter constellation we identifies solves the "relaxed" system (24). The solution $[\lambda = 0.47148, \alpha = 5.4795, \beta = 1]$ produces a better fit of (25) to the data. The implication is that, for this audit class, our model underpredicts slightly the reported number of employees.

C.1 Plotting Expression (25)

We want to work out expression (25) in the case of Power distribution. Since $\rho^{*-1}(r) = a + \frac{\alpha+1}{\alpha}(r-a)$, we have

$$\begin{aligned} & f\left(\rho^{*-1}(r)\right)\frac{\partial \rho^{*-1}(r)}{\partial r} \\ = & \left(\frac{1}{b-a}\right)^{\beta}\beta\left(\frac{\alpha+1}{\alpha}(r-a)\right)^{\beta-1}\frac{\alpha+1}{\alpha} \\ = & \left(\frac{\alpha+1}{\alpha}\right)^{\beta}\left(\frac{1}{b-a}\right)^{\beta}\beta(r-a)^{\beta-1} = \left(\frac{\alpha+1}{\alpha}\right)^{\beta} f(r) \end{aligned} \tag{34}$$

Then

$$\begin{aligned} & p^*(r)\cdot\left[(1-\lambda) f\left(\rho^{*-1}(r)\right)\frac{\partial \rho^{*-1}(r)}{\partial r} + \lambda f(r)\right] \\ = & \ p^*(r)\cdot\left[(1-\lambda)\left(\frac{\alpha+1}{\alpha}\right)^{\beta} + \lambda\right] f(r) \end{aligned}$$

where

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

and

$$f(r) = \left(\frac{1}{b-a}\right)^{\beta}\beta(r-a)^{\beta-1}.$$

D Details About Counterfactual Section I

In this appendix we describe how we compute $\widehat{\theta}$ and then to compute expressions 27 over 26.

$\widehat{\theta}$ is set so as to match the fraction of firms audited in our model to aggregate statistics available from INPS. In our model, the fraction of firms audited among those that report between a and M is constructed starting from (25) and is given by

$$\frac{\int_a^M p^*(r)\cdot\left[(1-\lambda)f\left(\rho^{*-1}(r)\right)\frac{\partial\rho^{*-1}(r)}{\partial r}+\lambda f(r)\right]\ dr}{\int_a^M\left[(1-\lambda)f\left(\rho^{*-1}(r)\right)\frac{\partial\rho^{*-1}(r)}{\partial r}+\lambda f(r)\right]\ dr}. \tag{35}$$

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 setting the "effective" probability of auditing to 5 percent. Therefore we set expression (35) equal to 5 percent. We substitute out for $p^*(r)$ and write this condition as

$$\frac{\int_a^M\left[1-\left(\frac{\alpha+1}{\alpha}\frac{r-a}{b-a}\right)^{\alpha}\right]\cdot\left[(1-\lambda)f\left(\rho^{*-1}(r)\right)\frac{\partial\rho^{*-1}(r)}{\partial r}+\lambda f(r)\right]\ dr}{\int_a^M\left[(1-\lambda)f\left(\rho^{*-1}(r)\right)\frac{\partial\rho^{*-1}(r)}{\partial r}+\lambda f(r)\right]\ dr}=\frac{\theta+t}{t}0.05 \tag{36}$$

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 $\widehat{\theta}\geq 0.33$.

D.1 Ancillary derivations used to compute expression 27

Expression (27) reads

$$\begin{aligned}
&\lambda\int_a^b tr\cdot f(r)\ dr\\
&+(1-\lambda)\int_a^M\left[tr+p^*(r)\left[(t+\theta)\left(\rho^{*-1}(r)-r\right)\right]\right]\cdot f\left(\rho^{*-1}(r)\right)\frac{\partial\rho^{*-1}(r)}{\partial r}\ dr\\
=\ &\lambda\int_a^b tr\cdot f(r)\ dr\\
&+(1-\lambda)\left(\frac{\alpha+1}{\alpha}\right)^{\beta}\int_a^M\left[tr+p^*(r)\left[(t+\theta)\left(\frac{\alpha+1}{\alpha}(r-a)-(r-a)\right)\right]\right]\cdot f(r)\ dr
\end{aligned}$$

Since $\rho^{*-1}(r) = a + \frac{\alpha+1}{\alpha}(r-a)$ and $f(\rho^{*-1}(r))\frac{\partial\rho^{*-1}(r)}{\partial r} = \left(\frac{\alpha+1}{\alpha}\right)^{\beta} f(r)$, (27) reads

$$\begin{aligned}
& \lambda \int_a^b tr \cdot f(r)\ dr \\
& + (1-\lambda)\int_a^M \left[tr + p^*(r)\left[(t+0.33)\left(\frac{\alpha+1}{\alpha}(r-a)-(r-a)\right)\right]\right]\cdot\left(\frac{\alpha+1}{\alpha}\right)^{\beta} f(r)\ dr \\
= \ & \lambda \int_a^b tr \cdot f(r)\ dr \\
& + (1-\lambda)\int_a^M \left[tr + p^*(r)\left[(t+0.33)(r-a)\left(\frac{\alpha+1}{\alpha}-1\right)\right]\right]\cdot\left(\frac{\alpha+1}{\alpha}\right)^{\beta} f(r)\ dr \\
= \ & \lambda \int_a^b tr \cdot f(r)\ dr \\
& + (1-\lambda)\int_a^M \left[tr + p^*(r)\left[(t+0.33)(r-a)\left(\frac{1}{\alpha}\right)\right]\right]\cdot\left(\frac{\alpha+1}{\alpha}\right)^{\beta} f(r)\ dr \\
= \ & \lambda t \int_a^b r \cdot f(r)\ dr + (1-\lambda)\, t \left(\frac{\alpha+1}{\alpha}\right)^{\beta}\int_a^M r\cdot f(r)\ dr \\
& + (1-\lambda)(t+0.33)\int_a^M p^*(r)\left[(r-a)\left(\frac{1}{\alpha}\right)\right]\cdot\left(\frac{\alpha+1}{\alpha}\right)^{\beta} f(r)\ dr
\end{aligned}$$

Lemma 4 $\int_a^M r\cdot f(r)\ dr = \left(\frac{\alpha}{\alpha+1}\right)^{\beta}\left[a + \frac{\alpha}{\alpha+1}\frac{\beta}{\beta+1}(b-a)\right]$

Proof.

$$\begin{aligned}
& \int_a^M r\cdot f(r)\ dr \\
= \ & \left(\frac{1}{b-a}\right)^{\beta}\beta\int_a^M r\cdot(r-a)^{\beta-1}\ dr \\
= \ & \left(\frac{1}{b-a}\right)^{\beta}\beta\left[\int_a^M (r-a)^{\beta}\ dr + \int_a^M a\cdot(r-a)^{\beta-1}\ dr\right] \\
= \ & \left(\frac{1}{b-a}\right)^{\beta}\beta\left[\int_0^{M-a} y^{\beta}\ dy + a\int_0^{M-a} y^{\beta-1}\ dy\right] \\
= \ & \left(\frac{1}{b-a}\right)^{\beta}\beta\left[\left.\frac{y^{\beta+1}}{\beta+1}\right|_{y=0}^{M-a} + a\left.\frac{y^{\beta}}{\beta}\right|_{y=0}^{M-a}\right] \\
= \ & \left(\frac{1}{b-a}\right)^{\beta}\beta\left[\frac{(M-a)^{\beta+1}}{\beta+1} + a\frac{(M-a)^{\beta}}{\beta}\right] \\
= \ & \left(\frac{1}{b-a}\right)^{\beta}\beta(M-a)^{\beta}\left[\frac{M-a}{\beta+1} + a\frac{1}{\beta}\right] \\
= \ & \left(\frac{M-a}{b-a}\right)^{\beta}\beta\left[\frac{M-a}{\beta+1} + a\frac{1}{\beta}\right].
\end{aligned}$$

Since $M = a + \frac{\alpha}{\alpha+1}(b-a)$, we can rewrite the above expression as

$$\begin{aligned}
& \left(\frac{\alpha}{\alpha+1}\right)^{\beta} \beta \left[\frac{\frac{\alpha}{\alpha+1}(b-a)}{\beta+1} + a\frac{1}{\beta}\right] \\
= & \left(\frac{\alpha}{\alpha+1}\right)^{\beta} \left[\beta\frac{\frac{\alpha}{\alpha+1}(b-a)}{\beta+1} + a\right] \\
= & \left(\frac{\alpha}{\alpha+1}\right)^{\beta} \left[a + \frac{\alpha}{\alpha+1}\frac{\beta}{\beta+1}(b-a)\right].
\end{aligned}$$

■

D.2 Audit class below 10

Solving equation (36) 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.

The first addend of (27) is the amount of taxes raised from honest firms; using Lemma 2, the first addend equals $\lambda t \left[a + \frac{\beta}{\beta+1}(b-a)\right] = \lambda t \cdot 3.9187 = \lambda \cdot 1.5675$. Combining Lemma 4 with (34), the second addend of (27) equals $(1-\lambda)\, t \left[a + \frac{\alpha}{\alpha+1}\frac{\beta}{\beta+1}(b-a)\right] = (1-\lambda)\, t \cdot 3.2516$; this is the amount of taxes paid by strategic firms.

Finally, the third addend of equation (27) is the money raised from audited cheaters and equals $(1-\lambda)(t+0.33)$ times the following:

$$\begin{aligned}
& \int_a^{10} \frac{0.4}{6.87+0.4}\left(1 - \left(\frac{\alpha+1}{\alpha}\frac{r-a}{b-a}\right)^{\alpha}\right) \\
& \times (r-a)\left(\frac{1}{\alpha}\right) \cdot \left(\frac{\alpha+1}{\alpha}\right)^{\beta}\left(\frac{1}{b-a}\right)^{\beta} \beta (r-a)^{\beta-1}\; dr.
\end{aligned}$$

We set the following values as definitions into the Scientific Word solver: $\alpha = 4.874; \beta = 0.48491; b = 12; a = 0$ and solve numerically. The integral is evaluated to be equal to 2.8132×10^{-2}. Therefore the third addend is equal to $(1-\lambda)(t+0.33)(0.02813)$. The total money raised from strategic firms is the sum of the second and third added, and after substituting $t = 0.4$ this amount equals

$$(1-\lambda)\, 1.3212.$$

D.3 Audit class 11-25

The relevant parameters for this class are $[a = 11, \lambda = 0.499\,37, \beta = 0.607\,88, \alpha = 4.\,684\,0]$. The highest true tax base in the audit class (unobserved because firms who report this much are not audited) equals $11 + (25 - 11)\ (\alpha + 1)/\alpha = 28 = b$.

Probability of being audited about 2%, double it to 4%.

Replace 0.05 with 0.04 in equation (36) and solve numerically to get $\widehat{\theta} = 8.447$.

The first addend in (27) is the amount of taxes raised from honest firms; using Lemma 2, it equals $\lambda t \left[a + \frac{\beta}{\beta+1}(b-a)\right] = \lambda t \cdot 17.\,427 = \lambda \cdot 6.\,970\,8$. The second term is the amount of taxes paid by strategic firms, and it equals $(1-\lambda)\, t \left[a + \frac{\alpha}{\alpha+1}\frac{\beta}{\beta+1}(b-a)\right] = (1-\lambda)\, t \cdot 16.\,296 = (1-\lambda) \cdot 6.\,518\,4$. (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 equal to $(1-\lambda)(t+0.33)$ times

$$\int_{11}^{25} \frac{0.4}{8.447+0.4}\left(1 - \left(\frac{\alpha+1}{\alpha}\frac{r-a}{b-a}\right)^{\alpha}\right) \times (r-a)\left(\frac{1}{\alpha}\right)\cdot\left(\frac{\alpha+1}{\alpha}\right)^{\beta}\left(\frac{1}{b-a}\right)^{\beta}\beta\,(r-a)^{\beta-1}\ dr.$$

Solving numerically yields $(1-\lambda)(t+0.33)\, 0.038$.

The total money raised from strategic firms in equilibrium is

$$\begin{aligned} & (1-\lambda)\left[t \cdot 16.\,296 + (t+0.33)(0.038)\right] \\ = & (1-\lambda)\left[(0.4)\cdot 16.\,296 + (0.4+0.33)(0.038)\right] \\ = & (1-\lambda)\cdot 6.\,546\,1 \end{aligned}$$

D.4 Audit class 26-50

Solving equation (36) numerically based on the calibrated parameters in Section 6, we get $\widehat{\theta} = 6.954$.

The first addend in (27) is the amount of taxes raised from honest firms; using Lemma 2, it equals $\lambda t \left[a + \frac{\beta}{\beta+1}(b-a)\right] = \lambda t \cdot\ 40.\,19 = \lambda \cdot\ 16.\,076$.

The second term is the amount of taxes paid by strategic firms, and it equals $(1-\lambda)\,t\left[a+\frac{\alpha}{\alpha+1}\frac{\beta}{\beta+1}(b-a)\right] = (1-\lambda)\,t\cdot 38.0 = (1-\lambda)\cdot 15.2$.

Finally, the third term is the money raised from audited cheaters and is equal to $(1-\lambda)(t+0.33)$ times

$$\int_{26}^{50} \frac{0.4}{6.954+0.4}\left(1-\left(\frac{\alpha+1}{\alpha}\frac{r-a}{b-a}\right)^{\alpha}\right) \\ \times (r-a)\left(\frac{1}{\alpha}\right)\cdot\left(\frac{\alpha+1}{\alpha}\right)^{\beta}\left(\frac{1}{b-a}\right)^{\beta}\beta\,(r-a)^{\beta-1}\;dr.$$

Solving numerically yields $(1-\lambda)(t+0.33)\,0.087$.

The total money raised from strategic firms in equilibrium is

$$\begin{aligned} & (1-\lambda)\left[t\cdot 38.0+(t+0.33)(0.087)\right] \\ = \; & (1-\lambda)\cdot 15.264 \end{aligned}$$

E Details on Counterfactual Section II

Assume F has a Power distribution. Fix B and consider the class of "simple" audit strategies, each of which is characterized by $\widehat{x}$, T and p. We want to get to a closed form expression for the revenue (30) as a function of (known parameters and) p alone. To this end, we need to express $\widehat{x}$ and T as a function of p. From the budget constraint (29) we get

$$(1-\lambda)\left(\frac{\widehat{x}-a}{b-a}\right)^{\beta}+\lambda\left(\frac{T-a}{b-a}\right)^{\beta}=\frac{B}{p}$$

Using (28) to substitute for $T-a$ yields

$$\begin{aligned} (1-\lambda)\left(\frac{\widehat{x}-a}{b-a}\right)^{\beta}+\lambda\left(\frac{p}{\widehat{\tau}}\frac{(\widehat{x}-a)}{b-a}\right)^{\beta} &= \frac{B}{p} \\ \left((1-\lambda)+\lambda\left(\frac{p}{\widehat{\tau}}\right)^{\beta}\right)\left(\frac{\widehat{x}-a}{b-a}\right)^{\beta} &= \frac{B}{p} \\ \left(\frac{\widehat{x}-a}{b-a}\right)^{\beta} &= \frac{B}{p\left((1-\lambda)+\lambda\left(\frac{p}{\widehat{\tau}}\right)^{\beta}\right)} \end{aligned}$$

Hence, we have

$$\widehat{x}-a = (b-a)\left(\frac{B}{p\left((1-\lambda)+\lambda\left(\frac{p}{\widehat{\tau}}\right)^{\beta}\right)}\right)^{1/\beta} \tag{37}$$

$$\widehat{x} = a+(b-a)\left(\frac{B}{p\left((1-\lambda)+\lambda\left(\frac{p}{\widehat{\tau}}\right)^{\beta}\right)}\right)^{1/\beta} \tag{38}$$

$$T-a = \frac{p}{\widehat{\tau}}(b-a)\left(\frac{B}{p\left((1-\lambda)+\lambda\left(\frac{p}{\widehat{\tau}}\right)^{\beta}\right)}\right)^{1/\beta} \tag{39}$$

Now let's turn to the revenue. Expression (30) contains an integral which we want to solve for analytically. From the proof of Lemma 4 we have

$$\begin{aligned}
\int_a^x (y-a)\cdot f(y)\ dy &= \left(\frac{x-a}{b-a}\right)^{\beta}\left[\frac{x\beta+a}{(\beta+1)}\right]-a\left(\frac{x-a}{b-a}\right)^{\beta} \\
&= \left(\frac{x-a}{b-a}\right)^{\beta}\left[\frac{x\beta+a}{(\beta+1)}-a\right] \\
&= \left(\frac{x-a}{b-a}\right)^{\beta}\left[\frac{x\beta-a\beta}{(\beta+1)}\right] \\
&= \left(\frac{x-a}{b-a}\right)^{\beta}\frac{\beta}{(\beta+1)}(x-a)\,.
\end{aligned}$$

Substituting into (30) yields

$$R=(1-\lambda)\left[ta+p\cdot(t+\theta)\,F(\widehat{x})\,\frac{\beta}{(\beta+1)}(\widehat{x}-a)+t\,(T-a)\,(1-F(\widehat{x}))\right]+\lambda t\mathbb{E}(X)\,.$$

Taking into account expressions (37)-(39), we have expressed the revenue as a function of p only. The term in brackets, corresponding to the revenue raised from strategic firms, is plotted in Figure **??**.

E.1 Computing the revenue raised with the extremal strategy in the class below 10

The extremal strategy audits with probability $t/\left(t+\widehat{\theta}\right)$ all firms who report less than T (a threshold yet to be determined) and does not audit any firm which 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, under the extremal strategy 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 (35), in equilibrium a fraction equal to

$$0.05 \int_a^M \left[(1-\lambda) f\left(\rho^{*-1}(r)\right) \frac{\partial \rho^{*-1}(r)}{\partial r} + \lambda f(r) \right] dr \tag{40}$$

of all firms in the audit class is audited. This equation reduces to

$$\begin{aligned} & 0.05 \left[(1-\lambda) \int_a^{10} \left(\frac{\alpha+1}{\alpha}\right)^\beta \left(\frac{1}{b-a}\right)^\beta \beta (r-a)^{\beta-1} \; dr + \lambda \left(\frac{M-a}{b-a}\right)^\beta \right] \\ = \; & 0.05 \left[(1-\lambda) + \lambda \left(\frac{10}{12}\right)^\beta \right] \\ = \; & 0.05 \left[(1-\lambda) + \lambda (0.915\,39) \right] = 0.05 \cdot 0.947\,41 = 0.047, \end{aligned}$$

where we substituted $\lambda = 0.621\,5$. The number 0.047 needs to equal to the fraction of firms audited under the extremal strategy. Under the extremal strategy, all firms (strategic or not) with tax base below T will report truthfully and be audited with probability $\frac{t}{t+\theta}$. No other firm will report in that range. Therefore, the fraction of firms audited under the extremal 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

$$T = a + (b-a) \left(\frac{t+\widehat{\theta}}{t} 0.047 \right)^{\frac{1}{\beta}} = 12 \left(0.047 \frac{0.4+6.87}{0.4} \right)^{\frac{1}{\beta}} = 8.\,671\,0.$$

The money raised from strategic firms under the extremal strategy equals the amount declared by firms with tax base below T, plus the amount declared by firms with tax base greater than T, which is exactly T. Formally,

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

E.2 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, The results for all audit classes are presented in the table below.

E.3 Audit class 11-25

According to equation (35), a fraction equal to

$$0.04\left((1-\lambda)+\lambda\left(\frac{M-a}{b-a}\right)^{\beta}\right) = 0.037$$

of all firms in the audit class is audited. This number needs to equal to the fraction of firms audited under the extremal strategy. Under the extremal strategy, the fraction of firms audited is given by the equation

$$\frac{t}{t+\widehat{\theta}}F(T) = \frac{t}{t+\widehat{\theta}}\left(\frac{T-a}{b-a}\right)^{\beta}.$$

Equating this to 0.037 and solving for T yields

$$T = a+(b-a)\left(\frac{t+\widehat{\theta}}{t}0.037\right)^{\frac{1}{\beta}} = a+(b-a)\left(\frac{0.4+8.447}{0.4}0.037\right)^{\frac{1}{\beta}} = 23.224.$$

The money raised from strategic firms under the extremal strategy equals

$$
\begin{aligned}
& (1-\lambda)\left[\int_a^T trf(r)\,dr + tT(1-F(T))\right] \\
= \;& (1-\lambda)\left[t\left(\frac{T-a}{b-a}\right)^{\beta}\beta\left[\frac{T-a}{\beta+1}+a\frac{1}{\beta}\right]+tT\left(1-\left(\frac{T-a}{b-a}\right)^{\beta}\right)\right] \\
= \;& (1-\lambda)\,t\,(12.784+4.219) = (1-\lambda)\,0.4\cdot(12.784+4.219) = (1-\lambda)\,6.8012
\end{aligned}
$$

where the first equality follows from Lemma 4.

So, summing up, the money raised from honest firms is 6.9708. That raised from strategic firms in equilibrium is 6.5461. That raised from strategic firms under the extremal strategy is 6.8012.

E.4 Audit class 26-50

In this audit class we focus on the parameter constellation corresponding to $\beta = 1$. According to equation (35), a fraction equal to

$$
0.046\left((1-\lambda)+\lambda\left(\frac{M-a}{b-a}\right)^{\beta}\right) = 0.042
$$

of all firms in the audit class is audited. This number needs to equal to the fraction of firms audited under the extremal strategy. Under the extremal strategy, the fraction of firms audited is given by the equation

$$
\frac{t}{t+\widehat{\theta}}F(T) = \frac{t}{t+\widehat{\theta}}\left(\frac{T-a}{b-a}\right)^{\beta}.
$$

Equating this to 0.042 and solving for T yields

$$
T = a+(b-a)\left(\frac{t+\widehat{\theta}}{t}0.042\right)^{\frac{1}{\beta}} = a+(b-a)\left(\frac{t+6.954}{t}0.042\right)^{\frac{1}{\beta}} = 47.914.
$$

The money raised from strategic firms under the extremal strategy equals

$$
\begin{aligned}
& (1-\lambda)\left[\int_a^T trf(r)\,dr + tT(1-F(T))\right] \\
= \;& (1-\lambda)\left[t\left(\frac{T-a}{b-a}\right)^{\beta}\beta\left(\frac{T-a}{\beta+1}+a\frac{1}{\beta}\right)+tT\left(1-\left(\frac{T-a}{b-a}\right)^{\beta}\right)\right] \\
= \;& (1-\lambda)\,t\,(28.537+10.917) = (1-\lambda)\,15.782
\end{aligned}
$$

where the first equality follows from Lemma 4.

So, summing up, the money raised from honest firms is 16.076. That raised from strategic firms in equilibrium is 15.264. That raised from strategic firms under the extremal strategy is 15.782.