Audit Rule Disclosure and Tax Compliance*

Enrico Di Gregorio Matteo Paradisi Elia Sartori

November 27, 2025

Abstract

Tax authorities typically concentrate enforcement resources on large businesses, yet small business evasion generates tax revenue losses that often exceed those from multinational profit shifting. We show that authorities can improve small business compliance by strategically disclosing audit-relevant information. We study audit rules that inform taxpayers that audit risk drops discontinuously above a threshold based on predicted revenues. A theoretical model identifies conditions under which disclosure outperforms secrecy and provides a test based on changes in the audit probability jump. We then use more than 26 million tax files (2007–2016) from Italy's Sector Studies, a policy with a disclosed threshold-based design. Taxpayers bunch sharply at the threshold, and bunching correlates with evasion proxies, evasion technologies, and tax incentives. Exploiting a staggered reform that increased the audit risk discontinuity, we find that compliance rises below the threshold and falls above it, yet average reported profits grow by 16.2% in treated sectors over six years. Interpreted through the model, this evidence implies that disclosure outperforms nondisclosure. Estimating the model, we show that by strengthening incentives at the threshold, the Tax Authority could cut the audit budget by up to two-thirds without reducing compliance, freeing resources to target larger firms, and that the reform brought the policy close to the optimum.

JEL Classification: D04, D22, H24, H25, H26, H32.
Keywords: tax compliance, enforcement, evasion, audit, disclosure, firm, bunching.

*Contact information: Enrico Di Gregorio, International Monetary Fund, HQ21900 Pennsylvania Avenue NWWashington, DC, edigregorio@imf.org; Matteo Paradisi, EIEF, via Sallustiana 62, 00187 Rome, matteo.paradisi@eief.it; Elia Sartori, CSEF and Università Federico II, Via Cintia complesso Monte S. Angelo, Edificio 3, 80126 Napoli, elia.sartori@unina.it. We thank Agenzia delle Entrate for granting access to the Sector Study database, and SOSE for their support in extracting and interpreting the raw data. All views expressed are solely our own. This project was made possible through Research Agreement R.U. 0521883.20-12-2018-U with Agenzia delle Entrate. Throughout the development of this project, we benefited from invaluable discussions with and comments from Alberto Alesina, Marcella Alsan, Lorenzo Casaburi, Raj Chetty, Claudio Di Gregorio, Gemma Dipoppa, Michele Fornino, Paola Giuliano, Edward Glaeser, Nathan Hendren, Anders Jensen, Felipe Jordán, Henrik Kleven, Kunal Mangal, Lukas Mayr, Tommaso Nannicini, Nicola Rosaia, Luisa Scarcella, Mark Shepard, Monica Singhal, Joel Slemrod, Stefanie Stantcheva, Giovanna Vallanti, Qiankun Zhong, and Gabriel Zucman. We also wish to thank Paolo Acciari, Marco Manzo, and Antonio Giagnoli at the Ministry of the Economy and Finance for providing us with property tax evasion estimates and ancillary data on tax litigation. Di Gregorio acknowledges the support of NBER and the Peter G. Peterson Foundation.

1 Introduction

Ensuring tax compliance among small businesses and the self-employed is historically a central challenge for tax agencies in developed and developing countries (Alm, Martinez-Vazquez, and Wallace, 2004). In the U.S., imperfect compliance among small businesses results in at least a 7-8% shortfall in total tax revenues.[1] Compared to profit shifting of U.S. multinationals, which amounts to 15% of corporate taxes and 1.32% of total revenues (Wier and Zucman, 2022), small businesses' evasion reduces government revenues by almost one order of magnitude more. Yet, tax authorities tend to skew their audit resources toward large firms so that more than 70 countries adopted special enforcement units for large taxpayers in the last 20 years (Almunia and Lopez-Rodriguez, 2018; Bachas, Fattal Jaef, and Jensen, 2019; Basri et al., 2019).[2] This might reflect a cost-effectiveness principle of tax administration, as auditors expect a relatively higher yield from auditing a large business rather than several small ones for any given budget (Boning et al., 2025). To the extent that tax agencies are unwilling or unable to distribute enforcement efforts equally across firm types, the identification of low cost strategies to promote small firms' voluntary tax compliance becomes essential to tax collection.

This paper provides the first evaluation of audit rule disclosure as a viable strategy to improve the incentives for tax base reporting by small businesses. We refer to audit rules as the criteria that tax agencies routinely adopt to guide audit case selection. Tax authorities seem to value the secrecy of these rules, which they generally keep from the public. Indeed, the choice of disclosure comes with a trade-off. On one hand, revealing what behaviors trigger a tax audit might nudge some taxpayers away from evasion. On the other, those who learn to be at lower risk of an audit might end up reporting a lower tax base. On the net, the impact of disclosure is ambiguous, and this paper aims to characterize and quantify this trade-off using evidence from a real-world policy.

We examine a specific case of audit rule disclosure in which the tax authority announces, prior to tax reporting, the precise threshold above which audit risk declines discontinuously. This threshold corresponds to a predicted revenue level for each firm, creating a firm-specific compliance target. A prominent example of such a rule is the Sector Studies (*Studi di Settore*,

[1] We sum the estimated yearly underreporting and non-filing among individual business income earners, self-employed, and small corporations, and divide by the total collected central government revenues for 2019 Internal Revenue Service (2022). Our inability to break down other tax gap items implies our estimates are lower bounds. We provide further details in section 3.

[2] In a 2019 survey from Italy, our empirical setting, the share of firms reporting any tax inspections over the previous 12 months was 9.9% among firms with less than 20 employees, and 18% among those with more than 100 (The World Bank, 2019).

henceforth SeS), an Italian audit system applied on small businesses and the self-employed.[3] To examine the compliance dynamics that result from this system, we access a novel, confidential database of over 26.6 million SeS files from the 2007-2016 tax period, including the previously unexploited universe of 2007-2010 files. This rich dataset includes small businesses earning less than 5.2 million euros in revenue each year.

Our focus on disclosed threshold rules is motivated by their widespread adoption across countries,[4] their capacity to leverage tax authorities' data and forecasting power to tailor compliance incentives, theoretical findings suggesting that threshold-based rules can maximize tax collection (Reinganum and Wilde, 1985; Sánchez and Sobel, 1993), and the intuition that tax liability notches create strong incentives even in the presence of low structural elasticities (Kleven and Waseem, 2013).[5] Despite their theoretical appeal, limited empirical attention has been devoted to the compliance effects of disclosed threshold rules. To the best of our knowledge, this paper provides the first evidence in this direction.[6]

We leverage the design of the SeS and a theoretical framework to evaluate whether disclosed audit rules based on statistical predictions can enhance tax compliance. We begin by developing a model of audit rule disclosure and deriving a test that assesses the desirability of threshold-based audit rules by examining firms' tax base responses to marginal changes in audit-related incentives. We then exploit a natural experiment that increased the audit risk discount at the disclosed threshold to empirically implement this test. Finally, we estimate the model and explore avenues for improving the current audit rule design.

Our theoretical framework is designed to evaluate the trade-offs involved in disclosing threshold-based audit rules. We model an economy consisting of firms within a given category (e.g., sector and location), where firms are risk-neutral and heterogeneous in both income and their propensity to evade taxes. Under an undisclosed rule, all firms in a category face a uniform audit probability. In contrast, the disclosed regime introduces a threshold: firms declaring revenues below it face a higher audit probability, while those declaring above it face a lower one. Given their evasion propensity, firms sort into four groups based on their true revenue levels. From lowest to highest, these groups are: (i) zero-declarers; (ii) declarers

[3] The Italian government estimates that these taxpayer categories accounted for up to 30.4% of all unpaid tax liabilities during the 2014-2016 period (Ministry of Economy and Finance, 2019).

[4] Among others, the Australia's periodic release of industry benchmarks by the Australian Tax Office (OECD, 2006), Greece's profit margin targets under its self-assessment program (Al-Karablieh, Koumanakos, and Stantcheva, 2021), Mexico's introduction of sector-specific effective tax rates in June 2021, and earlier presumptive taxation schemes like France's *forfait* and Israel's *tachshiv* (Thuronyi, 1996).

[5] Lazear (2006) show how disclosure may enhance aggregate outcomes in contexts beyond tax compliance.

[6] In this regard, our analysis diverges from studies on Large Taxpayer Units, which incentivize taxpayers to *underreport* income to avoid monitoring (Almunia and Lopez-Rodriguez, 2018; Bachas, Fattal Jaef, and Jensen, 2019; Basri et al., 2019), and from work on policies offering audit risk reductions tied to a common economy-wide threshold (Dwenger et al., 2016; Al-Karablieh, Koumanakos, and Stantcheva, 2021).

below the threshold; (iii) bunchers, who report exactly at the threshold; and (iv) declarers above the threshold.

We develop a test to assess whether disclosed threshold-based audit rules outperform undisclosed ones. The test relies on evaluating the revenue effects of marginal changes in the audit probability jump at the threshold—a policy variation that closely mirrors the introduction of a reward regime analyzed in our empirical analysis. We demonstrate that the revenue function is globally concave in the probability jump, provided the overall audit rate is low. Introducing a notch in the tax liability creates strong reporting incentives for taxpayers just below the threshold, which diminish as the audit discount grows. This pattern explains the decreasing marginal revenue gains from larger audit jumps and, in turn, the concavity of the revenue function. This concavity implies that if the increase in the probability jump induced by the reward regime led to higher revenues and a larger tax base, then the pre-policy threshold-based rule outperformed a flat audit regime using at least as many audits. This result holds in several extensions that allow for manipulation margins not targeted by the audit rule, for the manipulation of the audit threshold, for fixed costs of receiving an audit, and a non-flat audit probability in the undisclosed scenario. Interestingly, since disclosed rules work by incentivizing bunching, introducing any additional motive to bunch - e.g. adding fixed costs of receiving an audit, or allowing for additional manipulation margins - might improve the revenues of the Tax Authority.

We turn to the empirical section to validate some key predictions of the model and implement the test by leveraging the introduction of a reward regime as a natural experiment. In the data, we observe a significant spike in the distribution of declared revenues just above the disclosed revenue threshold. Consistent with the model's predictions, we test whether the extent of bunching correlates with various evasion margins and incentives to manipulate. We find that bunching is strongly associated with evasion in VAT, property taxes, income taxes, and with anonymous evasion reports. It is also more pronounced in upstream sectors, where evasion is more difficult due to third-party reporting, and in regions with higher personal income tax rates. Finally, in line with our theoretical predictions, bunching positively correlates with the share of firms declaring zero revenues in a given area.

To evaluate whether disclosed audit rules can improve tax compliance, we exploit a natural experiment that closely mirrors the marginal change in the audit probability jump analyzed in our theoretical model. Beginning in 2011, a staggered reform to the SeS—known as the "reward regime"—enhanced protections for firms that complied with SeS prescriptions and promised to devote more attention to those who did not. This reform widened the perceived audit risk gap around the presumed revenue threshold for firms subject to the new rules. Using a balanced panel from 2007 to 2016, our event-study design shows that firms in treated sectors adjusted their reported revenues closer to the SeS threshold. This behavioral shift occurred both below and above the threshold, supporting our model's prediction that disclosure-based policies create opposite incentives across taxpayers. On average, however, treated firms increased reported revenues by 12% and gross profits by 16.2% over six years. Interpreted through the lens of our model, this evidence suggests that disclosing audit risk discontinuities enabled the tax authority to expand the tax base among small businesses.

Exploiting the identifying variation generated by our natural experiment and data moments from the pre-policy period, we bring the model to the data. We estimate the structural primitives of the economy to match both the reduced-form responses and the distribution of declarations around the threshold prior to the Reward regime. The resulting estimates align with the tax authority's reported pre-reform probabilities and are consistent with the audit rules implemented under the Reward regime. We then use the estimated model to conduct several counterfactual exercises. First, we show that both the Reward regime and the pre-reform system substantially outperform an undisclosed audit rule that operates with the same audit budget. Next, we demonstrate that by further increasing the audit probability jump at the threshold, the Tax Authority could save up to two-thirds of the original SeS audit budget while keeping declared revenues unchanged. This implies that reallocating enforcement resources toward larger firms could be done with minimal compliance costs. Finally, we show that because the Reward regime moves toward a zero audit probability above the threshold, it approximates the optimal policy and yields declared revenues close to the best achievable within this class of rules, whereas the pre-policy rule performs significantly worse than the optimal policy for any plausible marginal audit cost.

Related literature: This paper provides several contributions to the literature on tax compliance and enforcement by studying the revenue effects of selectively disclosing audit rules based on statistical predictions. To our knowledge, this is the first study to examine such disclosure as a compliance tool. Unlike prior work on audit regimes such as Spain's Large Taxpayer Unit (Almunia and Lopez-Rodriguez, 2018) whereby corporations expect stricter enforcement when reporting above 6€ million, our setting encourages low-productivity firms to report more rather than less, consistently with the prescriptions of optimal audit theory.[7] Moreover, we focus on micro firms and the self-employed, who account for a significant share of the tax gap and are typically hard to monitor, making voluntary compliance schemes a promising enforcement strategy. Our approach also differs from studies of audit exemptions above common (firm-independent) thresholds (Dwenger et al., 2016; Al-Karablieh, Koumanakos, and Stantcheva, 2021), by studying the disclosure of firm-specific predictions.

[7] Basri et al. (2019) study a similar scheme with regional Medium Taxpayer Offices in Indonesia, but the exact formula behind firm assignment to these offices is not known.

In this sense, we build on findings from Paradisi and Sartori (2023) and align with recent work on machine learning in audit targeting (Battaglini et al., 2023).

A few papers study how incentives for taxpayers (Dunning et al., 2017; Carrillo, Castro, and Scartascini, 2017) and third parties (Naritomi, 2019; Choudhary and Gupta, 2019; Kumler, Verhoogen, and Frías, 2020) foster quasi-voluntary tax compliance. Unlike tax lotteries, tax amnesties, and temporary audit exemptions, we study a permanent incentive scheme that taxpayers can access autonomously by following predetermined prescriptions. SeS also differ from audit threat letters used in tax enforcement RCTs (Kleven et al., 2011; Pomeranz, 2015; Bérgolo et al., 2017), which do not aim to reveal the structure of the audit system and whose general equilibrium effects and scalability remain uncertain (Slemrod, 2019).

Italy provides a suitable setting for studying small firm tax compliance due to its large share of small businesses and self-employed, which often display high tax gaps (Arachi and Santoro, 2007). SeS taxpayers have been analyzed in both academic and policy work (Santoro, 2008; Santoro and Fiorio, 2011; Santoro, 2017; D'Agosto et al., 2017; Battaglini et al., 2020). This paper is the first to evaluate whether audit rule disclosure enhances the tax base and to assess the effects of the reward regime reform.

2 A model of misreporting with disclosed audit rules

2.1 Firm's Problem and Audit Rules

We analyze a tax evasion model where the economy consists of a "category" of firms, which may represent a sector, location, or any combination of observable characteristics. A continuum of firms earns income y drawn from a compact support $[0, \bar{y}]$ according to CDF $F(\cdot)$, which admits a continuous single-peaked density function $f(\cdot)$ bounded away from zero and which does not vary too rapidly relative to its level.[8] We treat income as exogenous, but we prove in Appendix A that production and declaration decisions are separable. Hence, one can interpret y as the optimal income of firms that are heterogeneous in productivity. Declared income d is taxed at linear rate τ, and any detected evasion is sanctioned at rate $\gamma > 1$.[9] Firms face an audit schedule $p : D \to [0, 1]$ that maps declared income into a probability of being audited. Given their income y, firms choose their reported income d to maximize

$$V(y) = \max_{d} y - \tau d - \tau\gamma \cdot p(d) \cdot (y - d) - c(y - d) \tag{1}$$

[8] Formally, we require that, for all y, $|\mathrm{d}\log(f(y))| < \eta$ for some $\eta > 0$. This assumption, while not necessary for comparative statics, is a sufficient — and overly conservative — condition for the validity of our concavity test. We will discuss it and provide weaker conditions while presenting the test (Section 2.4). Intuitively, this assumption implies that the income distribution is relatively dispersed.

[9] We assume that conditional on receiving an audit, the true income is discovered and the penalty γ is imposed without frictions.

for a manipulation cost function $c : \mathbb{R}^+ \to \mathbb{R}^+$, which is increasing and convex at a decreasing rate, i.e. it satisfies $c(0) = 0, c'(\cdot) \geq 0, c'(0) = 0, c''(\cdot) > 0, c'''(\cdot) \leq 0$.[10]

Since we focus on disclosed threshold-based audit rules, we consider an audit schedule that takes the following form

$$p(d, \hat{y}) = \begin{cases} p_H & d < \hat{y} \\ p_L & d \geq \hat{y} \end{cases} \tag{2}$$

with $p_H \geq p_L$. The case $p_L = p_H$ corresponds to a flat audit rule where all firms in a "category" face the same audit probability. If $p_L < p_H$ the Tax Authority grants an "audit discount" to all firms that declare (at or) above the threshold level $\hat{y}$. This threshold can be interpreted as a signal received by the Tax Authority regarding a firm's true income, potentially derived from a prediction model. A flat audit probability reflects a situation where audit rules and predicted incomes are not disclosed to taxpayers and kept "secret". For the moment, we treat $\hat{y}$, p_H, and p_L as exogenous parameters.

To streamline the exposition, the baseline model simplifies the firm's problem by abstracting from certain adjustment margins. However, our main results remain robust under less restrictive assumptions. In Section 2.5, we extend the framework to incorporate: (i) heterogeneous costs of income manipulation; (ii) manipulation of tax base components not directly targeted by the audit rule; (iii) manipulation of the threshold $\hat{y}$; and (iv) fixed costs associated with being audited.

We gather all proofs of the results presented in this Section in Appendix A.

2.2 Optimal declarations and comparative statics

If the audit rule were flat (*i.e.* $p(d) = p$), then the maximization problem (1) would reduce to choosing the amount of tax evasion that solves

$$\min_e e\tau (1 - p\gamma) - c(e).$$

This holds regardless of the firm's income level y. The interior solution to this problem is

$$e^I(p) = (c')^{-1} (\tau (1 - p\gamma)), \tag{3}$$

which represents the level of evasion chosen by firms for which $y - e^I(p) \geq 0$.[11] Since audit rules in (2) are non-constant, (3) does not directly apply to our setting. However, we can still

[10] The inequality on the second derivative is assumed to hold uniformly: there exists a positive constant $\delta > 0$ such that $c''(x) > \delta$ for all $x \in \mathbb{R}^+$. Convexity is a natural assumption that ensures a well-behaved declaration problem. We assume a diminishing cost curvature, avoiding scenarios with evasion capacity constraints.

[11] By the properties of c, we know that e^I is defined over positive domain, has positive image, and is increasing.

use this condition to characterize the declaration behavior. Specifically, a firm's behavior is characterized by three target evasion levels, e_H, e_L and $\tilde{e}$, where $e_i = e^I(p_i)$ in equation (3) for $i = H, L$, and $\tilde{e}$ solves

$$e_H \tau (1 - p_H \gamma) - c(e_H) = \tilde{e} \tau (1 - p_L \gamma) - c(\tilde{e}). \tag{4}$$

In words, e_i is the optimal evasion if the audit rule were flat at p_i, and $\tilde{e}$ is the level of evasion where the firm is indifferent between concealing $\tilde{e}$ income under p_L and choosing e_H under the (higher) audit probability p_H. We can now describe the declaration behavior of firms that solve problem (1) facing the class of audit rules (2).

Proposition 1. *$e_L \geq e_H \geq \tilde{e}$, with strict inequalities whenever $p_H > p_L$. If $e_H < \hat{y} + \tilde{e}$, the optimal declaration strategy partitions incomes into the following declaration areas*

$$d(y) = \begin{cases} 0 & y < y^{0H} \\ y - e_H & y^{0H} \leq y < y^{HB} \\ \hat{y} & y^{HB} \leq y < y^{BL} \\ y - e_L & y \geq y^{BL} \end{cases}, \qquad e(y) = \begin{cases} y & y < y^{0H} \\ e_H & y^{0H} \leq y < y^{HB} \\ y - \hat{y} & y^{HB} \leq y < y^{BL} \\ e_L & y \geq y^{BL} \end{cases},$$

where $y^{0H} = e_H$, $y^{HB} = \hat{y} + \tilde{e}$, and $y^{BL} = \hat{y} + e_L$.

Notice that e_i and $\tilde{e}$ depend solely on the probability levels and the cost function. These values — derived from (3) and (4) — induce the declaration partition described in Proposition 1.[12] Firms with low incomes, *i.e.* below the optimal evasion level in the p_H regime, declare no income and bunch at 0, as their optimal declaration $y - e_H$ is below the lower bound of zero. We define this region as the $\mathcal{O}$ area. Firms in the income region $\mathcal{H} = [e_H, \hat{y} + \tilde{e})$ declare as if the audit probability were flat at p_H. For some of them (those with income below $\hat{y}$) this is straightforward as all feasible declarations are associated with probability p_H. Firms with income above $\hat{y}$, instead, must choose between optimal behavior under the p_H regime or declaring at $\hat{y}$, which reduces both the evasion rate and the probability of being audited. The level of evasion $\tilde{e}$ determines the indifference between these two options. Firms in the region $\mathcal{B} = [\hat{y} + \tilde{e}, \hat{y} + e_L)$ bunch at declaration $\hat{y}$. Firms in $\mathcal{L} = (\hat{y} + e_L, \bar{y}]$ can behave optimally under a flat p_L audit rule and their declarations above the threshold are disciplined solely by the material cost of mirseporting.[13]

[12]Figure G2 provides a graphical representation of firms' optimal declarations.

[13]The Proposition assumes that $e_H < \hat{y} + \tilde{e}$, which ensures that the region $\mathcal{H}$ is well-defined. If this condition does not hold, taxpayers would switch from declaring zero to bunching at $\hat{y}$, which contradicts our data where we observe interior (non-zero) declarations below the threshold.

Proposition 1 allows us to analyze how firms' optimal declaration behavior responds to changes in the model's parameter, particularly the tax rate and the propensity to evade. To compare economies with varying propensities to evade, we parametrize the cost function as $c_\kappa(\cdot) = \kappa \cdot c(\cdot)$, where κ shifts the cost of evading at any evasion level. A lower κ implies a higher propensity to evade. Since we do not observe the evasion levels e_H, e_L and $\tilde{e}$ for each taxpayer in the data, we derive comparative statics on the masses of the declaration areas. These allow us to empirically test the model's predictions. We denote $M(i)$ the mass of area $i = \mathcal{O}, \mathcal{H}, \mathcal{B}, \mathcal{L}$. The following Proposition outlines the main comparative statics.

Proposition 2. *Let $\tilde{\tau} = \frac{\tau}{\kappa}$. Then, $\frac{dM(\mathcal{O})}{d\tilde{\tau}} \geq 0$, $\frac{dM(\mathcal{L})}{d\tilde{\tau}} \leq 0$ (where the inequalities are strict if $e_H < \bar{y}$ and $\hat{y} + e_L < \bar{y}$, respectively). Moreover, there always exists a threshold $\bar{m}$ such that if $M(\mathcal{L}) \geq \bar{m}$, then $\frac{dM(\mathcal{B})}{d\tilde{\tau}} > 0$.*

Given the definition of $\tilde{\tau}$, the comparative statics can be interpreted both as the effect of increasing the tax rate or as the effect of decreasing the propensity to evade (*i.e.*, increasing κ). The results for the $\mathcal{O}$ and $\mathcal{L}$ areas follow directly from the analysis of the evasion behavior. An increase in $\tilde{\tau}$ raises the incentives to misreport income, resulting in $\frac{\mathrm{d}e_H}{\mathrm{d}\tilde{\tau}}, \frac{\mathrm{d}e_L}{\mathrm{d}\tilde{\tau}}, \frac{\mathrm{d}\tilde{e}}{\mathrm{d}\tilde{\tau}} > 0$. Since the $\mathcal{O}$ area includes all firms with $y < e_H$, $M(\mathcal{O})$ increases as $\tilde{\tau}$ rises. Conversely, the $\mathcal{L}$ area, which consists of firms with $y > e_L$, decreases when $\tilde{\tau}$ increases.

The behavior of the bunching region is more complex. An increase in e_L expands the upper bound of the bunching region, while an increase in $\tilde{e}$ raises the lower bound, narrowing it. To assess the net effect on $M(\mathcal{B})$, we must weigh the positive effect of increasing e_L against the negative effect of increasing $\tilde{e}$, taking into account the densities at each bound. The proof shows that when $M(\mathcal{L})$ is sufficiently large,[14] the densities at both bounds are similar. In this case, it is sufficient to show that e_L grows faster than $\tilde{e}$. This result holds because changes in $\tilde{\tau}$ have a greater effect on the optimal declaration in $\mathcal{L}$ (see (3)) than on the indifference condition (4), where both sides of the equality shift in the same direction.

2.3 Government revenues and their response to audit reforms

As an initial step in studying the desirability of disclosed rules, we write the Tax Authority's revenues as a function of the policy parameters and we study how they are affected by the jump in probability at $\hat{y}$. We set aside revenues from audits and assume that the Authority allocates a limited amount of resources to audits within the sector. These assumptions are motivated by our focus on small businesses, where audit collections contribute negligibly to government revenues due to scarce resources and where audits are less effective because

[14]This occurs if and only if $\tilde{\tau}$ is small.

of the limited potential gains.[15] In this context, the Authority operates with a fixed audit threshold $\hat{y}$ and adjusts the size of the audit probability gap around it, starting from a low baseline μ. To formalize this, we parametrize the audit rule (2), assuming the Authority selects $\Delta \in [0, \min\{\mu, 1-\alpha\mu\}]$ and sets $p_L = \mu - \Delta$ and $p_H = \mu + \alpha\Delta$. If μ is small, it acts as the upper-bound on Δ. The parameter α determines how much p_H increases for each unit decrease in the audit probability above the threshold. The Authority takes μ, α, γ, τ, as well as the cost and income distributions as given and sets Δ to maximize revenues. The resulting revenue function is:[16]

$$R(\Delta) = \mathbb{E}\left[d(y, \Delta)\right] = \mathbb{E}\left[y\right] - \mathbb{E}\left[e(y, \Delta)\right].$$

We derive the following representation to show how widening the audit probability gap (*i.e.*, increasing Δ) impacts the Authority's objective.

Proposition 3. *The marginal revenue from a change in Δ is*

$$\frac{\mathrm{d}R(\Delta)}{\mathrm{d}\Delta} = -\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}\left(e_H - \tilde{e}\right) f\left(\hat{y} + \tilde{e}\right) - M(\mathcal{H})\frac{\mathrm{d}e_H}{\mathrm{d}\Delta} - M(\mathcal{L})\frac{\mathrm{d}e_L}{\mathrm{d}\Delta}. \tag{5}$$

Computing marginal revenues is straightforward.[17] There are two intensive margins, driven by the responses of firms in the $\mathcal{H}$ and $\mathcal{L}$ regions to changes in their respective audit probabilities. Both changes are weighted by the size of the respective region. Importantly, increasing Δ reduces the audit probability and increases evasion only in the $\mathcal{L}$ region, contributing negatively to (5). Additionally, declared income increases discontinuously for the lowest-income "buncher" that is indifferent between optimal evasion under p_H and declaring $\hat{y}$ to receive an audit discount. Hence, the extensive margin affecting total revenues involves the change of the threshold y^{HB}, which shifts one-to-one with $\tilde{e}$ and negatively in response to an increase in Δ. Multiplying the change in $\tilde{e}$ by the declaration gap $(e_H - \tilde{e})$ gives the gain in revenues from additional bunchers.

2.4 Testing improvements over a flat audit schedule

Our goal is to evaluate the effectiveness of disclosed audit rules that offer a reward to taxpayers who declare more than a specified threshold. To do this, we establish a benchmark audit rule where $\Delta = 0$. Under this rule, all declarations are audited with probability μ,

[15] This a scenario where the Authority allocates its budget across sectors, concentrating on those with higher income potential. The remaining sectors, which are of interest for this paper, receive limited resources and generate negligible audit revenues.

[16] The expectation operator integrates over the income distribution F. If evasion costs were heterogeneous, then $\mathbb{E}$ would integrate over the joint distribution $\tilde{F}$.

[17] Figure G2, Panel C, visually illustrates equation (5).

and, according to Proposition 1, firms with incomes below $e^I(\mu)$ declare zero, while those with incomes above $e^I(\mu)$ declare $y - e^I(\mu)$. We then explore the conditions under which the tax authority can improve upon this rule by offering an audit discount to declarations exceeding a threshold. Specifically, we develop a test that examines the change in average declarations in response to a perturbation in Δ when $\alpha > 0$. This policy shift is comparable to the natural experiment of implementing a reward regime.

Our test identifies scenarios where the average declaration behaves as a concave function of Δ: if this holds, observing an increase in average declarations when moving from a positive Δ_1 to a larger Δ_2 indicates that the initial disclosed rule outperforms a flat rule with $\Delta = 0$. Since the reward regime involves a shift in Δ starting from $\Delta > 0$, its causal effect on average declarations is sufficient to establish whether disclosing the dependency of audit rules on predicted income $\hat{y}$ is advantageous.

Theorem 4. ***[Concavity test]*** *Suppose μ is small. If the reward regime increases revenues, a disclosed rule outperforms the undisclosed one. Moreover, if the probability increase below the threshold induced by the reward regime is not excessive — specifically, if $\alpha < \frac{1}{F(\hat{y}+e^I(\mu))} - 1$ — then the disclosed rule not only outperforms the undisclosed one but also achieves this outcome while requiring less budget.*

We establish the first part of the Proposition by proving that, under the assumptions of our framework, the revenue function exhibits infinite concavity as $\Delta \to 0$, which ensures that it remains concave for small values of μ.[18] The second part of the Proposition reinforces the implications of our test by proving that, as long as α remains sufficiently small, disclosure not only improves revenues but it does so with reduced enforcement costs. A small α ensures that the reduction in p_L caused by the reward regime is significant relative to the increase in p_H. Given that the bound on α can approach or even exceed 1, this condition is likely to hold in practice. Indeed, the policy that we study is designed to protect taxpayers who report income at or above their presumed level $\hat{y}$.

The concavity of the objective function arises from the key tradeoff the Authority faces when increasing Δ, as highlighted by condition (5). On one hand, raising Δ boosts the intensive margin declarations of firms in the $\mathcal{H}$ area and incentivizes more firms to bunch

[18]Specifically, the sufficient conditions for infinite concavity at $\Delta = 0$ are that $f(\cdot)$ does not vary too rapidly relative to its level and that $c'''(\cdot) \leq 0$. These conditions are reasonable when the income distribution is regular and there are no evasion capacity constraints. Moreover, they are conservative, as shown in the proof where the requirement for infinite concavity is $-\frac{3}{2} - e^I \frac{f'\left(\hat{y}+e^I\right)}{f(\hat{y}+e^I)} + \frac{e^I}{2} \frac{c'''\left(e^I\right)}{c''(e^I)} < 0$. This condition is also satisfied if: (i) $f'\left(\hat{y}+e^I\right) > 0$ or $f(\cdot)$ is not too steep locally at $\hat{y}+e^I$; and (ii) $x\frac{c'''(x)}{c''(x)}$ uniformly bounded from above by a positive number. Therefore, as we discuss in a separate paragraph in the proof, a weakening of our assumptions would still preserve the validity of our test.

at $\hat{y}$, both of which contribute to higher declared revenues. However, the corresponding reduction in p_L lowers declared income for firms in the $\mathcal{L}$ area. Notably, the marginal change in $\tilde{e}$ approaches infinity when Δ is near zero, indicating that even a slight discontinuity in incentives can lead many firms to bunch, even when evasion elasticity is low (Kleven and Waseem, 2013). As the discontinuity in audit probability grows, the marginal effect on bunchers declarations declines, and so do the marginal gains from further increases in Δ. This pattern of diminishing marginal effects determines the concavity of the revenue function.

Our test in Theorem 4 establishes that it is sufficient to examine whether revenues improve with a marginal increase in Δ to determine if a discontinuity with the same α but a smaller Δ performed better than a flat rule at μ.[19] In the next Subsection we investigate how this claim remains robust to alternative setups, then we test it using data from the Italian SeS policy.

2.5 Extensions and robustness

Heterogeneous cost types: Our assumption of homogeneous evasion costs simplifies the analysis but does not fully capture the heterogeneity in firms' behaviors, as businesses differ in their material and moral deterrents to evasion. Moreover, this assumption is inconsistent with the presence of firms declaring just below the threshold in our empirical application (Section 4.2) as well in other contexts with tax notches (Kleven and Waseem, 2013). With heterogeneous costs, near-threshold declarations would come from firms with incomes just below $\hat{y}$ that face high costs of misreporting, leading them to (almost) truthfully declare. Hence, a model with heterogeneous costs is not only more realistic but necessary to capture some fundamental data patterns. Adding a second dimension of heterogeneity, firms are also characterized by a cost type κ, where $\tilde{F}(\cdot,\cdot)$ denotes the joint distribution of income and cost type. In this case, we represent costs as $c_\kappa(\cdot)$ to emphasize their dependency on cost type. Propositions 1 and 2 extend naturally to this setting, as they describe declaration behavior for a fixed κ and only require to recognize that both interior evasions $e_\kappa(p_i)$ and the threshold evasion $\tilde{e}_\kappa$ depend on κ and to interpret the income distribution as conditional on a specific κ. Aggregate moments can then be obtained by integrating over the marginal distribution of cost types $\tilde{f}_\kappa$. Under mild assumptions about $\tilde{F}$,[20] the marginal revenue expression in Proposition 3 remains robust, and the concavity condition of Theorem 4 also remains valid but must now hold in expectation (see Appendix A).

[19] This does not imply that revenues must be monotonic in Δ: if the increase in e_L and the share of firms in the $\mathcal{L}$ area remain steady across higher values of Δ, revenues may decrease for a sufficiently large Δ.

[20] We require that the support of income does not change with the cost type, so we can exchange the derivative of $R(\Delta)$ and the expected value operator.

Imperfect correspondence between the audit-targeted margin and the tax base: In our theoretical framework, audit rules impose a threshold on a specific margin—declared y—which corresponds with the tax base. While this assumption holds in certain contexts, the framework can be extended to situations where firms manipulate their tax base through margins not directly targeted by the audit rule. In our specific application, the Tax Authority discloses a threshold based on revenues, whereas the tax base is typically profits. There are exceptions, such as subgroups of businesses for whom only revenues are taxed, in which case our simplification is appropriate. However, in most cases, firms may reduce their tax liability by inflating costs, which fall outside the margin targeted by the disclosed audit rule.

Our insights extend directly to two specific cases: (i) when the targeted margin and the tax base are non-separable and linked through a deterministic function (e.g., costs are a function of revenues); and (ii) when the costs of manipulating the two margins are additively separable (e.g., inflating costs and reducing revenues are independent decisions). In case (i), generating revenue x requires a known cost $c(x)$, resulting in a profit of $x - c(x)$. This effectively reduces the problem to a single declaration margin, and our analysis applies directly after redefining the manipulation cost function accordingly. In case (ii), the availability of an additional margin to manipulate the tax base makes declaring above the threshold more attractive, as firms can further benefit from the reduced audit probability. Under separability, this additional margin does not influence behavior on the targeted margin conditional on being in regions $\mathcal{H}$ or $\mathcal{L}$, but it does amplify the incentive to bunch. In Appendix B.1, we show that the behavior of the marginal buncher remains qualitatively unchanged, and the sufficient conditions for government revenue concavity are identical to those in the baseline model. Since our results hold in both extreme cases—non-separable and fully separable manipulation margins—they should be robust to intermediate cases where manipulating the audit-targeted margin affects the marginal cost of manipulating other components of the tax base.

Manipulation of the threshold $\hat{y}$: Thus far we have assumed that the threshold $\hat{y}$ was exogenous. This assumption can be problematic in cases where the threshold itself is determined by some self-reported characteristics that can be manipulated by firms, e.g. in the context of SeS where $\hat{y}$ is partly based on reported cost items. In Appendix B.2 we show that relaxing this assumption does not qualitatively affect our main results. We assume that firms draw an exogenous threshold $\hat{y}_0$ that can be interpreted as the average predicted income within a firm's class. Firms can, however, manipulate this threshold by selecting a new value $\hat{y}$ at an increasing and convex cost. In this setup, firms choose both their income declaration and the extent to which they manipulate the threshold. By adjusting the threshold, firms shift the probability jump to a lower declared income, allowing them to face probability p_L while evading a larger fraction of their true income. We show that firms engage in threshold manipulation only if they also declare at the post-manipulation threshold. This expands the bunching region, which we now define as the mass of firms declaring at their chosen threshold. As disclosed rules work insofar as they incentivize bunching, this additional manipulation margin might benefit the authority because it makes bunching more desirable.

For the validity of our test, we need to ensure that the properties of government revenues as Δ vanishes remain the same.[21] We show that this is the case as the behavior of the marginal buncher is unchanged.

In the context of SeS, this model extension is conservative because it overlooks additional costs associated with adjusting $\hat{y}$. In SeS firms must reduce certain cost items in their declaration to reduce their presumed revenues and manipulate the threshold. Consequently, while they may report lower revenues when the threshold is lower, their tax base and in turn their tax liability increase due to reduced declared costs. Therefore, even if the threshold manipulation cost was small, it is not obvious that firms would find it beneficial to adjust the threshold.

Fixed cost of receiving an audit: A natural extension of our model introduces a fixed cost associated with undergoing an audit, capturing either tangible (e.g., foregone production) or intangible (e.g., psychological stress of being under scrutiny) costs. While firms' evasion behavior within regions $\mathcal{H}$ and $\mathcal{L}$ remains unchanged, the fixed cost makes the low-probability audit regime relatively more attractive (as it avoids this additional cost) and thus strengthens the incentives to bunch. Hence, as we formalize in Appendix B.3, extending the model to include this feature reinforces our results.

Non-flat probability under undisclosed rules: Our analysis assumes that firms perceive a flat audit probability independent of declared income in counterfactuals where the audit threshold is undisclosed. This assumption may overlook that firms could link larger levels of evasion to increased audit risks even without explicit knowledge of an audit threshold. If audit probabilities were non-constant, firms would solve a new optimization problem $\min_e e\tau\left(1-p\left(e\right)\gamma\right)-c\left(e\right)$, where $p\left(e\right)$ is an audit probability responsive to evasion levels. This introduces a wedge $e\tau p'\left(e\right)$ to their first-order condition (3), reflecting that the inframarginal units of evasion e are detected with higher probability $p'\left(e\right)>0$. Compared to a constant audit probability ($p'=0$), this wedge would reduce evasion, particularly when p' is large. However, we consider this adjustment to have limited empirical impact in our context, because the spread of audit likelihood across a wide range of feasible declarations in the absence of an explicit threshold would likely yield a small p'.[22] Additionally, a flat

[21] Results hold also in the extreme case where the manipulation cost is zero for a fraction of firms.

[22] For instance, a linearly decreasing audit probability would result in $p'\left(e\right)=\frac{p(0)}{\bar{y}}$, where $p\left(0\right)$ is the audit in the lowest possible declaration and $\bar{y}$ is the range of declarations. A large range of feasible declarations

audit probability is a reasonable approximation for our analysis since our model focuses on specific taxpayer "categories". While firms might gauge their audit probability as relatively high or low compared to other taxpayer categories, they generally lack detailed insights on how evasion *within* their category influences audit risks. Nonetheless, firms might still perceive a decreasing audit risk on both sides of the threshold. If such a gradient exists both after disclosure and in the counterfactual (*e.g.*, an audit schedule with negative slope and a drop in level at the threshold), then our simplification is unlikely to be restrictive. In this case, the additional wedge introduced by $p' \neq 0$ would be present in both the disclosed and undisclosed scenarios, implying that any distortion from a decreasing audit probability would in large part offset across the two settings.

3 Small businesses evasion and the Italian Sector Studies

A substantial share of tax liabilities goes uncollected due to evasion by small businesses. In Italy, estimates suggest that between 40% and 70% of income earned by individual businesses is not reported (Galbiati and Zanella, 2012; MEF, 2023). Comparable levels of underreporting are observed among similar taxpayer groups in Denmark and the UK (Kleven et al., 2011; HRMC, 2021). Figure 1 compares evasion rates as a share of collected revenues for small and large businesses across four countries: the US, Italy, UK, and Australia. In all cases, small businesses exhibit significantly higher evasion rates.[23] Both Italy and the US lose around 7% of their collected revenues to small business evasion, compared to just 1–2% from larger firms. Moreover, in every country except the UK, small business evasion exceeds the revenue losses from multinational profit shifting (Wier and Zucman, 2022). In the US, for example, small business evasion is almost six times larger than profit shifting. In sum, evasion by small businesses accounts for substantial revenue losses and exceeds in magnitude more widely studied forms of tax evasion and avoidance.

Disclosing audit rules: In response to large tax gaps, in 1998 the Italian government introduced an innovative auditing tool called *Studi di Settore* (Sector Studies, henceforth, SeS), targeting non-employee taxpayers with revenues up to 5.2€ million.[24] Since then, individual taxpayers, partnerships (pass-through businesses), and small corporations file every year towards their Sector Study, and are subject to tax audits ensuing from the analysis of the supplied information.

(large $\bar{y}$) would induce a low $p'(e)$. Moreover, Italian data on SeS suggest that $p(0)$ is small because the audit probability on those declaring on the lowest percentiles is around 2.5% (Battaglini et al., 2023).

[23] Our definition of small businesses includes the self-employed, passthrough businesses, and corporations below a certain level of assets, consistent with the taxpayer population covered by Italy's Sector Studies.

[24] During our primary sample period (2007–2010), taxpayers could opt out of SeS by entering a minimum taxpayer regime, provided they reported no more than 30,000€ in the prior tax year.

SeS provide taxpayers with a file-specific discontinuity in the probability of experiencing an audit based on reported revenues. In collaboration with SOSE, a publicly-owned analysis company, the Italian Revenue Agency (*Agenzia delle Entrate*) annually estimates sector-specific linear models of presumed revenue, drawing on previous declarations covering business turnover, operating costs, workforce composition, physical capital, input quantities, and the size and location of business premises. Each year, businesses report on these dimensions, enabling the model to generate a presumed revenue figure specific to their annual file. According to Law 146/1998, any revenue declaration falling below this presumed threshold may trigger a tax assessment.[25]

The timing and transparency of the SeS policy create incentives for taxpayers to align their reported revenues with the presumed revenue threshold. As illustrated in Figure G1, production activities for a given tax year are completed months before tax filing season, during which taxpayers fulfill both their tax and SeS requirements. Filing deadlines typically fall in June or by the end of September, approximately six months after production decisions are finalized for the relevant tax year.[26]

Taxpayers can learn their SeS thresholds at no cost. Each year, between February and May, the Revenue Agency releases *Gerico*, a free software tool that assists taxpayers with their SeS filings.[27] *Gerico* stores the sector-specific coefficients estimated by SOSE. By inputting relevant accounting and structural data, taxpayers can view their threshold before submitting the tax file and are allowed to adjust their declarations. While using *Gerico* is the quickest way to access the presumed revenue threshold, the Revenue Agency also annually publishes technical reports detailing the estimation procedure. Additionally, *Gerico* facilitates the dissemination of model updates, which are mandated by law to occur at least every three years, on a sector-specific schedule.[28]

Despite its sophistication, SeS is just one of several compliance tools available to the administration. Taxpayers can still trigger audits for reasons unrelated to their position relative to the SeS threshold.[29] Importantly for our analysis, the residual audit risk independent of

[25]The opening statement of Law 146/1998 makes it explicit: *"Tax assessments based on Sector Studies [...] shall apply to taxpayers [...] when declared revenues or remunerations are less than the revenues or remunerations which may be determined on the basis of such Studies"*.

[26]Deadlines are generally set for early summer for paper filers and early fall for those required to file online.

[27]A yearly press release announces *Gerico*'s availability on the Revenue Agency website, and as shown in Supplementary Material S.7, Google searches for "gerico" in Italy peak around the tax seasons.

[28]Model revisions involve re-estimating the presumed revenue functions using updated data, potentially impacting the selection of variables and the size of their associated coefficients.

[29]Italian tax enforcement also includes the *Guardia di Finanza*, a police force focused on fighting tax crimes. While they may use information from a taxpayer's SeS file to initiate investigations, their primary role is targeting tax-related criminal activity. Most standard file audits, however, are conducted by the Revenue Agency, which is also the only entity authorized to impose additional tax payments.

SeS remains stable around the presumed revenue threshold. Moreover, crossing this threshold confers no significant fiscal advantage aside from a reduction in audit risks, allowing us to attribute observed revenue responses directly to the audit incentives established by SeS.

Reward regime as a test for improved compliance: In Section 2, we introduced a test to evaluate whether implementing a disclosed threshold rule can enhance compliance. This test relies on a perturbation of the probability jump at the threshold and we exploit a reform of the SeS system to operationalize it. Beginning in 2011, the Italian government reinforced the discontinuity in incentives linked to SeS reporting through Law Decree 201/2011, which established the *regime premiale* (reward regime) to expand audit protections for SeS-compliant taxpayers. Table G1 compares the compliance requirements and benefits before and after this reform. The reward regime was gradually applied across SeS sectors at the beginning of each tax season.[30] It provided audit exemptions from additional investigations beyond SeS and shortened the audit statute of limitations by one year. To qualify for these benefits, businesses had to report revenues at or above the presumed level ("congruence" in the SeS framework) and fall within acceptable ranges on sector-specific accounting indicators ("normality" and "coherence"). Meanwhile, the reform encouraged tax authorities to increase enforcement among non-compliant businesses.

4 Data and Descriptive Facts

4.1 Administrative Data on Sector Studies

To analyze taxpayer behavior within the SeS framework, we use the universe of administrative SeS files from 2007 to 2010. We complement our main dataset with an unbalanced panel extending through 2016, including all taxpayers who filed continuously between 2008 and 2010. To our knowledge, this is the first paper to exploit all SeS files available in a given year. In total, the data encompasses nearly 26.7 million SeS declarations submitted by over 4.7 million Italian micro-businesses and self-employed individuals. Each tax year from 2007 to 2010 alone contributes more than 3.4 million files.[31]

The data provide detailed information on taxpayers' economic activity for the relevant tax year, covering reported revenues, gross profit or income, workforce size, wage expenditures, both a 6-digit industry code and an administrative SeS industry code, as well as business location details.[32] Importantly, each file includes the precise SeS presumed revenue threshold,

[30] Each tax season, the Revenue Agency released a list of sectors eligible for the new incentives. By 2016, most businesses in manufacturing, commerce, and services were covered, while professionals were largely excluded until a comprehensive SeS reform began in 2018.

[31] Supplementary Material S.5 provides an overview of this dataset.

[32] The majority of SeS filers are single-establishment businesses, with over 98% of the 2007-2010 files submitted by taxpayers who remain in their original province throughout the observed period. Approximately

enabling us to assess the relative distance between the taxpayer's declared revenues and the threshold disclosed by *Gerico* before filing.

SeS encompass a wide variety of firm types with different legal statuses. Nearly two-thirds of the 2007-2010 files come from individual businesses and self-employed professionals (64.8%), while the remainder includes partnerships (19.5%) and corporations (15.7%). A firm's geographic location, along with its legal status, determines its applicable profit tax regime: individuals and partnerships are subject to personal income taxes (PIT), with partnerships treated similarly to S-corporations in the U.S. for tax purposes, while corporations are subject exclusively to corporate income taxes.

Additional sources: A wide range of additional sources complements our administrative data.[33] We build a comprehensive database of evasion levels across regions and tax bases with information from the existing literature and administrative reports. Additionally, over 620,000 anonymous evasion reports submitted to *www.evasori.info* enable us to develop our own misreporting proxies for the years 2008-2011. We also gather data on local tax rates and tax litigation from the Ministry of Finance and the Economy. The Italian Institute of Statistics (*ISTAT*) provides input-output tables detailing sectoral exposure to final consumers, which we use to categorize industries as upstream or downstream along the supply chain.

4.2 Bunching at the SeS threshold and tax manipulation

We begin our empirical analysis by presenting descriptive findings on reporting behavior within the SeS system. This section has three main objectives: *i)* documenting the degree of bunching at the disclosed revenue threshold to demonstrate the saliency of the incentives created by the audit system, *ii)* identifying some correlations predicted by our framework between bunching behavior, various evasion indicators, and incentives to manipulate the tax base, and *iii)* testing our assumption that bunching reflects misreporting rather than production responses. Our findings indicate that taxpayers bunch with high frequency at their presumed revenue level. This behavior strongly correlates with both the ease and incentives for misreporting, but it shows no link to actual production adjustments.

Measuring bunching: Figure 2 Panel A presents the distribution of reported revenues relative to $\hat{y}$, leveraging the universe of SeS files submitted by single-sector businesses from 2007 to 2010. The horizontal axis represents the distance of reported revenues from the file's associated $\hat{y}$ (as a percentage of $\hat{y}$). A significant spike appears within 1 percentage point of

95% of these files correspond to one of the 110 provinces existing at the time, and 77% of personal income taxpayer files can be further assigned to over 8,000 municipalities. Municipalities are also grouped within 686 local labor markets as defined by ISTAT.

[33] We detail these sources in Supplementary Material S.6.

$\hat{y}$, suggesting that a substantial portion of taxpayers declare revenues at or just above the threshold to mitigate their audit risk.

To measure the extent of taxpayer bunching at the presumed revenue threshold, we build an empirical counterfactual distribution $g^c(\cdot)$ to capture the declaration patterns that would arise under a constant p_L probability. This approach leverages the fact that businesses declaring above $\hat{y}$ face probability p_L. We therefore infer from their behavior the counterfactual distribution following the method in Kleven and Waseem (2013), which requires selecting a polynomial order to fit the distribution and determining a threshold y^u that bounds the bunching region to the right of $\hat{y}$. Bunching is then quantified as the ratio of the observed distribution to the counterfactual distribution within the bunching region. We calculate standard errors through a bootstrap procedure with 1,000 iterations. Appendix C provides further details on the estimation.

Figure 2 Panel B illustrates our bunching estimate across all filers for the 2007-2010 tax period. In this example, the counterfactual distribution closely aligns with the empirical distribution to the right of $\hat{y}$ up to y^u. To the left of $\hat{y}$, the empirical distribution falls below the counterfactual, with the difference between them representing the "missing mass" attributable to bunching behavior. In our baseline, we observe significant bunching, quantified at 9.56 (bootstrap sd $=$ 0.61). This bunching leads to higher revenue reports relative to the counterfactual. The additional revenues from bunchers correspond to a uniform rightward shift of the counterfactual distribution equivalent to 1.13% and 3.05% of the observed mean and median revenues, respectively. Table G2 reports the sensitivity of our bunching estimates to the choice of polynomial order and y^u. Our baseline estimate lies on the lower end of the estimates distribution, reflecting our conservative definition of excess bunching, which attributes any extra mass to SeS incentives only within 1 percentage point of presumed revenues. The sizable bunching estimates allow us to conclude that the incentives created by the disclosed audit rule are salient to taxpayers and likely drive significant behavioral responses.

Consistently with our model, we provide two pieces of evidence suggesting that production responses are not the primary driver of the observed bunching. First, as shown in Figure 2, bunching is very sharp at the threshold. If firms were adjusting production, they would need near-perfect foresight about the location of $\hat{y}$, which is unlikely given the several-month delay between production decisions and the threshold disclosure. This evidence would also be hard to reconcile with reasonable output elasticities Best et al. (2015). Second, if production were influenced by SeS incentives, we would expect to see "learning" over time as businesses gradually align production to the threshold after each SeS model update. However, using sector-specific updates every three years, we find no evidence of increased alignment to the threshold over time following new SeS model introductions (Appendix D).

Bunching and evasion behavior: In the next step, we investigate whether bunching correlates with attitudes towards evasion and incentives to evade, as predicted by Proposition 2. Specifically, we study the correlation between bunching of SeS files within each of the 110 Italian provinces or 686 local labor markets (LLMs) in 2007-2010 and available local proxies of evasion across several tax bases, while controlling for regional fixed-effects and value added per capita.[34] Figure 3 documents a positive and often statistically significant correlation between the local bunching estimate and twelve different evasion measures, including one derived from scraping 620,000 anonymous evasion reports from the web. To summarize the magnitude of the correlation, we create an evasion index from a first principal component analysis of administrative-data-based proxies. A one standard deviation increase in this index is associated with a 0.5 standard deviation increase in bunching, indicating a strong relationship between bunching intensity and evasion behavior.

The positive association between evasion behavior and bunching across various metrics may stem from both the relative ease of misreporting (lower κ) and higher payoffs for evasion (higher τ). We provide two distinct sets of evidence in this direction. Consistent with the idea that bunching reflects an ease of misreporting, we observe that bunching is more pronounced in downstream sectors with limited third-party reporting, and in smaller and less complex legal forms of businesses (Figure 4). Moving from an upstream sector with zero share of sales to final consumers to a downstream one that only sells to final consumers is associated with a 5-point statistically significant increase in bunching over an average slightly above 9. Additionally, the bunching distribution among self-employed—who are subject to fewer accounting requirements—lies almost entirely to the right of the distribution for corporations, which have stricter standards.[35]

Furthermore, we investigate the correlation between bunching and the share of zero revenue declarers in a sector. An implication of Proposition 2 is that if a common driver—such as the evasion cost or the tax rate—affects $M(\mathcal{O})$ and $M(\mathcal{B})$, then the masses of the two areas should positively correlate. Figure 5 supports this prediction, showing that a 1 percentage point increase in the share of zero declarers is associated with a 0.5-point increase in bunching, against an average of 9.2 across LLMs.

5 Reward System: testing the effectiveness of disclosed audit rules

In this Section, we rely on a natural experiment *i)* to evaluate the impact of disclosure-based policies based on the theory we developed in Section 2, and *ii)* to study the effects of audit rule disclosure on a broader set of compliance margins. Conveniently, the staggered

[34] Supplementary Material S.8 provides further details on this exercise and additional correlations with evasion proxies.

[35] Supplementary Material S.8 provides additional correlations between bunching and tax incentives.

introduction of the 2011 reward regime (*"regime premiale"*) mirrors the logic of our original policy exercise where disclosure shapes reporting incentives in opposite directions depending on each taxpayer's position relative to the threshold.

Starting in 2011, the Italian government has promised stronger audit exemptions for taxpayers complying with SeS prescriptions, while intensifying enforcement for non-compliance. The new regime influenced audit risk perceptions in opposite ways depending on the location of the taxpayer relative to the presumed revenues threshold. Those planning on reporting revenues above the presumed level and meeting various accounting indicators defined by the tax authority received greater audit protection so that $p_{L,\text{Reward}} \leq p_{L,\text{Pre-Reward}}$. Conversely, non-compliant taxpayers faced increased audit probability. The combination of these measures implies that the perceived audit risk discontinuity at the threshold became more pronounced post-reform: $\Delta\,(\text{Reward}) \geq \Delta\,(\text{Pre-Reward})$.

We leverage the staggered inclusion of SeS sectors into the reward regime from 2011 to 2016 to evaluate the reform's effects. We focus on 155 treated sectors across manufacturing, commerce, services, and the skilled professions, using a balanced panel of businesses that consistently filed under SeS from 2007 to 2016.[36] Given that each SeS sector s joined the regime in a specific tax year t, we employ an event-study design to estimate the following specification:

$$y_{s,t} = \lambda_s + \gamma_t + \sum_{q=-k}^{+k'} \beta_q \cdot I\left(Q_{s,t} = q\right) + \sum_{r=2007}^{2016} \delta_r \cdot X_s \cdot I\left(t = r\right) + \varepsilon_{s,t}. \tag{6}$$

For each sector-by-tax year outcome $y_{s,t}$ analyzed below, coefficients β_q quantify the effect of sector inclusion into the reward regime during period q relative to the sector's entry year. These effects are identified under the assumption of parallel paths, meaning that, in the absence of the reform, outcomes in treated sectors would follow similar trends to those in not-yet-treated sectors. To enhance robustness, we include sector and tax year fixed effects (λ_s and γ_t), alongside a vector X_s of pre-treatment sectoral characteristics interacted with tax year dummy variables.[37] To ensure representativeness, we weight sectors by the number of SeS files submitted at the start of the sample period, capturing the behavior of the average taxpayer in our dataset. Standard errors are clustered at the sector level. Additionally, we conduct robustness checks using the de Chaisemartin and D'Haultfœuille

[36] Unlike other major SeS industries, the Revenue Agency has included only three out of twenty-four SeS sectors in the skilled professions by 2016 before a comprehensive SeS overhaul in 2018 (Figure G3). Our results may thus not be fully representative for all professional groups.

[37] These controls include dummy variables for the categories of manufacturing, commerce, services, and professions, as well as 2007–2010 averages of revenues, gross profits, employment cost incidence on turnover, and the annual growth rates of employment cost rates and revenues.

(2020) correction, addressing potential biases in two-way fixed-effects models when treatment effects are heterogeneous.

5.1 Distribution shifts caused by the introduction of the Reward System

Disclosure-based policies, such as the reward regime, shift perceived audit risk asymmetrically by lowering it above the disclosed threshold and raising it below. Consequently, reporting incentives change in opposite directions on either side of the threshold, and increased bunching may reflect taxpayers adjustments of opposite signs.

We begin by analyzing how small businesses adjust their declarations based on their position relative to the threshold. Taxpayers are grouped into six groups according to their distance from the presumed revenue threshold $\hat{y}$ in the year before their sector's reform: within 5, 5–10, and more than 10 percentage points below or above $\hat{y}$. Using the staggered rollout of the reward regime, we estimate equation (6) separately for each group, using the ratio of declared revenues to the firm-specific threshold as the outcome.[38]

Figure 6 Panel A confirms that small businesses react to the change in audit incentives, in line with our theoretical predictions. Taxpayers facing higher audit risk below the threshold increase compliance by declaring more, while those enjoying greater protection above the threshold reduce compliance. Firms originally reporting below the threshold raise their declarations by between 10 and 30 percentage points of $\hat{y}$, whereas those above reduce them by about 5 to 15 percentage points. As predicted, responses are more pronounced for firms farther from the threshold. These taxpayers are more likely to remain on their original side and can adjust more freely without crossing the threshold. Moreover, they also need to make larger adjustments to reach the threshold if they choose to bunch, further amplifying their observed response.

Additional evidence of these opposing compliance responses emerges when using a compliance indicator produced by the Tax Authority and linked to each firm's tax declaration. This binary index is designed to isolate the group of taxpayers whose declarations align with the expectations of the Tax Authority. Firms initially below the threshold increase their probability of compliance by between 20 and 40 percentage points, whereas those above the threshold see a decrease of about 30 percentage points.

We then further investigate the dynamics driving the opposing patterns of declaration behavior on either side of the cutoff. Specifically, we divide taxpayers into the $\mathcal{H}$ and $\mathcal{L}$ declaration groups defined in our theoretical framework. For each group, we track the distribution of filers across one-percentage-point bins relative to the presumed revenue threshold

[38] Since grouping is based on pre-reform declarations, concerns about mean reversion may arise. However, our difference-in-differences framework leverages untreated sectors as controls, allowing us to net out these dynamics.

over subsequent years. We estimate (6) using the share of taxpayers from a given area in a given bin as the outcome. This approach allows us to provide a detailed description of how the reform reshaped revenue declarations relative to the presumed threshold. Figure G4 illustrates these dynamics. Each panel displays the average of the six treatment coefficients β_q for each one-percentage-point bin, accompanied by their 95% confidence intervals. A green band in the background highlights the range where each group was located in the year prior to the introduction of the reward system. Bunching increases and is driven by businesses on both sides of the cutoff. Regardless of their initial position relative to $\hat{y}$, the reform's amplified audit risk gap at the threshold encourages more taxpayers to declare at or just above $\hat{y}$. The sharper decline in bin shares just below the threshold—particularly among those closer to it—reflects lower adjustment costs, as these taxpayers need to manipulate a smaller portion of their revenues to reach the threshold. A symmetric pattern appears above the threshold: businesses just above $\hat{y}$ bunch to a greater extent. In summary, the increase in bunching is primarily driven by taxpayers who were initially located closer to the threshold.

5.2 Improvements over flat rules: the effects of the Reward System

We now quantify the total period-by-period effects of the reward system on several reported margins, interpreting the results through the lens of our theoretical model. Figure 7, Panel A, displays the full set of β_q coefficients from (6), where the dependent variable is mean reported revenues by sector and tax year, expressed in logarithms. Prior to the reform, treated and control sectors follow similar reporting trends. After the reform, reported revenues in treated sectors rise sharply—by an average of 2.4% in the first year, reaching about 20.4% by year six. The steady growth of the coefficients suggests that it may take time for businesses to fully adjust to the new system.

Panel B shifts the focus to gross profits, capturing net reporting behavior. As with revenues, the stronger audit incentives introduced by the reward regime appear to have substantially expanded the reported tax base. On average, firms in treated sectors report gross profits 16.2% higher each year than those in sectors yet to adopt the reform. Cumulatively, this corresponds to an additional €33,671.77 in taxable profits per average treated business.[39]

The increase in profits is smaller than the rise in revenues because the policy targets only the revenue declaration margin while taxpayers can adjust other margins, such as costs, to offset the impact on the tax base. This behavior is consistent with evidence from Carrillo, Pomeranz, and Singhal (2017) in a different context. Our model extension shows that the concavity test remains valid even in this setting. To assess the extent to which inflated costs neutralize part of the increase in revenues, Panel C of Figure 7 examines the difference

[39] Figure G5 presents the estimates with outcomes expressed in euros.

between revenues and profits, an aggregate proxy for reported costs. Treated sectors report average costs 2% higher than control sectors in the first year post-reform, increasing to 20.7% by the final observed year. This suggests that while the reward system boosted reported revenues, it also spurred a proportional increase in reported costs, keeping profit rates stable. Nonetheless, because revenues rose overall, the net effect on the tax base remains positive.

Our results show that both revenues and profits—the main components of the tax base—increased after the reward regime's introduction. This has two key implications. First, disclosed audit rules could be locally improved by increasing the probability jump at the threshold. Second, by Theorem 4, the pre-reward regime rule outperformed a flat audit rule conducting a strictly larger number of audits. Overall, our evidence supports the desirability of SeS over flat, undisclosed rules in this context.

For consistency with the analysis in 5.1, we use a classical two-way fixed effects model as our baseline specification. To test the robustness of our findings, we also apply the estimator of de Chaisemartin and D'Haultfœuille (2020), which corrects for potential biases from heterogeneous treatment effects across cohorts and time. The results remain consistent, with similar patterns and magnitudes.

6 Checking for improvements and optimal disclosed rules

In this Section, we investigate how to improve on the existing disclosed rules in SeS and consider non-budget neutral policies. Our concavity test in Section 2.4 speaks about the possibility of disclosed rules to improve the revenues without increasing the budget. However, if we use an estimate of the marginal cost of running an audit, we can also study policies that are non-budget neutral and do create a tradeoff between budget increase and revenue increase. We study the desirability of adjusting the audit discount at $\hat{y}$ by estimating the primitives of the model using data moments derived from the natural experiment discussed in Section 5.

6.1 Model estimation

To bring our model to the data we first define the set of primitives to estimate, then we discuss the data moments that we employ in our estimation. Appendix F provides further details on the structural model and the estimation procedure.

Model primitives: We model firms as heterogeneous in both revenues and evasion costs, defined by type (y, κ). A fraction δ^H of firms are fully honest and always declare their true revenues, regardless of the audit environment. Firms types are distributed according to a truncated bivariate normal distribution over $([0, \bar{y}] \times [0, \infty))$, with mean vector (μ_y, μ_κ)

and variance-covariance matrix $\Sigma = \begin{pmatrix} \sigma_y^2 & \sigma_{y\kappa} \\ \sigma_{y\kappa} & \sigma_\kappa^2 \end{pmatrix}$. Hence, the model allows for correlation between a firm's revenue potential and its cost of evasion.

Firms choose their level of evasion optimally, given the tax rate τ, enforcement penalty γ, and a cost function $c(e,\kappa) = \frac{1}{\kappa\left(1+\frac{1}{\epsilon}\right)} (\kappa e)^{1+\frac{1}{\epsilon}}$, where e denotes the amount evaded. The parameter ϵ captures the elasticity of evasion to the net of tax liability and is held constant, while κ varies across firms.

To capture the impact of policy changes—specifically, the introduction of the Reward Regime—we allow the probabilities p_L and p_H to vary pre- and post-reform. Using the notation of our theory, the pre- and post-Reward regime equilibria are defined by μ, Δ, α, and Δ'.

Combining firm-level primitives and the audit policy, the full vector of structural primitives is

$$\theta = \left(\underbrace{\delta^H, \mu_y, \mu_\kappa, \sigma_y^2, \sigma_\kappa^2, \sigma_{y\kappa}, \epsilon}_{\text{Firm types}}, \underbrace{\mu, \Delta, \alpha, \Delta'}_{\text{Audit rules}} \right),$$

yielding a total of 11 parameters to estimate. The first sub-vector of primitives is sufficient to determine firms' evasion behavior under different audit regimes. By quantifying it, we will be able to conduct a counterfactual analysis of policy interventions.

Theoretical moments and data counterparts: The model provides a set of 11 theoretical moments derived from firm behavior under the pre- and post-Reward Regime environments. These moments include the mass of firms in the six regions studied in Figure 6 in the pre-policy period. We split the region just above $\hat{y}$ into two regions: one between $\hat{y}$ and 1 percentage point above it (the bunching area), and one between 1 and 5 percentage points of above $\hat{y}$. This provide a total of 7 regions. For a given region $\mathcal{A}$, the theoretical size of the region is $M(\mathcal{A}) = \mathbb{E}_\kappa \left[F_{y|\kappa}\left(\bar{y}^{\mathcal{A}}\right) - F_{y|\kappa}\left(\underline{y}^{\mathcal{A}}\right)\right]$ where $\underline{y}^{\mathcal{A}}$ and $\bar{y}^{\mathcal{A}}$ are the lowest and largest incomes declaring in region $\mathcal{A}$. Moreover, we employ the average declarations below $\hat{y}$ and above $\hat{y}$ before and after the policy.[40] We denote them by $\bar{d}^t_{\mathcal{A}}$ with $t \in \{Pre, Post\}$. Together, these 11 moments link firm behavior to the underlying structure of types, enforcement, and audit design.

We find a data counterpart to the theoretical moments using the observed declared revenues and their distribution around the threshold. The classification into the seven regions is straightforward since it is based on firms' pre-policy declared revenues that we observe. Given this classification, we quantify the masses of firms in each region, along with the average de-

[40] To exclude the bunching region and capture average declarations in the $\mathcal{L}$ area, we only consider firms declaring more than 5 p.p. above $\hat{y}$ to compute the average declaration above the threshold.

clared revenues above and below $\hat{y}$ before the policy is implemented. We express declared revenues in percentages of the threshold to be able to compare firms with varying thresholds within a given sector. To quantify the post-policy moments, we apply a difference-based approach. For each given area $\mathcal{A}$, we express the post-Reward Regime declared revenues as $\bar{d}_{\mathcal{A}}^{Post} = \bar{d}_{\mathcal{A}}^{Pre} + \delta\bar{d}_{\mathcal{A}}$, where $\delta\bar{d}_{\mathcal{A}}$ represents the causal effect of the policy on $\bar{d}_{\mathcal{A}}$. These effects are estimated using the identification strategy detailed in Section 5.[41]

Estimation: We estimate the model focusing on downstream sectors and using a simulated method of moments (SMM) in two stages. The estimator minimizes the mean squared distance between simulated and empirical moments. In the first stage, we estimate the primitives using an identity matrix with equal weights on all moments. Then, we construct a weighting matrix for the second stage based on the empirical variance-covariance matrix of theoretical moments estimated across multiple iterations of the economy estimated in the first step. This matrix is then used in the second stage to obtain efficient estimates. Details on the estimation procedure are provided in Appendix F.

Table 1 reports the estimated parameters together with key theoretical moments. The estimated probabilities of audit in the pre-Reward regime are $p_H = 0.148$ and $p_L = 0.118$. These values, and in particular the probability jump, are in line with the available evidence from administrative sources and policy reports, which place the corresponding probabilities around 0.105 and 0.07, respectively.[42] In our estimates, the Reward regime achieves higher declared revenues while operating with a lower enforcement budget. Interestingly, we find that $p_L = 0$ after the introduction of the Reward regime, exactly matching the design of the policy that sets the audit probability to zero above the threshold. Although this restriction was not directly targeted in estimation, the result is fully consistent with the institutional rule embedded in the Reward scheme.

6.2 Policy counterfactuals

Budget- or revenue-neutral modifications of the disclosed rule Our natural experiment reveals that the introduction of the Reward regime led to higher declared revenues. In our structural estimates, we further show that this improvement was accompanied by a reduction in audit costs. Using the calibrated model, we disentangle these two channels by conducting counterfactual policy exercises under alternative constraints. Specifically, we consider (i) policies that keep the total audit budget constant while optimizing the rule for revenue collection, and (ii) policies that maintain revenue performance constant while

[41] We summarize these time-varying effects into a single $\delta\bar{d}_{\mathcal{A}}$ per dependent variable by taking the average across the post-reform coefficients.

[42] Based on data on SeS from D'Agosto et al. (2017), we estimate that the share of taxpayers audited within SeS is 10.51% below the threshold and 7.12% above it. See Appendix F for further details.

minimizing audit budget. The first exercise measures how much additional revenue can be collected with the same enforcement resources, while the second quantifies potential budget savings that would deliver the same revenue target.

We find that increasing the audit risk discontinuity at the eligibility threshold can substantially enhance revenue performance even without raising the total audit budget. The revenue curve in Figure 8 Panel A shows the concavity discussed in Theorem 4. On the horizontal axis we vary p_L relative to its estimated level in the pre-reform scenario and we set p_H in such a way that the pre-policy budget clears in equilibrium. By steepening the audit jump, revenue collection can be improved by up to more than 3 percentage points of $\hat{y}$ relative to the pre-Reward regime. At a level of $\Delta = 0.03$, we recover the improvement in declared revenues under the Reward policy relative to the undisclosed scenario. Since the revenue function is monotonic in the audit discontinuity, even the pre-Reward regime can be improved through a more pronounced increase in audit risk at the threshold.

Conversely, the model allows us to evaluate budget savings that deliver unchanged declared revenues. Figure 8 Panel B displays the minimum audit budget required to reproduce the pre-Reward regime's revenue levels as a function of Δ. By increasing the audit discontinuity above the pre-Reward regime Δ, the same level of declarations can be sustained with more than 2/3 lower budget. This is a relevant result for policy design, as Tax Authorities typically allocate a disproportionate share of enforcement resources toward larger taxpayers, where audit returns are higher. Our simulations suggest that resources could be further reallocated toward these larger taxpayers without compromising overall compliance, provided that audit risk schedule is redesigned appropriately.

In the following section, we turn to the question of optimal policy design, characterizing the audit strategy that maximizes compliance given resource constraints, and assessing how it departs from the Reward regime currently in place.

Distance from the best disclosed policy We conclude by investigating how far the observed policy and its modifications are from the best disclosed policy. To this end, we solve the unconstrained problem of the Tax Authority

$$V^{\star}(\lambda) := \max_{p_L, p_H} R(p_L, p_H) - \lambda Q(p_L, p_H).$$

We quantify $V^{\star}(\lambda)$ and compare it to the post-Reward Regime scenario, the pre-policy one, and a regime with an undisclosed flat rule. The parameter λ captures the marginal cost of public enforcement and determines whether the Authority prioritizes additional revenue collection or budget savings. For high λ, the Authority optimally accepts somewhat lower revenues in exchange for substantial reductions in audit expenditures. Importantly, λ should be interpreted broadly: it reflects not only the direct cost of auditing small firms but also the opportunity cost of reallocating enforcement away from large taxpayers, whose audits yield higher returns. When this opportunity cost is substantial, the Authority has strong incentives to rely more heavily on discontinuous audit rules to economize on audits of small firms while reallocating them toward larger ones.

The Tax Authority faces a clear tradeoff in choosing p_L and p_H: it can strengthen the reward for declaring above $\hat{y}$ or, alternatively, raise the penalty for declaring below it. Both policies increase the incentive to report above the threshold, but they have different implications for the Authority's objective: increasing p_H raises enforcement costs, whereas reducing p_L lowers costs but increases evasion among firms declaring above $\hat{y}$.

In the unconstrained optimum, when the marginal audit cost λ is sufficiently small, the Authority sets $p_L = p_H = 1/\gamma$, the audit probability that induces truthful reporting from all firms. For intermediate and empirically relevant values of λ, the Authority sets $p_L = 0$, fully eliminating audits for firms declaring above $\hat{y}$, and lets p_H decline with λ (Figure 9, Panel A). In our estimated economy, increasing the discontinuity Δ is always beneficial: marginal revenue gains from a larger Δ dominate potential losses from the $\mathcal{L}$ region, even when audits are relatively cheap and Δ is therefore large. Once this dominance holds at a given λ, it continues to hold for all higher λ', because Δ falls with λ while losses in $\mathcal{L}$ remain roughly constant. For this reason, p_H always lies above p_L for intermediate values of λ. Finally, for very large λ, the optimal policy is to audit no firm.

Comparing the unconstrained optimum with the observed post-Reward regime, we find that the Reward policy performs strikingly close to the optimum (Panel B). The Reward regime sets p_L at its lower bound and reduces the total audit budget, which is precisely the adjustment our model identifies as welfare-improving relative to the pre-Reward regime. The objective values of the two regimes coincide for λ near the level at which the Reward regime's p_H becomes optimal, diverging only for higher λ. This intersection can be interpreted as the marginal audit cost that rationalizes the Reward regime as globally optimal. Even outside that range, the Reward regime remains within 10% of the full optimum, whereas the pre-Reward regime lies well below it, particularly for high λ, where optimal policy would require a smaller audit budget. A flat audit rule ($\Delta = 0$) performs even worse, falling between 10 and 40% below the optimum for intermediate levels of λ and confirming that a discontinuity in audit probabilities is a crucial feature of a well-designed policy.

7 Conclusions

Tax audits and their threat are a primary enforcement tool across developed and developing countries. The dissuasive power of audits, however, has hardly solved the long-standing problem of low compliance among micro to small businesses and the self-employed. We ask whether the strategic disclosure of audit selection criteria can improve the effectiveness of enforcement among these taxpayers. We answer our question by developing a theoretical model of audit disclosure, and implementing a test derived by the model using a quasi-experiment in the context of Sector Studies (SeS), an Italian policy informing small firms of their relative audit risk around a revenue threshold.

We develop a theory of optimal tax declarations of firms that face a discontinuous threshold-based audit probability, and we derive a test for the existence of improvements over flat audit rules. The test is based on studying the behavior of the revenue function in response to a marginal change in the audit probability jump at the threshold. Consistently with the model, the distribution of SeS files reveals that taxpayers are especially aware of and willing to adjust to clear audit risk signals. The extent of bunching is strongly related to several evasion proxies on other declaration margins, and seems to respond to the incentives to evade and the availability of evasion technologies. To implement the test for improvements over flat rules, we exploit a 2011 staggered reform that strengthened the original risk discontinuity at the disclosed SeS threshold. While taxpayers respond by bunching at the cutoff regardless of their relative position ahead of the reform, mean gross profits rise by 16.2% in treated sectors over the course of six years, suggesting that the pre-reform discontinuity performed strictly better than a counterfactual flat rule that used more audits.

Our work is encouraging as international attention grows on the importance of voluntary tax compliance and reliable tax collection for fiscal sustainability (OECD, 2017; IMF, 2021). Differently from tax lotteries and traditional tax amnesties, the disclosure framework we study grants broadly accessible and stable incentives to stimulate compliance. As tax agencies routinely define thresholds to target their audits, they might develop cost-effective communication strategies to nudge taxpayers around these cutoffs. At the same time, we are aware that net collection effects also depend on the quality of the ensuing audits once the pool of exempted taxpayers is defined. Although such effects are bounded to be of second-order importance in contexts where limited audit resources are allocated, we leave the study of realized audit collection in the presence of threshold-based rules to future research.

References

Al-Karablieh, Y., E. Koumanakos, and S. Stantcheva. 2021. "Clearing the Bar: Improving Tax Compliance for Small Firms through Target Setting." *Journal of International Economics*, 130.

Alm, J., J. Martinez-Vazquez, and S. Wallace. 2004. *Taxing the Hard-to-tax: Lessons from Theory and Practice.* Vol. 268, Elsevier, Amsterdam.

Almunia, M. and D. Lopez-Rodriguez. 2018. "Under the Radar: The Effects of Monitoring Firms on Tax Compliance." *American Economic Journal: Economic Policy*, 10: 1–38.

Arachi, G. and A. Santoro. 2007. "Tax Enforcement for SMEs: Lessons from theItalian Experience?" *eJournal of Tax Research (Michigan Issue)*, 5(2): 225–243.

Australian Taxation Office. 2024. "Tax gap." Accessed August 4, 2025.

Bachas, P., R. Fattal Jaef, and A. Jensen. 2019. "Size-Dependent Tax Enforcement and Compliance: Global Evidence and Aggregate Implications." *Journal of Development Economics*, 140: 203–222.

Basri, M. C., M. Felix, R. Hanna, and B. Olken. 2019. "Tax Administration vs. Tax Rates: Evidence from Corporate Taxation in Indonesia." *NBER Working Paper No. 26150.*

Battaglini, Marco, Luigi Guiso, Chiara Lacava, Douglas L. Miller, and Eleonora Patacchini. 2023. "Refining Public Policies with Machine Learning: The Case of Tax Auditing." *Working paper.*

Battaglini, M., L. Guiso, C. Lacava, and E. Patacchini. 2020. "Tax Professionals and Tax Evasion."

Bérgolo, M., R. Ceni, G. Cruces, M. Giaccobasso, and R. Perez-Truglia. 2017. "Tax Audits as Scarecrows: Evidence from a Large-Scale Field Experiment." *NBER Working Paper No. 23631.*

Best, Michael Carlos, Anne Brockmeyer, Henrik Jacobsen Kleven, Johannes Spinnewijn, and Mazhar Waseem. 2015. "Production versus Revenue Efficiency with Limited Tax Capacity: Theory and Evidence from Pakistan." *Journal of Political Economy*, 123(6): 1311–55.

Boning, William C., Nathaniel Hendren, Ben Sprung-Keyser, and Ellen Stuart. 2025. "A Welfare Analysis of Tax Audits Across the Income Distribution." *Quarterly Journal of Economics*, 140(1): 63–112.

Brosio, G., A. Cassone, and R. Ricciuti. 2002. "Tax Evasion across Italy: Rational Noncompliance or Inadequate Civic Concern?" *Public Choice*, 112: 259–273.

Carfora, A., R.V. Pansini, and S. Pisani. 2016. "Spatial Dynamics in Tax Gap Determinants." *Agenzia delle Entrate, Argomenti di Discussione.*

Carrillo, P. E., D. Pomeranz, and M. Singhal. 2017. "Dodging the Taxman: Firm Misreporting and Limits to Tax Enforcement." *American Economic Journal: Applied Economics*, 9(2): 144–64.

Carrillo, P. E., E. Castro, and C. Scartascini. 2017. "Do Rewards Work? Evidence from the Randomization of Public Works." *IDB Working Paper IDB-WP-794.*

Casaburi, L. and U. Troiano. 2016. "Ghost-House Busters: The Electoral Response to a Large Anti Tax Evasion Program." *The Quarterly Journal of Economics*, 131(1): 273–314.

Censis. 2003. "Capire il Sommerso: Supporto Conoscitivo ai Servizi per l'Impiego."

Choudhary, K. and B. Gupta. 2019. "Third-Party Audit and Tax Compliance – Evidence from a Notched Policy in India." *working paper.*

D'Agosto, E., M. Manzo, A. Modica, and S. Pisani. 2017. "Tax Audits and Tax Compliance - Evidence from Italy." *available at https://www.irs.gov/pub/irs-soi/17resconmodica.pdf.*

D'Agosto, E., M. Marigliani, and S. Pisani. 2014. "Asymmetries in the Territorial VAT Gap." *Agenzia delle Entrate, Argomenti di Discussione.*

de Chaisemartin, C. and X. D'Haultfœuille. 2020. "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects." *American Economic Review*, 110(9): 2964–96.

Dunning, T., F. Monestier, R. Pineiro, F. Rosenblatt, and G. Tunon. 2017. "Is Paying Taxes Habit Forming? Theory and Evidence from Uruguay." *Working paper.*

Dwenger, Nadja, Henrik Kleven, Imran Rasul, and Johannes Rincke. 2016. "Extrinsic and Intrinsic Motivations for Tax Compliance: Evidence from a Field Experiment in Germany." *American Economic Journal: Economic Policy*, 8(3): 203–232.

Galbiati, R. and G. Zanella. 2012. "The Tax Evasion Social Multiplier: Evidence from Italy."

Journal of Public Economics, 96: 485–494.

HM Revenue and Customs. 2025. "HMRC tax receipts and National Insurance contributions for the UK: statistics table." Last published July 22, 2025.

HM Revenue & Customs. 2025. "Measuring Tax Gaps Tables 2025." Last updated June 19, 2025. Accessed July 17, 2025.

HRMC. 2021. "Measuring Tax Gaps 2021 Edition - Tax Gap Estimates for 2019 to 2020." *available at https://www.gov.uk/government/statistics/measuring-tax-gaps/measuring-tax-gaps-2021-edition-tax-gap-estimates-for-2019-to-2020.*

IMF. 2021. *World Economic Outlook: Managing Divergent Recoveries.* Washington, DC, April.

Internal Revenue Service. 2022. "Tax gap estimates for tax years 2014–2016 (and projections for 2017–2019)." Accessed July 17, 2025.

Kleven, H. and M. Waseem. 2013. "Using Notches to Uncover Optimization Frictions and Structural Elasticites: Theory and Evidence from Pakistan." *Quarterly Journal of Economics*, 128: 669–723.

Kleven, H., M. Knudsen, C. Kreiner, S. Pedersen, and E. Saez. 2011. "Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark." *Econometrica*, 79(3): 651–692.

Kumler, T., E. Verhoogen, and F. Frías. 2020. "Enlisting Employees in Improving Payroll Tax Compliance: Evidence from Mexico." *The Review of Economics and Statistics*, 102(5): 881–896.

Lazear, E. 2006. "Speeding, Terrorism, and Teaching to the Test." *The Quarterly Journal of Economics*, 121(3): 1029–1061.

MEF. 2023. "Relazione sull'Economia Non Osservata e sull'Evasione Fiscale e Contributiva."

Ministero dell'Economia e delle Finanze. 2022. "Relazione sull'economia non osservata e sull'evasione fiscale e contributiva: Aggiornamenti per gli anni 2015–2020 a seguito della revisione dei conti nazionali apportata dall'ISTAT." Accessed July 17, 2025.

Ministry of Economy and Finance. 2019. "Relazione sull'Economia Non Osservata e sull'Evasione Fiscale e Contributiva."

Naritomi, J. 2019. "Consumers as Tax Auditors." *American Economic Review*, 109: 3031–3072.

OECD. 2006. "Strengthening Tax Audit Capabilities: Innovative Approaches to Improve the Efficiency and Effectiveness of Indirect Income Measurement Methods."

OECD. 2017. "Italy's Tax Administration: A Review of Institutional and Governance Aspects."

OECD. 2025. "Global Revenue Statistics." Last updated July 8, 2025. Accessed August 4, 2025.

Paradisi, Matteo and Elia Sartori. 2023. "Using Prediction Models to Design Tax Enforcement: Incentives vs Targeting." *Working paper.*

Pisani, S. and C. Polito. 2006. "Metodologia di Integrazione tra i Dati IRAP e Quelli di Contabilità Nazionale." *Agenzia delle Entrate, Documenti di Lavoro dell'Ufficio Studi.*

Pomeranz, D. 2015. "No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax." *American Economic Review*, 105: 2539–2569.

Reinganum, J. and L. Wilde. 1985. "Income Tax Compliance in a Principal-Agent Framework." *Journal of Public Economics*, 26(1): 1–18.

Sánchez, I. and J. Sobel. 1993. "Hierarchical Design and Enforcement of Income Tax Policies." *Journal of Public Economics*, 50(3): 345–369.

Santoro, A. 2008. "Taxpayers' Choices under Studi di Settore: What Do We Know and How Can We Interpret It?" *Giornale Degli Economisti e Annali Di Economia, Nuova Serie, 67, 2: 161-84.*

Santoro, A. 2017. "Do Small Businesses Respond to a Change in Tax Audit Rules? Evidence from Italy." *Public Finance Review*, 45: 792–814.

Santoro, A. and C. Fiorio. 2011. "Taxpayer Behavior When Audit Rules Are Known: Evidence from Italy." *Public Finance Review, 39:103–123.*

Slemrod, J. 2019. "Tax Compliance and Enforcement." *Journal of Economic Literature*, 57(4): 904–954.

The World Bank. 2019. "Enterprise Surveys." *available at http://www.enterprisesurveys.org.*

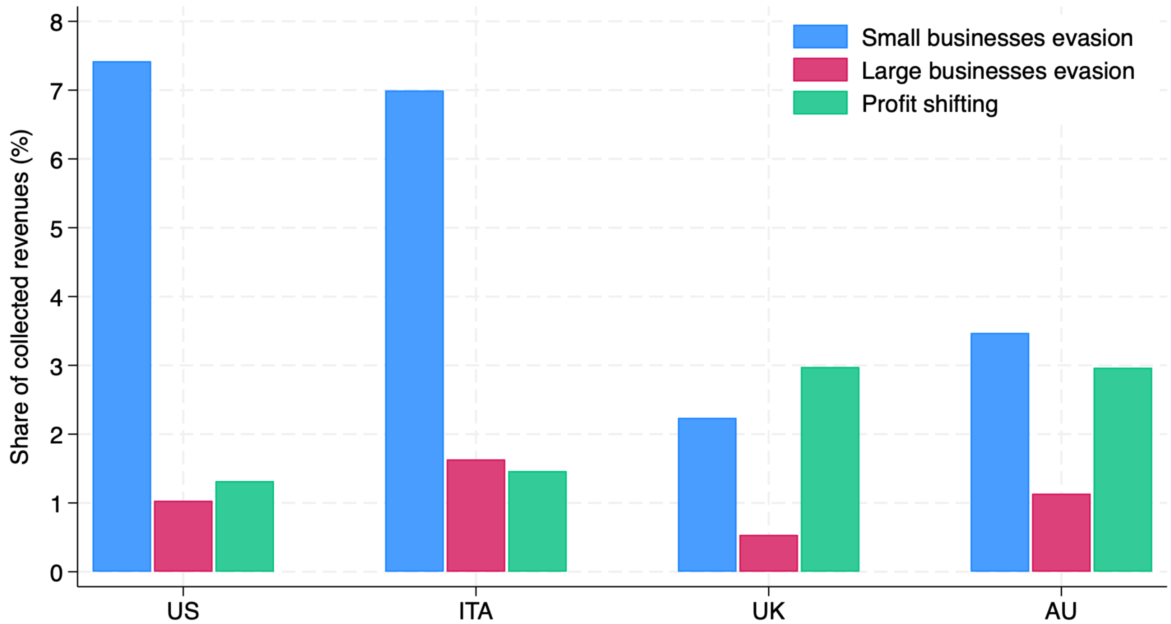
Thuronyi, V. 1996. *Presumptive Taxation.* in Tax Law Design and Drafting, International Monetary Fund.

Vallanti, G. and G. Gianfreda. 2020. "Informality, Regulation and Productivity: Do Small Firms Escape EPL through Shadow Employment?" *Small Business Economics.*

Wier, Ludvig and Gabriel Zucman. 2022. "Global Profit Shifting, 1975-2019." *NBER Working Paper No. 30673.*

Figures and Tables

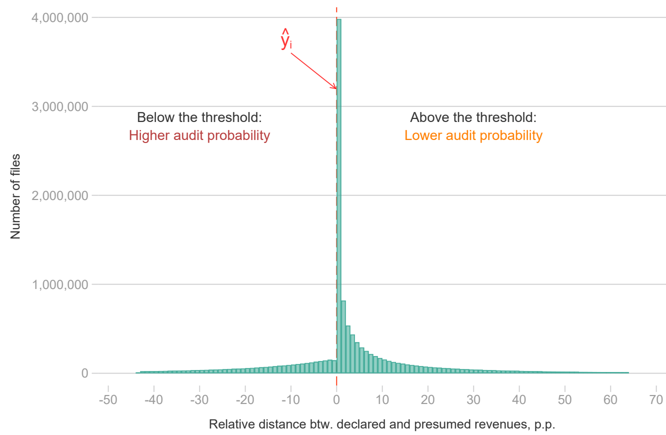
Figure 1: Tax gaps for small and large businesses across countries



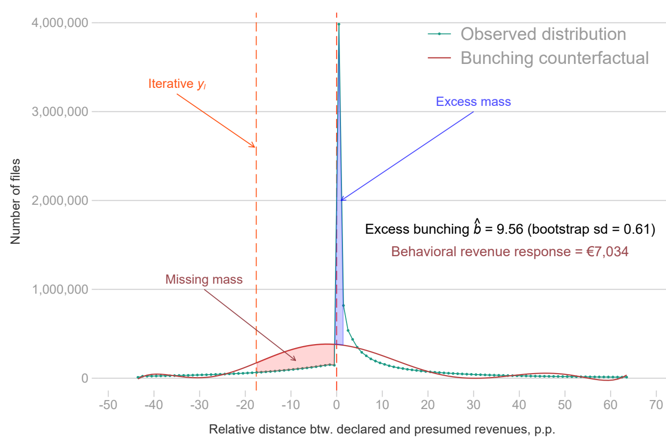
Notes: the Figure compares tax gaps between small and large businesses in four countries for 2019. Business income tax gaps are based on official reports from national tax authorities. Details on the definitions of small and large businesses, harmonization of the different sources, and methodology can be found in Supplementary Material S.4. Estimates of profit shifting by multinationals are from Wier and Zucman (2022). All quantities are expressed as a share of total collected central or federal government revenues.

Figure 2: Bunching in the universe of single-sector SeS filers, 2007-2010

Panel A: Reported revenues relative to SeS threshold

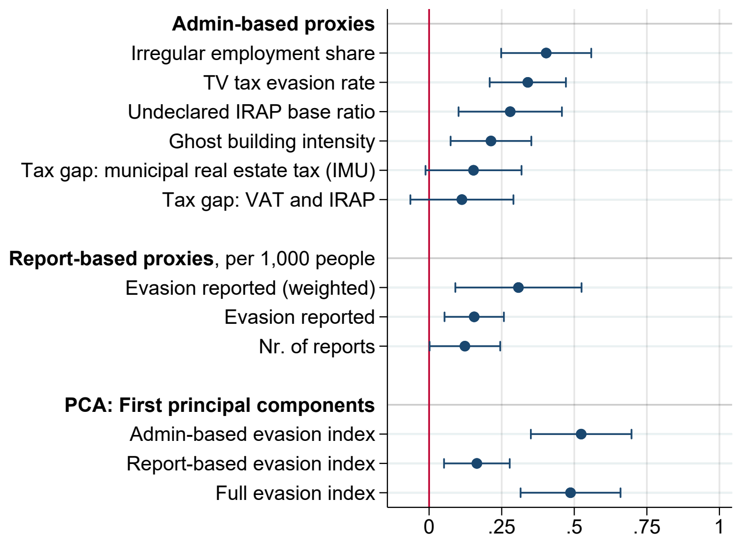


Panel B: Bunching estimates



Notes: the Figure presents the distribution of $(d_i - \hat{y}_i)/\hat{y}_i$, the relative distance between reported revenues d_i and presumed revenues $\hat{y}_i$, from each SeS file in the universe of single-sector businesses in the 2007-2010 tax years. Units on the horizontal axis are percentage points of each file's presumed revenues. We trim files reporting revenues below the 5$^{\text{th}}$ percentile or above the 95$^{\text{th}}$ percentile of relative distance from $\hat{y}$. This excludes taxpayers declaring zero revenues. Panel A displays the observed histogram of relative reported revenues. Panel B adds the smooth bunching counterfactual and presents the relevant estimates. The counterfactual density is estimated with an iterative procedure seeking to equate the excess mass above the threshold with the missing mass below it. The procedure stops with the definition of a lower bound y^l marked in Panel B with a dashed dark orange line. The smooth fit is obtained by estimating a regression with a 7$^{\text{th}}$-order polynomial in the bin order, and an upper bound set at the threshold bin (files with revenues falling within 1 percentage point above their presumed revenues). Excess bunching is the ratio of the excess mass and the height of the counterfactual at the threshold bin. Standard errors are computed with 1,000 bootstrap replications. The behavioral revenue response estimate comes from a corresponding bunching estimation where threshold distance is defined in Euro terms and bin width is equal to 500€.

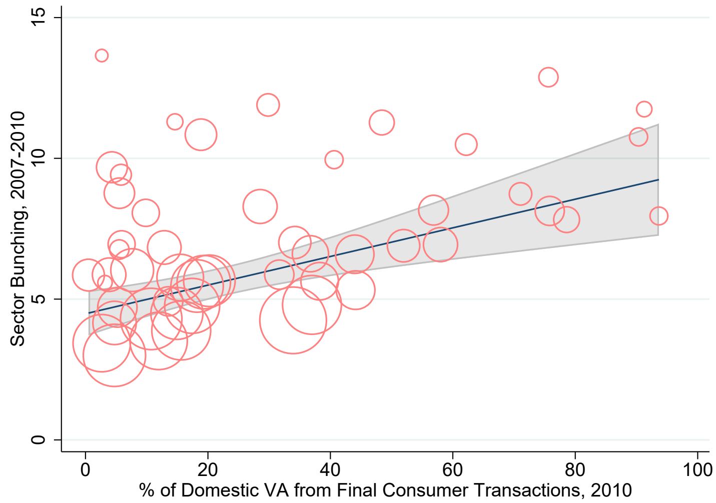
Figure 3: Provincial bunching correlates positively with local evasion



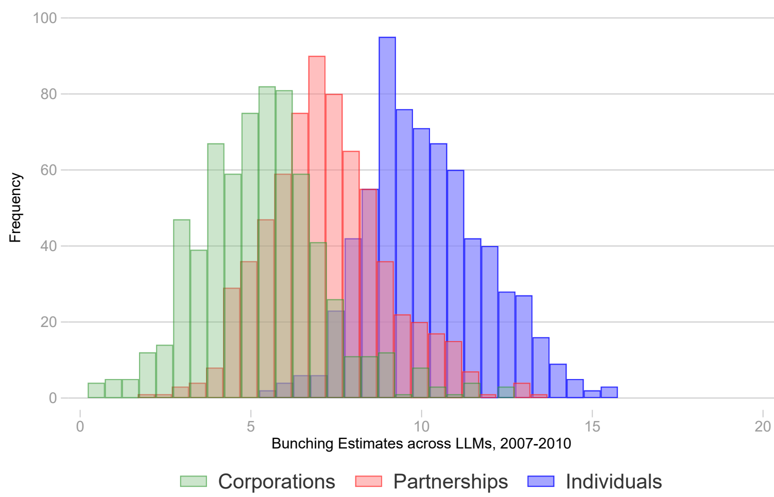
Notes: the Figure plots the standardized coefficients β and their 95% CIs from several regressions of SeS bunching on evasion proxies Evasion_i^j across 110 provinces i according to the model: $\text{Bunching}_i = \alpha + \beta \text{Evasion}_i^j + \gamma \log \text{VA pc}_i + \text{macroregion}_i + \varepsilon_i$. Standard errors are robust to heteroskedasticity. The sample includes each SeS file in the universe of single-sector businesses in the 2007-2010 tax years, except the top and bottom 5% in each province-level distribution that we trim to avoid irregularities in the estimation of the counterfactual. Bunching is computed at the province level following the procedure outlined in Section 4.2. Evasion proxies and their sources are described in Appendix S.6. The last three evasion proxies are the first principal components of the administrative-based, report-based, and all listed proxies, respectively. The first regression with our report-based proxies is weighted by the number of evasion reports from each province in 2008-2011.

Figure 4: Bunching tracks evasion potential: downstream sectors

Panel A: Correlation with share of sales to final consumers

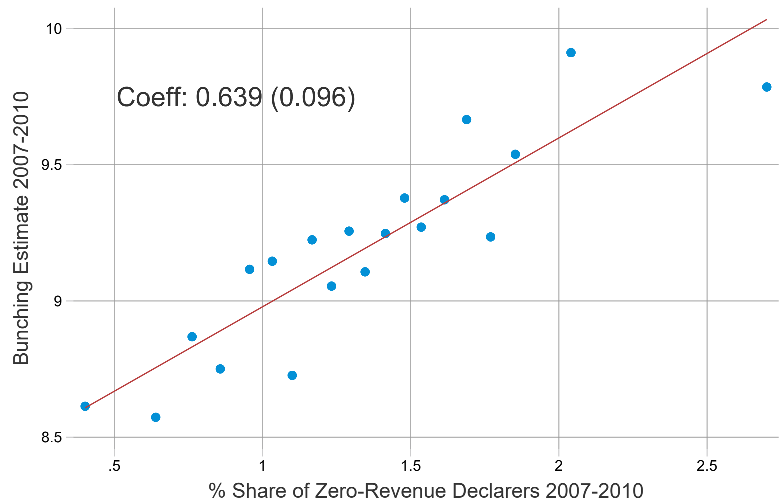


Panel B: Bunching across taxpayers with different legal complexity



Notes: this Figure' Panel A shows the sector-level scatterplot and linear fit of the relation between 2007-2010 bunching and the degree of relative exposure to the final consumer in 2010. Exposure is defined as a business sector's share of domestic value added (in 2010 current prices) that is determined by final consumption (see details in Appendix S.5). The sample consists of 51 1-digit and 2-digit ATECO sectors that we find both in the SeS database and *ISTAT*'s 2010-2013 input-output tables. Some sectors in this sample consist of one or more 2-digit sectors in the SeS data, in which case bunching is a weighted average of the 2-digit sector's bunching estimate, with weights equal to the sectors' number of 2007-2010 SeS files. We weight sectors by the mean presumed revenues associated to their 2007-2010 SeS files. The shaded area corresponds to a 95% confidence interval. The slope coefficient (robust standard error) from the corresponding weighted regression is 5.085 (1.106). Panel B plots the distribution of 2007-2010 bunching estimates computed at the LLM-level, separately for individuals (individual businesses and self-employed individuals), partnerships, and corporations. SeS taxpayers face increasing reporting and book-keeping requirements, with accounting complexity rising from a relatively low level in individually-owned activities to a progressively higher level among partnerships and corporations. We exclude 4% of estimates that are negative or in the 99$^{\text{th}}$ percentile of the distribution.

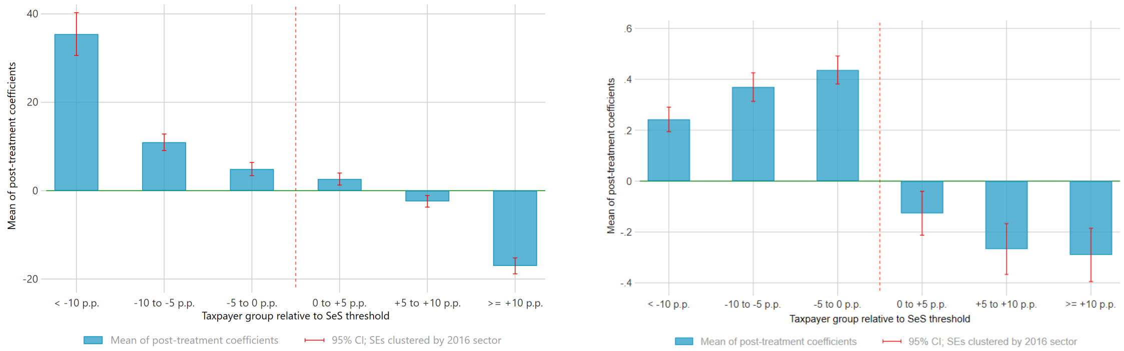
Figure 5: Provincial bunching correlates positively with share of zero declarers



Notes: the Figure correlates LLM-level shares of zero-revenue declarers with local SeS bunching. A binned scatterplot reports the slope coefficient and robust standard error from a regression of the form $\text{Bunching}_i = \alpha + \beta\text{Share zero declarers}_i^j + \gamma \log(\text{PIT base per taxpayer}_i) + \text{region}_i + \varepsilon_i$, including regional fixed effects and the logarithm of the average local PIT-base per individual taxpayer. The share of zero declarers is computed as the 2007-2010 local labor market share of SeS filers reporting exactly zero revenues. It ranges from 0 to 4.7%. The sample includes each SeS file in the universe of single-sector businesses in the 2007-2010 tax years, except the top and bottom 5% in each LLM-level distribution that we trim to avoid irregularities in the estimation of the counterfactual. Bunching is computed at the LLM level following the procedure outlines in Section 4.2.

Figure 6: Reward regime-induced declaration adjustments and compliance on the two sides of $\hat{y}$

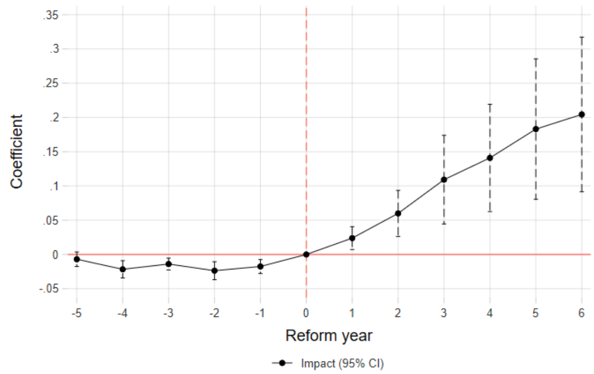
Panel A: Revenues as share of $\hat{y}$ Panel B: Compliance indicator



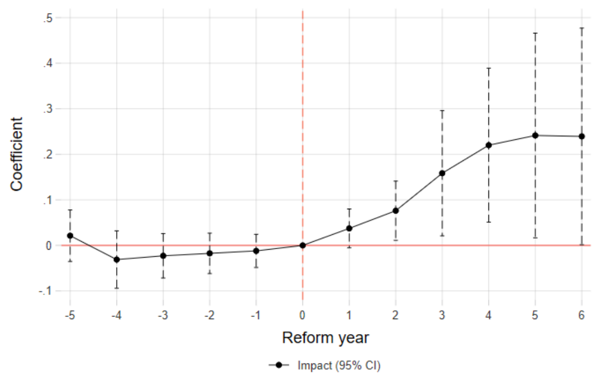
Notes: this Figure illustrates the effect of the reward regime on taxpayer declaration behavior above and below the presumed revenue threshold. Panels A and B show responses of taxpayers classified into six groups based on their distance from the threshold in the year prior to the regime's introduction. The dependent variables are: (i) revenues as a share of the firm-specific presumed revenues $\hat{y}$ (Panel A), and (ii) a compliance indicator reflecting the alignment of an individual declaration with benchmarks set by the Tax Authority (Panel B). In both panels, bars represent the average of six group-specific post-treatment coefficients from a staggered difference-in-differences based on the specification in (6). Whiskers represent 95% CIs of these linear combinations of coefficients. Standard errors are clustered at the sector level. The regressions are estimated on the sample of all SeS files from single-sector taxpayers continuously filing over the 2007-2016 period, aggregated by sector-year. Only sectors accessing the reward regime by 2016 are considered. Number of sector-years: 1550. Declared revenues are winsorized at the 99th percentile.

Figure 7: Reward regime effects on mean revenues and profits

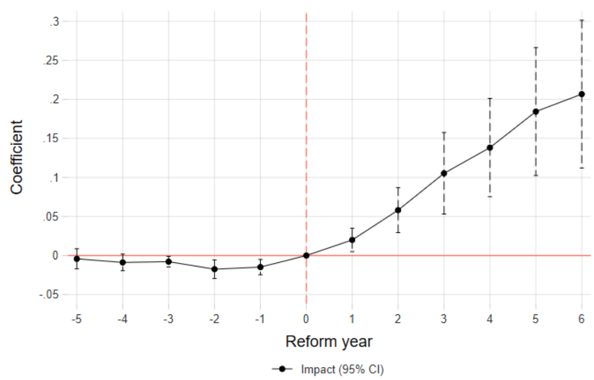
Panel A: Log-Mean reported revenues



Panel B: Log-Mean reported income



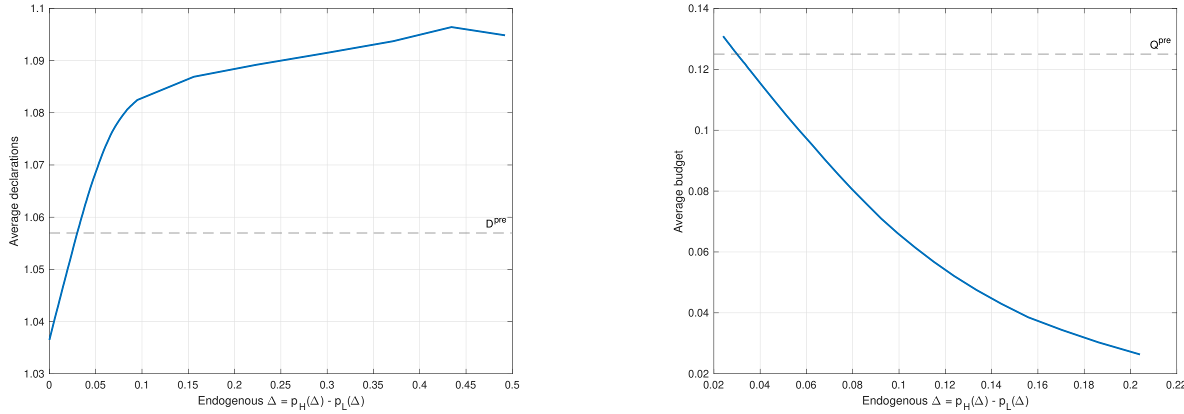
Panel C: Log-Mean reported log-costs



Notes: this Figure shows the effects of the reward regime's introduction in a sector on mean reported revenues (Panel A), mean gross profits (Panels B), and costs (Panel C) defined as the difference between reported revenues and gross profits. Dependent variables are expressed in logarithms. Whiskers represent 95% CIs. Effects are relative to the year before the advent of the reform in each sector, marked at year 0 by the red dashed vertical line. Estimates are based on our specification in (6). Standard errors are clustered at the sector level. The regressions are estimated on the sample of all Sector Study files from single-sector taxpayers continuously filing over the 2007-2016 period, aggregated by sector-year. Only sectors accessing the reward regime by 2016 are considered. Number of sector-years: 1550. Reported revenues are winsorized at the 99th percentile.

Figure 8: Policy counterfactuals: fixing budget and compliance

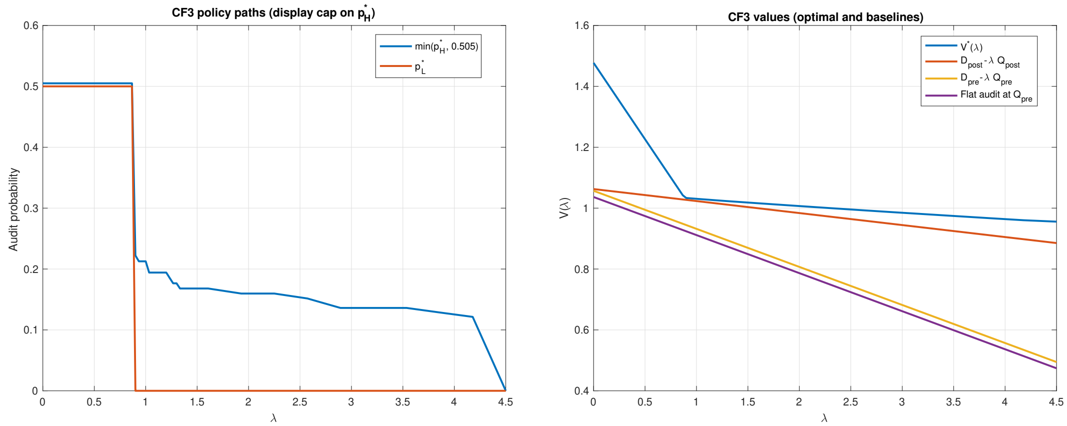
Panel A: Revenue improvements at fixed budget Panel B: Budget savings at fixed compliance



Notes: this Figure illustrates the revenue gains obtained with alternative policies while keeping the budget at the level estimated in the pre-policy scenario (Panel A), and the budget savings that alternative policies would achieve while keeping constant the total level of declared revenues in equilibrium (Panel B). The horizontal axis reports the endogenous Δ that arises when fixing a given p_L and searching for the p_H that clears the budget (Panel A) or keeps declared revenues constant (Panel B). The vertical axis reports average revenues as a share of $\hat{y}$ (Panel A) and share of audited taxpayers (Panel B). In Panel A, the horizontal dashed line represents the level of revenues achieved under the pre-policy scenario. The intersection between this line and the curve is the Δ estimated in the pre-policy scenario. In Panel B, the horizontal dashed line is the share of audited taxpayers under the pre-policy scenario.

Figure 9: Optimal policy: comparison with reward regime and pre-reform scenario

Panel A: Optimal p_H and p_L Panel B: Performance of different rules



Notes: this Figure illustrates the optimal policy for different levels of the marginal cost of running audits, and compares the returns from this policy to those achieved by the reward regime and the pre-reform scenario. Panel A shows the optimal level of the policy instruments (p_H and p_L) for different levels of the marginal cost of audit λ. Panel B shows the value of the government objective (declared revenues net of the cost of running audits) as a share of $\hat{y}$ for the different levels of λ. It compares four different rules: i) the optimal one, ii) the reward regime rule, iii) the pre-reward-regime rule, iv) the undisclosed rule using the same amount of budget as the pre-reward-regime one.

Table 1: SMM Estimates

Parameter/ Quantity	Estimate
p_H^{Pre}	0.148
p_L^{Pre}	0.118
p_H^{Post}	0.249
p_L^{Post}	0
Q^{Pre}	0.126
Q^{Post}	0.040
R^{Pre}	1.051
R^{Post}	1.068

Notes: this Table illustrates the estimates of the SMM procedure. We refer to the estimates for the pre-reward-regime scenario with the superscript Pre and those for the post-reform scenario with $Post$. Q represents the share of audited taxpayers, while R measures the average declared revenues as a share of $\hat{y}$.

Online Appendix

A Proofs of Propositions

Separability of Production and Declaration Decisions

We assume that y is exogenous, which is without loss of generality given that the declaration decision is separable from production. A model with endogenous production is equivalent to ours once y is interpreted as the optimal profit in a setting where firms differ in productivity. Formally, let $\pi(x, z)$ denote the profit of a firm choosing input x, with productivity z drawn from a distribution f_z. The firm solves:

$$V(z) = \max_{x,d} \pi(z, x) - \tau d - \tau\gamma \cdot p(d) \cdot (\pi(z, x) - d) - c(\pi(z, x) - d),$$

which delivers the following FOCs

$$x : \left[1 - \tau\gamma p - c'(\pi(z, x) - d)\right] \pi'(z, x) = 0$$

$$d : \tau(1 - \gamma p) - c'(\pi(z, x) - d) = 0.$$

Notice that the first factor in the FOC for x cannot be zero due to the FOC for d. It follows that the firm optimally chooses x such that $\pi'_x(z, x) = 0$. Let $x^*(z)$ be the solution to this FOC. The variable y used in our analysis corresponds to the resulting profit, $y = \pi(z, x^*(z)))$, and its distribution is given by $f(y) = f_z(z : \pi(z, x^*(z)) = y)$.

Proof of Proposition 1

A firm with income y solves

$$\max_d \left[\max\left\{u_H(y, d), u_L(y, d) \cdot \mathbb{I}(d > \hat{y})\right\}\right]$$

where

$$u_H(y, d) = y - \tau d - \tau\gamma p_H (y - d)^+ - \kappa c(y - d)$$
$$u_L(y, d) = y - \tau d - \tau\gamma p_L (y - d)^+ - \kappa c(y - d)$$

First notice that for all y, $u_i(y, d)$ is decreasing in d for $d \geq y$, so no firm over-reports. Maximizing $u_i(y, d)$ is the same as maximizing $\tilde{u}_i(e) = \tau(1 - \gamma p_i) e - \kappa c(e)$. By concavity of the objective, an interior maximum is characterized by the FOC

$$c'(e) = \frac{\tau(1 - \gamma p_i)}{\kappa}.$$

We denote e_i the solutions to these equations and use convexity of c to conclude that $e_L > e_H > 0$. We need to deal with corner cases. First, notice that $\tilde{u}_L(e)$ is valid only for declarations $d = y - e$ that lie above $\hat{y}$. Since, for all e, $\tilde{u}_H(e) < \tilde{u}_L(e)$, then whenever e_L is feasible (i.e., $y - e_L > \hat{y}$ and therefore $y > \hat{y} + e_L$) it solves the firm's problem.

For firms with $y < \hat{y}$ all candidate declarations (recall that over-reporting is suboptimal) are in the H region. As $\tilde{u}_H(e)$ is increasing below e_H, whenever e_H is not feasible (*i.e.* when $y - e_H < 0$, with 0 being the lower bound on feasible declarations), then $e = y$ (*i.e.* $d = 0$) is optimal. When instead $y > e_H$, then $d = y - e_H$ is optimal.

We are only left with solving the problem for firms with income $y \in [\hat{y}, \hat{y} + e_L]$. For y in that range, $\hat{y} = \arg\max_d u_L(y, d) \cdot \mathbb{I}(d \geq \hat{y})$ since the unconstrained maximum $y - e_L$ occurs at a point where the function already dropped to 0 and the objective is decreasing in the feasible domain. The maximum of $\tilde{u}_H(e)$ is instead e_H as this is feasible (recall we assumed $e_H < \hat{y}$). To find the global optimum we therefore need to compare the utility of evading $y - \hat{y}$ and facing p_L and of evading e_H and facing p_H. The former dominates iff

$$\tilde{u}_L(y - \hat{y}) > \tilde{u}_H(e_H) \equiv V_H$$

Notice that the RHS is flat in y, while the LHS is contiunous and increasing from $\tilde{u}_L(0) < V_H$ to $V_L \equiv \tilde{u}_L(e_L) > V_H$ (the latter inequality following by a simple envelope argument). Therefore there is a unique crossing, which is characterized by $V_H = \tilde{u}_L(\tilde{e})$, which yields to the condition

$$\frac{\tau}{\kappa}\left[e_H(1 - p_H\gamma) - \tilde{e}(1 - p_L\gamma)\right] = c(e_H) - c(\tilde{e}).$$

Summarizing, firms declare 0 if $y < e_H$, declare $y - e_H$ if $y \in [e_H, \hat{y} + \tilde{e}]$, bunch at $\hat{y}$ if $y \in [\hat{y} + \tilde{e}, \hat{y} + e_L]$, and declare $y - e_L$ above $\hat{y} + e_L$. This solution is valid if $e_H < \hat{y} + \tilde{e}$. Otherwise, the relevant deviation in the H region is to declare 0, there is no interior declarers in the H region and $\tilde{e}$ is defined by

$$(\hat{y} + \tilde{e}_0)\tau(1 - p_H\gamma) - \kappa c(\hat{y} + \tilde{e}_0) = \tilde{e}_0\tau(1 - p_L\gamma) - \kappa c(\tilde{e}_0).$$

Proof of Proposition 2

We derive first the comparative statics for the three evasion levels $e_H, e_L, \tilde{e}$. The FOC for interior evasions is $c'(e_i) = \tilde{\tau}(1 - \gamma p_i)$ for $i = H, L$, from which we have

$$\frac{\mathrm{d}e_i}{\mathrm{d}\tilde{\tau}} = \frac{1 - \gamma p_i}{c''(e_i)} = \frac{c'(e_i)}{\tilde{\tau}c''(e_i)} > 0.$$

Manipulating the equation that determines $\tilde{e}$ we obtain

$$\tilde{\tau}\left[e_H(1 - p_H\gamma) - \tilde{e}(1 - p_L\gamma)\right] = c(e_H) - c(\tilde{e}).$$

Using an envelope argument and rearranging we derive the following

$$\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\tilde{\tau}} = \frac{e_H(1 - p_H\gamma) - \tilde{e}(1 - p_L\gamma)}{\tilde{\tau}(1 - p_L\gamma) - c'(\tilde{e})} = \frac{c(e_H) - c(\tilde{e})}{\tilde{\tau}(c'(e_L) - c'(\tilde{e}))} > 0,$$

where the second equality uses the FOC for e_L and the definition of $\tilde{e}$, and the last inequality follows from the fact that $e_L > e_H > \tilde{e}$ and from the fact that $c(\cdot)$ and $c'(\cdot)$ are increasing.

Notice that $M(0) = F(e_H)$, $M(\mathcal{L}) = 1 - F(\hat{y} + e_L)$, $M(\mathcal{B}) = F(\hat{y} + e_L) - F(\hat{y} + \tilde{e})$. The comparative statics of e_H, e_L then imply that $M(0)$ is decreasing and $M(\mathcal{L})$ is increasing in $\tilde{\tau}$, strictly if the areas are non-degenerate which requires, respectively, that e_H and $\hat{y} + e_L$ are below $\bar{y}$. This proves the first two statements in the comparative statics part of the Proposition.

For $M(\mathcal{B})$, a quantitative assessment is needed since both the upper bound and the lower bound of the region decrease in $\tilde{\tau}$. The change in the size of the bunching region caused by a change in $\tilde{\tau}$ is

$$\frac{\mathrm{d}M(\mathcal{B})}{\mathrm{d}\tilde{\tau}} = f(\hat{y} + e_L)\frac{\mathrm{d}e_L}{\mathrm{d}\tilde{\tau}} - f(\hat{y} + \tilde{e})\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\tilde{\tau}}.$$

If $\hat{y} + e_L > \bar{y}$ (*i.e.* $M(\mathcal{L}) = 0$) then $f(\hat{y} + e_L) = 0$ and $\frac{\mathrm{d}M(\mathcal{B})}{\mathrm{d}\tilde{\tau}} < 0$ as only the lower-bound of the bunching area increases. Otherwise, the condition $\frac{\mathrm{d}M(\mathcal{B})}{\mathrm{d}\tilde{\tau}} > 0$ is equivalent to $\frac{f(\hat{y}+e_L)}{f(\hat{y}+\tilde{e})} > \frac{\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\tilde{\tau}}}{\frac{\mathrm{d}e_L}{\mathrm{d}\tilde{\tau}}}$. Since

$M(\mathcal{L})$ is monotonically decreasing in $\tilde{\tau}$ the condition $M(\mathcal{L}) > \bar{m}$ (stated in the Proposition) is equivalent to $\tilde{\tau} \leq \bar{\tau}$ for $\bar{\tau}$ such that $M(\mathcal{L})(\bar{\tau}) = \bar{m}$. Hence, we equivalently prove that the inequality is satisfied as $\tilde{\tau}$ vanishes, which is

$$\lim_{\tilde{\tau}\to 0} \frac{f(\hat{y}+e_L)}{f(\hat{y}+\tilde{e})} > \lim_{\tilde{\tau}\to 0} \frac{\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\tilde{\tau}}}{\frac{\mathrm{d}e_L}{\mathrm{d}\tilde{\tau}}}.$$

Using the FOC we get that $\lim_{\tilde{\tau}\to 0} c'(e_i) = 0$, implying that e_L, e_H (and a fortiori $\tilde{e}$ which is bounded above by both) converge to zero. This implies that $\lim_{\tilde{\tau}\to 0} \frac{f(\hat{y}+e_L)}{f(\hat{y}+\tilde{e})} = 1$. Regarding the right hand side, we show that its limit converges to a number below 1 by contradiction. We know that for all $\tilde{\tau} > 0$, $0 < \tilde{e}(\tilde{\tau}) < e_L(\tilde{\tau})$. Given its definition, $\frac{\mathrm{d}e_L}{\mathrm{d}\tilde{\tau}}$ is bounded as long as $c''(0) > 0$. Since $e_L > \tilde{e}$, by the fundamental theorem of calculus, $\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\tilde{\tau}}$ is also bounded. Because both $\frac{\mathrm{d}e_L}{\mathrm{d}\tilde{\tau}}$ and $\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\tilde{\tau}}$ are bounded and well-behaved, $\frac{\frac{\mathrm{d}}{\mathrm{d}\tilde{\tau}}\tilde{e}}{\frac{\mathrm{d}}{\mathrm{d}\tilde{\tau}}e_L}$ has a limit a. Suppose $a > 1$, then there would exist $\underline{\tau} > 0$ such that $\forall \tilde{\tau} \in [0, \underline{\tau}]$, $\frac{\frac{\mathrm{d}}{\mathrm{d}\tilde{\tau}}\tilde{e}}{\frac{\mathrm{d}}{\mathrm{d}\tilde{\tau}}e_L} > 1$, which in turns implies that

$$\tilde{e}(\underline{\tau}) = \tilde{e}(0) + \int_0^{\underline{\tau}} \frac{\mathrm{d}}{\mathrm{d}\tilde{\tau}}\tilde{e}(\tilde{\tau})\,\mathrm{d}\tilde{\tau} > e_L(0) + \int_0^{\underline{\tau}} \frac{\mathrm{d}}{\mathrm{d}\tilde{\tau}}e_L(\tilde{\tau})\,\mathrm{d}\tilde{\tau} = e_L(\underline{\tau}).$$

This is however a contradiction. Hence, it must be that $a < 1$ and that $\lim_{\tilde{\tau}\to 0} \frac{\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\tilde{\tau}}}{\frac{\mathrm{d}e_L}{\mathrm{d}\tilde{\tau}}} < 1$.

Proof of Proposition 3

Using the thesholds on income defined in Proposition 1, we define total revenues as follows

$$R(\Delta) = \int_{y^{0H}}^{y^{HB}} d_H(y)\, f(y)\,\mathrm{d}y + \int_{y^{HB}}^{y^{BL}} \hat{y} f(y)\,\mathrm{d}y + \int_{y^{BL}}^{\bar{y}} d_L(y)\, f(y)\,\mathrm{d}y$$

so marginal revenues are

$$\begin{aligned}
\frac{\mathrm{d}R(\Delta)}{\mathrm{d}\Delta} &= -\frac{\mathrm{d}y^{0H}}{\mathrm{d}\Delta}\left[d_H\left(y^{0H}\right)\right] f\left(y^{0H}\right) + \frac{\mathrm{d}y^{HB}}{\mathrm{d}\Delta}\left[d_H\left(y^{HB}\right) - \hat{y}\right] f\left(y^{HB}\right) \\
&\quad + \frac{\mathrm{d}y^{BL}}{\mathrm{d}\Delta}\left[\hat{y} - d_L\left(y^{BL}\right)\right] f\left(y^{BL}\right) + \int_{y^{0H}}^{y^{HB}} \frac{\mathrm{d}}{\mathrm{d}\Delta} d_H(y)\, f(y)\,\mathrm{d}y + \int_{y^{BL}}^{\bar{y}} \frac{\mathrm{d}}{\mathrm{d}\Delta} d_L(y)\, f(y)\,\mathrm{d}y \\
&= \overbrace{\underbrace{\frac{\mathrm{d}y^{HB}}{\mathrm{d}\Delta}}_{<0}\left[\underbrace{d_H\left(y^{HB}\right) - \hat{y}}_{<0}\right]}^{>0} f\left(y^{HB}\right) + M(\mathcal{H})\underbrace{\frac{\mathrm{d}}{\mathrm{d}\Delta}d_H(y)}_{>0} + M(\mathcal{L})\underbrace{\frac{\mathrm{d}}{\mathrm{d}\Delta}d_L(y)}_{<0}
\end{aligned}$$

where the last equality exploits the fact that $d_H\left(y^{0H}\right) = 0$ and $d_L\left(y^{BL}\right) = \hat{y} + e_L$ by definition, that for a given cost function $\frac{\mathrm{d}}{\mathrm{d}\Delta}d_H(y)$ and $\frac{\mathrm{d}}{\mathrm{d}\Delta}d_L(y)$ are constant across ys, and defines $M(\mathcal{H}) = F\left(y^{HB}\right) - F\left(y^{0H}\right)$ and $M(\mathcal{L}) = 1 - F\left(y^{BL}\right)$. To obtain the expression in the statement, notice further that

$$d_H\left(y^{HB}\right) - \hat{y} = y^{HB} - e_H - \hat{y} = \hat{y} + \tilde{e} - e_H - \hat{y} = \tilde{e} - e_H,$$

that $\frac{\mathrm{d}d_i(y)}{\mathrm{d}\Delta} = -\frac{\mathrm{d}e_i}{\mathrm{d}\Delta}$ for $i = H, L$, and $\frac{\mathrm{d}y^{HB}}{\mathrm{d}\Delta} = \frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}$.

Proof of Theorem 4

First, we derive the following two Lemmas that provide useful results to write the marginal revenue expressions.

Lemma 5. *It holds*

$$\frac{dy^{HB}}{d\Delta} = \frac{d\tilde{e}}{d\Delta} = -\frac{\tau\gamma\left[\tilde{e}\left(\Delta\right)+\alpha e_H\left(\Delta\right)\right]}{c'\left(e_L\left(\Delta\right)\right)-c'\left(\tilde{e}\left(\Delta\right)\right)}, \tag{A.1}$$

$$\frac{d^2\tilde{e}\left(\Delta\right)}{d\Delta^2} = \frac{\frac{(\tau\gamma\alpha)^2}{c''(e_H)}-\frac{d\tilde{e}}{d\Delta}\left[2\tau\gamma-\frac{d\tilde{e}}{d\Delta}c''\left(\tilde{e}\right)\right]}{c'\left(e_L\right)-c'\left(\tilde{e}\right)}, \text{ and } \lim_{\Delta\to 0}\frac{dy^{HB}}{d\Delta} = -\infty.$$

Proof. From the indifference condition we have $V^H\left(y^{HB}, d^H\left(y^{HB}\right)\right) = V^L\left(y^{HB}, \hat{y}\right)$ that solves

$$V^H\left(y\right) \equiv \max_d \left(y-d\right)\tau\left(1-\gamma p_H\left(\Delta\right)\right)-c\left(y-d\right) = \left(y-\hat{y}\right)\tau\left(1-\gamma p_L\left(\Delta\right)\right)-c\left(y-\hat{y}\right) \equiv V^B\left(y\right).$$

So, we want $\frac{\mathrm{d}y^{HB}}{\mathrm{d}\Delta}$ where y^{HB} solves $V^H\left(y\right)-V^B\left(y\right)=0$ (for detailed derivations see Supplementary Material S.1). This is

$$\frac{\mathrm{d}}{\mathrm{d}\Delta}y^{HB} = -\frac{\tau\gamma\left[\tilde{e}\left(\Delta\right)+\alpha e_H\left(\Delta\right)\right]}{c'\left(e_L\left(\Delta\right)\right)-c'\left(\tilde{e}\left(\Delta\right)\right)}.$$

Because both the numerator and denominator are positive, we obtain $\frac{\mathrm{d}}{\mathrm{d}\Delta}y^{HB} < 0$. The second derivative is

$$\frac{\mathrm{d}^2\tilde{e}\left(\Delta\right)}{\mathrm{d}\Delta^2} = \frac{\frac{(\tau\gamma\alpha)^2}{c''(e_H)}-\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}\left[2\tau\gamma-\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}c''\left(\tilde{e}\right)\right]}{c'\left(e_L\right)-c'\left(\tilde{e}\right)}.$$

In addition, $\lim_{\Delta\to 0}\frac{\mathrm{d}}{\mathrm{d}\Delta}y^{HB} = -\infty$ because i) the numerator converges to a finite number since $\lim_{\Delta\to 0}\tau\gamma\left[\tilde{e}\left(\Delta\right)+\alpha e_H\left(\Delta\right)\right] = \tau\gamma\left(1+\alpha\right)e^I$, ii) the denominator converges to zero since when $\lim_{\Delta\to 0}p_H - p_L = 0$ we get $d^{HB}\to\hat{y}$ as we prove below, and costs (and marginal costs) are continuous functions, which implies that $\lim_{\Delta\to 0}c'\left(e_H\right)-c'\left(\tilde{e}\left(\Delta\right)\right) = 0$.

We are left to prove that $\lim_{\Delta\to 0}d^H\left(y^{HB}\right)-\hat{y} = 0$. To do that, recall that $V^H\left(y^{HB}, d^H\left(y^{HB}\right)\right) = V^L\left(y^{HB}, \hat{y}\right)$ since $V^H\to V^L$ as a function when $\Delta\to 0$ (because $p_L\to p_H\to\mu$). Then, the equality can be satisfied only if $d^H\left(y^{HB}\right) = \hat{y}$. $\square$

Lemma 6. *The limit of the bunching component of marginal revenues for $\Delta\to 0$ is finite and reads*

$$\lim_{\Delta\to 0}\frac{dy^{HB}}{d\Delta}\left[d^H\left(y^{HB}\right)-\hat{y}\right]f\left(y^{HB}\right) = \frac{\tau\gamma\left(1+\alpha\right)e^I\left(\mu\right)}{c''\left(e^I\left(\mu\right)\right)}f\left(\hat{y}+e^I\left(\mu\right)\right).$$

Proof. Using (A.1), we have

$$\frac{\mathrm{d}y^{HB}}{\mathrm{d}\Delta}\left[\tilde{e}\left(\Delta\right)-e_H\left(\Delta\right)\right] = \frac{\tau\gamma\left[\tilde{e}\left(\Delta\right)+\alpha e_H\left(\Delta\right)\right]}{c'\left(e_L\left(\Delta\right)\right)-c'\left(\tilde{e}\left(\Delta\right)\right)}\left(e_H\left(\Delta\right)-\tilde{e}\left(\Delta\right)\right)$$

as $\Delta\to 0$ both the numerator and the denominator go to 0. Since $\tau\gamma\left(\tilde{e}+\alpha e_H\right)$ converges to $\tau\gamma\left(1+\alpha\right)e^I\left(\mu\right)$ for $\Delta\to 0$, we use L'Hopital to conclude the following (see full derivation in

Supplementary Material S.1)

$$\lim_{\Delta\to 0}\frac{\mathrm{d}y^{HB}}{\mathrm{d}\Delta}\left[\tilde{e}\left(\Delta\right)-e_H\left(\Delta\right)\right]=\tau\gamma e^I\left(\mu\right)\left(1+\alpha\right)\lim_{\Delta\to 0}\frac{\frac{\mathrm{d}}{\mathrm{d}\Delta}e_H-\frac{\mathrm{d}}{\mathrm{d}\Delta}\tilde{e}}{\frac{\mathrm{d}}{\mathrm{d}\Delta}\left\{c'\left(e_L\right)-c'\left(\tilde{e}\right)\right\}}=\frac{\tau\gamma e^I\left(\mu\right)\left(1+\alpha\right)}{c''\left(e^I\right)},$$

where we used that, local to 0, $\frac{\mathrm{d}}{\mathrm{d}\Delta}\tilde{e}$ diverges (Lemma 5) while

$$\frac{\mathrm{d}e_H}{\mathrm{d}\Delta}=-\frac{\gamma\tau\frac{\mathrm{d}p_H}{\mathrm{d}\Delta}}{c''\left(e^I\left(p_H\right)\right)}=-\frac{\gamma\tau\alpha}{c''\left(e_H\right)}\to-\frac{\gamma\tau\alpha}{c''\left(e^I\left(\mu\right)\right)}$$

remains bounded. □

Having proved the two Lemmas, the rest of the proof is conceptually simple (we directly differentiate the marginal revenue function), but involves many algebraic steps. We express

$$MR\left(\Delta\right)=B\left(\Delta\right)+L\left(\Delta\right)+H\left(\Delta\right)$$

where

$$B\left(\Delta\right)=-\frac{\mathrm{d}\tilde{e}\left(\Delta\right)}{\mathrm{d}\Delta}\cdot\left(e_H\left(\Delta\right)-\tilde{e}\left(\Delta\right)\right)\cdot f\left(\hat{y}+\tilde{e}\left(\Delta\right)\right),\ L\left(\Delta\right)=-\frac{\mathrm{d}e_L\left(\Delta\right)}{\mathrm{d}\Delta}\cdot\left(1-F\left(\hat{y}+e_L\left(\Delta\right)\right)\right),$$

$$H\left(\Delta\right)=-\frac{\mathrm{d}e_H\left(\Delta\right)}{\mathrm{d}\Delta}\cdot\left(F\left(\hat{y}+\tilde{e}\left(\Delta\right)\right)-F\left(e_H\left(\Delta\right)\right)\right).$$

By direct differentiation, we obtain

$$\begin{aligned}-\frac{\mathrm{d}^2R}{\mathrm{d}\Delta^2}=&\frac{\mathrm{d}^2\tilde{e}\left(\Delta\right)}{\mathrm{d}\Delta^2}\cdot\left(e_H\left(\Delta\right)-\tilde{e}\left(\Delta\right)\right)\cdot f\left(\hat{y}+\tilde{e}\left(\Delta\right)\right)+\frac{\mathrm{d}\tilde{e}\left(\Delta\right)}{\mathrm{d}\Delta}\frac{\mathrm{d}e_H\left(\Delta\right)}{\mathrm{d}\Delta}\cdot f\left(\hat{y}+\tilde{e}\left(\Delta\right)\right)-\left(\frac{\mathrm{d}\tilde{e}\left(\Delta\right)}{\mathrm{d}\Delta}\right)^2f\left(\hat{y}+\tilde{e}\left(\Delta\right)\right)\\&+\left(\frac{\mathrm{d}\tilde{e}\left(\Delta\right)}{\mathrm{d}\Delta}\right)^2\left(e_H\left(\Delta\right)-\tilde{e}\left(\Delta\right)\right)\cdot f'\left(\hat{y}+\tilde{e}\left(\Delta\right)\right)+\frac{\mathrm{d}^2e_L\left(\Delta\right)}{\mathrm{d}\Delta^2}\cdot\left(1-F\left(\hat{y}+e_L\left(\Delta\right)\right)\right)-\left(\frac{\mathrm{d}e_L\left(\Delta\right)}{\mathrm{d}\Delta}\right)^2f\left(\hat{y}+e_L\left(\Delta\right)\right)\\&+\frac{\mathrm{d}^2e_H\left(\Delta\right)}{\mathrm{d}\Delta^2}\cdot\left(F\left(\hat{y}+\tilde{e}\left(\Delta\right)\right)-F\left(e_H\left(\Delta\right)\right)\right)+\frac{\mathrm{d}e_H\left(\Delta\right)}{\mathrm{d}\Delta}\frac{\mathrm{d}\tilde{e}\left(\Delta\right)}{\mathrm{d}\Delta}f\left(\hat{y}+\tilde{e}\left(\Delta\right)\right)-\left(\frac{\mathrm{d}e_H\left(\Delta\right)}{\mathrm{d}\Delta}\right)^2f\left(e_H\left(\Delta\right)\right).\end{aligned}\tag{A.2}$$

Using that $\frac{\mathrm{d}^2e_i(\Delta)}{\mathrm{d}\Delta^2}$, $\frac{\mathrm{d}e_i(\Delta)}{\mathrm{d}\Delta}$ are bounded (by c''', c'', respectively) in the limit as $\Delta\to 0$, the second derivative (A.2) evaluates to

$$\begin{aligned}-\lim_{\Delta\to 0}\frac{\mathrm{d}^2R}{\mathrm{d}\Delta^2}=&\frac{\mathrm{d}^2\tilde{e}\left(\Delta\right)}{\mathrm{d}\Delta^2}\cdot\left(e_H\left(\Delta\right)-\tilde{e}\left(\Delta\right)\right)\cdot f\left(\hat{y}+\tilde{e}\left(\Delta\right)\right)-\left(\frac{\mathrm{d}\tilde{e}\left(\Delta\right)}{\mathrm{d}\Delta}\right)^2f\left(\hat{y}+\tilde{e}\left(\Delta\right)\right)\\&+\left(\frac{\mathrm{d}\tilde{e}\left(\Delta\right)}{\mathrm{d}\Delta}\right)^2\left(e_H\left(\Delta\right)-\tilde{e}\left(\Delta\right)\right)\cdot f'\left(\hat{y}+\tilde{e}\left(\Delta\right)\right)+2\frac{\mathrm{d}e_H\left(\Delta\right)}{\mathrm{d}\Delta}\frac{\mathrm{d}\tilde{e}\left(\Delta\right)}{\mathrm{d}\Delta}f\left(\hat{y}+\tilde{e}\left(\Delta\right)\right).\end{aligned}$$

Now, using the fact that we can write $\frac{\mathrm{d}^2\tilde{e}(\Delta)}{\mathrm{d}\Delta^2}$ as a polynomial function of $\frac{\mathrm{d}\tilde{e}(\Delta)}{\mathrm{d}\Delta}$, we rewrite this expression as a polynomial in $\frac{\mathrm{d}\tilde{e}(\Delta)}{\mathrm{d}\Delta}$ (full derivation in Supplementary Material S.1)

$$-\lim_{\Delta\to 0}\frac{\mathrm{d}^2R}{\mathrm{d}\Delta^2}=\lim_{\Delta\to 0}\left(\left(\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}\right)^2\alpha_2+\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}\alpha_1+\alpha_0\right),\tag{A.3}$$

where

$$\alpha_0 = \frac{(\tau\gamma\alpha)^2}{c''(e_H)} \frac{(e_H - \tilde{e}) \cdot f(\hat{y} + \tilde{e})}{c'(e_L) - c'(\tilde{e})}, \ \alpha_1 = 2\tau\gamma f(\hat{y} + \tilde{e}) \left(-\frac{(e_H - \tilde{e})}{c'(e_L) - c'(\tilde{e})} - \frac{\alpha}{c''(e_H)} \right),$$

$$\alpha_2 = f(\hat{y} + \tilde{e}) \left((e_H - \tilde{e}) \frac{f'(\hat{y} + \tilde{e})}{f(\hat{y} + \tilde{e})} + c''(\tilde{e}) \frac{(e_H - \tilde{e})}{c'(e_L) - c'(\tilde{e})} - 1 \right).$$

We derive the following Lemma.

Lemma 7. *R is concave if and only if $\lim_{\Delta \to 0} \alpha_1 + \lim_{\Delta \to 0} \frac{d\tilde{e}}{d\Delta} \alpha_2 < 0$.*

Proof. Using the fact that $\lim_{\Delta \to 0} \frac{(e_H - \tilde{e})}{c'(e_L) - c'(\tilde{e})} = \frac{1}{c''(e^I)}$, $\lim_{\Delta \to 0} e_H, e_L, \tilde{e} = e^I(\mu)$, and the definitions of the α coefficients in (A.3), we have

$$\lim_{\Delta \to 0} \alpha_0 = \frac{(\tau\gamma\alpha)^2}{c''(e^I)^2} f\left(\hat{y} + e^I\right), \ \lim_{\Delta \to 0} \alpha_1 = -f\left(\hat{y} + e^I\right) \frac{2\tau\gamma(1 + \alpha)}{c''(e^I)}, \ \lim_{\Delta \to 0} \alpha_2 = 0.$$

Recall that $\lim_{\Delta \to 0} \frac{d\tilde{e}}{d\Delta} = -\infty$. Since the limits of both α_0 and α_1 are finite this means that α_0 is irrelevant in the limit. Because it is multiplied by a term that converges to 0, we need however to study the term in $\left(\frac{d\tilde{e}}{d\Delta}\right)^2$. Specifically, we study $\lim_{\Delta \to 0} \frac{d\tilde{e}}{d\Delta} \alpha_2$ to understand if the term in $\left(\frac{d\tilde{e}}{d\Delta}\right)^2$ dominates over that in α_1. Indeed, if $\frac{d\tilde{e}}{d\Delta} \alpha_2$ is finite, then $-\lim_{\Delta \to 0} \frac{d^2 R}{d\Delta^2} = \lim_{\Delta \to 0} \frac{d\tilde{e}}{d\Delta} \left(\alpha_1 + \lim_{\Delta \to 0} \frac{d\tilde{e}}{d\Delta} \alpha_2\right)$ and the revenue function is (infinitely) concave/convex depending on the sign of $\lim_{\Delta \to 0} \alpha_1 + \lim_{\Delta \to 0} \frac{d\tilde{e}}{d\Delta} \alpha_2$. Since $\frac{d\tilde{e}}{d\Delta}$ is (infinitely) negative, R is concave if and only if $\lim_{\Delta \to 0} \alpha_1 + \lim_{\Delta \to 0} \frac{d\tilde{e}}{d\Delta} \alpha_2 < 0$.

□

We hence prove that the condition in Lemma 7 holds: $\lim_{\Delta \to 0} \alpha_1 + \lim_{\Delta \to 0} \frac{d\tilde{e}}{d\Delta} \alpha_2 < 0$. Since we already derived $\lim_{\Delta \to 0} \alpha_1$, we must find an expression for $\lim_{\Delta \to 0} \frac{d\tilde{e}}{d\Delta} \alpha_2$. We summarize here the results, but full derivations are in Supplementary Material S.1. We have

$$\begin{aligned} \lim_{\Delta \to 0} \frac{d\tilde{e}}{d\Delta} \alpha_2 &= \lim_{\Delta \to 0} -\frac{\tau\gamma[\tilde{e} + \alpha e_H]}{c'(e_L) - c'(\tilde{e})} f(\hat{y} + \tilde{e}) \left[(e_H - \tilde{e}) \frac{f'(\hat{y} + \tilde{e})}{f(\hat{y} + \tilde{e})} + c''(\tilde{e}) \frac{(e_H - \tilde{e})}{c'(e_L) - c'(\tilde{e})} - 1 \right] \\ &= -\tau\gamma e^I (1 + \alpha) f\left(\hat{y} + e^I\right) \left[\lim_{\Delta \to 0} \frac{(e_H - \tilde{e})}{c'(e_L) - c'(\tilde{e})} \frac{f'(\hat{y} + \tilde{e})}{f(\hat{y} + \tilde{e})} + \lim_{\Delta \to 0} \frac{c''(\tilde{e})(e_H - \tilde{e}) - c'(e_L) + c'(\tilde{e})}{[c'(e_L) - c'(\tilde{e})]^2} \right]. \end{aligned}$$

We investigate the second term in the bracket using L'Hopital

$$\lim_{\Delta \to 0} \frac{c''(\tilde{e})(e_H - \tilde{e}) - c'(e_L) + c'(\tilde{e})}{[c'(e_L) - c'(\tilde{e})]^2} = -\frac{\frac{c'''(e^I)}{c''(e^I)} e^I + 1}{2e^I c''(e^I)}.$$

So, we have

$$\begin{aligned} \lim_{\Delta \to 0} \frac{d\tilde{e}}{d\Delta} \alpha_2 &= -\tau\gamma e^I (1 + \alpha) f\left(\hat{y} + e^I\right) \left[\frac{1}{c''(e^I)} \frac{f'\left(\hat{y} + e^I\right)}{f\left(\hat{y} + e^I\right)} - \frac{\frac{c'''(e^I)}{c''(e^I)} e^I + 1}{2e^I c''(e^I)} \right] \\ &= -\tau\gamma(1 + \alpha) f\left(\hat{y} + e^I\right) \left[\frac{2e^I f'\left(\hat{y} + e^I\right) - \frac{c'''(e^I)}{c''(e^I)} e^I f\left(\hat{y} + e^I\right) - f\left(\hat{y} + e^I\right)}{2c''(e^I) f\left(\hat{y} + e^I\right)} \right], \end{aligned}$$

which is finite by boundedness away from zero of c'' and f. Hence, according to Lemma 7, concavity of the revenue is determined by the sign of

$$\lim_{\Delta\to 0}\left[\alpha_1+\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}\alpha_2\right]=-\frac{\tau\gamma\left(1+\alpha\right)f\left(\hat{y}+e^I\right)}{c''\left(e^I\right)}\left[2+\frac{2e^If'\left(\hat{y}+e^I\right)-\left(\frac{c'''\left(e^I\right)}{c''\left(e^I\right)}e^I+1\right)f\left(\hat{y}+e^I\right)}{2f\left(\hat{y}+e^I\right)}\right]$$

$$\propto -\frac{3}{2}-e^I\frac{f'\left(\hat{y}+e^I\right)}{f\left(\hat{y}+e^I\right)}+\frac{e^I}{2}\frac{c'''\left(e^I\right)}{c''\left(e^I\right)}<0.$$

where the final inequality follows from our assumptions on the income distribution and the shape of the cost function. Lemma 7 identifies a sufficient condition to have infinite concavity at (and, by continuity, in a neighborhood of) 0. If μ is small, this means the function is concave in the relevant domain $[0,\mu]$ for Δ. As marginal revenues are decreasing, observing a positive increment in the reward system means the pre-reward system performs better than the flat rule μ. This concludes the proof to the first statement in the Theorem.

We now turn to the proof of the second part of the Theorem, which states that in the same region where revenue is concave, the enforcement budget decreases with Δ. This implies that if revenues increase, they do so while simultaneously reducing the budget required for enforcement.

Define the budget function Q as the mapping from an audit rule to the share of audited taxpayers, given by

$$\begin{aligned}Q&=\left(\mu+\alpha\Delta\right)F\left(y^{HB}\right)+\left(\mu-\Delta\right)\left[1-F\left(y^{HB}\right)\right]\\&=\left(\mu+\alpha\Delta\right)F\left(\hat{y}+\tilde{e}\left(\Delta\right)\right)+\left(\mu-\Delta\right)\left[1-F\left(\hat{y}+\tilde{e}\left(\Delta\right)\right)\right].\end{aligned}\tag{A.4}$$

We now complete the proof of the second part of the Theorem by identifying the condition under which a local perturbation of the flat audit rule saves budget.

Lemma 8. *If* $\alpha<\frac{1}{F(\hat{y}+e(\mu))}-1$ *then* $\lim_{\Delta\to 0}\frac{dQ}{d\Delta}<0$.

Proof. Differentiating (A.4),

$$\begin{aligned}\frac{\mathrm{d}Q}{\mathrm{d}\Delta}&=\alpha F\left(\hat{y}+\tilde{e}\left(\Delta\right)\right)+\left(\mu+\alpha\Delta\right)\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}f\left(\hat{y}+\tilde{e}\left(\Delta\right)\right)-\left[1-F\left(\hat{y}+\tilde{e}\left(\Delta\right)\right)\right]-\left(\mu-\Delta\right)\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}f\left(\hat{y}+\tilde{e}\left(\Delta\right)\right)\\&=\alpha F\left(\hat{y}+\tilde{e}\left(\Delta\right)\right)-\left[1-F\left(\hat{y}+\tilde{e}\left(\Delta\right)\right)\right]+\left(\alpha+1\right)\Delta\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}f\left(\hat{y}+\tilde{e}\left(\Delta\right)\right)\end{aligned}$$

so

$$\lim_{\Delta\to 0}\frac{\mathrm{d}Q}{\mathrm{d}\Delta}=\alpha F\left(\hat{y}+e\left(\mu\right)\right)-\left[1-F\left(\hat{y}+e\left(\mu\right)\right)\right]+\left(\alpha+1\right)f\left(\hat{y}+e\left(\mu\right)\right)\lim_{\Delta\to 0}\Delta\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}.\tag{A.5}$$

Finally, using L'Hopital

$$\lim_{\Delta\to 0}\Delta\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}=\lim_{\Delta\to 0}-\Delta\frac{\tau\gamma\left[\tilde{e}+\alpha e_H\right]}{c'\left(e_L\right)-c'\left(\tilde{e}\right)}\lim_{\Delta\to 0}-\frac{\Delta\tau\gamma\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}}{-\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}c''\left(\tilde{e}\right)}=\lim_{\Delta\to 0}\frac{\Delta\tau\gamma}{c''\left(\tilde{e}\right)}=0,$$

which, plugged back in (A.5), yields the statement of the Lemma. ∎

Weaker sufficient conditions for concavity: The proof of Lemma 7 relies on the baseline assumptions on the income distribution and the cost function. However, they are needed only insofar as they ensure that $-\frac{3}{2}-e^I\frac{f'(\hat{y}+e^I)}{f(\hat{y}+e^I)}+\frac{e^I}{2}\frac{c'''(e^I)}{c''(e^I)}<0$, which is the necessary condition. Our assumptions can be relaxed while still ensuring that the necessary condition is satisfied. Specifically, the condition that $f(\cdot)$ does not vary too rapidly relative to its level can be relaxed as follows. First, it is actually sufficient that the density does not *decrease* too rapidly as $f'(\cdot) > 0$ would reinforce the inequality. Second, even the assumption that the density does not decrease too rapidly does not need to hold uniformly as per our assumption, but only at the point $\hat{y} + e^I$. For example, a unimodal distribution, even though it decreases very rapidly in some portions, would still satisfy the condition provided that $\hat{y} + e^I$ is relatively small, *i.e.* in the increasing portion of the density. Finally, one could also relax the assumption on $c'''(\cdot) \le 0$ as long as $x\frac{c'''(x)}{c''(x)}$ is uniformly bounded from above.

Extension with heterogeneous costs: In this extension, marginal revenues are obtained by integrating (5) over the costs, and the aggregate revenue $R(\Delta)$ is concave if

$$\mathbb{E}_\kappa \left[-\frac{3}{2} - e^I_\kappa \frac{\tilde{f}'_{y|\kappa}\left(\hat{y} + e^I_\kappa\right)}{\tilde{f}_{y|\kappa}\left(\hat{y} + e^I_\kappa\right)} + \frac{e^I_\kappa}{2} \frac{c'''_\kappa\left(e^I_\kappa\right)}{c''_\kappa\left(e^I_\kappa\right)} \right] > 0.$$

The concavity condition is generally met if the argument of the expectation holds across all cost types. When income is independent of cost type, it suffices for this condition to be satisfied by the cost type choosing the highest evasion.

B Model Extensions

B.1 Imperfect correspondence between the audit target and the tax base

Consider a model where firms have two margins: a targeted margin y (*e.g.*, revenues) and a margin to manipulate the tax base x (*e.g.*, costs) that is not targeted by the audit rule. The problem of the firm reads

$$\max_{d_y, d_x} y - x - \tau\left(d_y - d_x\right) - \tau\gamma p\left(\hat{y}, y\right)\left(y - x - \left(d_y - d_x\right)\right) - c_y\left(y - d_y\right) - c_x\left(d_x - x\right).$$

Notice that this problem for flat a rule decouples and $V(p) = V_y(p) + V_x(p)$ where

$$V_i(p) = \max_{e_i} \tau\left[1 - \tau\gamma p\right] e_i - c_i\left(e_i\right).$$

Under a flat rule at p, the FOCs read

$$\tau\left(1 - \gamma p\right) = c'_y\left(e_y\right) \text{ and } \tau\left(1 - \gamma p\right) = c'_x\left(e_x\right),$$

where e_y and e_x are the optimal levels of evasion on the two margins. Notice that the implied evasions in the different margins might lead the firm to declare negative profits in case $y - x < e_y\left(p_H\right) + e_x\left(p_H\right)$. In some instances, negative profits generate a tax credit at the same tax rate τ and the solution to the firm's problem is straightforward so that there is no zero-declarers area. If firms are constrained to declare weakly positive profits they will declare $d_y = d_x = d$ at a level that minimizes total evasions costs, namely that solves

$$\min_d c_y\left(y - d_y\right) + c_x\left(x - d_x\right).$$

The optimal declaration equalizes marginal costs $c'_y\left(y - d^\star\right) = c'_x\left(d^\star - x\right)$ and we call $d^0(y, x)$ the solution to this problem. All the firms equating revenues and cost declarations are zero profit declarers and they have revenues below $y^{0H} = x + e_y\left(p_H\right) + e_x\left(p_H\right)$. The optimal declaration is therefore

$$(d_y, d_x) = \begin{cases} (d^o(y,x), d^o(y,x)) & \text{if } y < y^{0H} \\ (y - e_y(p_H), e_x(p_H) + x) & \text{if } y^{0H} < y < \hat{y} + \tilde{e}_y \\ (\hat{y}, e_x(p_L) + x) & \text{if } \hat{y} + \tilde{e}_y < y < \hat{y} + e_y(p_L) \\ (y - e_y(p_L), e_x(p_L) + x) & \text{if } y > \hat{y} + e_y(p_L) \end{cases}.$$

The indifferent firm between declaring in $\mathcal{H}$ and bunching at $\hat{y}$ solves the following

$$\begin{aligned} -\tau(y - e_y(p_H) - (e_x(p_H) + x)) - \tau\gamma p_H(e_y(p_H) + e_x(p_H)) - c_y(e_y(p_H)) - c_x(e_x(p_H)) = \\ -\tau(y - \tilde{e}_y - (e_x(p_L) + x)) - \tau\gamma p_L(y - \hat{y} + e_x(p_L)) - c_y(y - \hat{y}) - c_x(e_x(p_L)), \end{aligned}$$

which we can rewrite as

$$\tau(1 - \gamma p_L)(\tilde{e}_y) - c_y(\tilde{e}_y) = V_y(p_H) + \underbrace{V_x(p_H) - V_x(p_L)}_{\leq 0}.$$

Notice that the function $\tau(1 - \gamma p_L)(e) - x_y(e)$ is increasing in $[0, e_L]$ (it is a concave function below its max), therefore the solution to the equation is increasing in the RHS and $V_x(p_H) - V_x(p_L) \leq 0$ implies that $\tilde{e}_y$ is below $\tilde{e}$ in the baseline version of the model (when cost manipulation is not available). Hence, firms have greater incentives to bunch compared to our baseline specification. Despite the greater incentives to bunch, we will show that the behavior of $\tilde{e}_y$ is qualitatively similar to before in the limit $\Delta \to 0$.

When considering the effect of Δ on the revenues of the Tax Authority, we first notice that given a true cost x, the firm's behavior in the $\mathcal{H}$ and $\mathcal{L}$ areas is unchanged and therefore $\left\{\frac{\mathrm{d}e_i(p_H)}{\mathrm{d}\Delta}, \frac{\mathrm{d}e_i(p_L)}{\mathrm{d}\Delta}\right\}_{i=y,x}$ are finite for $\Delta \to 0$. We thus need to focus on marginal bunchers, and evaluate if their gains remain bounded in the limit. The gain of a marginal buncher is $\tilde{e}_y + e_x(p_L) - (e_y(p_H) + e_x(p_H))$. We study the limit behavior of

$$\begin{aligned} \left(\frac{\mathrm{d}\tilde{e}_R}{\mathrm{d}\Delta}\right) \cdot (\tilde{e}_R + e_c(p_L) - (e_r(p_H) + e_c(p_H))) = \\ -\tau\gamma(\alpha e_{r,H} + \alpha e_{c,H} + e_{c,L} + \tilde{e}) \frac{(\tilde{e}_R + e_c(p_L) - (e_r(p_H) + e_c(p_H)))}{x_r'(e_L) - x_r'(\tilde{e}_R)}. \end{aligned}$$

We can apply l'Hopital rule to claim $\lim_{\Delta \to 0}\left(\frac{\mathrm{d}\tilde{e}_R}{\mathrm{d}\Delta}\right) \cdot (\tilde{e}_R + e_c(p_L) - (e_r(p_H) + e_c(p_H)))$ is

$$-\tau\gamma\left((1+\alpha)e_r^I + (1+\alpha)e_c^I\right) \lim_{\Delta \to 0} \frac{\left(\frac{\mathrm{d}}{\mathrm{d}\Delta}\tilde{e}_R + \frac{\mathrm{d}}{\mathrm{d}\Delta}e_c(p_L) - \frac{\mathrm{d}}{\mathrm{d}\Delta}(e_r(p_H) + e_c(p_H))\right)}{\frac{\mathrm{d}}{\mathrm{d}\Delta}e_L x_r''(e_L) - \frac{\mathrm{d}}{\mathrm{d}\Delta}\tilde{e}_R x_r''(\tilde{e}_R)}$$

using that all $\frac{\mathrm{d}e_i(p)}{\mathrm{d}\Delta}$, $i = y, x$ are finite. The limit simplifies to

$$\lim_{\Delta \to 0}\left(\frac{\mathrm{d}\tilde{e}_R}{\mathrm{d}\Delta}\right) \cdot (\tilde{e}_R + e_c(p_L) - (e_r(p_H) + e_c(p_H))) = \tau\gamma(1+\alpha)\left(e_r^I + e_c^I\right)\frac{1}{x_r''(e_r^I)},$$

which is finite.

With this insight we can write the government revenue $x-$wise using total evasions, defining $e_i \coloneqq (e_y(p_i) + e_x(p_i))$, $i \in \{H, L\}$ and $\tilde{e} \coloneqq \tilde{e}_y + e_x(p_L)$. Notice that the properties of the $e_i, \tilde{e}$ are the same as before: $e_L > e_H > \tilde{e}$ and e_i has finite derivative, while $\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}$ is infinite but stabilized by $\tilde{e} - e_H$ (as proved above). Hence, the expression for the $c-$wise marginal revenue are identical to those in the proof of Theorem 4, with the only difference that the term $B(\Delta)$ in the marginal Tax Authority revenue is proportional to $\frac{\mathrm{d}}{\mathrm{d}\Delta}\tilde{e}_y \cdot (\tilde{e} - e_H)$ so there is a slight asymmetry in that only changes in evasions in y matter to determine the change in the mass of bunchers (because bunchers declare x in the interior). However, this difference is immaterial in the limit as $\Delta \to 0$ because the term in $\frac{\mathrm{d}}{\mathrm{d}\Delta}e_L$ vanishes. The arguments above hold for any x, and we can use the joint distribution of (y, x) to aggregate across xs and obtain the sufficient condition of Theorem 4 in expectation.

B.2 Threshold manipulation

Consider the following modification of our baseline model, in which firms draw a threshold $\hat{y}_0$ when reporting their characteristics truthfully, but can manipulate the threshold by paying a increasing and convex cost $\tilde{c}(\cdot)$ where $\tilde{c}(0) = 0, \tilde{c}'(\cdot) \geq 0, \tilde{c}'(0) = 0, \tilde{c}''(\cdot) > 0$, and $\tilde{c}''(\cdot)$ is uniformly bounded away from zero. We assume $\tilde{c}'(0) = 0$ so that we consider a conservative case that departs more from our baseline model and where the manipulation of the threshold is more attractive.[43] The non-manipulated threshold $\hat{y}_0$ is common across all firms in a given class or sector and, similarly to $\hat{y}$ in our baseline model, one can interpret it as the unconditional average in the sector. By paying $\tilde{c}(\hat{y}_0 - \hat{y})$, firms can face an audit probability that jumps at $\hat{y}$ instead of $\hat{y}_0$. Our baseline model assumed $\tilde{c}'(\cdot) = \infty$. Firms now choose jointly the misreported threshold $\hat{y}$ and declaration d to solve

$$\max_{d,\hat{y}} y - \tau d - \tau\gamma p(d, \hat{y})(y - d)^+ - c(y - d) - \tilde{c}(\hat{y}_0 - \hat{y})$$

where $p(d, \hat{y})$ is the audit probability (2) that jumps from p_H to p_L at $\hat{y}$. We make a few prelimiary remarks that characterize the optimal behavior of firms. All proofs of the following Lemmata are in Supplementary Material S.2.

Lemma 9. *At the optimum:*

1. *A firm that manipulates the original threshold $\hat{y}_0$ always declares at the manipulated threshold $\hat{y}$: if $\hat{y} \neq \hat{y}_0$, then $d = \hat{y}$.*
2. *Firms with high enough income do not manipulate the threshold: if $y > \hat{y}_0 + e_L$ then $d = y - e_L$ and $\hat{y} = \hat{y}_0$.*
3. *Firms that optimally decide to face probability p_H will not manipulate the threshold: if $p(d, \hat{y}) = p_H$ for optimal choices, then $d \in \{0, y - e_H\}$ and $\hat{y} = \hat{y}_0$.*

This Lemma has important implications as it allows to collapse the optimal declaration and the threshold manipulation into a single dimensional problem. Specifically, using $p(\hat{y}, \hat{y}) = p_L$ we simplify the firm's problem to

$$V^B(y) = \max_{\hat{y}} y - \tau\hat{y} - \tau\gamma p_L (y - \hat{y})^+ - c(y - \hat{y}) - \tilde{c}(\hat{y}_0 - \hat{y}). \tag{B.1}$$

Since cost functions are convex, the problem is globally concave and its solution is characterized by the first order condition

$$\tau(1 - \gamma \cdot p_L) = c'(y - \hat{y}) + \tilde{c}'(\hat{y}_0 - \hat{y}). \tag{B.2}$$

[43] Since most of our results rely on characterizing the behavior of revenues local to a flat ($\Delta = 0$) rule, we need to make threshold manipulation attractive even in that case. Thus, we require $\tilde{c}'(0) = 0$, which implies that it is almost free to do the first bit of manipulation. If this is not the case, as it might be for many relevant situations, then our analysis goes through unchanged as threshold manipulation would not be used in the $\Delta \to 0$ limit.

Denote $\hat{y}(y)$ the solution to this problem. Since firms have two margins of manipulation, it is sufficient that both marginal costs at zero manipulation equal zero to ensure that both margins are used at the optimum. Assuming $\tilde{c}'(0) > 0$ would imply that threshold manipulation occurs only if its marginal benefit is large enough, a condition that fails in the limit $\Delta \to 0$, which is key to our analysis of revenue concavity. Thus, the case where $\tilde{c}'(0) = 0$ represents the extension in which the baseline model's results are not directly applicable. Nonetheless, we demonstrate that the concavity result still holds even in this case.

The following Lemma derives some properties of $\hat{y}(y)$. To draw comparative statics that vary the cost of manipulation, in what follows we parametrize the total manipulation cost as $\breve{\kappa} \cdot \tilde{c}(\cdot)$ and change the parameter $\breve{\kappa} \in [0, +\infty]$ so that $\breve{\kappa} = 0$ is the case in which manipulation is free, while $\breve{\kappa} = +\infty$ is the case of no manipulation (*i.e.* our baseline model).

Lemma 10. *The optimal threshold $\hat{y}(y)$ is increasing in y and such that $\hat{y}(\hat{y}_0 + e_L) = \hat{y}_0$. Therefore, $\hat{y}(y) < \hat{y}_0$ for $y < \hat{y}_0 + e_L$. For all y, $\lim_{\breve{\kappa}\to\infty} \hat{y}(y) = \hat{y}_0$ and $\lim_{\breve{\kappa}\to 0} \hat{y}(y) = y - e_L$. Finally, $V^B(y)$ is increasing in y and decreasing in $\breve{\kappa}$.*

The results of Lemma 10 can be summarized as follows. If manipulation was prohibitively costly, then firms would not manipulate the threshold and bunching would only occur at the natural threshold $\hat{y}_0$. If manipulation was free, then all firms would place the threshold at the level that allows them to pursue the optimal evasion in the p_L regime. In intermediate cases where manipulation has a positive and bounded marginal cost, we have that *i*) conditional on bunching, the threshold is lower than $\hat{y}_0$ (the more so the cheaper it is to manipulate), and *ii*) the highest y that bunches is independent of the manipulation cost and equal to $\hat{y}_0 + e_L$ and its optimal threshold is $\hat{y}(\hat{y}_0 + e_L) = \hat{y}_0$. We finally need to address how the *lowest* y that bunches behaves. Towards this end, first notice that by Lemma 9 firms that declare in the $\mathcal{H}$ audit region evade e_H and do not manipulate the threshold. Therefore, their value is V^H, the same as in our baseline model. It reads

$$V^H = \max_d y - \tau d - \tau\gamma p_H (y-d) - c(y-d) = y - \tau(y - e_H) - \tau\gamma p_H e_H - c(e_H).$$

Denote with $\tilde{y}^{HB}$ the lowest y buncher. Its indifference condition is therefore[44]

$$V^H\left(\tilde{y}^{HB}\right) = V^B\left(\tilde{y}^{HB}\right).$$

Let $\breve{e} = \tilde{y}^{HB} - \hat{y}\left(\tilde{y}^{HB}\right)$ be the evasion of this buncher. It satisfies

$$e_H\tau(1-\gamma p_H) - c(e_H) = \tau(1-\gamma p_L)\breve{e} - c(\breve{e}) - \breve{\kappa}\tilde{c}\left(\hat{y}_0 - \tilde{y}^{HB} + \breve{e}\right).$$

Absent $\tilde{c}$, or when $\breve{e} = \tilde{y}^{HB} - \hat{y}_0$ we obtain the same equation for $\tilde{e}$ as in the baseline version of the model. However, now we have the term in $\tilde{c}\left(\hat{y}_0 - \tilde{y}^{HB} + \breve{e}\right)$, which arises due to the cost of moving the threshold to the level that allows to evade $\breve{e}$. Compared to the baseline model, there is a notable difference: bunchers declare (and bunch) at a lower level, but the share of bunchers is higher. Hence, it is not ex-ante obvious that threshold manipulation is revenue detrimental for a Tax Authority. We now derive some comparative statics on the size of the bunching region.

Lemma 11. *The bunching region contracts with $\breve{\kappa}$ and expands with Δ. In particular, $\frac{d}{d\Delta}\tilde{y}^{HB} = -\frac{\tau\gamma(\breve{e}+\alpha e_H)}{\breve{\kappa}\tilde{c}'(\hat{y}_0 - \hat{y}(\tilde{y}^{HB}))}$.*

[44] Notice that this indifference condition has at most one solution as the difference $V^H(y) - V^B(y)$ is monotonically decreasing in y because, using the expressions that we will derive in the proof to Lemma 11, $\frac{\mathrm{d}V^H(y)}{\mathrm{d}y} - \frac{\mathrm{d}V^B(y)}{\mathrm{d}y} = -\breve{\kappa}\tilde{c}'(\hat{y}_0 - \hat{y}(y)) < 0$.

In the limit economies with no manipulation and free manipulation, bunching behaves as follows.

Lemma 12. *When $\tilde{\kappa} = 0$ then $V^H(0) < V^B(0)$ and so all firms bunch at their manipulated threshold, while if $\tilde{\kappa} = +\infty$ then $\tilde{y}^{HB} = \hat{y}_0 + \tilde{e}$ where $\tilde{e}$ is the evasion of the marginal buncher in our baseline model. As $\Delta = 0$ instead, the possibility of threshold manipulation becomes immaterial for any $\tilde{\kappa}$. As a consequence, $\hat{y}\left(\tilde{y}^{HB}\right) = \hat{y}_0$ when $\Delta = 0$, and $\lim_{\Delta\to 0} \frac{d}{d\Delta}\tilde{y}^{HB} = -\infty$.*

As it is cheaper to manipulate the threshold, more firms will do so. In the polar case where costs are negligible then all firms will manipulate and a "fake threshold" economy arises where all firms declare the threshold that allows them to optimally evade under the p_L schedule. If instead threshold manipulation is infinitely costly, we are back to our baseline model where the threshold is exogenous. When $\Delta = 0$ firms do not manipulate since they face a flat rule and manipulation would imply paying a cost for no advantage.

Concavity test under threshold manipulation

We now study how the concavity test that we derived in the baseline model holds in the extension that allows for threshold manipulation. We start from the polar case where manipulation is free ($\tilde{\kappa} = 0$) and then move to the case of a positive and finite marginal cost of manipulation.

Adding a share of free manipulators only attenuates the effect: Consider an economy where a share of firms ν has no manipulation cost. These firms manipulate their threshold to $\hat{y}(y) = y - e_L$ and declare $d = \hat{y}(y)$ (Lemma 10). Hence, the Tax Authority's revenues on these firms are $\tilde{R} = \mathbb{E}(y) - e_L$ and $\frac{\mathrm{d}\tilde{R}^2}{\mathrm{d}^2\Delta} = -\frac{\mathrm{d}^2 e_L}{\mathrm{d}^2\Delta}$, where $\frac{\mathrm{d}e_L}{\mathrm{d}\Delta} = \frac{\tau\gamma}{c''(e_L)}$ and $\frac{\mathrm{d}^2 e_L}{\mathrm{d}\Delta^2} = -(\tau\gamma)^2 \frac{c'''(e_L)}{(c''(e_L))^3}$, which implies that $\frac{\mathrm{d}\tilde{R}^2}{\mathrm{d}^2\Delta} \propto c'''(e_L)$.

Considering the entire economy, and letting R represent the revenues from all firms that do not manipulate the threshold, then

$$\frac{\mathrm{d}^2}{\mathrm{d}\Delta^2}R^\nu = (1-\nu)\frac{\mathrm{d}^2}{\mathrm{d}^2\Delta}R - \nu\frac{\mathrm{d}^2 e_L}{\mathrm{d}^2\Delta} = (1-\nu)\frac{\mathrm{d}^2}{\mathrm{d}^2\Delta}R + \nu(\tau\gamma)^2\frac{c'''(e_L)}{(c''(e_L))^3}.$$

The second term remains bounded in the limit $\Delta \to 0$. Hence, we can disregard it when Δ is small by the same reason we disregarded the terms $\frac{\mathrm{d}^2 e_i}{\mathrm{d}^2\Delta}$ in the proof of Theorem 4. It follows that revenues remain concave in Δ for Δ small and under the conditions stated in the Theorem statement. Adding the share of free manipulators only increases the number of taxpayers that reduce their declarations when Δ increases. Hence, it makes an increase in Δ less likely to generate revenues.

An Intermediate Manipulation Representative Economy: Suppose instead that we allow for an interior cost of threshold manipulation $\tilde{\kappa}$. In that case,

$$R(\Delta) = \int_{e_H}^{\tilde{y}^{HB}} (y - e_H)\,\mathrm{d}F + \int_{\tilde{y}^{HB}}^{\hat{y}_0 + e_L} \hat{y}(y)\,\mathrm{d}F + \int_{\hat{y}_0 + e_L}^{\bar{y}} (y - e_L)\,\mathrm{d}F.$$

Hence, an increase in Δ causes the following change in government revenues:

$$\frac{\mathrm{d}}{\mathrm{d}\Delta}R(\Delta) = H(\Delta) + L(\Delta) + B(\Delta)$$

where

$$H(\Delta) = -\frac{\mathrm{d}e_H}{\mathrm{d}\Delta}\left(F\left(\tilde{y}^{HB}\right) - F(e_H)\right); \quad L(\Delta) = -\frac{\mathrm{d}e_L}{\mathrm{d}\Delta}\left(1 - F\left(\hat{y}_0 + e_L(\Delta)\right)\right)$$

$$B(\Delta) = \frac{\mathrm{d}\tilde{y}^{HB}(\Delta)}{\mathrm{d}\Delta} f\left(\tilde{y}^{HB}(\Delta)\right)\left(\check{e}(\Delta) - e_H(\Delta)\right) + \int_{\tilde{y}^{HB}(\Delta)}^{\hat{y}_0 + e_L(\Delta)} \frac{\partial \hat{y}(y)}{\partial \Delta} \mathrm{d}F.$$

So the second derivative of revenues is

$$\begin{aligned}
\frac{\mathrm{d}^2}{\mathrm{d}\Delta^2} R(\Delta) = &- \left(\frac{\mathrm{d}^2 e_H}{\mathrm{d}\Delta^2}\left(F\left(\tilde{y}^{HB}\right) - F(e_H)\right) - \left(\frac{\mathrm{d}e_H}{\mathrm{d}\Delta}\right)^2 f(e_H) + \frac{\mathrm{d}e_H}{\mathrm{d}\Delta}\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta} f\left(\tilde{y}^{HB}\right)\right) \\
&- \left(\frac{\mathrm{d}^2 e_L(\Delta)}{\mathrm{d}\Delta^2} \cdot \left(1 - F(\hat{y}_0 + e_L(\Delta))\right) - \left(\frac{\mathrm{d}e_L(\Delta)}{\mathrm{d}\Delta}\right)^2 f(\hat{y}_0 + e_L(\Delta))\right) \\
&+ \frac{\mathrm{d}^2 \tilde{y}^{HB}(\Delta)}{\mathrm{d}\Delta^2} f\left(\tilde{y}^{HB}(\Delta)\right)\left(\check{e}(\Delta) - e_H(\Delta)\right) + \left(\frac{\mathrm{d}\tilde{y}^{HB}(\Delta)}{\mathrm{d}\Delta}\right)^2 f'\left(\tilde{y}^{HB}(\Delta)\right)\left(\check{e}(\Delta) - e_H(\Delta)\right) \\
&+ \frac{\mathrm{d}\tilde{y}^{HB}(\Delta)}{\mathrm{d}\Delta} f\left(\tilde{y}^{HB}(\Delta)\right)\left(\frac{\mathrm{d}\check{e}(\Delta)}{\mathrm{d}\Delta} - \frac{\mathrm{d}e_H(\Delta)}{\mathrm{d}\Delta} - \frac{\partial \hat{y}\left(\tilde{y}^{HB}\right)}{\partial \Delta}\right) + \int_{\tilde{y}^{HB}(\Delta)}^{\hat{y}_0 + e_L(\Delta)} \frac{\partial^2 \hat{y}(y)}{\partial \Delta^2} \mathrm{d}F
\end{aligned}$$

where we exploit $\frac{\partial \hat{y}(\hat{y}_0 + e_L(\Delta))}{\partial \Delta} = 0$. For the same arguments as in the proof of Theorem 4, $\frac{\mathrm{d}e_i(\Delta)}{\mathrm{d}\Delta}$, $\frac{\mathrm{d}^2 e_i(\Delta)}{\mathrm{d}\Delta^2}$, $i = H, L$ are bounded and $\frac{\partial^2 \hat{y}(y)}{\partial \Delta^2}$ is also bounded. To prove it, we differentiate the FOC in (B.2) and get

$$\frac{\partial \hat{y}(y)}{\partial \Delta} = -\frac{\gamma\tau}{c''(y - \hat{y}(y)) + \tilde{\kappa}\tilde{c}''(\hat{y}_0 - \hat{y}(y))}; \quad \frac{\partial^2 \hat{y}(y)}{\partial \Delta^2} = (\gamma\tau)^2 \frac{c'''(y - \hat{y}(y)) + \tilde{\kappa}\tilde{c}'''(\hat{y}_0 - \hat{y}(y))}{\left(c''(y - \hat{y}(y)) + \tilde{\kappa}\tilde{c}''(\hat{y}_0 - \hat{y}(y))\right)^3}.$$

Given that $c''(\cdot)$ and $\tilde{c}''(\cdot)$ are uniformly bounded away from zero, then $\frac{\partial^2 \hat{y}(y)}{\partial \Delta^2}$ is bounded. So, our analysis can focus on the following terms when considering the limit for $\Delta \to 0$

$$\begin{aligned}
\lim_{\Delta \to 0} \frac{\mathrm{d}^2}{\mathrm{d}\Delta^2} R(\Delta) \propto\ & \frac{\mathrm{d}^2 \tilde{y}^{HB}(\Delta)}{\mathrm{d}\Delta^2} f\left(\tilde{y}^{HB}(\Delta)\right)\left(\check{e}(\Delta) - e_H(\Delta)\right) - \frac{\mathrm{d}e_H}{\mathrm{d}\Delta}\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta} f\left(\tilde{y}^{HB}\right) \\
&+ \left(\frac{\mathrm{d}\tilde{y}^{HB}(\Delta)}{\mathrm{d}\Delta}\right)^2 f'\left(\tilde{y}^{HB}(\Delta)\right)\left(\check{e}(\Delta) - e_H(\Delta)\right) \\
&+ \frac{\mathrm{d}\tilde{y}^{HB}(\Delta)}{\mathrm{d}\Delta} f\left(\tilde{y}^{HB}(\Delta)\right)\left(\frac{\mathrm{d}\check{e}(\Delta)}{\mathrm{d}\Delta} - \frac{\mathrm{d}e_H(\Delta)}{\mathrm{d}\Delta} - \frac{\partial \hat{y}\left(\tilde{y}^{HB}\right)}{\partial \Delta}\right).
\end{aligned}$$

We establish a couple of useful results.

Lemma 13. $\lim_{\Delta \to 0} \frac{d\tilde{y}^{HB}}{d\Delta}\left(\check{e}(\Delta) - e_H(\Delta)\right) = \frac{\tau\gamma(1+\alpha)e^I(\mu)}{c''(e^I(\mu))}$ *and* $\frac{d^2\tilde{y}^{HB}(\Delta)}{d\Delta^2} = \frac{a_0 + a_1 \frac{d\tilde{y}^{HB}}{d\Delta} + a_2\left(\frac{d\tilde{y}^{HB}}{d\Delta}\right)^2}{\tilde{c}'(\hat{y}_0 - \hat{y}(\tilde{y}^{HB}))}$ *for some* a_0, a_1, a_2.

Using Lemma 13, the second derivative of revenues is

$$\lim_{\Delta\to0}\frac{\mathrm{d}^2}{\mathrm{d}\Delta^2}R(\Delta)\propto\left(\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}\right)^2\left[f'\left(\tilde{y}^{HB}\right)(\check{e}-e_H)+f\left(\tilde{y}^{HB}\right)\left(1-\frac{\partial\hat{y}\left(\tilde{y}^{HB}\right)}{\partial y}\right)\right.$$
$$\left.-f\left(\tilde{y}^{HB}\right)(\check{e}-e_H)\frac{\tilde{c}''\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)}{\tau\gamma\tilde{c}'\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)}\frac{\partial\hat{y}\left(\tilde{y}^{HB}\right)}{\partial\Delta}c''\left(\tilde{y}^{HB}-\hat{y}\left(\tilde{y}^{HB}\right)\right)\right]$$
$$-\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}\left[-2\frac{\tilde{c}''\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)}{\tilde{c}'\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)}\frac{\partial\hat{y}\left(\tilde{y}^{HB}\right)}{\partial\Delta}f\left(\tilde{y}^{HB}\right)(\check{e}-e_H)+\frac{\mathrm{d}e_H}{\mathrm{d}\Delta}f\left(\tilde{y}^{HB}\right)\right.$$
$$\left.+\left(\frac{\mathrm{d}e_H}{\mathrm{d}\Delta}+2\frac{\partial\hat{y}\left(\tilde{y}^{HB}\right)}{\partial\Delta}\right)f\left(\tilde{y}^{HB}\right)\right]+\frac{1}{\tilde{\kappa}}\frac{\left(\frac{(\tau\gamma\alpha)^2}{c''(e_H)}+\tau\gamma\frac{\partial\hat{y}\left(\tilde{y}^{HB}\right)}{\partial\Delta}\right)}{\tilde{c}'\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)}f\left(\tilde{y}^{HB}\right)(\check{e}-e_H),$$

which we can write as

$$\lim_{\Delta\to0}\frac{\mathrm{d}^2}{\mathrm{d}\Delta^2}R(\Delta)=\lim_{\Delta\to0}\left(\left(\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}\right)^2\alpha_2+\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}\alpha_1+\alpha_0\right).$$

Lemma 14. *We have that* $\lim_{\Delta\to0}\alpha_0$ *is finite,* $\lim_{\Delta\to0}\alpha_1>0$ *(and finite),* $\lim_{\Delta\to0}\alpha_2=0$.

Recall that $\lim_{\Delta\to0}\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}=-\infty$. Therefore, following the same argument as in Lemma 7, revenues are concave if and only if $\lim_{\Delta\to0}\alpha_1+\lim_{\Delta\to0}\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}\alpha_2>0$. Since the limit of α_1 is positive (see above), we have to study $\lim_{\Delta\to0}\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}\alpha_2$.

Lemma 15. $\lim_{\Delta\to0}\frac{d\tilde{y}^{HB}}{d\Delta}\alpha_2=\frac{\tau\gamma(1+\alpha)e^I(\mu)}{c''(e^I(\mu))}f'\left(\hat{y}_0+e^I(\mu)\right)-\tau\gamma(1+\alpha)\left[\frac{\frac{c'''\left(e^I(\mu)\right)}{c''\left(e^I(\mu)\right)}\frac{\tilde{\kappa}\tilde{c}''(0)e^I(\mu)}{c''(e^I(\mu))+\tilde{\kappa}\tilde{c}''(0)}+1}{2c''(e^I(\mu))}\right]$.

It follows that

$$\lim_{\Delta\to0}\alpha_1+\lim_{\Delta\to0}\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}\alpha_2=\frac{2\gamma\tau}{c''\left(e^I\right)+\tilde{\kappa}\tilde{c}''(0)}f\left(\hat{y}_0+e^I\right)\left[1+\frac{\tilde{\kappa}\tilde{c}''(0)}{c''\left(e^I\right)}\right]+2\frac{\tau\gamma\alpha}{c''\left(e^I(\mu)\right)}f\left(\hat{y}_0+e^I\right)$$
$$+\frac{\tau\gamma(1+\alpha)e^I}{c''\left(e^I\right)}f'\left(\hat{y}_0+e^I\right)-\tau\gamma(1+\alpha)f\left(\hat{y}_0+e^I\right)\left[\frac{\frac{c'''\left(e^I\right)}{c''\left(e^I\right)}\frac{\tilde{\kappa}\tilde{c}''(0)e^I}{c''\left(e^I\right)+\tilde{\kappa}\tilde{c}''(0)}+1}{2c''\left(e^I\right)}\right]$$
$$\propto\frac{3}{2}+\frac{f'\left(\hat{y}_0+e^I\right)}{f\left(\hat{y}_0+e^I\right)}e^I-\frac{1}{2}\frac{c'''\left(e^I\right)}{c''\left(e^I\right)}\frac{\tilde{\kappa}\tilde{c}''(0)e^I}{c''\left(e^I\right)+\tilde{\kappa}\tilde{c}''(0)}$$

We must require therefore that

$$\frac{3}{2}+\frac{f'\left(\hat{y}_0+e^I\right)}{f\left(\hat{y}_0+e^I\right)}e^I-\frac{e^I}{2}\frac{c'''\left(e^I\right)}{c''\left(e^I\right)}\frac{\tilde{\kappa}\tilde{c}''(0)}{c''\left(e^I\right)+\tilde{\kappa}\tilde{c}''(0)}>0,$$

which is identical to the condition in Theorem 4 up to an adjustment term that depends on $\tilde{\kappa}$. Importantly, our statement on sufficient conditions still holds. Notice that when $\tilde{\kappa}\to+\infty$ the condition converges to the one derived in the baseline model.

B.3 Fixed cost of receiving an audit

Denote $a > 0$ the fixed cost of receiving an audit. Notice that behavior inside the $\mathcal{H}$ and $\mathcal{L}$ areas is unchanged. The value of the firm is now

$$V(p) = \max_{e_i} \tau (1 - \tau\gamma p) e_i - c(e_i) - pa,$$

so the indifference condition for the marginal buncher reads

$$\tau (1 - \tau\gamma p_H) e_H - c(e_H) - p_H a = \tau (1 - \tau\gamma p_L) \tilde{e} - c(\tilde{e}) - p_L a,$$

which we rewrite as

$$\tau [1 - \tau\gamma p_L] \tilde{e} - c(\tilde{e}) = V_H - a(p_H - p_L).$$

Given the exact same structure as in Section B.1, we conclude that the fixed cost makes the $\mathcal{L}$ region more amenable. Because the change in $a(p_H - p_L)$ caused by an increase in Δ is finite, $\lim_{\Delta \to 0} \frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta} (\tilde{e} - e_H)$ is finite and the sufficient conditions or Theorem 4 still hold.

C Bunching Estimation

We define bunching as the excess mass in the observed revenue declaration distribution relative to a counterfactual distribution that would arise with a constant p_L probability. Since businesses declaring above $\hat{y}$ face probability p_L, our strategy uses their distribution to infer this counterfactual. Empirically, we follow the approach in Kleven and Waseem (2013), which relies on a flexible polynomial and excludes an area $[y_l, y_u]$ around $\hat{y}$ from the density distribution estimation. We bin the data in segments whose length is 1 percentage point of $\hat{y}$ and run the following regression for the number of SeS files c in each bin j

$$c_j = \sum_{i=1}^{K} \beta_i \left(y_j\right)^i + \sum_{h=y_l}^{y_u} \gamma_h \mathbb{1}\left(y_j = h\right) + \varepsilon_j, \tag{C.1}$$

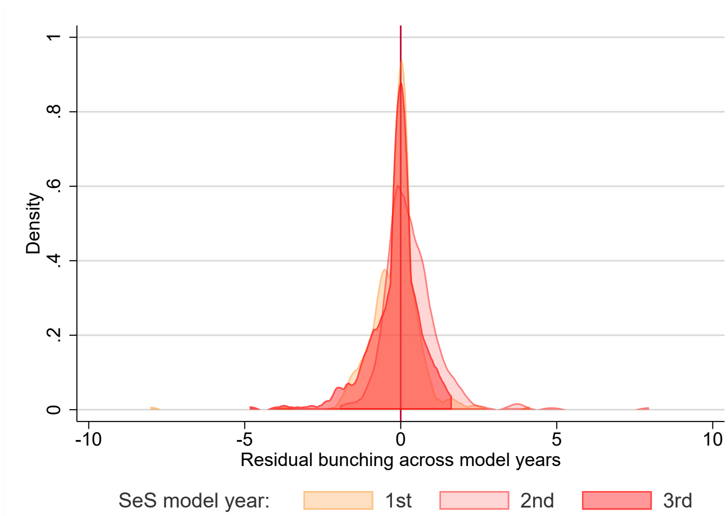
where i indicates the polynomial degree in the first sum. We use a 7^{th} degree polynomial in our baseline estimates. To avoid irregularities coming from the far tails of the distribution, we exclude files with reported revenues below the 5^{th} percentile or above the 95^{th} percentile of relative distance from $\hat{y}$. These restrictions automatically drop files with zero reported revenues, which account for slightly less than 2% of all 2007-2010 files. The excluded segment $\left[y^l, y^u\right]$ is the area affected by bunching responses. Bin dummies for $y \in \left[y^l, y^u\right]$ ensure that the excess mass at $\hat{y}$ does not affect the counterfactual distribution fit. While our preferred choice is to set y^u visually at the first bin above $\hat{y}$, we choose y^l using an iterative procedure. The latter searches for the bin that generates an estimated counterfactual with a missing mass below $\hat{y}$ equal to the excess mass above $\hat{y}$. Using the estimated counterfactual, we can compute the excess mass as the ratio between the excess (relative to counterfactual) observed number of SeS files and the average level of the counterfactual in the segment $[\hat{y}, y^u]$. We will refer to this relative excess mass as a bunching estimate, or $\hat{B}$.

We compute standard errors to bunching estimates using a semi-parametric bootstrap procedure. Equation (C.1) provides the structure for our routine. In every bootstrap iteration we draw with replacement from the residuals $\hat{\varepsilon}_j = c_j - \hat{c}_j$, where $\hat{c}_j = \sum_{i=1}^{K} \hat{\beta}_i \left(y_j\right)^i + \sum_{i=y_l}^{y_u} \hat{\gamma}_i \mathbb{1}\left(y_j = i\right)$, and $\left(\hat{\beta}, \hat{\gamma}\right)$ are the estimated coefficients from the specification in (C.1). We use the residuals to build a new number of taxpayers in each bin j so that in iteration r the number of taxpayers in bin j is $c_j^r = \hat{c}_j + \varepsilon_j^r$ and ε_j^r is the residual drawn for bin j in iteration r. We use the new vector $\left(c_j^r\right)_{j \in J}$ as the dependent variable when re-estimating (C.1) and we employ the resulting $\left(\hat{c}_j^r\right)_{j \in J}$ as the counterfactual needed to compute a bunching quantity $\hat{B}^r$. We repeat this routine for 1,000 iterations. Confidence intervals on $\hat{B}$ can be computed by taking the 2.5^{th} and the 97.5^{th} percentiles of the bunching estimate distribution across all iterations, while the bunching standard deviation is simply the standard deviation of the same empirical distribution.

D Reporting VS Production Responses

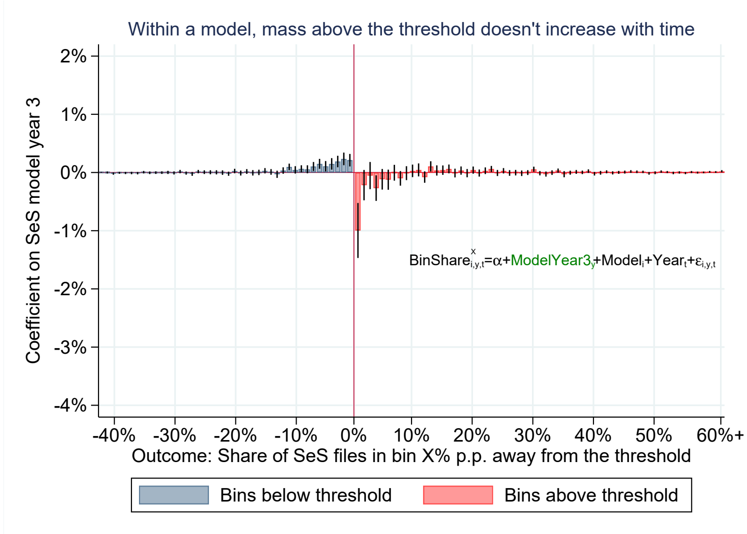
We assess the learning-to-adjust hypothesis discussed in Section 4.2 in two ways. First, we estimate bunching for every sector and year, and residualize these estimates by sector and year fixed effects. Figure D1 plots the residual bunching distributions for the first, second, and third year of application of a given revenue prediction model for any given sector. Despite the potential for learning, bunching residuals aren't significantly higher for the later years of application of the same model. Second, we split SeS files in one percentage point bins of distance from the presumed revenues for each sector-year. If production adjustment takes place over time, we expect mass gains in the bins just above the SeS threshold. For each bin, we thus regress its file share on a dummy for the last year of application of the relevant model, along with sector and calendar year fixed effects. Figure D2 plots the coefficient on the third-year dummy for each bin around the SeS threshold. We don't find any evidence that the bins just above the threshold gain mass by the end of a model's application, as most coefficients are negative but small or insignificant in size.

Figure D1: Bunching evolution within SeS models, 2007-2010



Notes: this Figure plots the distribution of 2007-2010 bunching residuals from a regression of the form: $\text{bunching}_{i,t} = \alpha + \beta_i + \gamma_t + \varepsilon_{i,t}$, where the unit of observation is a SeS model-year, and we include fixed effects for each SeS model i and calendar year t. By SeS model we refer to the three-year application of a given SeS estimation model, inclusive of the presumed revenues function, to a given business sector defined by the SeS. We thus plot three residual distributions, separately for the first, second, or third year of application of a given SeS model. Only positive bunching estimates are employed. Regression sample is of size 762 and excludes SeS model-years with negative bunching estimates.

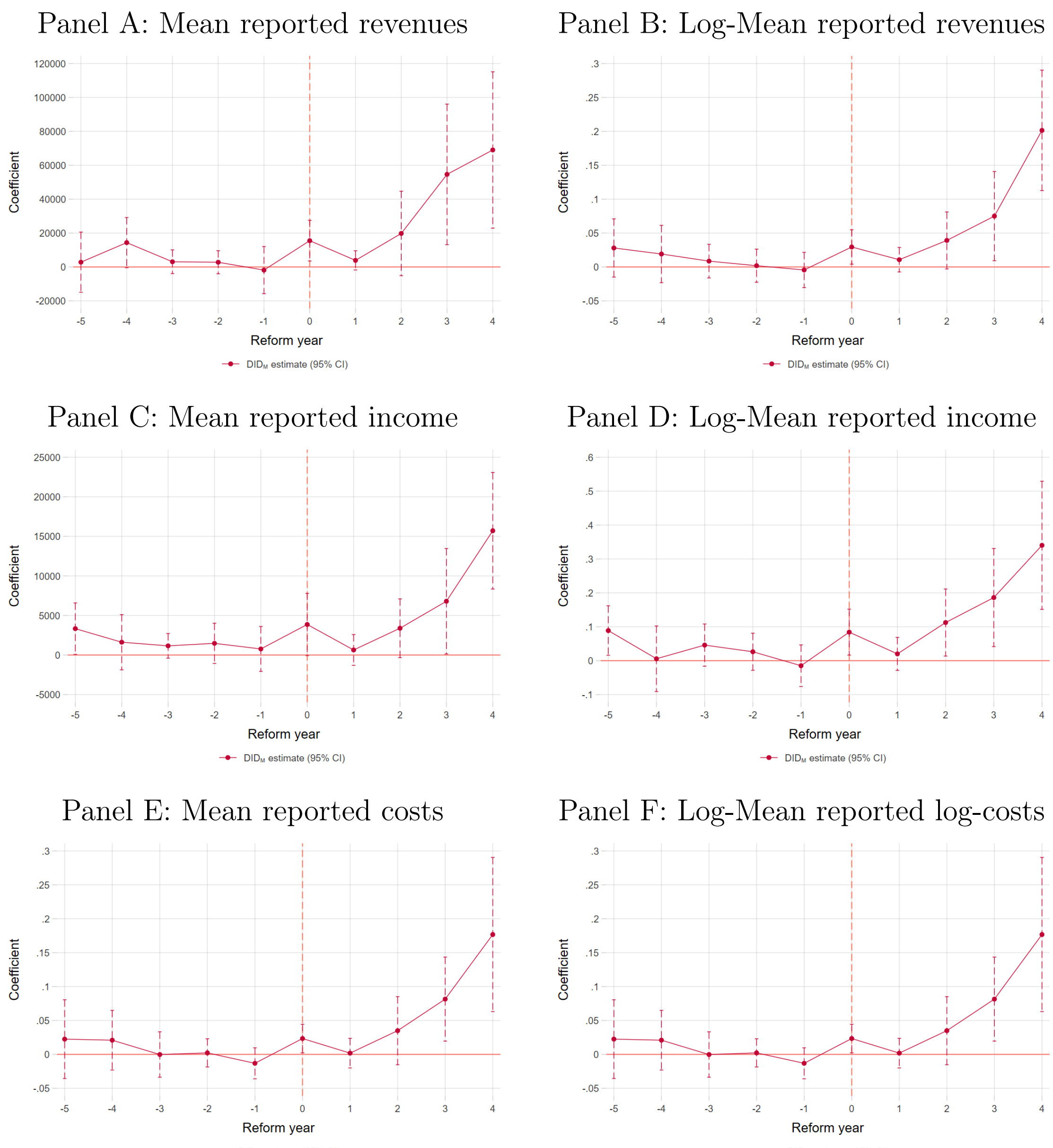
Figure D2: Bin share effect of the last year of application of a SeS model



Notes: this Figure provides the coefficient plot from several regressions as the one printed above. Specifically, we observe SeS models i and consider whether they are being applied for the s^{th} year (that is, first, second, or third year) during calendar year t. Across all SeS model-years, we regress the share of SeS files at each one percentage point of distance X from the SeS presumed revenue threshold on a dummy for the third (last) year of application of a SeS estimation model, controlling for SeS model and calendar year fixed effects, and clustering standard errors by SeS model. We then plot the coefficient associated to the third (last) year dummy at each point of distance below (in blue) and above the threshold (in red), along with its 95% CIs. To compute the bin shares, we consider the sample of SeS taxpayers continuously filing over 2007-2016.

E Robustness of Event Study Estimates

Figure F1: Robustness: de Chaisemartin and D'Haultfoeuille (2020) correction



Notes: this Figure shows the effects of the reward regime's introduction in a sector from the estimator proposed by de Chaisemartin and D'Haultfoeuille (2020) (DID_M). The regressions are estimated on the sample of all Sector Study files from single-sector taxpayers continuously filing over the 2007-2016 period, aggregated by sector-year. Only sectors accessing the reward regime by 2016 are considered. Number of sector-years: 1550. Reported revenues are winsorized at the 99th percentile. Controls and weighting are as defined in the discussion of Eq. (6). We mark the relative year before the reform with the red dashed vertical line at year 0. Estimation is performed using the *dynamic* and *placebo* options of the authors' supplied Stata package *did_ multiplegt*. Estimation requirements allow us to compute only up to four post-treatment effects. Standard errors are computed with a bootstrap procedure with 50 replications.

F Structural Estimation

F.1 Model Primitives

To simulate our economy, we need two sets of model primitives. First, we need primitives on firms income and preferences for evasion. Second, we need the primitives of the audit system.

Firms' types: Each firm is defined by a revenue and cost type combination (y, κ). We let δ^H denote the share of always honest firms (those that would declare their true revenues under any audit system). Conditional on having a positive level of revenues, firm types are distributed according to a joint truncated normal distribution. The support of the marginal distribution on y is $[0, \bar{y}]$ where $\bar{y}$ is the maximum revenues observed in the data, while the support of the marginal distribution on κ is $[0, +\infty]$. We can write this truncated joint normal distribution as a function of some primitives as follows: $\mathcal{N}(\mu, \mathbf{\Sigma}) := N\left[\begin{pmatrix} \mu_y \\ \mu_\kappa \end{pmatrix}, \begin{pmatrix} \sigma_y^2 & \sigma_{y\kappa} \\ \sigma_{y\kappa} & \sigma_\kappa^2 \end{pmatrix}\right]$. Hence, μ_y and μ_κ denote the average revenues and cost parameter conditional on having positive revenues, σ_y^2 and σ_κ^2 are the variances, and $\sigma_{y\kappa}$ measures the covariance between revenues and the cost of evasion.

The optimal evasion level for each firm is determined by τ and γ that we calibrate separately and by the cost of evading. We assume the following functional form for this cost

$$c(e, \kappa) = \frac{1}{\kappa\left(1 + \frac{1}{\epsilon}\right)} (\kappa e)^{1+\frac{1}{\epsilon}}.$$

While we let κ change across taxpayers, we assume a unique elasticity ϵ across taxpayers in the same macro-sector.

Audit rules The audit system is characterized by a threshold $\hat{y}$ and by the audit probabilities above and below $\hat{y}$. Since we will compare the pre- and post-Reward Regime economies, we will need to estimate p_L and p_H before and after the policy. Using the notation from the rest of our theory, we need to quantify μ, Δ, α, and Δ'.

Primitive vector: Putting all primitives together, we derive the following vector θ

$$\theta = \left(\underbrace{\delta^H, \mu_y, \mu_\kappa, \sigma_y^2, \sigma_\kappa^2, \sigma_{y\kappa}, \epsilon}_{\text{Firm types}}, \underbrace{\mu, \Delta, \alpha, \Delta'}_{\text{Audit rules}}\right).$$

So we have 11 primitives to estimate.

F.2 Theoretical Moments

For our estimation, we exploit the following theoretical moments. First, we use the size of the masses in the six regions studied in Figure G4 in the pre-policy period (denoted by the superscript *Pre*). We split the region just above $\hat{y}$ into two regions: one between $\hat{y}$ and 1 percentage point above it (the bunching area), and one between 1 and 5 percentage points of above $\hat{y}$. This provide a total of 7 regions defined formally by the following expression

$$M^{Pre}(\mathcal{A}) = \mathbb{E}_k\left[F_{y|\kappa}\left(\bar{y}^{\mathcal{A}}\right)\right] - \mathbb{E}_\kappa\left[F_{y|\kappa}\left(\underline{y}^{\mathcal{A}}\right)\right].$$

Moreover, we employ the average declarations below $\hat{y}$ and above $\hat{y}$ before and after the policy[45]

$$\bar{d}^{Pre}_{\mathcal{A}} = \mathbb{E}\left[y|y \in \mathcal{A}\right] - \mathbb{E}_k\left[e^{Pre}_{A,\kappa}\right].$$

Similarly, we can define similar theoretical expressions for the moments $\bar{d}^{Post}_{\mathcal{A}}$ that refer to the same quantities after the introduction of the Reward Regime. Hence, we have a total of 11 theoretical expressions. In the next Subsection, we discuss how we measure in the data the moments just described.

F.3 Data moments

Declarations normalization: To implement our estimation, we pool together the taxpayers in the downstream sector. For this reason, there is variation in the presumed revenue threshold across units of observation. To deal with it, we normalize each taxpayer's declaration with respect to its specific threshold, and express all the formulas in percentage deviation from the threshold. Formally, we transform the observed declarations into $\tilde{d} = \frac{d}{\hat{y}}$ and we generate a set of bins $\mathcal{I} = \left\{0, \frac{i}{100 \cdot \hat{y}}, 1, 1 + \frac{i}{100}\frac{\bar{y}}{\hat{y}}\right\}_{i \in \mathbb{N}_{<100}}$.

Moments quantification: We measure the data moments as follows. First, measuring the masses of taxpayers in the data is straightforward since the seven groups are defined based on their declared revenues before the policy, which we can observe. Similarly, we can easily measure the average declarations conditional on being above and below $\hat{y}$. Due to the normalization, these declarations are expressed in percentage points of the threshold. To quantify the declarations after the implementation of the Reward Regime, we exploit the fact that for a given declaration $\bar{d}_{\mathcal{A}}$ it holds that $\bar{d}^{Post}_{\mathcal{A}} = \bar{d}^{Pre}_{\mathcal{A}} + \delta\bar{d}_{\mathcal{A}}$, where $\delta\bar{d}_{\mathcal{A}}$ is the causal effect of the policy on $\bar{d}_{\mathcal{A}}$ that we can estimate using the identification strategy outlined in Section 5. We summarize the time-varying effects from equation (6) into a single $\delta\bar{d}_{\mathcal{A}}$ per dependent variable by taking the average across the post-reform coefficients.

F.4 Estimation

Our estimation algorithm searches for the vector of primitives that minimizes the distance between our theoretical and empirical moments. Theoretical moments are determined by a procedure in five steps outlined below.

Step 1: We calibrate the tax rate and the penalty rate ex ante. The tax rate is set at 24% to match the corporate tax rate (IRES), while we use $\gamma = 2$.[46]

Step 2: We start by computing the three key levels of evasion e_L, e_H, $\tilde{e}$ for each firm type. Specifically, given a firm type, we solve the firm's problem taking as input the cost elasticity ϵ and the audit system parameters. While e_L and e_H solve the FOC in (3), $\tilde{e}$ solves the indifference condition in (4).

Step 3: Based on the optimal declarations, we classify each firm type into one of the declaration groups using the results of Proposition 1.

[45] To exclude the bunching region and capture average declarations in the $\mathcal{L}$ area, we only consider firms declaring more than 5 p.p. above $\hat{y}$ to compute the average declaration above the threshold.

[46] For γ, we select the lower bound of the penalty range set by law in the sample period, when administrative fines - corresponding to $\gamma - 1$ in our model - could vary between 100% and 200% of any detected evasion. We thus set $\gamma = 2$.

Step 4: Given the optimal declarations and a distribution of firm types $F(\cdot,\cdot)$, we find the associated distribution of declared revenues. We calibrate $\bar{y}$ at the 95^{th} percentile of observed declarations, which corresponds to 2.5 times $\hat{y}$. Given our assumptions, in the truncated region $[0,\bar{y}] \times [0,\infty)$ the joint distribution of firm types is

$$f(y,\kappa) = \frac{\phi_{\boldsymbol{\mu},\boldsymbol{\Sigma}}(y,\kappa)}{\mathbb{P}_{\mu,\Sigma}([0,\bar{y}] \times [0,\infty))}$$

where $\phi_{\boldsymbol{\mu},\boldsymbol{\Sigma}}$ is the pdf of the $\mathcal{N}(\mu,\boldsymbol{\Sigma})$ distribution and

$$\mathbb{P}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}([0,\bar{y}] \times [0,\infty)) = \int_0^{\bar{y}} \int_0^{\infty} \phi_{\boldsymbol{\mu},\boldsymbol{\Sigma}}(y,\kappa)\,\mathrm{d}y\mathrm{d}\kappa$$

is a positive constant. Notice that there exists a finite $\bar{\kappa}$ such that firms with $\kappa > \bar{\kappa}$ are effectively honest (*i.e.* their declaration will fall in the same revenue bin as their true revenues). The latter are a fraction $\frac{\mathbb{P}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}([0,\bar{y}]\times[\bar{k},\infty))}{\mathbb{P}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}([0,\bar{y}]\times[0,\infty))}$ of the not always honest declarers that earn positive revenues. An important assumption that we maintain throughout our estimation is that the share of always honest firms is equally spread across all revenue types. We draw a sample of 200,000 firms in the simulated economy.

Step 5: Using the distribution of declared revenues $g(\cdot)$, we can quantify (as a function of the primitives) the masses in the seven regions before the policy, and the average declaration in each region ($\mathcal{H}$ and $\mathcal{L}$) before and after the policy. This provides us our theoretical moments.

Minimizing the distance between theoretical and empirical moments: Given our theoretical and data moments, we search for a vector of primitives that minimizes the following loss function:

$$m(\theta)' \, \Omega m(\theta),$$

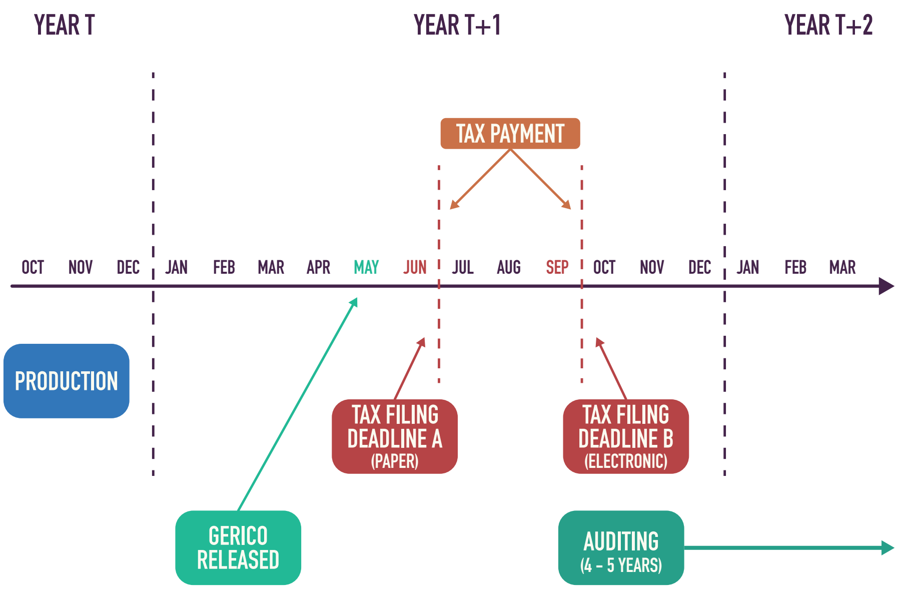
where m is the vector collecting all differences between theoretical and empirical moment, and Ω is a weighting matrix. Each element of m is $m_i(\theta) - \tilde{m}_i$, where $\tilde{m}_i$ is the data moment i and $m_i(\theta)$ is its theoretical counterpart, which is simulated from the vector of primitives θ. We run the estimation in two steps. The first step uses an identity weighting matrix $\Omega^1 = I$ that gives equal weight to each moment and estimates a vector of primitives $\hat{\theta}^1$. Then, we make B draws of 200,000 firms from the economy characterized by $\hat{\theta}^1$. We simulate a total of $B = 200$ iterations. Next, we compute the variance-covariance matrix of the theoretical moments across iterations and we use it to generate the vector of weights for the second step of our estimation. Specifically, we define the weight matrix in step 2 as $\Omega^2 = V_b\left(m^b\left(\hat{\theta}^1\right)\right)$, where $m^b\left(\hat{\theta}^1\right)$ is the vector of theoretical moments quantified in iteration b.

Comparing estimated p_H and p_L to the observed ones: We compare our estimates of p_H and p_L to the pre-reward-regime ones by re-elaborating statistics provided by D'Agosto et al. (2017). From their Table 1, we derive information on the number of audited and non-audited SeS taxpayers in the years 2007 to 2010. Table 4 instead provides the information on the share of audited and non-audited taxpayers that had declared above their threshold. Combining the information from the two tables we are therefore able to estimate the probability that a taxpayer declaring above or below the threshold receives a tax audit. These are respectively 7.127% and 10.517%, which we then compare to the p_L and p_H estimated by our SMM.

G Additional Tables and Figures

G.1 Additional Figures

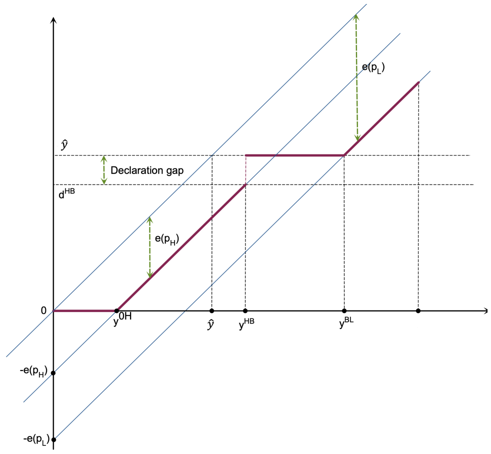
Figure G1: Timeline of Sector Studies Reporting



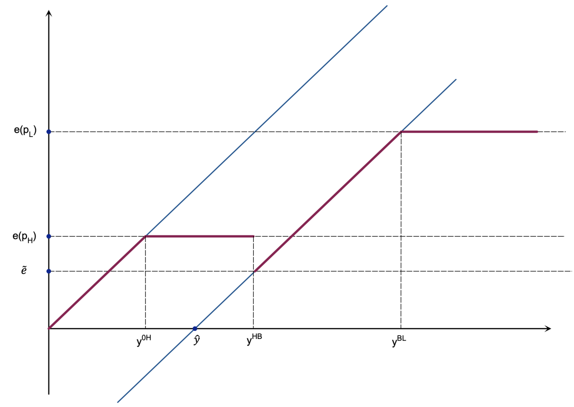
Notes: this Figure outlines the timeline of production and tax enforcement events from the perspective of taxpayers. Businesses generating revenues during year T are required to file their tax returns as well as their separate SeS file during the following year $T+1$. SeS filing follows the tax filing cycle. During our sample period, the taxpayers in our data file and pay their taxes either in June or in September, depending on whether they file on paper or electronically, respectively. At the beginning of every filing season, the Italian Revenue Agency releases *Gerico* to help with SeS filing and allowing taxpayers to compute their presumed revenues and a broader set of accounting indicators. After submission, auditing of SeS files and tax returns can take place over the following 4 to 5 years.

Figure G2: Optimal Declaration and Evasion

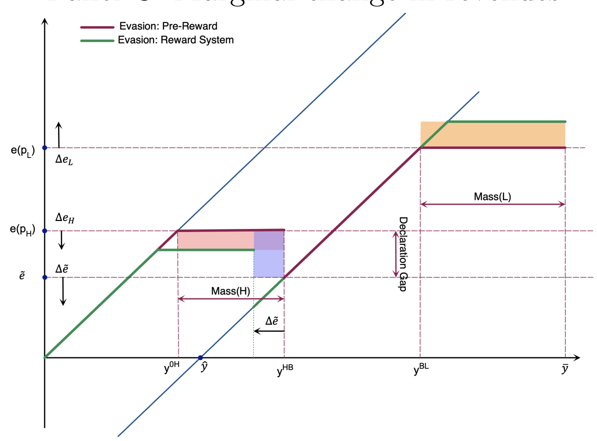
Panel A: Declaration Partition



Panel B: Evasion Partition



Panel C: Marginal change in revenues



Notes: this Figure outlines the optimal declaration (Panel A) and evasion (Panel B) for all levels of income (on the x-axis) and for a fixed cost function such that, given a perceived probability of audit, all incomes evade the same amount. Panel C shows graphically the marginal change in revenues caused by an increase in Δ, as described by Proposition 3. In Panel A, the blue diagonal lines represent (from top to bottom) the honest declaration pattern (45-degree line); the declaration pattern with a constant p_L; the declaration pattern with a constant p_H. The purple line shows the equilibrium pattern of declarations. In Panel B, the diagonal blue lines represent (from top to bottom) the full evasion pattern (45-degree line); and the evasion patter if firms declared $\hat{y}$ (parallel to 45-degree intersecting x-axis at $\hat{y}$). The purple line shows the equilibrium pattern of evasion. In Panel C, the blue lines are identical to the ones in Panel B, the purple line represents the evasion pattern before the change in Δ, while the green line is the evasion pattern after the increase in Δ.

G.2 Additional Tables

Table G1: Sector Studies compliance benefits, before and after 2011

SeS required condition			**Audit exemption benefits**	
Congruence	*Normality*	*Coherence*	*Before 2011*	*Since 2011*
✓			**No SeS audits** (revenues)	
	✓	✓	**No SeS audits** (costs, inputs)	
✓	✓		**No analytic-inductive audits** *up to* $e \leq 40\% y$, $e \leq €50{,}000$	
✓	✓	✓		**1. No analytic-inductive audits** *up to any amount* **2. No synthetic audits** *up to* $\pi(s)-\pi \leq 33\% \cdot \pi(s)$ **3. Shorter statute of limitation**

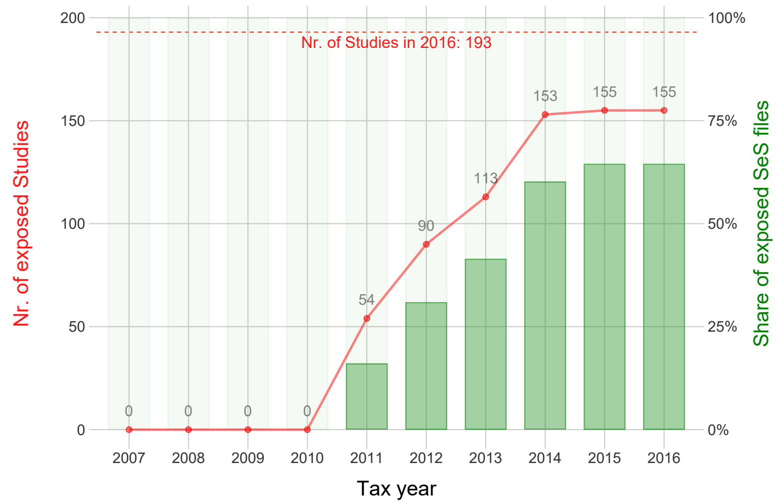
Notes: the Table reports the main tax audit and assessment benefits from being congruous, coherent, and normal by the definitions provided by Sector Studies, before and after the introduction of the 2011 reward regime. Congruence refers to the condition of reporting revenues at or above the level presumed by *Gerico*. Normality and coherence refer to the condition of reporting a number of accounting and economic indicators within sector-specific acceptable ranges as determined by *Gerico*. Notation: e refers to undeclared amounts, y to revenues, π to gross profits or income, and $\pi(s)$ to synthetically-determined income. The statute of limitation to inspect an eligibile taxpayer's file drops by one year since 2011.

Table G2: Bunching estimates by polynomial order and upper bound

	POLYNOMIAL ORDER							
UPPER BOUND	3	4	5	6	7	8	9	10
0	15.38	14.95	11.53	11.42	**9.56**	9.15	8.47	7.69
1	18.57	18.77	14.45	14.01	12.27	11.63	10.74	9.95
2	21.36	20.66	16.86	16.25	14.6	13.03	12.75	11.45
3	23.73	23.91	19.55	18.33	16.85	14.57	14.77	13.24
4	15.07	24.78	21.83	19.93	18.91	16.39	16.32	14.91
5	17.17	26.59	24.34	21.31	20.4	17.99	17.43	16.71

Notes: the Table reports various bunching estimates computed with the SeS files submitted by the universe of single-sector businesses in the 2007-2010 tax years. Estimates are reported for each combination of two parameters choices. First, the upper bound y_u of the area affected by excess bunching, identified by the floor of the relevant bin in percentage points of presumed revenues. For reference, upper bound 0 indicates we limit the bunching area to the one-percentage-point bin including the presumed revenues threshold. Second, the polynomial order, that is the degree of the polynomial in bin order used to estimate the smooth bunching counterfactual. We select 0 for the upper bound and 7 for the polynomial order in our baseline estimate, highlighted in orange. In all estimations, bin width is fixed at one percentage point of presumed revenues.

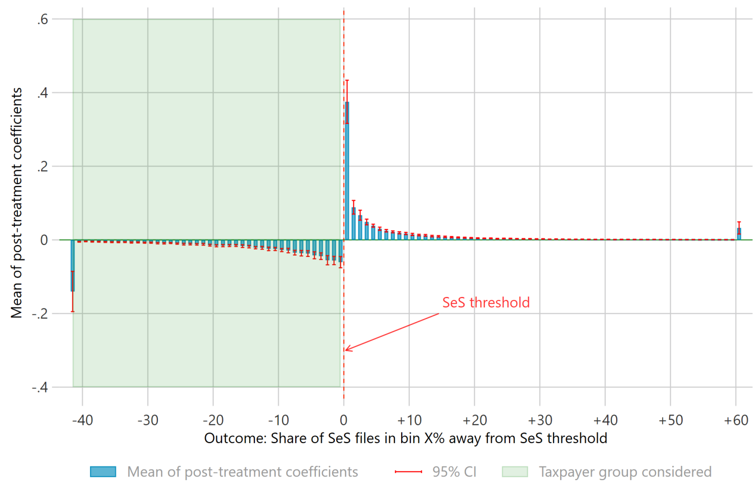
Figure G3: Reward regime: staggered introduction, 2011-2016



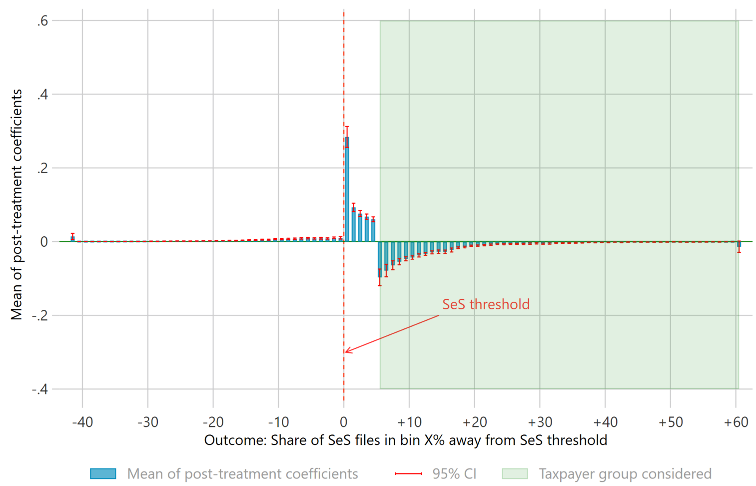
Notes: the Figure shows the staggered introduction of the 2011 reward regime among existing Sector Studies (for brevity, referred to as sectors). The red line displays the number of sectors with access to the regime in each year up to 2016 (scale on the left vertical axis). The dark green bars reflect the share of all files with access to the reward regime in each year (scale on the right vertical axis). The share is computed over the population of files from single-sector, continuous filers over 2007-2016. For simplicity, we code five sectors with partial access as having full regime access. The horizontal dashed line represents the total number of sectors in 2016.

Figure G4: Reward regime-induced adjustments, by presumed revenues distance before the reform

Panel A: $\mathcal{H}$ declarers



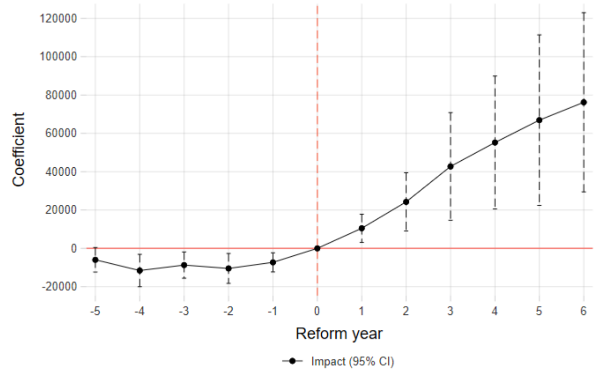
Panel B: $\mathcal{L}$ declarers



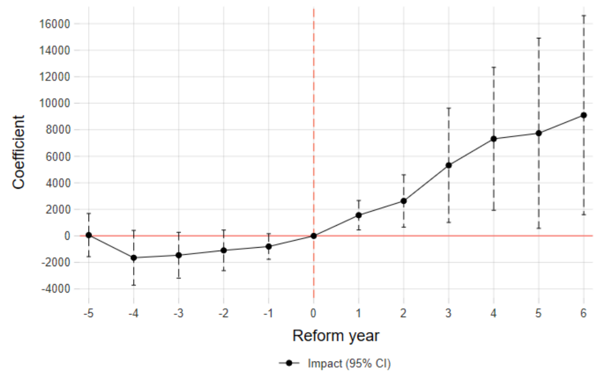
Notes: this Figure illustrates the effect of the reward regime on taxpayer declaration behavior for the $\mathcal{H}$ (below $\hat{y}$) and $\mathcal{L}$ (above $\hat{y}$) areas defined in our model. Taxpayers are assigned to these areas based on their pre-reform declarations. Here, the dependent variable is the share of taxpayers in the reference group declaring within each one-percentage-point bin. The original location of each group is marked by a green band. In both panels, bars represent the average of six group-specific post-treatment coefficients from a staggered difference-in-differences based on the specification in (6). Whiskers represent 95% CIs of these linear combinations of coefficients. Standard errors are clustered at the sector level. The regressions are estimated on the sample of all SeS files from single-sector taxpayers continuously filing over the 2007-2016 period, aggregated by sector-year. Only sectors accessing the reward regime by 2016 are considered. Number of sector-years: 1550. Declared revenues are winsorized at the 99th percentile.

Figure G5: Reward regime effects on mean revenues and profits (in Euros)

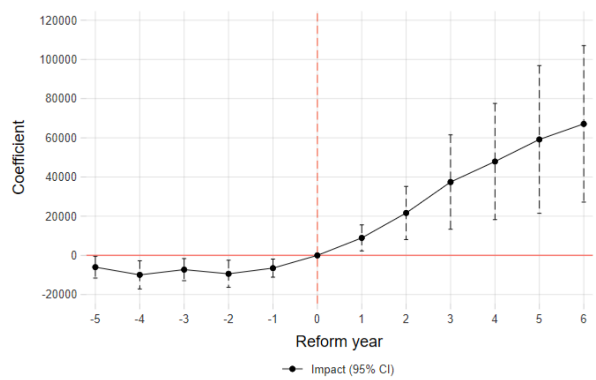
Panel A: Mean reported revenues



Panel B: Mean reported income



Panel C: Mean reported costs



Notes: this Figure shows the effects of the reward regime's introduction in a sector on mean reported revenues (Panel A), mean gross profits (Panels B), and costs (Panel C) defined as the difference between reported revenues and gross profits. Dependent variables are expressed in Euro terms. Whiskers represent 95% CIs. Effects are relative to the year before the advent of the reform in each sector, marked at year 0 by the red dashed vertical line. Estimates are based on our specification in (6). Standard errors are clustered at the sector level. The regressions are estimated on the sample of all Sector Study files from single-sector taxpayers continuously filing over the 2007-2016 period, aggregated by sector-year. Only sectors accessing the reward regime by 2016 are considered. Number of sector-years: 1550. Reported revenues are winsorized at the 99th percentile.

Supplementary Material

S.1 Omitted Computations from Main Propositions

Computations of Lemma 5

From the indifference condition we have $V^H\left(y^{HB}, d^H\left(y^{HB}\right)\right) = V^L\left(y^{HB}, \hat{y}\right)$ that solves

$$V^H(y) \equiv \max_d (y-d)\,\tau\left(1-\gamma p_H(\Delta)\right) - c(y-d) = (y-\hat{y})\,\tau\left(1-\gamma p_L(\Delta)\right) - c(y-\hat{y}) \equiv V^B(y).$$

So, we want $\frac{\mathrm{d}y^{HB}}{\mathrm{d}\Delta}$ where y^{HB} solves $V^H(y) - V^B(y) = g(y,\Delta) = 0$. Hence,

$$\frac{\mathrm{d}}{\mathrm{d}\Delta}y^{HB} = -\frac{\frac{\partial}{\partial\Delta}g(y,\Delta)}{\frac{\partial}{\partial y}g(y,\Delta)},$$

and

$$\begin{aligned}\frac{\partial}{\partial\Delta}g(y,\Delta) &= \frac{\partial}{\partial\Delta}\left[V^H(y) - V^B(y)\right] = -\tau\gamma\frac{\partial p_H}{\partial\Delta}(y-d) + \tau\gamma\frac{\partial p_L}{\partial\Delta}(y-\hat{y}) \\ &= \tau\gamma\left[\frac{\partial p_L}{\partial\Delta}(y-\hat{y}) - \frac{\partial p_H}{\partial\Delta}(y-d)\right] < 0.\end{aligned}$$

It follows that $\frac{\mathrm{d}}{\mathrm{d}\Delta}y^{HB} \propto \frac{\partial}{\partial y}g(y,\Delta)$, which is

$$\begin{aligned}\frac{\partial}{\partial y}g(y,\Delta) &= \frac{\partial}{\partial y}\left[V^H(y) - V^B(y)\right] = \tau(1-\gamma p_H) - c'(y-d) - \left[\tau(1-\gamma p_L) - c'(y-\hat{y})\right] \\ &= \tau(1-\gamma p_H - 1 + \gamma p_L) + c'(y-\hat{y}) - c'(y-d) \\ &= -\tau\gamma(p_H - p_L) + c'(y-\hat{y}) - c'(y-d)\end{aligned}$$

Evaluating at $y = y^{HB}$, $d = d^{HB}$ we get

$$\left.\frac{\partial}{\partial y}g(y,\Delta)\right|_{y=y^{HB}} = -\tau\gamma(p_H - p_L) + c'\left(y^{HB} - \hat{y}\right) - c'\left(y^{HB} - d^{HB}\right).$$

We know that $\tilde{e} = y^{HB} - \hat{y}$ and $e_H = y^{HB} - d^{HB}$, and that $c'(\tilde{e}) - c'(e_H) < 0$ since $e_H > \tilde{e}$ and costs are convex. The expression for the comparative statics is

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\Delta}y^{HB} &= -\frac{\tau\gamma\left[\frac{\partial p_L}{\partial\Delta}(y-\hat{y}) - \frac{\partial p_H}{\partial\Delta}(y-d)\right]}{-\tau\gamma(p_H - p_L) + c'(y^{HB} - \hat{y}) - c'(y^{HB} - d^{HB})} \\ &= -\frac{\tau\gamma\left[\tilde{e}(\Delta) + \alpha e_H(\Delta)\right]}{\tau\gamma(p_H - p_L) + c'(e_H(\Delta)) - c'(\tilde{e}(\Delta))} \\ &= -\frac{\tau\gamma\left[\tilde{e}(\Delta) + \alpha e_H(\Delta)\right]}{c'(e_L(\Delta)) - c'(\tilde{e}(\Delta))},\end{aligned}$$

where the second equality uses the definitions of $\tilde{e}$ and e_H, and the fact that $\frac{\partial p_H}{\partial\Delta} = \alpha$ and $\frac{\partial p_L}{\partial\Delta} = -1$, while the second equality uses the fact that FOCs for the firm deliver $\tau\gamma(p_H - p_L) = c'(e_L) - c'(e_H)$. Because both the numerator and denominator are positive, we obtain $\frac{\mathrm{d}}{\mathrm{d}\Delta}y^{HB} < 0$. The second derivative is

$$
\begin{aligned}
\frac{\mathrm{d}^2\tilde{e}(\Delta)}{\mathrm{d}\Delta^2} &= -\frac{\tau\gamma\left[\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}+\alpha\frac{\mathrm{d}e_H}{\mathrm{d}\Delta}\right]\left[c'(e_L)-c'(\tilde{e})\right]}{\left[c'(e_L)-c'(\tilde{e})\right]^2}+\frac{\tau\gamma\left[\tilde{e}+\alpha e_H\right]\left[c''(e_L)\frac{\mathrm{d}e_L}{\mathrm{d}\Delta}-c''(\tilde{e})\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}\right]}{\left[c'(e_L)-c'(\tilde{e})\right]^2}\\
&= \frac{\tau\gamma\left[\tau\gamma\left(\tilde{e}+\alpha e_H\right)\right]}{\left[c'(e_L)-c'(\tilde{e})\right]^2}-\frac{\tau\gamma\alpha\frac{\mathrm{d}}{\mathrm{d}\Delta}e_H}{\left[c'(e_L)-c'(\tilde{e})\right]}-\frac{\tau\gamma\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}}{\left[c'(e_L)-c'(\tilde{e})\right]}-\frac{\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}\tau\gamma\left[c''(\tilde{e})\left[\tilde{e}+\alpha e_H\right]\right]}{\left[c'(e_L)-c'(\tilde{e})\right]^2}\\
&= -\frac{\tau\gamma\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}}{c'(e_L)-c'(\tilde{e})}-\frac{\tau\gamma\alpha\frac{\mathrm{d}}{\mathrm{d}\Delta}e_H}{c'(e_L)-c'(\tilde{e})}-\frac{\tau\gamma\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}}{c'(e_L)-c'(\tilde{e})}+\frac{\left(\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}\right)^2c''(\tilde{e})}{c'(e_L)-c'(\tilde{e})}\\
&= \frac{\frac{(\tau\gamma\alpha)^2}{c''(e_H)}-\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}\left[2\tau\gamma-\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}c''(\tilde{e})\right]}{c'(e_L)-c'(\tilde{e})}
\end{aligned}
$$

Computations of Theorem 4

Using the fact that we can write $\frac{\mathrm{d}^2\tilde{e}(\Delta)}{\mathrm{d}\Delta^2}$ as a polynomial function of $\frac{\mathrm{d}\tilde{e}(\Delta)}{\mathrm{d}\Delta}$, we rewrite the expression fo $\frac{\mathrm{d}^2R}{\mathrm{d}\Delta^2}$ as a polynomial in $\frac{\mathrm{d}\tilde{e}(\Delta)}{\mathrm{d}\Delta}$

$$
\begin{aligned}
-\lim_{\Delta\to 0}\frac{\mathrm{d}^2R}{\mathrm{d}\Delta^2} &= \lim_{\Delta\to 0}\Bigg(\frac{\frac{(\tau\gamma\alpha)^2}{c''(e_H)}-\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}\left[2\tau\gamma-\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}c''(\tilde{e})\right]}{c'(e_L)-c'(\tilde{e})}\cdot(e_H-\tilde{e})\cdot f(\hat{y}+\tilde{e})-\left(\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}\right)^2f(\hat{y}+\tilde{e})\\
&\quad+\left(\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}\right)^2(e_H-\tilde{e})\cdot f'(\hat{y}+\tilde{e})\\
&\quad-2\frac{\gamma\tau\alpha}{c''(e_H)}\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}f(\hat{y}+\tilde{e})\Bigg)\\
&= \lim_{\Delta\to 0}\Bigg(\frac{(\tau\gamma\alpha)^2}{c''(e_H)}\frac{(e_H-\tilde{e})\cdot f(\hat{y}+\tilde{e})}{c'(e_L)-c'(\tilde{e})}-\frac{\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}2\tau\gamma\cdot(e_H-\tilde{e})\cdot f(\hat{y}+\tilde{e})}{c'(e_L)-c'(\tilde{e})}\\
&\quad+\frac{\left(\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}\right)^2c''(\tilde{e})(e_H-\tilde{e})\cdot f(\hat{y}+\tilde{e})}{c'(e_L)-c'(\tilde{e})}-\left(\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}\right)^2f(\hat{y}+\tilde{e})\\
&\quad+\left(\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}\right)^2(e_H-\tilde{e})\cdot f'(\hat{y}+\tilde{e})-2\frac{\gamma\tau\alpha}{c''(e_H)}\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}f(\hat{y}+\tilde{e})\Bigg)\\
&= \lim_{\Delta\to 0}\left(\left(\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}\right)^2\alpha_2+\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}\alpha_1+\alpha_0\right)
\end{aligned}
$$

where

$$
\alpha_0=\frac{(\tau\gamma\alpha)^2}{c''(e_H)}\frac{(e_H-\tilde{e})\cdot f(\hat{y}+\tilde{e})}{c'(e_L)-c'(\tilde{e})},\ \alpha_1=2\tau\gamma f(\hat{y}+\tilde{e})\left(-\frac{(e_H-\tilde{e})}{c'(e_L)-c'(\tilde{e})}-\frac{\alpha}{c''(e_H)}\right),
$$

$$
\alpha_2=f(\hat{y}+\tilde{e})\left((e_H-\tilde{e})\frac{f'(\hat{y}+\tilde{e})}{f(\hat{y}+\tilde{e})}+c''(\tilde{e})\frac{(e_H-\tilde{e})}{c'(e_L)-c'(\tilde{e})}-1\right).
$$

We further show that

$$\begin{aligned}
\lim_{\Delta\to 0}\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}\alpha_2 &= \lim_{\Delta\to 0} -\frac{\tau\gamma\left[\tilde{e}+\alpha e_H\right]}{c'\left(e_L\right)-c'\left(\tilde{e}\right)} f\left(\hat{y}+\tilde{e}\right)\left[\left(e_H-\tilde{e}\right)\frac{f'\left(\hat{y}+\tilde{e}\right)}{f\left(\hat{y}+\tilde{e}\right)}+c''\left(\tilde{e}\right)\frac{\left(e_H-\tilde{e}\right)}{c'\left(e_L\right)-c'\left(\tilde{e}\right)}-1\right]\\
&= \lim_{\Delta\to 0} -\tau\gamma\left[\tilde{e}+\alpha e_H\right] f\left(\hat{y}+\tilde{e}\right)\left[\frac{\left(e_H-\tilde{e}\right)}{c'\left(e_L\right)-c'\left(\tilde{e}\right)}\frac{f'\left(\hat{y}+\tilde{e}\right)}{f\left(\hat{y}+\tilde{e}\right)}+\frac{c''\left(\tilde{e}\right)\left(e_H-\tilde{e}\right)-c'\left(e_L\right)+c'\left(\tilde{e}\right)}{\left[c'\left(e_L\right)-c'\left(\tilde{e}\right)\right]^2}\right]\\
&= -\tau\gamma e^I\left(1+\alpha\right) f\left(\hat{y}+e^I\right)\left[\lim_{\Delta\to 0}\frac{\left(e_H-\tilde{e}\right)}{c'\left(e_L\right)-c'\left(\tilde{e}\right)}\frac{f'\left(\hat{y}+\tilde{e}\right)}{f\left(\hat{y}+\tilde{e}\right)}+\lim_{\Delta\to 0}\frac{c''\left(\tilde{e}\right)\left(e_H-\tilde{e}\right)-c'\left(e_L\right)+c'\left(\tilde{e}\right)}{\left[c'\left(e_L\right)-c'\left(\tilde{e}\right)\right]^2}\right].
\end{aligned}$$

We investigate the second term in the bracket using L'Hopital

$$\begin{aligned}
\lim_{\Delta\to 0}\frac{c''\left(\tilde{e}\right)\left(e_H-\tilde{e}\right)-c'\left(e_L\right)+c'\left(\tilde{e}\right)}{\left[c'\left(e_L\right)-c'\left(\tilde{e}\right)\right]^2} &= \lim_{\Delta\to 0}\frac{\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}c'''\left(\tilde{e}\right)\left(e_H-\tilde{e}\right)+c''\left(\tilde{e}\right)\left(\frac{\mathrm{d}e_H}{\mathrm{d}\Delta}-\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}\right)-c''\left(e_L\right)\frac{\mathrm{d}e_L}{\mathrm{d}\Delta}+c''\left(\tilde{e}\right)\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}}{2\left[c'\left(e_L\right)-c'\left(\tilde{e}\right)\right]\left(\frac{\mathrm{d}e_L}{\mathrm{d}\Delta}c''\left(e_L\right)-\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}c''\left(\tilde{e}\right)\right)}\\
&= \lim_{\Delta\to 0}\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}\frac{c'''\left(\tilde{e}\right)\left(e_H-\tilde{e}\right)}{2\left[c'\left(e_L\right)-c'\left(\tilde{e}\right)\right]\left(\frac{\mathrm{d}e_L}{\mathrm{d}\Delta}c''\left(e_L\right)-\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}c''\left(\tilde{e}\right)\right)}\\
&+\lim_{\Delta\to 0}\frac{c''\left(\tilde{e}\right)\frac{\mathrm{d}e_H}{\mathrm{d}\Delta}-c''\left(e_L\right)\frac{\mathrm{d}e_L}{\mathrm{d}\Delta}}{2\left[c'\left(e_L\right)-c'\left(\tilde{e}\right)\right]\left(\frac{\mathrm{d}e_L}{\mathrm{d}\Delta}c''\left(e_L\right)-\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}c''\left(\tilde{e}\right)\right)}\\
&= \lim_{\Delta\to 0}\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}\frac{c'''\left(\tilde{e}\right)\left(e_H-\tilde{e}\right)}{2\left[c'\left(e_L\right)-c'\left(\tilde{e}\right)\right]\left(\gamma\tau-\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}c''\left(\tilde{e}\right)\right)}\\
&-\frac{\tau\gamma\left(1+\alpha\right)}{2}\lim_{\Delta\to 0}\frac{1}{\left[c'\left(e_L\right)-c'\left(\tilde{e}\right)\right]\left(\gamma\tau-\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}c''\left(\tilde{e}\right)\right)}\\
&= \lim_{\Delta\to 0}\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}\frac{c'''\left(\tilde{e}\right)\left(e_H-\tilde{e}\right)}{2\left[c'\left(e_L\right)-c'\left(\tilde{e}\right)\right]\left(-\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}c''\left(\tilde{e}\right)\right)}\\
&-\frac{\tau\gamma\left(1+\alpha\right)}{2}\lim_{\Delta\to 0}\frac{1}{\tau\gamma\left[\tilde{e}+\alpha e_H\right]c''\left(\tilde{e}\right)}\\
&= -\lim_{\Delta\to 0}\frac{c'''\left(\tilde{e}\right)\left(e_H-\tilde{e}\right)}{2\left[c'\left(e_L\right)-c'\left(\tilde{e}\right)\right]c''\left(\tilde{e}\right)}-\frac{1}{2e^I c''\left(e^I\right)}\\
&= -\frac{\frac{c'''\left(e^I\right)}{c''\left(e^I\right)}}{2c''\left(e^I\right)}-\frac{1}{2e^I c''\left(e^I\right)} = -\frac{\frac{c'''\left(e^I\right)}{c''\left(e^I\right)}e^I+1}{2e^I c''\left(e^I\right)},
\end{aligned}$$

where in the fourth equality we exploited that $i)$ $\lim_{\Delta\to 0}\frac{h(\Delta)}{2\left[c'\left(e_L\right)-c'\left(\tilde{e}\right)\right]\left(\gamma\tau-\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}c''\left(\tilde{e}\right)\right)} = \lim_{\Delta\to 0} -\frac{h(\Delta)}{2\left[c'\left(e_L\right)-c'\left(\tilde{e}\right)\right]\left(\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}c''\left(\tilde{e}\right)\right)}$ for any h such that $\lim_{\Delta\to 0} h\left(\Delta\right)$ is finite and $ii)$ the definition of $\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}$. The sixth equality uses again that $\lim_{\Delta\to 0}\frac{\left(e_H-\tilde{e}\right)}{c'\left(e_L\right)-c'\left(\tilde{e}\right)} = \frac{1}{c''\left(e^I\right)}$.

So, we have

$$\begin{aligned}
\lim_{\Delta\to 0}\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}\alpha_2 &= -\tau\gamma e^I\left(1+\alpha\right) f\left(\hat{y}+e^I\right)\left[\frac{1}{c''\left(e^I\right)}\frac{f'\left(\hat{y}+e^I\right)}{f\left(\hat{y}+e^I\right)}-\frac{\frac{c'''\left(e^I\right)}{c''\left(e^I\right)}e^I+1}{2e^I c''\left(e^I\right)}\right]\\
&= -\tau\gamma\left(1+\alpha\right) f\left(\hat{y}+e^I\right)\left[\frac{2e^I f'\left(\hat{y}+e^I\right)-\frac{c'''\left(e^I\right)}{c''\left(e^I\right)}e^I f\left(\hat{y}+e^I\right)-f\left(\hat{y}+e^I\right)}{2c''\left(e^I\right) f\left(\hat{y}+e^I\right)}\right],
\end{aligned}$$

which is finite by boundedness away from zero of c'' and f. Hence, according to Lemma 7, concavity of the revenue is determined by the sign of

$$
\begin{aligned}
\lim_{\Delta\to 0}\alpha_1+\lim_{\Delta\to 0}\frac{\mathrm{d}\tilde{e}}{\mathrm{d}\Delta}\alpha_2 &= -f\left(\hat{y}+\tilde{e}\right)\frac{2\tau\gamma\left(1+\alpha\right)}{c''\left(e^I\right)}\\
&\quad -\tau\gamma\left(1+\alpha\right)f\left(\hat{y}+e^I\right)\left[\frac{2e^If'\left(\hat{y}+e^I\right)-\frac{c'''\left(e^I\right)}{c''\left(e^I\right)}e^If\left(\hat{y}+e^I\right)-f\left(\hat{y}+e^I\right)}{2c''\left(e^I\right)f\left(\hat{y}+e^I\right)}\right]\\
&= -f\left(\hat{y}+e^I\right)\frac{\tau\gamma\left(1+\alpha\right)}{c''\left(e^I\right)}\left[2+\left[\frac{2e^If'\left(\hat{y}+e^I\right)-\frac{c'''\left(e^I\right)}{c''\left(e^I\right)}e^If\left(\hat{y}+e^I\right)-f\left(\hat{y}+e^I\right)}{2f\left(\hat{y}+e^I\right)}\right]\right]\\
&\propto -\frac{3f\left(\hat{y}+e^I\right)+2e^If'\left(\hat{y}+e^I\right)-\frac{c'''\left(e^I\right)}{c''\left(e^I\right)}e^If\left(\hat{y}+e^I\right)}{2f\left(\hat{y}+e^I\right)}\\
&= -\frac{3}{2}-e^I\frac{f'\left(\hat{y}+e^I\right)}{f\left(\hat{y}+e^I\right)}+\frac{e^I}{2}\frac{c'''\left(e^I\right)}{c''\left(e^I\right)}.
\end{aligned}
$$

S.2 Additional Proofs

We collect here the proofs of the Lemmata used in the discussion of model extensions.

Proof of Lemma 9

Point 1 is intuitive: if firms manipulate the threshold, they do so to bunch at that level. Indeed, if they declared less than $\hat{y}$, then they would face audit probability p_H, which they could achieve by keeping $\hat{y}_0$ and saving $\tilde{c}\left(\hat{y}_0-\hat{y}\right)>0$. On the other hand, if they declared $d>\hat{y}$, then by decreasing their declared tax base to the threshold $\hat{y}$ they would face the same probability p_L and save $\tilde{c}\left(\hat{y}_0-\hat{y}\right)-\tilde{c}\left(\hat{y}_0-d\right)>0$.

Point 2 states that firms who can choose the unconstrained optimal evasion e_L will not manipulate the threshold. This is because the firm's value when facing p_L is maximized at $d=y-e_L$, which is also feasible if $y>\hat{y}_0+e_L$. Hence, firms will optimally save the threshold manipulation cost.

Point 3 states that if a firm optimally chooses to declare below $\hat{y}$ and face p_H, then it will also choose $\hat{y}=\hat{y}_0$. This is because setting $\hat{y}\neq\hat{y}_0$ is costly and does not increase the firm's value, so it will never be optimal. These firms will optimally declare either e_H or 0 (if $y<e_H$), as proved in Proposition 1.

Proof of Lemma 10

The fact that $\hat{y}\left(y\right)$ increases in y is easily derived from the FOC in (B.2) by exploiting the convexity of $c\left(\cdot\right)$. By Lemma 9, we know that every firm with $y>\hat{y}_0+e_L$ will optimally declare $y-e_L$. The optimal manipulation of the firm with $y=\hat{y}_0+e_L$ is zero. This is because if it declared $d\left(y\right)=\hat{y}_0$, the FOC would be satisfied as $\tau\left(1-\gamma\cdot p_L\right)=c'\left(e_L\right)$ and the level of evasion e_L would allow the firm to face probability p_L. If this firm manipulated the first Euro of threshold, then it would not satisfy the optimality condition anymore. Therefore, the firm chooses to not manipulate and save that cost. Since $\hat{y}$ is increasing in y, and since $y=\hat{y}_0+e_L$ declares $\hat{y}_0$, then all $y<\hat{y}_0+e_L$ will declare below $\hat{y}_0$. Therefore, at the optimum the only way to satisfy (B.2) is to choose $d\left(y\right)=y-\hat{y}<e_L$ and manipulate the threshold to remain above it and face probability p_L.

Limit results are immediate to prove. If the cost of threshold manipulation was infinite, the firm would not manipulate. This is the case investigated in our baseline model. If instead manipulation was free, the FOC in (B.2) would collapse into the one in (3) evaluated at $p = p_L$. As a result, all firms will declare $d(y) = y - e_L$ and set $\hat{y}(y) = y - e_L$.

The last result in the Lemma is a straightforward application of the envelope theorem on (B.1). We have that

$$\frac{\partial V^B(y)}{\partial y} = 1 - \tau\gamma p_L - c'(y - \hat{y}(y)) = 1 - \tau + \tilde{\kappa}\tilde{c}'(\hat{y}_0 - \hat{y}(y)) > 0,$$

where the second line substituted the FOC for $\hat{y}$. Moreover, we have

$$\frac{\partial V^B(y)}{\partial \tilde{\kappa}} = -\tilde{c}'(\hat{y}_0 - \hat{y}(y)) < 0.$$

Proof of Lemma 11

We prove the comparative statics showing that $\frac{\mathrm{d}}{\mathrm{d}\tilde{\kappa}}\tilde{y}^{HB} > 0$, $\frac{\mathrm{d}}{\mathrm{d}\tilde{\kappa}}\tilde{y}^{BL} = 0$, while $\frac{\mathrm{d}}{\mathrm{d}\Delta}\tilde{y}^{HB} < 0$, $\frac{\mathrm{d}}{\mathrm{d}\Delta}\tilde{y}^{BL} > 0$. Recall that $\tilde{y}^{BL} = \hat{y}_0 + e_L$ which is independent of $\tilde{\kappa}$ and increasing in Δ as $\frac{\mathrm{d}}{\mathrm{d}\Delta}e_L > 0$. To prove the results for the lower bound of the bunching region we differentiate the indifference condition $V^H(\tilde{y}^{HB}) = V^B(\tilde{y}^{HB})$ which implicitly defines it. For a generic parameter ν, it holds

$$\frac{\mathrm{d}}{\mathrm{d}\nu}\tilde{y}^{HB} = \frac{\frac{\partial}{\partial\nu}V^B(\tilde{y}^{HB}) - \frac{\partial}{\partial\nu}V^H(\tilde{y}^{HB})}{\frac{\partial}{\partial y}V^H(\tilde{y}^{HB}) - \frac{\partial}{\partial y}V^B(\tilde{y}^{HB})}.$$

Recall that $V^H(y) = \max_d y - \tau d - \tau\gamma(\mu + \alpha\Delta)(y - d) - c(y - d)$ while $V^B(y) = \max_{\hat{y}} y - \tau\hat{y} - \tau\gamma(\mu - \Delta)(y - \hat{y})^+ - c(y - \hat{y}) - \tilde{\kappa}\tilde{c}(\hat{y}_0 - \hat{y})$. By envelope,

$$\frac{\partial}{\partial y}V^B(y) = 1 - \tau + \tilde{\kappa}\tilde{c}'(\hat{y}_0 - \hat{y}(y)), \quad \frac{\partial}{\partial y}V^H(y) = 1 - \tau,$$

where we substitute the FOCs $\tau(1 - \gamma(\mu - \Delta)) = c'(\check{e}) + \tilde{\kappa}\tilde{c}'(\hat{y}_0 - \tilde{y}^{HB} + \check{e})$ and $\tau(1 - \gamma(\mu + \alpha\Delta)) = c'(e_H)$, respectively. Moreover,

$$\frac{\partial}{\partial\tilde{\kappa}}V^B(y) = -\tilde{c}'(\hat{y}_0 - \hat{y}(y)), \quad \frac{\partial}{\partial\tilde{\kappa}}V^H(y) = 0, \quad \frac{\partial}{\partial\Delta}V^H(y) = -\tau\gamma\alpha e_H, \quad \frac{\partial}{\partial\Delta}V^B(y) = \tau\gamma(y - \hat{y}(y)).$$

Substituting the relevant partial derivatives and evaluating them at $y = \tilde{y}^{HB}$, we finally obtain

$$\frac{\mathrm{d}}{\mathrm{d}\tilde{\kappa}}\tilde{y}^{HB} = \frac{-\tilde{c}'(\hat{y}_0 - \hat{y})}{1 - \tau - [1 - \tau + \tilde{\kappa}\tilde{c}'(\hat{y}_0 - \tilde{y}^{HB} + \check{e})]} = \frac{\tilde{c}'(\hat{y}_0 - \hat{y})}{\tilde{\kappa}\tilde{c}'(\hat{y}_0 - \tilde{y}^{HB} + \check{e})} > 0,$$

and

$$\frac{\mathrm{d}}{\mathrm{d}\Delta}\tilde{y}^{HB} = \frac{\tau\gamma\check{e} + \tau\gamma\alpha e_H}{1 - \tau - ([1 - \tau\gamma(\mu - \Delta)] - c'(y - \hat{y}(y)))} = -\frac{\tau\gamma(\check{e} + \alpha e_H)}{\tilde{\kappa}\tilde{c}'(\hat{y}_0 - \hat{y}(\tilde{y}^{HB}))} < 0.$$

This proves the desired result.

Proof of Lemma 12

The fact that $V^H(0) < V^B(0)$ when $\tilde{\kappa} = 0$ follows from the observation that when manipulation is free $V^B(y) = V^L(y) > V^H(y)$, where the equality follows from Lemma 10 and the inequality holds for any y since V^L dominates point-wise. Therefore, all firms will decide to manipulate the threshold to declare $d(y) = y - e_L$ and will all bunch at their individual threshold. The result for $\tilde{\kappa} = \infty$ instead is straightforward from the fact that with an infinite cost of manipulation the problem collapses to our baseline model. In this case, as showed in Proposition 1, the marginal buncher has income equal to $\hat{y}_0 + \tilde{e}$.

The second part of the Lemma follows from noticing that the indirect utility from declaring in $\mathcal{H}$ is larger than the one from manipulating if $\Delta = 0$. In this case $p_H, p_L = p$ and we have

$$V^B(y) = \max_d y - \tau d - \tau\gamma p (y-d)^+ - c(y-d) - \tilde{\kappa}\tilde{c}(\hat{y}_0 - d),$$

$$V^H(y) = \max_d y - \tau d - \tau\gamma p (y-d) - c(y-d).$$

Hence, it is clear that $V^H(y) \geq V^B(y)$ since the two are the same up to the term $\tilde{\kappa}\tilde{c}(\hat{y}_0 - d)$, with equality holding only if the maximizer of V^H is $d^\star = \hat{y}_0$. Hence, $\tilde{y}^{HB}$ is such that $\hat{y}\left(\tilde{y}^{HB}\right) = \hat{y}_0$. Using i) the continuity of $\hat{y}(\Delta)$, $\tilde{y}(\Delta)$, ii) the fact that $\hat{y}(y) < \hat{y}_0$ (as proved in Lemma 10), iii) $\tilde{c}'(0) = 0$, we conclude that the denominator in $\frac{\mathrm{d}}{\mathrm{d}\Delta}\tilde{y}^{HB}$ (derived in Lemma 11) converges to 0 from above, while the numerator converges to a constant.[47] It follows that $\lim_{\Delta\to 0}\frac{\mathrm{d}}{\mathrm{d}\Delta}\tilde{y}^{HB} = -\infty$, which proves the result.

Proof of Lemma 13

Using the expression for $\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}$ derived in the proof to Lemma 11, we get

$$\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}(\check{e} - e_H) = -\frac{\tau\gamma(\check{e} + \alpha e_H)}{\tilde{\kappa}\tilde{c}'\left(\hat{y}_0 - \hat{y}\left(\tilde{y}^{HB}\right)\right)}(\check{e} - e_H).$$

Notice that $\lim_{\Delta\to 0}\hat{y}\left(\tilde{y}^{HB}\right) = \hat{y}_0$, so since $\tilde{c}(0) = 0$ the limit is indeterminate and we need to apply de L'Hopital's rule. We therefore have

$$\lim_{\Delta\to 0}\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}(\check{e} - e_H) = \lim_{\Delta\to 0} -\frac{\tau\gamma(\check{e} + \alpha e_H)}{\tilde{\kappa}\tilde{c}'\left(\hat{y}_0 - \hat{y}\left(\tilde{y}^{HB}\right)\right)}(\check{e} - e_H) = -\frac{\tau\gamma(1+\alpha)e^I(\mu)}{\tilde{\kappa}}\cdot\lim_{\Delta\to 0}\frac{\frac{\mathrm{d}}{\mathrm{d}\Delta}(\check{e} - e_H)}{\frac{\mathrm{d}}{\mathrm{d}\Delta}\tilde{c}'\left(\hat{y}_0 - \hat{y}\left(\tilde{y}^{HB}\right)\right)}$$

where we exploit the fact that $\lim_{\Delta\to 0}\check{e} = \lim_{\Delta\to 0} e_H = e^I(\mu)$. Also, notice that

$$\frac{\mathrm{d}\hat{y}\left(\tilde{y}^{HB}\right)}{\mathrm{d}\Delta} = \frac{\partial\hat{y}\left(\tilde{y}^{HB}(\Delta), \Delta\right)}{\partial\Delta} + \frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}\frac{\partial\hat{y}\left(\tilde{y}^{HB}(\Delta), \Delta\right)}{\partial y},$$

$$\frac{\mathrm{d}\check{e}}{\mathrm{d}\Delta} = \frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta} - \frac{\mathrm{d}\hat{y}\left(\tilde{y}^{HB}\right)}{\mathrm{d}\Delta} = \frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}\left(1 - \frac{\partial\hat{y}\left(\tilde{y}^{HB}(\Delta), \Delta\right)}{\partial y}\right) - \frac{\partial\hat{y}\left(\tilde{y}^{HB}(\Delta), \Delta\right)}{\partial\Delta}.$$

[47] Notice that the result hinges on the assumption $\tilde{c}'(0) = 0$, which is the interesting case where threshold manipulation is active even local to a flat rule. If this was not the case, the results of the main analysis would go through unchanged as any threshold manipulation would be suboptimal when $\Delta \to 0$.

Denote $\delta_1 = \frac{\partial \hat{y}\left(\tilde{y}^{HB}(\Delta),\Delta\right)}{\partial y}$ and $\delta_0 = \frac{\partial \hat{y}\left(\tilde{y}^{HB}(\Delta),\Delta\right)}{\partial \Delta}$, we have $\frac{\mathrm{d}\hat{y}\left(\tilde{y}^{HB}\right)}{\mathrm{d}\Delta} = \delta_0 + \delta_1 \frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}$, $\frac{\mathrm{d}\check{e}}{\mathrm{d}\Delta} = (1-\delta_1)\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta} - \delta_0$. So the limit becomes

$$\lim_{\Delta\to 0} \frac{\frac{\mathrm{d}}{\mathrm{d}\Delta}\check{e} - \frac{\mathrm{d}}{\mathrm{d}\Delta}e_H}{-\frac{\mathrm{d}\hat{y}\left(\tilde{y}^{HB}\right)}{\mathrm{d}\Delta}\tilde{c}''\left(\hat{y}_0 - \hat{y}\left(\tilde{y}^{HB}\right)\right)} = \frac{(1-\delta_1)\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta} - \delta_0 + \frac{\gamma\tau\alpha}{c''(e^I(\mu))}}{-\left(\delta_0 + \delta_1 \frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}\right)\tilde{c}''(0)} \propto -\frac{(1-\delta_1)\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}}{\delta_1 \tilde{c}''(0)\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}} = -\frac{1-\delta_1}{\delta_1 \tilde{c}''(0)}$$

where the last line focuses only on the terms in $\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}$ since $\lim_{\Delta\to 0}\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}$ is not finite. To derive $\delta_1 = \frac{\partial \hat{y}\left(\tilde{y}^{HB}(\Delta),\Delta\right)}{\partial y}$, we differentiate the FOC for $\hat{y}$ and get

$$\frac{\partial \hat{y}\left(\tilde{y}^{HB}\right)}{\partial y} = \frac{c''\left(y - \hat{y}\left(\tilde{y}^{HB}\right)\right)}{c''\left(y - \hat{y}\left(\tilde{y}^{HB}\right)\right) + \tilde{\kappa}\tilde{c}''\left(\hat{y}_0 - \hat{y}\left(\tilde{y}^{HB}\right)\right)} < 1.$$

We now notice that $\lim_{\Delta\to 0}\delta_1 = \lim_{\Delta\to 0}\frac{\partial \hat{y}\left(\tilde{y}^{HB}(\Delta),\Delta\right)}{\partial y} = \lim_{\Delta\to 0}\frac{c''\left(\tilde{y}^{HB} - \hat{y}_0\right)}{c''\left(\tilde{y}^{HB} - \hat{y}_0\right) + \tilde{\kappa}\tilde{c}''(0)}$ and $\lim_{\Delta\to 0}\left(\tilde{y}^{HB} - \hat{y}_0\right) = e^I(\mu)$. So by replacing this limit, as well as the expression for $\frac{\mathrm{d}}{\mathrm{d}\Delta}e_H$ in the formula for $\lim_{\Delta\to 0}\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}(\check{e} - e_H)$ that we derived above, we get the desired result

$$\lim_{\Delta\to 0}\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}(\check{e} - e_H) = -\frac{\tau\gamma(1+\alpha)e^I(\mu)}{\tilde{\kappa}} \cdot \lim_{\Delta\to 0}\frac{\frac{\mathrm{d}}{\mathrm{d}\Delta}(\check{e} - e_H)}{\frac{\mathrm{d}}{\mathrm{d}\Delta}\tilde{c}'\left(\hat{y}_0 - \hat{y}\left(\tilde{y}^{HB}\right)\right)} - \frac{\tau\gamma(1+\alpha)e^I(\mu)}{\tilde{\kappa}} \cdot \frac{1-\delta_1}{\delta_1\tilde{c}''(0)} = \frac{\tau\gamma(1+\alpha)e^I(\mu)}{c''(e^I(\mu))}.$$

Notice that this expression is independent of the manipulation cost $\tilde{c}$ since in the limit there is no manipulation and the marginal cost for the first unit of threshold manipulation is zero. This expression also coincides with the one derived in the model without threshold manipulation.

To prove the second result in the Lemma, we proceed by differentiating the expression for $\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}$ derived in Lemma 11. We have

$$\begin{aligned}
\frac{\mathrm{d}^2}{\mathrm{d}\Delta^2}\tilde{y}^{HB} &= -\frac{\tau\gamma}{\tilde{\kappa}}\frac{\frac{\mathrm{d}(\check{e}+\alpha e_H)}{\mathrm{d}\Delta}\tilde{c}'\left(\hat{y}_0 - \hat{y}\left(\tilde{y}^{HB}\right)\right) + \frac{\mathrm{d}\hat{y}\left(\tilde{y}^{HB}\right)}{\mathrm{d}\Delta}\tilde{c}''\left(\hat{y}_0 - \hat{y}\left(\tilde{y}^{HB}\right)\right)(\check{e}+\alpha e_H)}{\tilde{c}'\left(\hat{y}_0 - \hat{y}\left(\tilde{y}^{HB}\right)\right)^2} \\
&= \frac{\frac{1}{\tilde{\kappa}}\left(\frac{(\tau\gamma\alpha)^2}{c''(e_H)} + \tau\gamma\frac{\partial \hat{y}\left(\tilde{y}^{HB}\right)}{\partial\Delta}\right) + \frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}\frac{\tilde{c}''\left(\hat{y}_0 - \hat{y}\left(\tilde{y}^{HB}\right)\right)}{\tau\gamma}\frac{\partial \hat{y}\left(\tilde{y}^{HB}\right)}{\partial\Delta}\left[2\tau\gamma - \frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}c''\left(\tilde{y}^{HB} - \hat{y}\left(\tilde{y}^{HB}\right)\right)\right]}{\tilde{c}'\left(\hat{y}_0 - \hat{y}\left(\tilde{y}^{HB}\right)\right)},
\end{aligned}$$

which conforms with the statement letting

$$a_0 = \frac{1}{\tilde{\kappa}}\left(\frac{(\tau\gamma\alpha)^2}{c''(e_H)} + \tau\gamma\frac{\partial \hat{y}\left(\tilde{y}^{HB}\right)}{\partial\Delta}\right),\ a_1 = 2\tilde{c}''\left(\hat{y}_0 - \hat{y}\left(\tilde{y}^{HB}\right)\right)\frac{\partial \hat{y}\left(\tilde{y}^{HB}\right)}{\partial\Delta},$$

$$a_2 = -\frac{\tilde{c}''\left(\hat{y}_0 - \hat{y}\left(\tilde{y}^{HB}\right)\right)}{\tau\gamma}\frac{\partial \hat{y}\left(\tilde{y}^{HB}\right)}{\partial\Delta}c''\left(\tilde{y}^{HB} - \hat{y}\left(\tilde{y}^{HB}\right)\right).$$

Proof of Lemma 14

Given the following expression for the second derivative of revenues

$$
\begin{aligned}
\lim_{\Delta\to 0}\frac{\mathrm{d}^2}{\mathrm{d}\Delta^2}R(\Delta) \propto \left(\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}\right)^2 & \left[f'\left(\tilde{y}^{HB}\right)(\check{e}-e_H)+f\left(\tilde{y}^{HB}\right)\left(1-\frac{\partial\hat{y}\left(\tilde{y}^{HB}\right)}{\partial y}\right)\right. \\
& \left. -f\left(\tilde{y}^{HB}\right)(\check{e}-e_H)\frac{\tilde{c}''\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)}{\tau\gamma\tilde{c}'\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)}\frac{\partial\hat{y}\left(\tilde{y}^{HB}\right)}{\partial\Delta}c''\left(\tilde{y}^{HB}-\hat{y}\left(\tilde{y}^{HB}\right)\right)\right] \\
& -\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}\left[-2\frac{\tilde{c}''\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)}{\tilde{c}'\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)}\frac{\partial\hat{y}\left(\tilde{y}^{HB}\right)}{\partial\Delta}f\left(\tilde{y}^{HB}\right)(\check{e}-e_H)\right. \\
& \left. +\frac{\mathrm{d}e_H}{\mathrm{d}\Delta}f\left(\tilde{y}^{HB}\right)+\left(\frac{\mathrm{d}e_H}{\mathrm{d}\Delta}+2\frac{\partial\hat{y}\left(\tilde{y}^{HB}\right)}{\partial\Delta}\right)f\left(\tilde{y}^{HB}\right)\right] \\
& +\frac{1}{\tilde{\kappa}}\frac{\left(\frac{(\tau\gamma\alpha)^2}{c''(e_H)}+\tau\gamma\frac{\partial\hat{y}\left(\tilde{y}^{HB}\right)}{\partial\Delta}\right)}{\tilde{c}'\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)}f\left(\tilde{y}^{HB}\right)(\check{e}-e_H),
\end{aligned}
$$

the αs limits are

$$
\begin{aligned}
\lim_{\Delta\to 0}\alpha_0 &= -\frac{1}{\tilde{\kappa}}\left(\frac{(\tau\gamma\alpha)^2}{c''\left(e^I(\mu)\right)}+\tau\gamma\frac{\partial\hat{y}\left(\tilde{y}^{HB}\right)}{\partial\Delta}\right)\frac{\tilde{\kappa}}{c''\left(e^I(\mu)\right)}f\left(\hat{y}_0+e^I(\mu)\right) \\
&= -\left(\frac{(\tau\gamma\alpha)^2}{c''\left(e^I(\mu)\right)}-\frac{(\tau\gamma)^2}{c''\left(e^I(\mu)\right)+\tilde{\kappa}\tilde{c}''(0)}\right)\frac{f\left(\hat{y}_0+e^I(\mu)\right)}{c''\left(e^I(\mu)\right)}
\end{aligned}
$$

$$
\begin{aligned}
\lim_{\Delta\to 0}\alpha_1 &= 2\frac{\tilde{\kappa}\tilde{c}''(0)}{c''\left(e^I(\mu)\right)}\frac{\gamma\tau}{c''\left(e^I(\mu)\right)+\tilde{\kappa}\tilde{c}''(0)}f\left(\hat{y}_0+e^I(\mu)\right)+\frac{\tau\gamma\alpha}{c''\left(e^I(\mu)\right)}f\left(\hat{y}_0+e^I(\mu)\right) \\
&+\left(\frac{\tau\gamma\alpha}{c''\left(e^I(\mu)\right)}+2\frac{\gamma\tau}{c''\left(e^I(\mu)\right)+\tilde{\kappa}\tilde{c}''(0)}\right)f\left(\hat{y}_0+e^I(\mu)\right) \\
&= \frac{2\gamma\tau}{c''\left(e^I(\mu)\right)+\tilde{\kappa}\tilde{c}''(0)}f\left(\hat{y}_0+e^I(\mu)\right)\left[1+\frac{\tilde{\kappa}\tilde{c}''(0)}{c''\left(e^I(\mu)\right)}\right]+2\frac{\tau\gamma\alpha}{c''\left(e^I(\mu)\right)}f\left(\hat{y}_0+e^I(\mu)\right)>0
\end{aligned}
$$

$$
\begin{aligned}
\lim_{\Delta\to 0}\alpha_2 &= f\left(\hat{y}_0+e^I(\mu)\right)\lim_{\Delta\to 0}\left[\left(1-\frac{\partial\hat{y}\left(\tilde{y}^{HB}\right)}{\partial y}\right)\right. \\
&\left. -(\check{e}-e_H)\frac{\tilde{c}''\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)}{\tau\gamma\tilde{c}'\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)}\frac{\partial\hat{y}\left(\tilde{y}^{HB}\right)}{\partial\Delta}c''\left(\tilde{y}^{HB}-\hat{y}\left(\tilde{y}^{HB}\right)\right)\right] \\
&= f\left(\hat{y}_0+e^I(\mu)\right)\left[\frac{\tilde{\kappa}\tilde{c}''(0)}{c''\left(e^I(\mu)\right)+\tilde{\kappa}\tilde{c}''(0)}-\tilde{\kappa}\frac{\tilde{c}''(0)}{c''\left(e^I(\mu)\right)+\tilde{\kappa}\tilde{c}''(0)}\right]=0.
\end{aligned}
$$

where we used the fact that $\lim_{\Delta\to 0}\frac{\frac{\mathrm{d}}{\mathrm{d}\Delta}(\check{e}-e_H)}{\frac{\mathrm{d}}{\mathrm{d}\Delta}\tilde{c}'(\hat{y}_0-\hat{y}(\tilde{y}^{HB}))}=-\frac{\tilde{\kappa}}{c''(e^I(\mu))}$.

Proof of Lemma 15

We have

$$\lim_{\Delta\to 0}\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}\alpha_2=\lim_{\Delta\to 0}\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}\Bigg[f'\left(\tilde{y}^{HB}\right)(\check{e}-e_H)+f\left(\tilde{y}^{HB}\right)\left(1-\frac{\partial\hat{y}\left(\tilde{y}^{HB}\right)}{\partial y}\right)$$

$$-f\left(\tilde{y}^{HB}\right)(\check{e}-e_H)\frac{\tilde{c}''\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)}{\tau\gamma\tilde{c}'\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)}\frac{\partial\hat{y}\left(\tilde{y}^{HB}\right)}{\partial\Delta}c''\left(\tilde{y}^{HB}-\hat{y}\left(\tilde{y}^{HB}\right)\right)\Bigg]$$

$$=\frac{\tau\gamma(1+\alpha)e^I(\mu)}{c''\left(e^I(\mu)\right)}f'\left(\hat{y}_0+e^I(\mu)\right)$$

$$-f\left(\hat{y}_0+e^I(\mu)\right)\lim_{\Delta\to 0}\frac{\tau\gamma(\check{e}+\alpha e_H)}{\tilde{\kappa}\tilde{c}'\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)}\left[\frac{\tilde{c}''\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)\left[\tilde{\kappa}+\frac{(\check{e}-e_H)}{\tilde{c}'\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)}c''\left(\tilde{y}^{HB}-\hat{y}\left(\tilde{y}^{HB}\right)\right)\right]}{c''\left(\tilde{y}^{HB}-\hat{y}\left(\tilde{y}^{HB}\right)\right)+\tilde{\kappa}\tilde{c}''\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)}\right]$$

where we used $\lim_{\Delta\to 0}\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}\left(\check{e}(\Delta)-e_H(\Delta)\right)=\frac{\tau\gamma(1+\alpha)e^I(\mu)}{c''(e^I(\mu))}$ from Lemma 13. We now focus on the second term of the above expression

$$\lim_{\Delta\to 0}\frac{\tau\gamma(\check{e}+\alpha e_H)}{\tilde{\kappa}\tilde{c}'\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)}\left[\frac{\tilde{c}''\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)\left[\tilde{\kappa}+\frac{(\check{e}-e_H)}{\tilde{c}'\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)}c''\left(\tilde{y}^{HB}-\hat{y}\left(\tilde{y}^{HB}\right)\right)\right]}{c''\left(\tilde{y}^{HB}-\hat{y}\left(\tilde{y}^{HB}\right)\right)+\tilde{\kappa}\tilde{c}''\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)}\right]=$$

$$\frac{\tau\gamma(1+\alpha)e^I(\mu)\tilde{c}''(0)}{c''\left(e^I(\mu)\right)+\tilde{\kappa}\tilde{c}''(0)}\lim_{\Delta\to 0}\Bigg[\frac{\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}\left[-\tilde{\kappa}\delta_1\tilde{c}''\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)+(1-\delta_1)c''\left(\tilde{y}^{HB}-\hat{y}\left(\tilde{y}^{HB}\right)\right)+(1-\delta_1)(\check{e}-e_H)c'''\left(\tilde{y}^{HB}-\hat{y}\left(\tilde{y}^{HB}\right)\right)\right]}{-2\tilde{\kappa}\tilde{c}''\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)\left(\delta_0+\delta_1\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}\right)\tilde{c}'\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)}$$

$$+\frac{-\tilde{\kappa}\delta_0\tilde{c}''\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)+\left(-\delta_0+\frac{\gamma\tau\alpha}{c''(e_H)}\right)c''\left(\tilde{y}^{HB}-\hat{y}\left(\tilde{y}^{HB}\right)\right)-(\check{e}-e_H)\delta_0c'''\left(\tilde{y}^{HB}-\hat{y}\left(\tilde{y}^{HB}\right)\right)}{-2\tilde{\kappa}\tilde{c}''\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)\left(\delta_0+\delta_1\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}\right)\tilde{c}'\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)}\Bigg]$$

we show that both the numerator and denominator inside the bracket converge to a number. The numerator of the first term is

$$\lim_{\Delta\to 0}\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}\left[-\tilde{\kappa}\delta_1\tilde{c}''\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)+(1-\delta_1)c''\left(\tilde{y}^{HB}-\hat{y}\left(\tilde{y}^{HB}\right)\right)+(1-\delta_1)(\check{e}-e_H)c'''\left(\tilde{y}^{HB}-\hat{y}\left(\tilde{y}^{HB}\right)\right)\right]=$$

$$\lim_{\Delta\to 0}\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}(\check{e}-e_H)(1-\delta_1)c'''\left(\tilde{y}^{HB}-\hat{y}\left(\tilde{y}^{HB}\right)\right)=\frac{\tau\gamma(1+\alpha)e^I(\mu)}{c''\left(e^I(\mu)\right)}\frac{\tilde{\kappa}\tilde{c}''(0)}{c''\left(e^I(\mu)\right)+\tilde{\kappa}\tilde{c}''(0)}c'''\left(e^I(\mu)\right).$$

The numerator of the second term is

$$\lim_{\Delta\to 0}\left[-\tilde{\kappa}\delta_0\tilde{c}''\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)+\left(-\delta_0+\frac{\gamma\tau\alpha}{c''(e_H)}\right)c''\left(y-\hat{y}\left(\tilde{y}^{HB}\right)\right)-(\check{e}-e_H)\delta_0c'''\left(y-\hat{y}\left(\tilde{y}^{HB}\right)\right)\right]=\gamma\tau(1+\alpha)$$

The denominator is

$$\lim_{\Delta\to 0}-2\tilde{\kappa}\tilde{c}''\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)\left(\delta_0+\delta_1\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}\right)\tilde{c}'\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)=-2\tilde{\kappa}\tilde{c}''(0)\lim_{\Delta\to 0}\delta_1\frac{\mathrm{d}\tilde{y}^{HB}}{\mathrm{d}\Delta}\tilde{c}'\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)$$

$$=\frac{2\tau\gamma(1+\alpha)e^I(\mu)\tilde{c}''(0)c''\left(e^I(\mu)\right)}{c''\left(e^I(\mu)\right)+\tilde{\kappa}\tilde{c}''(0)}.$$

Putting everything together

$$\lim_{\Delta\to 0}\frac{\tau\gamma\left(\check{e}+\alpha e_H\right)}{\tilde{\kappa}\tilde{c}'\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)}\left[\frac{\tilde{c}''\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)\left[\tilde{\kappa}+\frac{(\check{e}-e_H)}{\tilde{c}'(\hat{y}_0-\hat{y}(\tilde{y}^{HB}))}c''\left(y-\hat{y}\left(\tilde{y}^{HB}\right)\right)\right]}{c''\left(\tilde{y}^{HB}-\hat{y}\left(\tilde{y}^{HB}\right)\right)+\tilde{\kappa}\tilde{c}''\left(\hat{y}_0-\hat{y}\left(\tilde{y}^{HB}\right)\right)}\right]$$

$$=\tau\gamma\left(1+\alpha\right)\left[\frac{\frac{c'''\left(e^I(\mu)\right)}{c''(e^I(\mu))}\frac{\tilde{\kappa}\tilde{c}''(0)e^I(\mu)}{c''(e^I(\mu))+\tilde{\kappa}\tilde{c}''(0)}+1}{2c''\left(e^I\left(\mu\right)\right)}\right].$$

This gives the desired expression.

S.3 Additional Theoretical Results

A general test for the desirability of disclosed rules

We consider a test involving a local perturbation of the policy, which yields the following result.

Theorem 16. *If*

$$e^I\left(\mu\right)>\frac{1-F\left(\hat{y}+e^I\left(\mu\right)\right)}{f\left(\hat{y}+e^I\left(\mu\right)\right)} \tag{S.3.1}$$

then a threshold-based audit rule exists that outperforms a flat audit rule by generating higher revenues with a smaller budget.

Proof. Putting the results of Lemmata 5 and 6 together, we obtain

$$\begin{aligned}\lim_{\Delta\to 0}R'\left(\Delta\right)&=\frac{\tau\gamma\left(1+\alpha\right)e^I\left(\mu\right)}{c''\left(e^I\left(\mu\right)\right)}f\left(y^{HB}\right)-\left(1-F\left(\hat{y}+e^I\left(\mu\right)\right)\right)\frac{\tau\gamma}{c''\left(e^I\left(\mu\right)\right)}\\&+\left[F\left(\hat{y}+e^I\left(\mu\right)\right)-F\left(e^I\left(\mu\right)\right)\right]\frac{\tau\gamma\alpha}{c''\left(e^I\left(\mu\right)\right)}\\&=\frac{\tau\gamma}{c''\left(e^I\left(\mu\right)\right)}\left[\left(1+\alpha\right)\left[e^I\left(\mu\right)f\left(\hat{y}+e^I\left(\mu\right)\right)-\left(1-F\left(\hat{y}+e^I\left(\mu\right)\right)\right)\right]+\alpha\left[1-F\left(e^I\left(\mu\right)\right)\right]\right]\end{aligned}$$

which, in the case $\alpha=0$ simplifies to $e^I\left(\mu\right)f\left(\hat{y}+e^I\left(\mu\right)\right)-\left(1-F\left(\hat{y}+e^I\left(\mu\right)\right)\right)$. Therefore, if

$$e^I\left(\mu\right)>\frac{1-F\left(\hat{y}+e^I\left(\mu\right)\right)}{f\left(\hat{y}+e^I\left(\mu\right)\right)}$$

then marginally decreasing the audit probability for declarations above $\hat{y}$ improves revenues. Since the number of audits run also decreases (by quantity $1-F\left(\hat{y}+e^I\left(\mu\right)\right)$), we have our desideratum. □

This Theorem tests whether the marginal revenues from equation (5) are positive when $\Delta=0$ in an environment where $\alpha=0$. This scenario corresponds to the Authority marginally reducing the audit probability above the threshold $\hat{y}$, while keeping the probability below the threshold constant at μ. This policy clearly saves budget, as no firm faces a higher number of audits. To assess its impact on revenues, note that $\alpha=0$ eliminates the second term in the decomposition of (5) (since $\frac{\mathrm{d}e_H}{\mathrm{d}\Delta}=0$). As a result, the effect on revenues depends on the balance between losses in the $\mathcal{L}$ region and gains from the bunching margin. Near $\Delta=0$, $\tilde{e}$ decreases with infinite pace from its original level $e^I\left(\mu\right)$. Consequently, the bunching region grows quickly for a small audit premium, which offsets the fact that $\tilde{e} \to e^I(\mu)$, leading to a finite, positive limit on marginal revenues from the extensive margin.[48] Condition (S.3.1) ensures that this finite limit more than compensates for the revenue loss from the marginal increase in evasion above the threshold.

The interpretation of condition (S.3.1) is straightforward. The hazard rate reflects the ratio between losses and gains as $\Delta \to 0$. The term $1 - F\left(\hat{y} + e^I(\mu)\right)$ represents the proportion of firms that reduce their declarations when Δ increases, while $f\left(\hat{y} + e^I(\mu)\right)$ measures the mass of marginal bunchers who increase their declarations on the extensive margin. The level of evasion, on the other hand, indicates how much bunching can enhance compliance: when evasion is low, there is less room for firms to adjust their reported revenues. Since $e^I(\mu)$ decreases with μ, this condition is satisfied when μ is sufficiently low for a fixed $\hat{y}$. This result implies that threshold-based audit discounts are particularly effective when the number of audits assigned to a firm class is small.

Although conceptually insightful, Theorem 16 is of limited practical interest. Condition (S.3.1) is challenging to quantify empirically, as it requires knowledge of the true income distribution and $e^I(\mu)$, which is typically unobservable. Moreover, the assumption that $\alpha = 0$ may be restrictive, as any reduction in p_L will likely be accompanied by an increase in audits in the $\mathcal{H}$ region.

The test in (S.3.1) is conducted with $\hat{y}$ held constant, but if the authority could also choose $\hat{y}$, an improvement would certainly be achievable. Condition (S.3.1) only requires that evasion under the flat rule exceeds the hazard rate at the "degenerate bunching point" $\hat{y} + e^I(\mu)$. By moving this point to the upper end of the support, where the hazard rate approaches zero, we ensure that this condition is satisfied. Indeed, the hazard rate approaches zero as the cdf converges to 1 (*i.e.*, as y approaches $\bar{y}$.), while the pdf remains bounded away from zero, as per our initial assumption. This reasoning leads to the following corollary.

Corollary 17. *If the authority can set $\hat{y}$ arbitrarily large, then they can always improve over any flat rule using less budget.*

Proof. The Corollary follows from the fact that $h(y) = \frac{1-F(y)}{f(y)}$ decreases to zero as y approaches $\bar{y}$. This implies that, if the authority is given a flat rule μ, they can always set a threshold such that condition (S.3.1) is satisfied and a marginal increase in the probability of audit above the threshold raises revenues. □

[48] Lemma 6 provides a formal proof.

S.4 Details on Tax Gaps and Profit Shifting in Figure 1

This section documents the methodology used to construct Figure 1 quantifying tax revenue losses due to evasion and avoidance in four countries: the United States, Italy, United Kingdom, and Australia. The estimates refer to the year 2019 (the most recent available) and distinguish three sources of revenue loss: (i) tax evasion by small businesses, (ii) tax evasion by large businesses, and (iii) profit shifting by multinational firms. All values are expressed as a share of total central government tax revenue, allowing for a consistent cross-country comparison despite differences in tax systems and reporting methods.

Small and Large Business Income Tax Evasion: The standard metric used to measure tax evasion is the *tax gap*, defined as the difference between the total theoretical tax liability (under full compliance) and the actual amount collected. We focus on the gross tax gap, which is the gap before amendments, enforced collections, and late payments. We retrieve the gross tax gap and, where necessary, isolate the underreporting component, excluding the underpayments (*i.e.*, reported taxes that are not collected).

Because no harmonized international estimates of tax gaps exist, we rely on official national sources, which vary across countries in tax structure, estimation method, and reporting conventions. For each country and category (*e.g.*, Italian small businesses), we compute a gross income tax gap measure and scale it by total central government tax revenue in 2019 (OECD (2025)), to allow for cross-country and cross-statistic comparability.

To harmonize definitions, we adopt a consistent breakdown between "small" and "large" businesses. The small business income tax gap includes both the corporate income tax (CIT) gap for small firms and the personal income tax (PIT) gap on business income earned by sole proprietors, partnerships, and self-employed individuals. The large business tax gap includes only the CIT gap attributable to large firms. Importantly, thresholds for "small" and "large" businesses are not uniform: they depend on country-specific classifications and may differ substantially. When no disaggregation between small and large is available, we make conservative choices and attribute the full corporate tax evasion to large businesses. Conceptual and methodological differences across countries are detailed in Table S1, which reports, for each country, the relevant data sources, the definitions adopted for "small" and "large" business taxpayers, and any additional notes.

Profit shifting: The data source for profit shifting is Wier and Zucman (2022), which provides comparable estimates across countries. They define profit shifting as "*a tax-motivated and artificial transfer of paper profits within a multinational firm from high-tax countries to low-tax locales*". Based on this definition, they measure profit shifting to tax havens as "*the amount of multinational profits booked by companies in these havens above and beyond what can be explained by real economic activity (such as capital, labour, country characteristics, industry composition, and research and development [R&D] spending)*" (p. 3). Their estimates of revenue losses due to profit shifting are reported as a share of corporate income tax (CIT) revenue.[49] To ensure consistency with the rest of our analysis, we convert these losses into a share of total central government tax revenue by multiplying the original estimate by the ratio of CIT revenue to total central government revenue in each country in 2019 (OECD (2025)).

[49] Appendix Table B in Wier and Zucman (2022).

Table S1: Definitions and sources: small and large business gross income tax gaps

Country	Source	Notes	Small Business Tax Gap	Large Business Tax Gap
United States	IRS (2022)	No yearly estimates available, only average projections for 2017–2019.	CIT gap for "small" firms (assets $<$ \$10M) PIT gap from business income.	CIT gap for "large" firms (assets $>$ \$10M).
Italy	MEF (2022)	No firm-size disaggregation for CIT (IRES) is available.	– PIT (IRPEF) gap for self-employed and individual business income earners; – CIT (IRES) gap not included → lower bound.	CIT (IRES) gap for all firms (small and large combined) → upper bound.
United Kingdom	HMRC (2025) Fiscal year 19-20	We disregard HMRC's aggregate customer segmentation and isolate only the relevant components.	– CIT gap for "small" firms (turnover <£10M; $<$ 20 empl.) – PIT (Self-Assessment)[a] for business taxpayers and partnerships.	CIT gap for "medium" and "large" firms (turnover >£10M; $>$ 20 empl.)
Australia	ATO (2024) Fiscal year 19-20	Estimates are segmented by customer group, not by legal form.	ATO "Small business income" tax gap (<A\$10M turnover; includes sole traders, small companies, trusts, and partnerships).[b]	ATO "Medium business income" + "Large corporate groups" tax gap (>A\$10M turnover).

Notes: this Table presents an overview of the sources, definitions and methodological choices behind the estimates of Figure 1.

[a] Under the Self-Assessment (SA) system, HMRC reports a combined tax gap covering income tax, national insurance contributions (NICs), and capital gains. To isolate the income tax component, we scale the reported estimate by the share of SA income tax receipts in total SA receipts (i.e., income tax plus NICs). While separate data for SA capital gains are unavailable, their contribution is likely negligible. In the fiscal year 2019–2020, this share is approximately 90% (HMRC, 2025).

[b] Reported on Schedules C (sole proprietorship), E (rents, S-corps, partnerships), and F (farming) in IRS Form 1040.

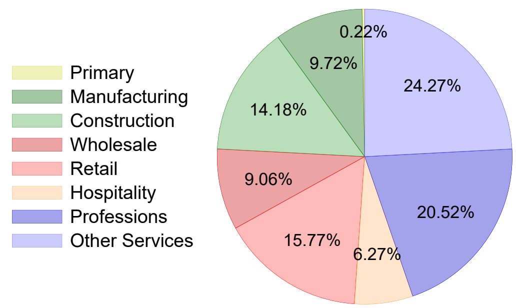
S.5 Details on SeS Data

Figure S1: SeS dataset overview

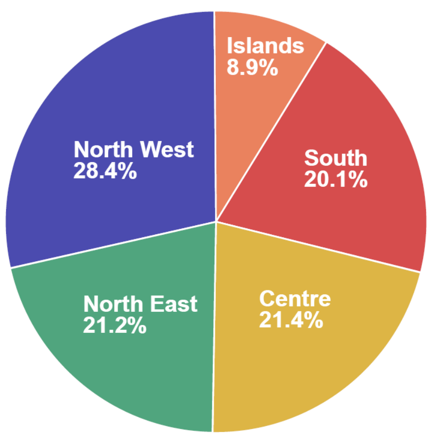
Panel A: Dataset Structure

Tax Year	Observations
2007	3, 753, 997
2008	3, 475, 482
2009	3, 470, 191
2010	3, 482, 862
2011	2, 472, 183
2012	2, 318, 416
2013	2, 142, 884
2014	2, 016, 286
2015	1, 849, 767
2016	1, 700, 551

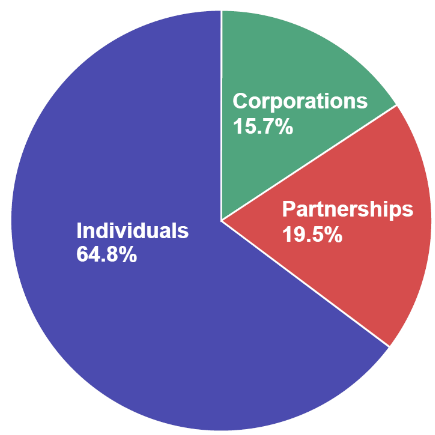
Panel B: 2007-2010 macro-sectors



Panel C: 2007-2010 geographic composition

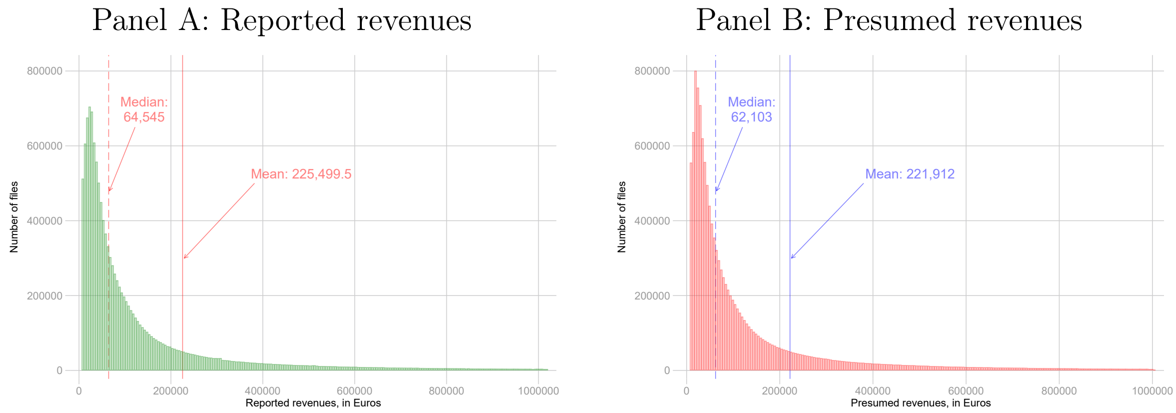


Panel D: 2007-2010 legal status



Notes: this Figure provides an overview of our Sector Studies (SeS) database. Italian businesses and the self-employed file for SeS if they generate no more than 5.2€ million in a given year. Panel A shows the total number of files we access for each of the 2007-2016 tax years. The first four years (in blue) consists of the universe of files submitted by SeS taxpayers in that period. The following years (in green) consist of the files submitted by taxpayers who continuously filed for SeS over 2008-2010. Hence, the sample size decreases as we move to the end of our sample period. The following panels break down the 2007-2010 universe along three dimensions. Panel B shows the relative distribution of SeS files across eight macro-sectors defined by the authors. Panel C shows the breakdown across the five NUTS-1 macro-regions of Italy. Panel D shows the three-way split between individuals, partnerships (akin to U.S. S-corps for tax purposes), and corporations. Italian individual taxpayers and partnerships are subject to the personal income tax, while corporations are subject to the corporate income tax.

Figure S2: Distribution of reported revenues and presumed revenues, 2007-2010



Notes: this Figure shows the distribution of the revenues reported by taxpayers in their SeS files (Panel A) and the revenues presumed by *Gerico* using the relevant sector-specific prediction function and the information imputed by the taxpayer (Panel B). The data consists of the universe of SeS files submitted in the 2007-2010 period, trimmed at 5$^{\text{th}}$ and 95$^{\text{th}}$ percentile of the respective distributions. In the left panel, this excludes about 2% of files which report 0 revenues. SeS technical details: reported revenues include so-called spontaneous revenue adjustment to the SeS presumed revenues available to SeS filers upon submitting. Presumed revenues include any SeS recession corrective available to taxpayers in that tax year.

Figure S3: Reward regime: balanced vs. unbalanced samples, 2007-2010

Variable	**Year**	**Balanced 2007-2016**	***Obs.***	**Unbalanced**	***Obs.***	**Sig.**
Declared revenues (€,000)	2007	241.42	*1,412,980*	184.07	*2,181,464*	***
	2008	245.61	*1,412,980*	208.71	*1,896,637*	***
	2009	229.36	*1,412,980*	202.17	*1,890,103*	***
	2010	235.52	*1,412,980*	198.17	*1,902,521*	***
Gross profits (€,000)	2007	44.25	*1,412,973*	22.68	*2,181,435*	***
	2008	43.59	*1,412,980*	21.86	*1,896,637*	***
	2009	40.42	*1,412,980*	20.04	*1,890,103*	***
	2010	42.11	*1,412,980*	21.63	*1,902,521*	***
Congruous, normal, coherent	2007	52.0%	*1,411,105*	36.5%	*2,174,708*	***
	2008	40.4%	*1,411,316*	24.7%	*1,892,864*	***
	2009	47.1%	*1,411,926*	29.8%	*1,882,565*	***
	2010	52.4%	*1,407,532*	34.3%	*1,893,273*	***

Notes: the Table reports summary statistics for single-sector taxpayers from the 2007-2016 balanced panel used in the reward regime analysis and the remaining taxpayers in each year of our universe period (2007-2010). Congruence, normality, and coherence are the SeS conditions ultimately required to access the reward regime within those sectors progressively included starting from 2011. Columns 3 and 5 report mean values for each sample-year combination. The last column reports, for each variable-year combination, the p-value from an unequal variances test for the equality of variable means across the two samples. *** denotes 1% significance of mean differences. In line with the rest of the reward regime analysis, declared revenues are winsorized at the 99th percentile of the global distribution.

S.6 Additional Data Sources

S.6.1 Local evasion proxies

We construct a broad dataset of local evasion proxies for Italian regions, provinces, and municipalities, depending on data availability. Since the definition and true extent of evasion and underreporting are elusive, we gather several sources from the administrative and economic literature, as well as a large number of citizen-supplied evasion reports submitted to the private online platform at *evasori.info* over four years. Below, we list the sources of the variables we generate, along with their original level of disaggregation. We include all relevant references in our bibliography, and refer to them for further details.

Irregular employment share: Average share of irregular employment for the years 1999 and 2000. *ISTAT* estimates for 103 provinces reported in Table 3 of Censis (2003). Provincial estimates are obtained by *ISTAT* applying at the provincial level the coefficients of a region-level, step-wise regression of irregular employment shares on contextual factors. Significant factors from the region-level regression include unemployment rates, relative relevance of foreign trade and the construction sector, the frequency of workplace injuries, per capita firm registration rates, and population aging.

TV tax evasion rate: Ratio between the number of 2014 TV subscriptions and the 2011 Census number of resident households. Municipal-level estimates are available online at *twig.carto.com* and are based on the TV subscription records with the Italian public TV service *RAI*. Provincial and LLM estimates are a weighted average of the municipal-level estimates, using the number of resident households as weights.

Undeclared IRAP base ratio: Ratio between undeclared and declared IRAP tax bases, 1998-2002. IRAP is the regional tax on productive activities. Its tax base is essentially given by business revenues minus operating costs, with the general exception of employee-related expenses. Estimates for 103 provinces from Table A1 in Pisani and Polito (2006). Estimation relies on a comparison between the local valued added at factor prices reported by *ISTAT* and the local reported tax base for IRAP. We additionally define a regional *IRAP base gap* from Table 31 in the same source as the ratio between the undeclared IRAP base and the sum of the declared and undeclared IRAP base. We compute the declared base dividing the undeclared base by the reported intensity of underreporting.

Ghost-building intensity: Ratio of the number of land registry parcels found with unregistered buildings to the total number of land registry parcels. Municipal-level estimates were produced by the Agenzia del Territorio as a result of a 2007 aerial-photograph and land-mapping exercise. More details are provided in Casaburi and Troiano (2016). Provincial and LLM estimates are a weighted average of the municipal-level estimates, using the number of land registry parcels as weights.

Tax gap: municipal real estate tax (IMU): Ratio between the tax gap and the potential tax base for the 2012 municipal property tax (*imposta municipale unica*, or IMU). We use the first year of IMU implementation, covering all residential units, land holdings, and other buildings. Estimates for 108 provinces based on underlying municipal estimates are provided to the authors by the Ministry of the Economy and Finance. Provinces in the Trentino-Alto Adige region are excluded due to the presence of a different type of real estate tax.

Tax gap: VAT and IRAP: Combined estimates for VAT and IRAP tax gaps, 2007-2010. Estimates for 106 provinces are computed by the Italian Revenue Agency and reported as Table 3 in Vallanti and Gianfreda (2020). Gaps are computed as the difference between the revenues expected by and actually reported to the tax authority, divided by the expected revenues. Estimation of the potential tax base involves both a “top-down”’ approach, comparing the national accounts with tax collection data, as well as a “bottom-up” approach, relying on audit data.

Concealed income share: Ratio of the difference between the average taxable income attested by the Italian Tax Police auditors and the average taxable income reported by taxpayers as a percentage of the average attested taxable income, 1987. Regional estimates come from Table 2 in Galbiati and Zanella (2012) and rely on the universe of audits on individual businesses and the self-employed carried out by the Italian Tax Police for the 1987 tax year.

PIT evasion index: Personal income tax evasion index, computed as the ratio of taxed income and taxable income, late 1980s. Regional estimates come from Table 1 in Brosio, Cassone, and Ricciuti (2002) and draw from Ragazzi (1993).

VAT evasion index: Ratio between taxed value added and taxable value, late 1980s. Regional estimates come from Table 1 in Brosio, Cassone, and Ricciuti (2002) and draw from the analysis of the commerce sector in Cerea (1992).

VAT base gap: Ratio between the VAT base gap and the VAT base theoretical liability (including that from the General Government), averaged over 2007-2010. Regional estimates come from Table B.3 in D’Agosto, Marigliani, and Pisani (2014) (VAT base gap propensity).

Total tax gap ratio: 2001-2011 median of the ratio between the difference of the potential tax yield and the actual tax revenues, and the total voluntary returns, for several taxes under the duty of the Italian Revenue Agency. Taxes considered include the VAT, personal income taxes, corporate income taxes, and IRAP. Regional estimates come from Table 1 in Carfora, Pansini, and Pisani (2016).

Evasion reports from *evasori.info*: In 2008, a computer science professor started an online initiative to raise awareness on the diffusion of evasion behaviors, launching the website *evasori.info*. Through this platform, business customers can anonymously report the location, amount, and sector of any evasion instance they encounter in their daily life in Italy. Most commonly these are missing receipts for modest amounts, but they might reflect more sizable underreporting, as in the case of salaries paid out to irregular workers. *evasori.info* thus provides an independent repository for crowd-sourced and fine-grained repository of information on evasion in Italy.

Coherently with the civic engagement spirit of the initiative, the website provides access to the individual reports via a dedicated API available at `evasori.info/api`. We write a Python script to download all reports submitted between 2008 and 2011, and summarize the obtained information in Table \ref{table:evasori_info}.

We then develop two province-level measures of evasion intensity based on these reports. One is the raw count of reports submitted from each province throughout our sample period, divided by the 2011 Census population. The other is the 2008-2011 total volume of reported evasion divided by the 2011 Census population. We then rescale each measure in terms of 1,000 inhabitants.

S.6.2 Other data sources

Personal income tax data: 2007-2010 data for the national progressive PIT rate schedule and the municipal PIT surcharge rates come from the website of the Ministry of the Economy and Finance (`finanze.gov.it`). Separate files from the same source report the number of individuals filing for the PIT at the municipal level in each tax year, as well as their total reported PIT base. Regional surcharges are instead desumed from the instruction tables attached to the PIT returns for the relevant time period.

For our correlational analysis, we construct a 2007-2010 LLM-level weighted average of the municipal PIT surchage rates in two steps. In the first step, we take the LLM-year average of all municipality-years with a recorded PIT surcharge, weighting each observation by the number of individuals filing for the PIT in that municipality-year. In the second step, we take the simple within-LLM average of the yearly averages obtained in the first step.

Local value added and population data: We draw from *ISTAT*'s online database at `dati.istat.it` to gather information about Italy's provinces. Province-level value added per capita comes from the national accounts tables (*Principali aggregati territoriali di Contabilità Nazionale*). For our correlational analysis, we average the yearly estimates over 2007-2010 for each province. 2011 Census estimates for the provincial resident population are available at `dati-censimentopopolazione.istat.it`.

Input-output tables: We compute measures of sector-level exposure to the final consumer drawing from *ISTAT*'s input-output tables for the 2010-2013 period. We retrieve the relevant data at `https://www.istat.it/it/archivio/195028`. We rely on the symmetric table for 63 1-digit 2-digit sectors, which we are able to match with 51 corresponding sectors with data in the SeS database. The table reports the total value of final uses at 2010 current prices. We build our estimates of the share of domestic value added from final consumer transactions as the sector-specific ratio of final consumer spending and the difference between total uses and exports.

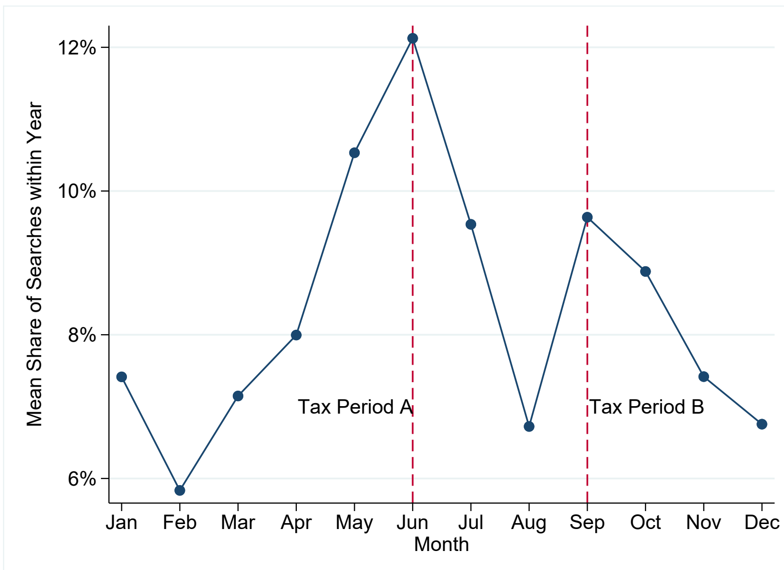
Tax litigation: We capture a component of the cost of engaging with the tax administration with the average length of litigation at the provincial tax court level. Data come from the annual reports on the state of tax litigation and the tax courts released by Ministry of the Economy and Finance and available at `finanze.gov.it`. We gather the province court-level estimates of the average duration of adjudicated cases. Each year, the Ministry estimates this duration as the ratio between the number of days - summed across all cases - it takes to adjudicate each case since the appeal is filed with the court, and the number of adjudicated cases during the year. For each province, we take a simple average of the mean litigation length in each year for the 2009-2012 period.

Beyond the provincial level, litigation can move to the regional level and at the level of the Supreme Court of Cassation (the highest civil court in Italy). By the Ministry's reports, provincial litigation is on average between one third and one half longer than regional litigation in the 2009-2012 period.

S.7 Knowledge of the threshold: Gerico's software Google Searches

Audit rules are effectively disclosed if taxpayers are aware of their functioning. In the context of SeS, we claim that businesses know the threshold at which the probability of audit jumps discontinously because they can learn it at no cost. Indeed, ahead of the tax season, the Revenue Agency releases a freely downloadable software that assists taxpayers in preparing their SeS file. The software is called *Gerico*. We look for evidence of taxpayers awareness about it by looking at Google searches for the word "*gerico*" over the 2004-2017 period in Italy. Figure S1 shows that searches spike in June and September, which are the two tax periods in each tax year.

Figure S1: Google searches for "gerico" spike in tax periods, 2004-2017



Notes: this Figure shows the month-by-month average intensity of Google searches for "*gerico*" over the 2004-2017 period in Italy. This time frame fully includes our SeS sample period, which stretches over the 2007-2016 tax years and the 2008-2017 filing years. Month-level data come from `trends.google.com`. Searches in off-peak months are partly explained by the fact that the actual filing deadlines are postponed in some years due to administrative constraints.

S.8 Bunching Correlation With Evasion and Tax Incentives

To the extent that bunching at the SeS threshold reflects a reporting response, we should observe higher bunching in contexts where underreporting of real economic activity is more intense, either because of higher payoffs to evasion or because of a relative ease of misreporting. We thus study the correlation between bunching of SeS files for each of the 110 Italian provinces in 2007-2010 and available local proxies of evasion across several tax bases. Figure S1, Panel A provides summary statistics and a map of province-level bunching, while Panel B shows the local labor market (LLM) patterns. At both levels of aggregation, bunching is both sizable and heterogeneous across geographical units. We regress the bunching estimates for all local areas i on one evasion proxy Evasion^j at a time according to the following model:

$$\text{Bunching}_i = \alpha + \beta \text{Evasion}_i^j + \gamma \log \text{VA pc}_i + \text{macroregion}_i + \varepsilon_i,$$

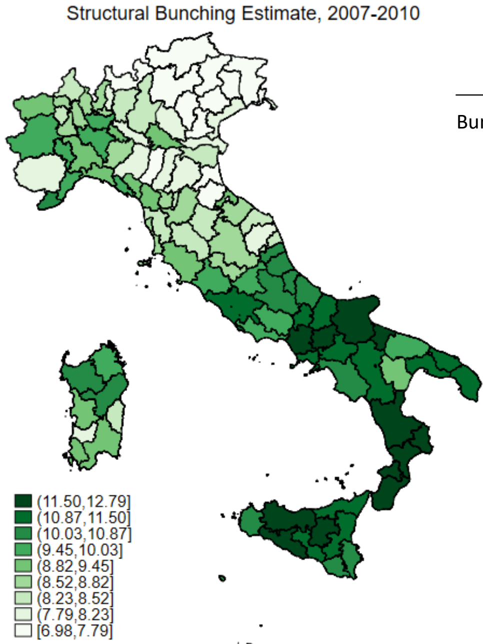
where we introduce fixed effects for the five NUTS-1 macroregions (North West, North East, Center, South, and the Islands) and the logarithm of value added per inhabitant to control for relative provincial prosperity. We use several definitions of i depending on the level of observation of the relevant evasion proxy. Figure S2 disaggregates our analysis whenever a fine evasion measure is available. We show that the correlation between bunching and misreporting holds even at the level of the 686 local labor markets (LLMs) defined by *ISTAT* in 2001, controlling for twenty regional fixed effects and the logarithm of the local PIT base per taxpayer reported by resident individuals.

We also find a positive correlation between bunching and the incentives as well as the opportunities for underreporting. Figure S3 displays a positive and significant conditional correlation between LLM bunching and the weighted average of municipal PIT surcharge rates.[50]

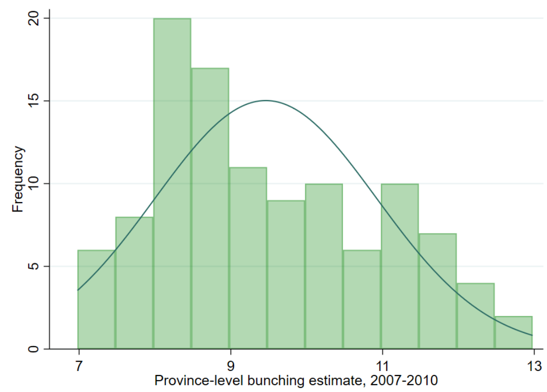
[50] Municipalities can impose a surcharge rate of less than 1% on top of the national personal income tax schedule.

Figure S1: Local heterogeneity in bunching estimates, 2007-2010

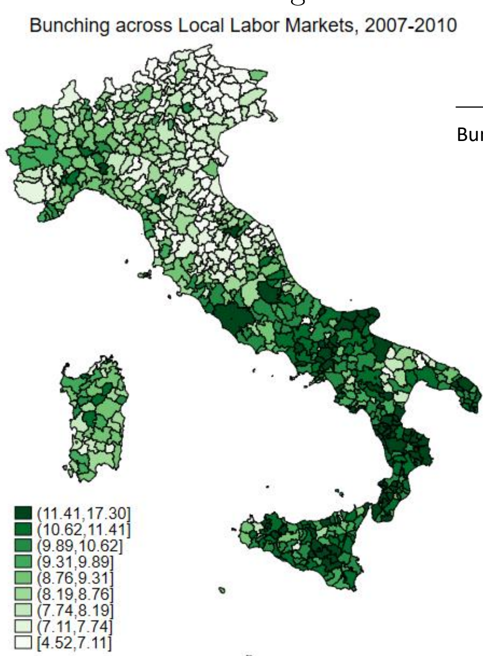
Panel A: Bunching across 110 provinces



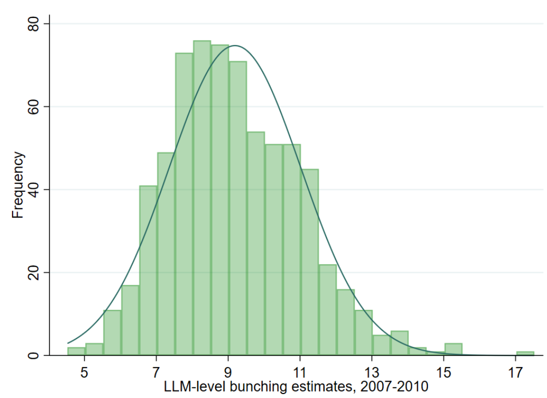
	N. Obs.	Mean	Median	SD	Min	Max
Bunching	110	9.46	9.23	1.46	6.98	12.79



Panel B: Bunching across 686 local labor markets (2010 LLMs)



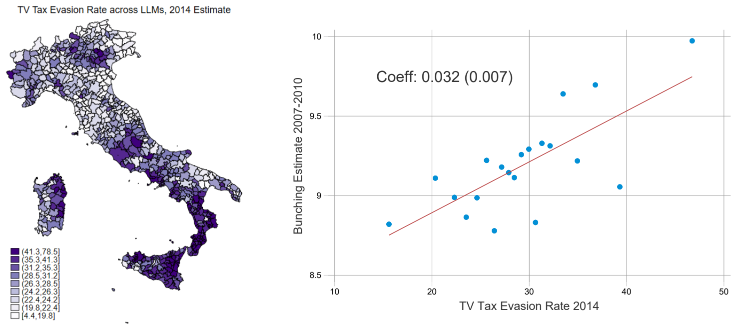
	N. Obs.	Mean	Median	SD	Min	Max
Bunching	686	9.19	8.99	1.83	4.52	17.30



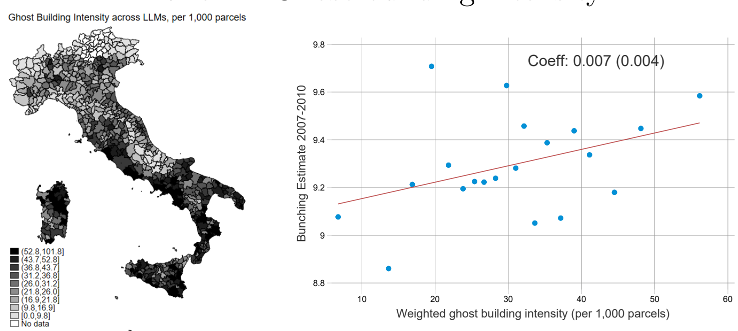
Notes: this Figure plots and summarizes our estimates of bunching at the SeS presumed revenues at the level of the Italian provinces (Panel A) and 2001 LLMs (Panel B). The sample includes each SeS file in the universe of single-sector businesses in the 2007-2010 tax years, except the top and bottom 5% in each province- or local-labor-market-level distribution that we trim to avoid irregularities in the estimation of the counterfactual. Bunching is computed at the local level following the procedure outlines in Section 4.2.

Figure S2: LLM bunching correlates positively with local evasion

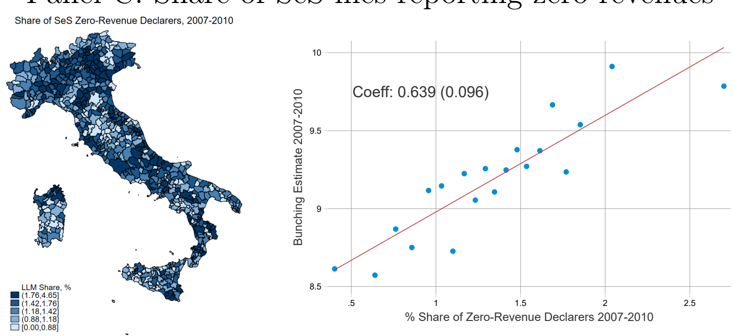
Panel A: TV tax evasion rate



Panel B: Ghost-building intensity

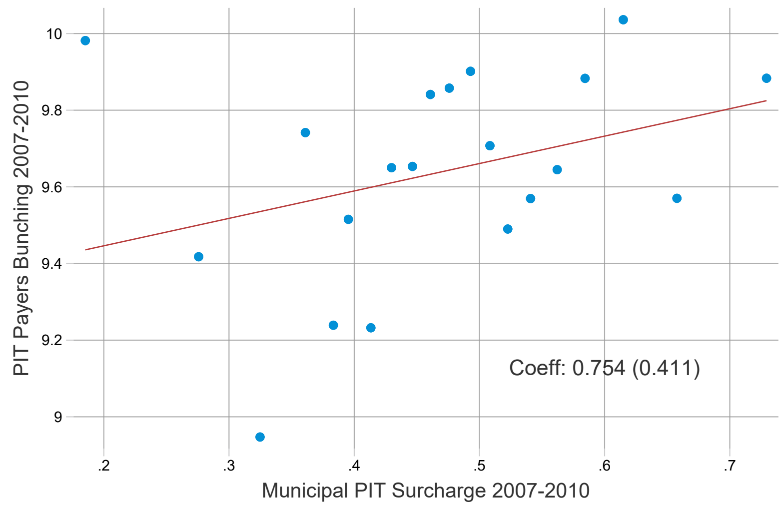


Panel C: Share of SeS files reporting zero revenues



Notes: the Figure maps three LLM-level estimates of behaviors that are plausibly related to evasion or misreporting, and correlates each with local SeS bunching. The three evasion proxies are defined in Appendix S.6. On the right, binned scatterplots report the main slope coefficient and robust standard error from a regression of the form $\text{Bunching}_i = \alpha + \beta Evasion_i^j + \gamma \log(\text{PIT base per taxpayer}_i) + \text{region}_i + \varepsilon_i$, including regional fixed effects and the logarithm of the average local PIT-base per individual taxpayer. Panel A: 2014 TV tax evasion estimates from 8,044 municipalities, weighted by 2011 resident households. Panel B: 2007 ghost-building intensity data from 7,744 municipalities, weighted by number of land registry parcels. Panel C: the 2007-2010 local labor market share of SeS filers reporting exactly zero revenues, which ranges from 0 to 4.7%. The sample includes each SeS file in the universe of single-sector businesses in the 2007-2010 tax years, except the top and bottom 5% in each LLM-level distribution that we trim to avoid irregularities in the estimation of the counterfactual. Bunching is computed at the LLM level following the procedure outlines in Section 4.2.

Figure S3: Bunching tracks evasion incentives: municipal taxes



Notes: this Figure provides a binned scatterplot and the linear fit for the relation between SeS bunching among PIT-payers and the weighted average of municipal PIT surcharges for the 2007-2010 period at the LLM level. We also report the main slope coefficient and robust standard error from a regression of the form $\text{Bunching}_{i,p} = \alpha + \beta(\text{PIT surcharge}_i) + \gamma\text{Litigation}_p + \delta\log(\text{PIT base per taxpayer}_i) + \text{region}_i + \varepsilon_{i,j}$, including the 2009-2012 mean length of litigation at the tax court of province p, regional fixed effects, and the logarithm of the average local PIT-base per individual taxpayer in LLM i. Municipal PIT surcharges don't exceed the national PIT schedule rates by more than 0.8%. Regional PIT surcharge variation is captured by regional fixed effects. The sample includes each SeS file in the universe of single-sector businesses in the 2007-2010 tax years, except the top and bottom 5% in each LLM-level distribution that we trim to avoid irregularities in the estimation of the counterfactual. Bunching is computed at the LLM level following the procedure outlines in Section 4.2.