AUDIT RULE DISCLOSURE AND TAX COMPLIANCE

ENRICO DI GREGORIO⋆ AND MATTEO PARADISI†

May 31, 2022

Abstract. We show that tax authorities can stimulate tax compliance by strategically releasing audit-relevant information. We rely on the Sector Studies, an Italian policy disclosing to small firms and the self-employed that audit risk drops above file-specific revenue thresholds. This allows us to pursue two empirical strategies, leveraging more than 26 million Sector Study files submitted between 2007 and 2016. First, we estimate a structural model to match the heterogeneous bunching we observe on the low-risk side of the disclosed thresholds. Relative to scenarios where these thresholds are secret, we determine that disclosure results on average in 6.3-7.7% higher declared revenues, but modest welfare costs. Second, we exploit a staggered Sector Studies reform that widens the initial audit risk discontinuity. In line with our theory, taxpayers who benefit from greater audit exemptions above the threshold tend to reduce their relative compliance, while those originally below the threshold improve it. However, mean reported profits increase by 16.2% in treated sectors over six years.

JEL Classification: D04, D22, H24, H25, H26, H32

Keywords: tax compliance, enforcement, evasion, audit, disclosure, firm, bunching

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 in light of our research proposal "Sanzioni o Incentivi? L'Effetto dell'Introduzione del Regime Premiale in Italia" dated July 10, 2017. Throughout the development of this project, we benefited from invaluable discussions with and comments from Alberto Alesina, Marcella Alsan, Leonardo Bursztyn, Michela Carlana, Lorenzo Casaburi, Raj Chetty, Claudio Di Gregorio, Gemma Dipoppa, Michele Fornino, Paola Giuliano, Edward Glaeser, Nathan Hendren, Anders Jensen, Felipe Jordán, Henrik Kleven, Alejandro Lagomarsino, Kunal Mangal, Lukas Mayr, Tommaso Nannicini, Nicola Rosaia, Elia Sartori, Luisa Scarcella, Jesse Shapiro, 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. All mistakes are our own.

⋆Corresponding author. National Bureau of Economic Research. Email: enrico.digre@gmail.com.

†Einaudi Institute of Economics and Finance (EIEF). Email: matteo.paradisi@eief.it.

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 et al., 2004). In Italy as in Denmark, the undeclared share of individual income not subject to third-party reporting may well exceed 40% (Kleven et al., 2011; Galbiati and Zanella, 2012), while 43% of the UK tax gap accrues to small firms (HRMC, 2021). In the U.S., imperfect compliance among small businesses results in at least 6.3-8.3% of the total tax liability not being collected without costly enforcement.[1] Yet, tax authorities tend to skew their audit resources toward large firms (Almunia and Lopez-Rodriguez, 2018; Bachas et al., 2019; Basri et al., 2019).[2] This might reflect a cost-effectiveness principle of tax administration, as enforcers expect a relatively higher yield from auditing a large business rather than several small ones for any given budget. 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. These criteria help to split taxpayers into high and low audit risk pools. 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 effect of disclosure is ambiguous ex ante. We set out to characterize and quantify the involved trade-off in a real-world setting.

We study a specific case of audit rule disclosure: ahead of firm reporting, a tax authority reveals the exact location of a threshold above which audit risks drop discretely. We focus on revealed threshold rules for two main reasons. The first is their real-world diffusion. As other government agencies, tax authorities routinely target their actions based on sharp cut-offs.[3] In most cases involving disclosure, tax authorities reveal such cut-offs to affect compliance

[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 true tax liability, separately for 2008-2010 and 2011-2013 (IRS, 2016; IRS, 2019). Our inability to break down other tax gap items implies our estimates are lower bounds.

[2] In a 2019 survey from Italy, our study 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).

[3] As recently as 2017, half of 58 surveyed national tax authorities reported using evasion risk profiling with administrative data to automate case selection, implying the widespread reliance on threshold-based criteria (OECD, 2019). The adoption of industry benchmarks is a forerunner. OECD (2006) describes how tax agencies in countries such as Australia, Germany, New Zealand and the U.S. would audit taxpayers falling outside of sector-specific cutoffs in the distribution of key accounting ratios.

incentives. Contemporary examples include selective audit rule disclosure in Australia, Greece, Mexico, France and Israel.[4] The second reason is that early results in optimal audit theory suggest threshold rules maximize tax collection. In particular, if authorities can commit to an audit strategy ahead of reporting by risk neutral taxpayers, the tax revenue-maximizing rule generally involves a disclosed threshold below which audit risk is discretely higher (Reinganum and Wilde, 1985; Sánchez and Sobel, 1993).[5] In the data, little attention has been devoted to the voluntary compliance effects of disclosed threshold rules.

To our knowledge, we offer the first evidence on the effects of disclosing a threshold above which firms secure a partial audit exemption.[6] We rely on the Sector Studies (*Studi di Settore*, henceforth SeS), an Italian audit system dedicated to small firms and the self-employed.[7] SeS estimate a sector-specific presumed revenue function drawing from the detailed information that businesses submit each year. Just ahead of the tax season, the Italian Revenue Agency provides firms with a software to file the required information and compute the presumed revenues associated with their declaration. The law exempts taxpayers declaring at least the presumed revenue amount from audits stemming from the SeS system. To study the ensuing compliance dynamics, we access a novel confidential database of more than 26.6 million SeS files from the 2007-2016 tax period, including the previously unexploited universe of 2007-2010 files. This rich source of data covers small businesses earning less than €5.2 million in revenues in any given year regardless of incorporation status, location, and sector.

We leverage the disclosure design in the SeS to derive two consistent findings. First, we estimate a structural model to show that presumed revenue disclosure raises mean reported revenues relative to several counterfactuals where firms perceive a constant risk. Second, we exploit a natural experiment to show that tax authorities can expand the reported tax base in the short run by widening the perceived risk gap at a previously disclosed threshold.

We begin with a conceptual framework serving two purposes: i) decompose the audit risk and revenue dynamics behind the effect of disclosure, and ii) clarify how bunching analysis helps us to pin down this effect. Our data provides us with the mean revenues reported under disclosure, when firms perceive higher risk below the presumed revenue level and lower risk

[4]More in detail, we refer to the periodic release of industry benchmarks by the Australian Tax Office (OECD, 2006), the publication of profit margin targets as part of Greece's self-assessment program (Al-Karablieh et al., 2021), Mexico's introduction of effective tax rates by sector in June 2021, and France's *forfait* and Israel's *tachshiv* as early presumptive taxation schemes (Thuronyi, 1996).

[5]More generally, Lazear (2006) discusses how disclosure might improve aggregate outcomes in policy realms other than tax compliance. From high-stake school tests to car speeding and terrorism prevention, authorities can disclose the content of enforcement to concentrate the monitored population's incentives for compliance when enforcement resources and the cost of misbehavior are low. Wong (2021) provides an application on the control of funds misappropriation in the Indian bureaucracy.

[6]In this sense, we differ from the recent literature on Large Taxpayer Units, which induce taxpayers to *reduce* their tax base to escape additional monitoring.

[7]By the estimates of the Italian government, these taxpayer categories are responsible for as much as 30.4% of all unpaid tax liabilities in the 2014-2016 period in the form of non-compliance with personal income tax obligations alone (Ministry of Economy and Finance, 2019).

above it. When presumed revenues are undisclosed, firms perceive an intermediate, locally constant audit risk. As a result, disclosure might reduce revenues among those whose perceived risk drops relative to the counterfactual, and raise them among those whose perceived risk rises. Our goal is to compare the mean revenues we observe under disclosure with the mean revenues reported in sensible constant risk scenarios.

To reconstruct suitable policy counterfactuals from the observed revenue distribution, we build on the logic in Kleven and Waseem (2013) and Aghion et al. (2017). We estimate a smooth counterfactual of relative reported revenues from the portion of the observed distribution above the threshold. This first instrumental counterfactual approximates the set of decisions that firms would make if they all perceived the same low risk prevailing above the threshold. Comparing the observed and estimated distributions at the threshold provides us a way to assess how firms respond to perceived changes in audit risk. In the data, we find that a large number of taxpayers bunch at their SeS presumed revenues, and we show that the extent of bunching correlates positively with evasion levels and incentives. Based on this evidence, we set up a structural model of revenue underreporting to describe the indifference condition of the last firm to bunch when audit risks rise discretely below the threshold. Focusing on the vast share of SeS firms subject to personal income taxation over 2007-2010, we then estimate the model parameters exploiting the available heterogeneity in bunching, local personal income tax rates, and evasion potential across places, sectors and firm sizes.

While our structural estimates suggest that evasion is large at plausible audit risk levels, disclosure might play a role in reducing it in a cost-effective manner.[8] When we let the counterfactual audit risk vary between the high and low risks that we estimate around the SeS threshold, disclosure raises mean reported revenues between 6.3% and 7.7% of counterfactual mean revenues, respectively. This exercise accounts for the potential revenue gains and losses that disclosure might generate among taxpayers who perceive an audit risk hike or reduction. The finding that gains seem to exceed losses reflects the possibility that, while observed bunching is large, the firms' implied elasticity to the tax incentives is relatively low. As such, the bunching we observe is more likely the product of audit risk changes inducing relatively more firms to increase their revenues.

Our result would be confirmed by any model specification where the intensive margin elasticity of revenue manipulation is not the sole cause of bunching. These alternative models could feature audit costs independent of evasion and a margin of presumed revenue manipulation (such as reported cost underreporting), which might reduce the estimated elasticity of revenue manipulation and in turn the estimated loss from disclosure. For this reason, even if our model is parsimonious in the way firms can respond to policy incentives,

[8] On average, we estimate that revenue underreporting alone is larger than one third of median reported profits when all firms perceive the low audit risk prevailing above the SeS presumed revenues.

it delivers conservative estimates of the compliance effect of disclosure relative to several realistic modelling choices.

The implied return on SeS disclosure for the tax administration is also economically large: the administrative cost of setting up and updating the SeS system is a fraction of the estimated tax gains. Relatedly, a marginal value of public funds (MVPF) analysis (Hendren and Sprung-Keyser, 2020) suggests that the average welfare cost of disclosure tends to be modest, as firms perceive opposite policy changes on the two side of the threshold.

Even if SeS disclosure might raise reported revenues, firms could still keep their tax liability constant by reporting higher costs (Carrillo et al., 2017b). Given that our model is not well suited to study compliance margins other than revenues, we rely on a natural experiment that closely mimics our policy counterfactual logic to get a more complete view of tax base dynamics. Starting in 2011, a staggered reform to the SeS known as "reward regime" extended the protections provided to those in line with SeS prescriptions, and promised to devote more attention to those who did not comply. This should widen the audit risk gap perceived around presumed revenues for firms exposed to the new rules. Using a balanced panel for the 2007-2016 decade, our event-study design shows that taxpayers in treated sectors report revenues closer to the SeS threshold. Both firms below and above the threshold ahead of the reform display this adjustment, confirming our intuition that disclosure-based policies might offer opposite incentives to different taxpayers. However, reported gross profits rise on average by 16.2% over six years in treated sectors. This might point to a desirable feature of audit rules with predicted revenue thresholds: since higher reported costs translate into higher predicted revenues, the scope for cost adjustments to offset any reported revenue increase is limited. Together, our structural and quasi-experimental evidence suggests that tax authorities can expand the tax base by disclosing and strengthening an audit risk discontinuity.

This paper provides several contributions to the study of tax compliance among small businesses. Our work is most closely related to a growing tax enforcement literature evaluating the relative merits of different collection strategies. We believe this to be the first paper to study the use of selective audit rule disclosure as a revenue-enhancing strategy. Almunia and Lopez-Rodriguez (2018) analyze the case of a Large Taxpayer Unit in Spain, whereby corporations expect stricter enforcement when reporting above €6 million.[9] Instead, the disclosure process in our setting encourages seemingly low-productivity businesses to report more, rather than less, revenues. In addition, the bulk of our data comes from micro firms and the self-employed, which are typically hard to monitor and for whom voluntary compliance schemes may compensate for the tax agency's inability to ramp up audits.

A few recent papers highlight the role of incentives for taxpayers (Dunning et al., 2017; Carrillo et al., 2017a; Al-Karablieh et al., 2021) and third parties (Naritomi, 2019; Choudhary

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

and Gupta, 2019; Kumler et al., 2020) in stimulating quasi-voluntary tax compliance. Differently from common tax lotteries, tax amnesties, and temporary audit exemptions, we examine the permanent introduction of compliance incentives which taxpayers can access autonomously by following predetermined prescriptions.[10]

SeS disclosure is also distinct from the one implied by audit threat letters, the hallmark of tax enforcement randomized control trials (Kleven et al., 2011; Pomeranz, 2015; Bérgolo et al., 2017; Carrillo et al., 2017b). Unlike in SeS, the goal of these interventions is not to reveal the structure of the audit system to all taxpayers. Therefore, their general equilibrium effects and whether their threat credibility can scale up remain uncertain (Slemrod, 2019).

Methodologically, our structural model fits in the recent public finance literature on bunching (Saez, 2010; Chetty et al., 2011; Chetty et al., 2013) with "notched" incentives (Kleven and Waseem, 2013). Provided that only audit risks change at the SeS threshold, and that the scheme reveals its exact location only after the end of the production year, our setting allows to credibly separate production and reporting responses to the notch. This constitutes an advance relative to study settings where taxpayers can fully adjust their production to non-idiosyncratic thresholds that are stable over time. Relative to original bunching set-ups, where the goal is the estimation of a single elasticity parameter, we study reporting decisions which depend both on unknown audit risk perceptions and the degree of responsiveness to tax incentives. We thus exploit variation in bunching responses, policy parameters such as tax rates, and local differences in auditing activity to jointly estimate a broader set of primitives as in Aghion et al. (2017), for the explicit purpose of policy evaluation.

Lastly, Italy provides a suitable setting to the study of small firm tax compliance. The Italian economy is abundant in small businesses and self-employed individuals, so its enforcement experience might yield useful insights for economies with a similar production structure (Arachi and Santoro, 2007).[11] SeS taxpayers are the object of theoretical and empirical work by both academics and practitioners (Santoro, 2008; Santoro and Fiorio, 2011; Santoro, 2017; D'Agosto et al., 2017; Battaglini et al., 2020). This paper is the first to ask whether disclosure can prove tax base-enhancing. As a result, our original contributions include the application of bunching techniques to the universe of SeS filers and the structural estimation of the compliance and welfare effects of disclosure among SeS personal income taxpayers. In the process, we provide novel estimates of perceived audit risks and firm responsiveness to audit incentives. We are also the first to assess the dynamic impact of the 2011 reward regime as a stimulus to voluntary tax compliance for all firm types. Starting in 2018, the new regime's

[10]In this sense, SeS also differ from the case of the Greek self-assessment program in Al-Karablieh et al. (2021), where voluntary participation comes to a substantial halt with the beginning of the national fiscal crisis starting in 2010.

[11]To the extent that small firms display similar tax gaps across countries, our results might also apply to other contexts since the composition of the productive landscape in Italy might explain the seemingly large size of its shadow economy (Zanella, 2012).

logic inspired a comprehensive overhaul of SeS. We provide a blueprint for the future analysis of this reform and voluntary compliance incentives in enforcement more in general.

2. Conceptual framework

We build a model to study when the disclosure of a threshold-based audit rule can improve reporting behavior among firms. The framework clarifies what are the sufficient statistics that determine the revenues effects of such policy. We also discuss which moments help identifying these statistics in the data.

2.1. **Firm problem.** Consider a class of firms where each firm produces revenues $y^\star$, and can report an amount $y \in [0, y^\star]$ evading $e(y^\star) = y^\star - y$.[12] Firms are heterogeneous in production and we denote this heterogeneity with a continuum of firm types ξ that guarantees a smooth distribution of $y^\star$. Under a policy regime denoted by θ, each firm faces some perceived audit probability schedule π_θ and an income tax schedule, so that evasion generates an expected avoided tax liability $T_\theta(e)$ that depends on evasion e, the perceived audit schedule, and the income tax schedule.[13] Firms are risk neutral and maximize their expected value. Assuming separability between revenue production and reporting decisions, we focus on the latter, so that firms maximize

$$V(y^\star, y, T_\theta(y^\star - y)) = T_\theta(y^\star - y) - g(y^\star - y), \tag{2.1}$$

where $g(\cdot)$ is a convex cost of evasion that captures the organizational hurdles and psychological costs of concealing production, separately from any administrative penalty applied upon detection. We assume that $T_\theta(e)$ is continuously differentiable when the probability schedule π_θ is flat, so that the firm's optimality condition guarantees an interior solution that solves

$$T'_\theta(e_\theta(y^\star)) = g'(e_\theta(y^\star)). \tag{2.2}$$

It follows that a smooth distribution of $y^\star$ maps into a smooth declared revenue distribution $H_\theta(y_\theta(y^\star))$. We define average reported revenues in counterfactual θ as $\bar{y}_\theta = E_{y^\star}[y_\theta(y^\star)]$.

2.2. **Audit policy counterfactuals.** We study a specific case of audit rule disclosure: before tax season, the tax authority reveals the exact location of a firm class-specific revenue threshold $\hat{y}$ above which audit risks drop discretely.[14] We denote with p_H and p_L the perceived

[12] We assume that production and declaration decisions are separable. One can show that separability arises in a model where input costs are fully deductible and production and evasion costs are separable. Tax declarations typically occur months after production, suggesting that this is a reasonable assumption. Appendix E discusses empirical evidence supporting separability. We also implicitly assume that firms never report more than their earned revenues. Galbiati and Zanella (2012) show that over-reporting is rare in Italian data.

[13] Firm take perceived probabilities as given. This assumption rules out coordination behavior across firms to decrease the expected tax liability.

[14] In the case we study, the Italian tax authority estimates a different audit-relevant revenue function for each of several narrowly defined business sectors.

probabilities of audit below and above the threshold, respectively, so that $\Delta p = p_H - p_L > 0$ (Figure 1). In the disclosure counterfactual D we have

$$\pi_D = \begin{cases} p_H & y < \hat{y} \\ p_L & y \geqslant \hat{y}. \end{cases}$$

When facing such probability schedule, firms report on average $\bar{y}_D$. Because $T_D(\cdot)$ changes discretely at $\hat{y}$, we expect to observe a number of firms bunching at that point to reduce their audit risk. These are businesses that in a scenario H with constant probability p_H would declare $y_H < \hat{y}$, but for whom $V(y^\star, y_H(y^\star), T_H(y^\star - y_H(y^\star))) \leqslant V(y^\star, \hat{y}, T_D(y^\star - \hat{y})))$.

Consider two alternative counterfactuals. First, a policy counterfactual scenario without audit rule disclosure that we label C. We assume that in regime C firms would perceive a locally constant audit risk p_C, and would declare on average $\bar{y}_C$.[15] Relative to this counterfactual, disclosure can at most reduce the risk perceived by firms with $y \geqslant \hat{y}$, and it can weakly raise perceived risks among firms with $y < \hat{y}$ as they are revealed to be a preferred audit target. Hence, we focus on counterfactuals C where $p_C \in [p_L, p_H]$ and therefore $\Delta p_C = p_C - p_L \leqslant \Delta p$. Second, we define an instrumental counterfactual L that would arise with a constant audit risk p_L, with average revenues $\bar{y}_L$. We will refer to this counterfactual as *bunching counterfactual* since the smooth revenue distribution under this scenario will be used to create our empirical bunching measures.

2.3. **The revenue effects of disclosure.** Using average revenues in the three counterfactuals C, D, and L, we can characterize the revenues effect of audit rule disclosure. Consider moving from a situation with a constant audit risk p_C to one with a jump Δp in audit risk at $\hat{y}$. The total revenue effect of this policy change, $\bar{y}_D - \bar{y}_C$, can be decomposed into two components. First, we label *probability reduction effect* (PRE) the loss in mean reported revenues $\bar{y}_C - \bar{y}_L$ that occurs because audit exemptions uniformly reduce perceived risks by Δp_C and firms adjust from a policy counterfactual with constant risk p_C to the bunching counterfactual with constant risk p_L. Figure 2 depicts this effect as a left shift of the distribution of declared revenues. Second, we call the *incentive provision effect* (IPE) the reported revenue gain $\bar{y}_D - \bar{y}_L$ induced by the audit risk hike Δp below the threshold, with firms raising mean reported revenues from $\bar{y}_L$ to the level $\bar{y}_D$ that we observe under disclosure.[16] In Figure 2, this is represented by a shift to the right in the distribution of declared revenues by businesses

[15]If the size of the firm class is small enough this simplification does not come at large costs. Indeed, by suitably defining firm classes, this constant-p_C representation can be rationalized even when there is progressivity in audit policies such that larger businesses face a larger probability of audit. Moreover, our results will show revenue improvements for any constant p_C, making our estimates conservative relative to a case where p_C increases within each firm class.

[16]Gains stem not only from the possibility that perceived audit risks rise below the threshold, but also from firms learning that reporting at the threshold reduces the probability of receiving an audit.

declaring $y < \hat{y}$ in the L counterfactual, with many of them jumping at $\hat{y}$ and creating a bunching mass.

Audit rule disclosure raises reported revenues if and only if:

$$\bar{y}_D > \bar{y}_C = \bar{y}_L + \text{PRE}. \tag{2.3}$$

In the data, we observe $\bar{y}_D$ and $\bar{y}_L$, but we cannot directly measure PRE.[17] We use the theory to guide us in determining the primitives needed to quantify it. Let us define the aggregate elasticity of declared revenues $\varepsilon^{agg}(\bar{y}_\theta) = E_{y^\star}[\frac{\partial y_\theta(y^\star)}{\partial T'_\theta(y^\star - y_\theta(y^\star))} \cdot \frac{T'_\theta(y^\star - y_\theta(y^\star))}{\bar{y}_\theta}] < 0$. We assume linearity of $T_\theta(e)$ in a p-constant counterfactual, so that $T'_\theta(e) = \tau - \tau\gamma p_\theta$, where τ is a flat tax rate on income and $\gamma > 1$ is a penalty rate on detected evasion.[18] For every additional Euro of evasion, a firm saves τ in taxes, but with some probability will pay back the evaded amount increased by a penalty. Hence, the change in the marginal expected avoided tax liability between counterfactuals C and L is $T'_C(e) - T'_L(e) = -\tau\gamma\Delta p_C$. Using these definitions, we derive the following result.

Result 1. Given a class of firms, a counterfactual with constant probability p_C, and assuming a constant aggregate elasticity $\varepsilon^{agg}(\bar{y}_\theta) = \varepsilon^{agg}$ for any $\bar{y}_\theta$, the disclosure of an audit schedule π_D with threshold $\hat{y}$ is revenue improving if

$$\frac{\bar{y}_D}{\bar{y}_L} > \frac{1}{1 + \frac{\varepsilon^{agg} \cdot \gamma \cdot \Delta p_C}{1 - \gamma \cdot p_C}}. \tag{2.4}$$

Appendix A derives the result. This condition has an intuitive interpretation. The left-hand side is the gross size of the IPE, that is, the gross percentage gain in reported revenues from the incentive provision effect of disclosure. The right-hand side provides the gross size of the PRE, or the gross percentage loss in reported revenues arising if all taxpayers perceived an audit risk reduction equal to Δp_C. Since both percentage changes are expressed with respect to $\bar{y}_L$, they are directly comparable. Fixing the left-hand side (or, relatedly, given a distribution under disclosure and a bunching counterfactual), condition (2.4) also suggests that highly responsive firms and larger audit risk drops above the threshold lower the potential for disclosure to improve revenue reporting. While the left-hand side mostly depends on the audit risk drop Δp, the right-hand-side depends on ε^{agg} and on the audit risk drop $\Delta p_C \leqslant \Delta p$. It follows that revenue improvements occur as long as the risk drop at the threshold is large enough compared to the responsiveness of revenue declarations.

We can use the condition in (2.4) to guide our search for the relevant parameters to test for the revenue potential of a disclosure policy in the data. We access the rich administrative

[17] Section 5 describes how we estimate a counterfactual distribution based on p_L, from which we derive $\bar{y}_L$.

[18] We assume linearity of income tax rates within firm classes, but we do not restrict the heterogeneity in levels of τ across classes. Hence, the model can accommodate applications to non-linear tax schedules.

database of the Sector Studies, a long-standing Italian audit system which relies on disclosure to encourage tax compliance among small firms. We first use bunching and a structural model to reconstruct suitable policy counterfactuals and compute the revenue effect of disclosure in the Sector Studies. Next, we exploit a Sector Study reform as a natural experiment to evaluate the effect of disclosure-based policies on compliance margins that we cannot model, such as gross profits.

3. Disclosing audit rules: the Italian Sector Studies

In 1998 the Italian government implemented a novel auditing tool known as *Studi di Settore* or Sector Studies (henceforth, SeS), targeting non-employee taxpayers generating no more than €5.2 million in revenues.[19] Since then, individuals, partnerships (pass-through businesses), and small corporations file every year in compliance with their Sector Study, and might be subject to tax audits ensuing from the analysis of the supplied information.

SeS generate a file-specific discontinuity in the probability that taxpayers experience an audit on reported revenues. *Agenzia delle Entrate* (the Italian Revenue Agency), in collaboration with *SOSE*, a publicly-owned analysis company, estimates sector-specific linear models of presumed revenues using past declarations on business turnover, operating costs, workforce details, physical capital, input quantities, the size and location of their premises (Figure 3). Every year, businesses must report on these dimensions of their activity, allowing the model to determine a level of presumed revenues idiosyncratic to that year's file through the previously estimated coefficients. As specified in the instituting Law 146/1998, declaring less than the presumed revenue amount provides the Revenue Agency with a motive to initiate a tax assessment.[20]

The policy's design encourages taxpayers to adjust their reporting behavior to the presumed revenues for two reasons. The first is the timing of filing (Figure 4). For any given tax year, production ends months before tax season, when taxpayers fulfill both their tax and SeS obligations. Filing deadlines are generally set in June or by the end of September, at least half a year after production decisions have been made for the relevant tax year.[21] The second reason is the system's transparency. Taxpayers can learn about their SeS threshold at no cost. Before the tax season, between February and May, the Revenue Agency releases a

[19]In our main sample period, 2007-2010, taxpayers could seek exemption from SeS by opting into a minimum taxpayer regime, with eligibility conditional on reporting up to €30,000 in the previous tax year. To our knowledge, there is no clear evidence that audit risks change systematically at this threshold.

[20]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".*

[21]Although the exact tax days often change across years, deadlines are generally set in the early summer and in the fall for taxpayers filing on paper or required to do so online, respectively.

freely downloadable software that assists taxpayers in preparing their SeS file.[22] The software, known as *Gerico*, stores the coefficients associated to any sector-specific presumed revenue function estimated by *SOSE*. Upon imputation of the relevant accounting and structural information, *Gerico* informs the taxpayer of their threshold value before they submit their file, allowing for adjustments. Working through the software provides the fastest way to learn one's threshold.[23]

Compliance with SeS presumed revenues is one of many audit selection criteria considered by the administration. As a result, taxpayers can trigger an audit for reasons unrelated to SeS behavior, regardless of where they locate relative to the threshold. Crucial to our analysis, the residual audit risk independent of SeS filing stays constant around presumed revenues. Moreover, reporting above the cutoff provides no specific fiscal benefit other than a relative reduction in audit risks. This allows us to attribute the observed revenue responses solely to the audit incentives generated by SeS.[24]

3.1. **Reward regime.** Starting in 2011, the Italian government reinforced the discontinuity in incentives associated to SeS reporting. Law Decree 201/2011 instituted what is commonly referred to as *regime premiale* or reward regime, which sought to extend a set of ancillary audit protections for taxpayers complying with SeS prescriptions. We compare the pre- and post-reform regimes in Table A2. Introduced in a staggered manner across SeS sectors, the new regime promised audit exemptions from additional investigation sources other than SeS and shortened the statute of limitations of audits by one year.[25] To access these benefits, businesses would not only report revenues at or above the presumed level (a condition known as congruence in the SeS framework), but also fall within acceptable ranges of several sector-specific accounting indicators (two conditions known as normality and coherence). The reform otherwise encouraged the tax administration to boost enforcement among non-compliant firms through dedicated audit plans and detailed analysis of their financial relationships with third parties.

[22]A yearly press release announces the availability of *Gerico*'s free download on the Revenue Agency website. Figure A3 in the Appendix shows that Google searches for the word "gerico" in Italy peak around the two tax seasons.

[23]The details regarding the estimation procedure are published by the Revenue Agency in dedicated yearly technical reports. *Gerico* allows the timely dissemination of the estimation models' updates, which the law requires at least once every three years according to a sector-specific calendar. Model revisions involve re-estimating the sector-specific presumed revenue functions with relatively more recent data. The process may thus affect both the selection of relevant input variables as well as the size of the associated coefficients.

[24]The Italian tax enforcement system further includes *Guardia di Finanza*, a police force tasked with fighting tax crimes. Although they can rely on information from a taxpayer's SeS file to initiate an audit, their investigative activity focuses on tax-related crimes. The Revenue Agency runs most of the ordinary file auditing, and is the only agency with the power to request additional tax payments.

[25]Inclusion would happen at the beginning of each tax season for the previous calendar year, with the Revenue Agency releasing the updated list of sectors to benefit from the new incentives. A majority of businesses in manufacturing, commerce, and services were included by the 2016 tax year, when our data period ends. Professionals were mostly excluded until a more organic transformation of the SeS system starting with the 2018 tax year.

4. Data sources

To examine the role of SeS disclosure for taxpayer behavior, we access two complementary confidential administrative datasets. The first consists of the universe of SeS files for the 2007-2010 tax years. The second is a 2007-2016 unbalanced panel of all taxpayers who have continuously filed between 2008 and 2010. To our knowledge, this is the first paper to exploit all SeS files available in any given year. Put together, the raw data covers almost 26.6 million SeS declarations submitted by over 4.7 million Italian micro businesses and self-employed. Each of the tax years between 2007 and 2010 alone generates more than 3.4 million files. Appendix B offers an overview.

The data provide detailed information about the taxpayer's economic activity for the relevant tax year, including their reported revenues, gross profit or income, the size of the workforce, the wage bill, a number of cost items, and the surface area of their premises. Crucially, each file comes with the exact value of the associated SeS threshold. This allows us to assess the relative distance between the revenues declared by the taxpayer and those presumed and disclosed by *Gerico* before filing.

A snapshot of the context in which taxpayers operate comes from the files' information on their business sector and location. Sectors are identified both by the standard 6-digit industry code (Eurostat NACE Rev.2), as well as by the administrative SeS code of reference. Reported locations have special relevance since the vast majority of SeS filers are single-establishment businesses with low spatial mobility.[26] Over the 2007-2010 period, all files are associated to one of twenty administrative regions, and about 95% of all files are associated to at least one of 110 provinces. In addition, we are able to assign 77% of the subset of personal income taxpayers to one of more than 8,000 municipalities and 686 local labor markets as defined by the Italian Institute of Statistics (*ISTAT*).[27] We will exploit these links to perform analysis at different levels of geography.

SeS cover a broad spectrum of firm types with diverse legal status. Almost two thirds of 2007-2010 files come from individual businesses and self-employed professionals (64.8%). The rest pertain to partnerships (19.5%) and corporations (15.7%). Along with the geographic location, the legal status of a firm determines its profit tax regime. Personal income taxes (PIT) are paid by individuals and partnerships, with the latter akin to U.S. S-corps for tax purposes. Corporate income taxes affect corporations only. In our structural analysis, we rely on the tax heterogeneity generated by municipal and regional surcharges applied on top of national personal income tax rates. Given that corporations face a single flat corporate

[26] More than 98% of 2007-2010 files are submitted by taxpayers who never move out of their original province over the observed period.

[27] To avoid risks of single outing, the Revenue Agency forbids the disclosure of a taxpayer's location when there are no more than three establishments in their same sector in a given municipality. Given the extremely low mobility of our taxpayers, we impute a taxpayer's municipality for a given SeS file using their location reported in any of their other SeS files in our data.

tax rate, we exclude them from the structural estimation of disclosure effects on reported revenues, but include them in the remaining parts of the analysis.

We complement our SeS data with a wide range of contextual information detailed in Appendix C. We devote special attention to consulting available sources on tax evasion across space and time in Italy, including both official government statistics and online crowd-sourced reports of fiscal malfeasance witnessed in the country during our study period.

5. Bunching at the presumed revenue threshold

This section examines firm bunching at the SeS presumed revenue threshold. We start our empirics by measuring bunching for two reasons. First, bunching analysis allows us to gauge the magnitude and patterns of firm responses to SeS disclosure. Second, bunching will provide a crucial moment in our structural estimation.

In a disclosed regime the expected avoided tax liability $T_D(\cdot)$ jumps discontinuously at the threshold $\hat{y}$ creating an incentive for a subset of the taxpayers to report revenues at $\hat{y}$ and reduce their audit risk. This is a subset of the businesses who would have declared $y < \hat{y}$ with a constant probability p_L. Therefore, we expect an excess mass of individuals at $\hat{y}$ relative to the baseline revenue distribution where audit risks are constant at p_L, the scenario with declared revenues distribution $H_L(\cdot)$ that we labelled *bunching counterfactual*.

Figure 5 Panel A shows the distribution of reported revenues around $\hat{y}$, leveraging the universe of SeS files submitted by single-sector businesses for the 2007-2010 tax years. The horizontal axis represents the distance of reported revenues from the file's associated $\hat{y}$ in percentage terms of $\hat{y}$ itself.[28] There is a significant spike in the distribution within 1 percentage point of $\hat{y}$, consistent with a large share of taxpayers declaring at or slightly above $\hat{y}$ to avoid audits.

To quantify the extent to which taxpayers bunch at the presumed revenue level, we build an empirical counterfactual bunching distribution $\hat{H}_L(\cdot)$. Since businesses declaring $y_D > \hat{y}$ 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:

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

[28]We rely on the relative distance from the threshold for illustrative purposes only. In our structural analysis, we model taxpayers responses based on their absolute distance from the threshold, that is in Euro terms. We then split taxpayers into groups of relatively similar turnover, and compute bunching based on their absolute distance from presumed revenues.

where i indicates the polynomial degree in the first sum. We use a 7^{th} degree polynomial in our baseline estimates and provide estimates with different degrees for robustness below.[29] 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 $\hat{b}$, from now on. We compute the standard deviation of our bunching estimates using a bootstrap procedure with 1,000 replications. Appendix D outlines the steps of this procedure.

Figure 5 Panel B exemplifies our bunching estimation in the universe of filers for the 2007-2010 tax period. The counterfactual closely follows the empirical distribution on the right of $\hat{y}$ up until y^u, which we use to delimit the excess mass. On the left of $\hat{y}$ the empirical distribution lies below the counterfactual and the difference between the two is the missing mass generated by bunchers. We find substantial bunching equal to 9.56 (bootstrap sd = 0.61) in our baseline. Bunching induces higher revenue reports relative to the L counterfactual. The extra revenues reported by bunchers only are equivalent to a uniform right-shift of the $H_L(\cdot)$ distribution for an amount equal to 1.13% and 3.05% of the observed mean and median revenues, respectively. Table A3 reports the sensitivity of our bunching estimates to the choice of polynomial order and upper bound y^u. Our baseline estimate lies on the lower end of the estimates distribution. This follows from our conservative definition of excess bunching, since we attribute any excess mass to SeS incentives only if within 1 percentage point of presumed revenues.

The bunching measured in the data might stem from adjustments to firm production or reporting as taxpayers try to comply with their disclosed level of presumed revenues. Appendix E shows that our bunching estimates correlate positively and significantly with several measures of the intensity, incentives, and potential for tax base misreporting across places and sectors. In addition, three pieces of evidence suggest that production responses are second order: i) bunching is sharp at the threshold, ii) there is a time lag between the production year and the moment when the threshold is revealed, iii) individuals do not increasingly locate at the threshold over time in the years following a new SeS model introduction (Appendix E). We use this evidence to motivate the separability between the reporting and production margins that we assume in our theoretical model and estimation.

[29] 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.

6. Identification of the model primitives

In Section 2 we highlighted the role of behavioral parameters in determining the compliance effects of audit rule disclosure. Building on our model of firm's reporting decisions, we now develop an identification strategy to estimate related sufficient statistics in the data.

Our data allows us to observe firm reports with a disclosure policy in place. Using our theory and the bunching approach discussed in Section 5, we derived a counterfactual revenue distribution in a constant-p_L scenario. Next, we design a strategy that relies on the moments of these two distributions and on firms' optimality conditions.

6.1. **Changes in incentives and bunching.** We carry the linearity assumption on $T(e)$ from Section 2 and define $T(e) = (\tau - \tau\gamma p)\cdot e$. Because $T(e)$ is continuously differentiable when p is constant, the optimal choice of declared revenues in the L counterfactual is determined by $T'(y^\star - y_L(y^\star)) = g'(y^\star - y_L(y^\star))$. Given that $y^\star$ has a smooth distribution, the reported revenue distribution $H_L(y_L(y^\star))$ is also smooth. The introduction of SeS creates a "notch", *i.e.* a discrete increase in firm's expected tax liabilities below $\hat{y}$. In the standard interpretation of Kleven and Waseem (2013), both the average and marginal tax rate on manipulated revenues change at $\hat{y}$ (see Figure 6). The expected avoided tax liability becomes $T(e) = (\tau - \tau\gamma p_L)\cdot e - \tau\cdot\gamma\cdot\Delta p\cdot e\cdot I(y < \hat{y})$. Because of the discontinuous nature of these incentives, a subset of firms bunches at the threshold $\hat{y}$ by reducing revenue manipulation relative to the baseline scenario with p_L.[30]

Two types of firms characterize the bunching area. First, firms that would already report revenues equal to $\hat{y}$ or higher in the L scenario would not change their choice under the discontinuous schedule with p_H below the threshold. Second, firms who would declare less than $\hat{y}$ with p_L, and would instead jump to $\hat{y}$ under the discontinuous risk schedule. Firms that bunch to $\hat{y}$ are included in the interval $[\hat{y} - \Delta\hat{y}, \hat{y})$ in the L scenario. We call marginal buncher the firm that would declare revenues equal to $\hat{y} - \Delta\hat{y}$ with a flat probability p_L and is indifferent between declaring $\hat{y}$ and its interior solution y_H when the audit rule is disclosed. Overall, bunching under disclosure produces a new distribution with mean reported revenues $\bar{y}_D > \bar{y}_L$.

6.2. **Indifference condition.** We exploit the marginal buncher's indifference condition to link the model primitives to observable data moments that we derive from bunching estimation. Consider the marginal firm reporting $\hat{y} - \Delta\hat{y}$ when audit risk is constant at p_L, and denote its true revenues with $\hat{y}^* - \Delta\hat{y}^*$. When the audit rule is disclosed, this firm is indifferent

[30] The notch also creates a dominated area below $\hat{y}$ in the interval $\left[\hat{y} - \Delta\hat{y}^D, \hat{y}\right)$. The intuition is that firms that are close enough but below $\hat{y}$ can achieve two goals by reducing manipulation: decrease their manipulation cost and reduce the expected cost of auditing. It follows that the dominated area should have zero mass after disclosure. Frictions such as fixed moral costs from evasion, taxpayer knowledge limits, and low audit rule salience may all contribute to the existence of mass in the dominated area.

between choosing $e_H(\hat{y}^* - \Delta\hat{y}^*)$, the interior solution under the higher audit risk p_H, and a level of manipulation $\hat{e} = \hat{y}^* - \Delta\hat{y}^* - \hat{y}$ such that the firm locates at the threshold.

We define the value of the marginal firm at the notch $\hat{y}$ with $V^N \equiv V(\hat{e}, \hat{y}^* - \Delta\hat{y}^*, T_D(\hat{e}))$, and its value at the interior solution with high audit risk as $V^H \equiv V(e_H, \hat{y}^* - \Delta\hat{y}^*, T_H(e_H))$. Next, we assume the following iso-elastic form for manipulation costs:[31]

$$g(e) = \frac{k_e}{1+\frac{1}{\varepsilon_e}} \cdot \left(\frac{e}{k_e}\right)^{1+\frac{1}{\varepsilon_e}}, \tag{6.1}$$

where k_e is a scale factor and ε_e is the elasticity of manipulation with respect to $\tau^e \equiv T'(e_L) = \tau - \tau\gamma p_L$, the implicit tax rate on evasion. Along with the perceived audit risks, these two will be the structural parameters of interest.

Define $\Delta\tau^e \equiv T'(e_H) - T'(\hat{e}) = -\tau\gamma\Delta p$. As described in Appendix F, we can now use the value of the firm in (2.1), the optimality condition in (2.2), and the manipulation costs in (6.1) to express the indifference condition $V^N = V^H$ for any $\hat{y}$ as follows:[32]

$$\tau^e\left[-\frac{\Delta\hat{y}}{k_e} + (\tau^e)^{\varepsilon_e}\right] - \frac{\varepsilon_e}{1+\varepsilon_e}\left[-\frac{\Delta\hat{y}}{k_e} + (\tau^e)^{\varepsilon_e}\right]^{1+\frac{1}{\varepsilon_e}} - \frac{1}{1+\varepsilon_e}\left[\tau^e + \Delta\tau^e\right]^{1+\varepsilon_e} = 0. \tag{6.2}$$

The expression in (6.2) depends on the following primitives: k_e and ε_e, which capture the role in the manipulation decision of a firm's scale of operations and its cost function curvature, respectively; p_L and Δp, or equivalently p_L and p_H, which reflect the level of audit risk perceived on both sides of the SeS threshold. In addition, the condition rests on a set of observables, including the tax rate τ, the penalty rate γ, and the marginal buncher's revenue response $\Delta\hat{y}$, which we define as the length of the bunching segment.

6.3. **Alternative models.** Before moving to estimation, we discuss two alternative ways to generate bunching in a misreporting model: i) firms expect audit costs independent of evasion, ii) firms manipulate their presumed revenues by underreporting their costs. Appendix G discusses these alternatives more in depth. Crucially, these models have similar implications for the estimation of disclosure effects. By providing an additional motive for bunching, they would reduce the scope for the type of revenue misreporting elasticity that we estimate through equation (6.2).

[31] Convexity of $g(.)$ can reflect either of two processes: first, sustaining progressively higher levels of evasion may require disproportionate coordination between a business owner and their partners (suppliers and buyers) or their employees. Second, the guilt or potential shame from defying the law may increase disproportionately with the extent of wrongdoing. While we simplify $g(.)$ to have a common elasticity to guarantee its estimation, our structural analysis allows for heterogeneity by assigning different elasticity parameters across business sectors.

[32] $\hat{y}$ might not be taken as given by the firm, which could manipulate the information provided to the government in order to lower $\hat{y}$. If this is the case, the model above would still allow us to identify the structural manipulation elasticity as long as the cost of manipulating inputs is separable from the cost of manipulating revenues. Otherwise, we would obtain an upper bound of the true revenue elasticity, which delivers conservative estimates of the revenue improvements from audit rule disclosure.

Including a business cost of engaging with the tax administration in an audit would amount to adding an intercept in $T(\cdot)$ with no consequences for the intensive margin of evasion.[33] These costs would induce taxpayers below $\hat{y}$ to boost their report to reduce the expected tax liability by a multiple of Δp. Similarly, suppose firms could reduce $\hat{y}$ by decreasing reported business costs. By doing so, they would increase the tax base. In addition, if this extra margin of manipulation existed, the observed bunching would not only be the byproduct of the revenue elasticity ε.

Because the manipulation margins in these alternatives do not respond to uniform changes in a flat probability of audit, these models reduce the estimated revenue loss from the PRE compared to our baseline model. Since we quantify the IPE (revenue gains) directly in the data free of structural assumptions, we conclude that these models would deliver larger estimates of the revenue effects of disclosure.[34] For this reason, even if our model is parsimonious in the way it allows firms to respond to policy incentives, it delivers conservative estimates for the purposes of our analysis, providing a lower bound on the revenue gains from disclosure.

7. Structural estimation

7.1. **Simulated method of moments.** By the model in Section 6, different combinations of the values in the parameter set $K = \{k_e, \varepsilon_e, p_L, p_H\}$ result in different revenue responses through (6.2). We label the theoretical response by the marginal buncher in a given group of taxpayers as $\Delta\hat{y}^{Theory}$. For every taxpayer group, we can also quantify an empirical counterpart $\Delta\hat{y}^{Data}$, which can be determined from bunching estimates through the following formula:

$$B = \int_{\hat{y}-\Delta\hat{y}}^{\hat{y}} \hat{h}_L(y_L)\, dy_L \approx h_L(\hat{y}) \cdot \Delta\hat{y}, \tag{7.1}$$

where B denotes the excess mass at the bunching point $\hat{y}$, and in order to approximate the formula we assume that the counterfactual distribution $\hat{h}_L(y_L)$ is constant over the segment $(\hat{y} - \Delta\hat{y}, \hat{y})$. We approximate excess bunching with $\hat{b} = \frac{B}{\hat{h}_L(\hat{y})}$ as anticipated in Section 5. Therefore, finding the monetary value of revenue responses $\Delta\hat{y}$ simply requires to rescale bunching estimates by the width of the bins involved in the estimation.

Structural estimation follows a simulated GMM approach. The procedure iterates through suitable values of the parameter set $K = \{k_e, \varepsilon_e, p_L, p_H\}$ to jointly determine $\Delta\hat{y}^{Theory}$ as a solution to the indifference condition in (6.2). Of all candidate $\Delta\hat{y}^{Theory}$, we choose those that most closely match the corresponding $\Delta\hat{y}^{Data}$ according to a quadratic loss minimization:

$$\min_K L(K) = \sum_{i=1}^{N} \left(\Delta\hat{y}_i^{\,Theory} - \Delta\hat{y}_i^{\,Data}\right)^2. \tag{7.2}$$

[33] The expected avoided tax would be $T(e) = (\tau - \tau\gamma p) \cdot e + p \cdot a$, where a is the cost of undergoing an audit.

[34] More precisely, ignoring an additional margin of manipulation while fixing the size of revenue gains would result in larger estimates for ε, a smaller audit risk gap Δp, or both. In our framework, these both result in losses being relatively larger than gains from disclosure.

To estimate the relevant parameters, we need to ensure that their number does not exceed the number of empirical moments, so that $N \geqslant |K|$. We do so by appropriately choosing a sensible set of restrictions as discussed below.

To define our empirical moments, we focus on taxpayers who are subject to personal income taxation (PIT). While corporate income tax (CIT) rates are homogeneous throughout the country, PIT rates display considerable variation across regions and municipalities as local administrations can impose surcharges on the national tax rate schedule.[35] We thus divide PIT-payers into plausibly homogeneous groups along three dimensions: *i*) geography, based on twenty NUTS-2 administrative regions; *ii*) business sectors, grouping industries into manufacturing and construction, wholesale, retail, the professions, and a residual service category, including hospitality, restaurants, and other personal services; and *iii*) turnover size, clustering taxpayers by comparable scales of operations in terms of presumed revenue terciles within each region-sector pair. This results in $N = 300$ groups of PIT-payers, for which we estimate the revenue responses to the SeS threshold over the 2007-2010 tax period.[36]

Next, we impose restrictions on the parameters we wish to estimate in a way that reflects plausible fundamentals of the evasion and auditing process among SeS filers. First, we assign a pair $\{p_L, p_H\}$ to each of the twenty regions in the country. This approximates the decentralized nature of auditing activity in Italy, with the Italian Revenue Agency and *Guardia di Finanza* (Tax Police) relying on provincial and regional divisions to carry out day-to-day enforcement operations in pursuit of national enforcement goals.[37] Second, we let the manipulation elasticity ε_e vary across business sectors only. This choice is meant to capture the broad differences in the responsiveness to manipulation incentives across industries already documented in the evasion literature.[38] Third, we assign a common k_e to all groups in the same relative presumed revenue tercile. This parameter provides a role for turnover size in the manipulation cost function $g(.)$. The restriction should thus allow the data to reveal whether revenue concealment is more or less costly by scale of operations. After all these restrictions are imposed, we are left with $|K| = 48$ parameters to estimate.

Lastly, we calibrate the policy parameters τ and γ based on the prevailing local and national laws. We approximate τ as the average ratio between the PIT due across all SeS filers in a given taxpayer group and their reported gross profit. As tax rates vary, the economic incentives tied to crossing $\hat{y}$ change from group to group even aside from audit risks. For γ, we select the lower bound of the penalty range set by law. In our sample period, net

[35]For the 2007-2010 tax years, surcharges did not exceed the national PIT rates by more than 0.8% and 1.7% at the municipal and regional level, respectively, in line with the caps set by national law.

[36]For consistency with our theoretical model, each group's underlying bunching estimation is based on absolute deviations of reported revenues from presumed revenues, with SeS files grouped in bins of size €500.

[37]Galbiati and Zanella (2012) discuss how the allocation of auditing resources across local offices in Italy tends to be sticky. As a result, firms might expect stable differences in enforcement risk based on their location.

[38]For example, Pomeranz (2015), Almunia and Lopez-Rodriguez (2018), and Naritomi (2019) point to structural differences in exploiting evasion opportunities along the supply chain in VAT systems.

administrative fines on top of the recovered tax liability (that is, $\gamma - 1$ in our model) could vary between 100% and 200% of any detected evasion. We thus set $\gamma = 2$.[39]

Estimation is made possible by leveraging both the variation in revenue responses across taxpayer groups as well as the heterogeneity in tax incentives and in the ease of revenue manipulation faced by the businesses in our data. The intuition behind the identification of the parameters rests on the way that the identifying variation in bunching and taxes combines with the restrictions. Consider two groups of PIT-payers defined by their location, sector, and relative scale, with high and low bunching respectively. Assume that each administrative region corresponds to a separate enforcement environment due to local tax office staffing, budgets, and strategies. Next, focus on two alternative scenarios. First, assume that the two groups belong to different regions. If the two groups have similar evasion cost structures and face similar tax levels, then their bunching difference might provide us with information about their audit risk perceptions. Second, we can fix the enforcement environment and exploit the variation in taxes. If the two groups belong to the same region and have similar evasion cost structures, but face different tax incentives, relating the observed differences in bunching and taxes might yield insights on the firms' responsiveness parameters. Iterating this logic across all firm groups and restriction combinations, we may recover all parameter estimates.

7.2. **Structural estimates and equilibrium evasion.** Results from our structural estimation are summarized in Table 3. On average, SeS taxpayers behave as if perceiving an audit risk change $\Delta p = 4.8$ percentage points around the SeS presumed revenue threshold. While we cannot assess this figure against individual-level audit data, our result is close to the probability change implied by the aggregate data released by the Italian Revenue Agency. For the 2007-2010 tax period, audit tabulations in D'Agosto et al. (2017) suggest that small businesses with revenues above the SeS threshold could benefit from an audit risk reduction of about 3.4 percentage points. This corresponds to realized audit risks equal to 7.1% and 10.5% above and below the SeS revenue threshold, respectively, which are slightly smaller but in the same order of magnitude as the corresponding perceived risks that we estimate to be at $p_L = 10.8\%$ and $p_H = 15.6\%$ on average across regions.

Our parameter estimates help to portray the extent of revenue underreporting among the self-employed and small firms in our sample. Since our goal is to quantify the reported revenue effect of disclosure, we first compute the equilibrium amount of revenue evasion that would obtain in each taxpayer group if all SeS filers perceived a similar constant risk of audit. We set this risk to p_L, depending on the regional estimate, and follow Appendix F to obtain equilibrium evasion as:

$$e_L = k_e \left(\tau - \tau\gamma p_L\right)^{\varepsilon_e} . \tag{7.3}$$

[39]Given that we exclude from the analysis any taxpayer reporting $y = 0$, we don't consider the higher fines associated to missing declarations, ranging from 120% to 240% of any uncovered tax liability.

We then normalize each group's e_L by its median reported gross profits, so as to convey the magnitude of revenue underreporting relative to a representative, observable, tax-relevant quantity.[40] We report the average evasion rates across structural groups in Table 4. If all taxpayers perceived the relatively low risk p_L absent disclosure (*i.e.* an upper bound to manipulation), revenue underreporting would be on average in excess of one third of median reported profits (37.3%).[41]

We plot the distribution of estimated evasion rates in Figure 7. Different levels of expected tax savings from evasion as well as different estimated responsiveness to underreporting incentives produce considerable heterogeneity in the incidence of manipulation across structural groups. Consistent with the literature on the distribution of evasion in Italy, Appendix H shows that our model correctly predicts higher evasion intensity among Southern regions and downstream sectors.

8. Disclosure effects I: counterfactual evaluation

We now assess the reported revenue effects of audit rule disclosure in the SeS. To do so, we compare the mean reported revenues under disclosure and under secrecy, that is, if all taxpayers perceived the same constant audit risk. We find that revealing the exact location of the SeS threshold improves revenue compliance on average. Under our modeling assumptions, the analysis suggests that disclosure is likely to succeed when taxpayers are relatively more sensitive to audit incentives than tax rates, and that the average welfare cost of disclosure is likely limited.

8.1. **Revenue effects of disclosure.** In line with Section 2, we evaluate the effects of disclosure by comparing three mean reported revenue levels. We label $\bar{y}_D$ the average revenues reported across our 300 structural groups of PIT-payers over 2007-2010.[42] We further label $\bar{y}_C$ the revenues reported on average in the policy counterfactuals where perceived audit risks are constant at p_C, and $\bar{y}_L$ the special case where $p_C = p_L$.[43]

We generate policy counterfactuals by computing $\bar{y}_C$ with progressively higher levels of audit risk relative to that perceived in the bunching counterfactual. This allows for disclosure to have both incentive provision (revenue gains) and probability reduction (revenue losses)

[40] The tax evasion literature often resorts to computing a potential tax base to use as denominator, adjusting reported quantities for different sources of misreporting. Our model on the other hand doesn't directly capture cost manipulation. We thus solely rely on gross profits rather than adding assumptions on the nature of input misreporting for the sake of this exercise.

[41] We provide the corresponding absolute evasion distribution in Figure A15. As shown in Table 4, a PIT-payer underreports about €7,170 in revenues on average across all groups.

[42] Averages are weighted by the relative share of SeS files contributing to the analysis in each group.

[43] Counterfactual averages are obtained in three steps. First, we compute the average level of true revenues $\bar{y}_j^* = \bar{y}_{j,L} + e_{j,L}$ for each structural group j as the sum of the average revenues reported in the bunching counterfactual and the average equilibrium evasion when $p_C = p_L$. Note that by our model true revenues are invariant to audit risks. Second, we obtain $\bar{y}_{j,C}$ by subtracting the equilibrium evasion with p_C from $\bar{y}_j^*$. Third, we average over all js based on their relative numerosity in the counterfactual scenario under study.

effects, as audit risks rise or drop around the threshold, respectively. To mimic the possibility that pre-disclosure conditions differ across groups, we consider each group's estimated regional p_L and p_H, and let $p_C = p_L + \Delta p_C$. $\Delta p_C > 0$ reflects a progressively higher share of the local difference between p_H and p_L, the two extreme of what we label as the sensible counterfactual range.

Figure 8, Panel A delivers our main result for PIT-payers over 2007-2010. Relative to the mean revenues reported in the relevant counterfactuals, disclosing the location of SeS presumed revenues raises mean reported revenues by 7.7% when $p_C = p_L$ and by 6.3% when $p_C = p_H$, the highest risk within the sensible range (the blue band in the graph).

By construction, the higher the perceived audit risk is ahead of disclosure, the smaller the benefit from revealing the SeS audit selection criteria. In the language of Section 2, this corresponds to a larger probability reduction effect or reported revenue loss. While in our analysis the incentive provision effect or revenue gains are constant at $\bar{y}_D - \bar{y}_L$, revenue losses $\bar{y}_C - \bar{y}_L$ increase with the length of the horizontal shift that the policy counterfactual has to undergo to match the bunching counterfactual. We assess the growing size of losses in terms of the gains in Figure 8, Panel B. Despite its steady growth along the sensible counterfactual range, we estimate that the probability reduction effect of disclosure overturns less than 20% of the positive incentive provision effect.

8.2. **Discussion and interpretation.** The conceptual framework in Section 2 helps to interpret the finding that disclosure improves compliance. The policy structure we study assigns a different relative importance to elasticities and perceived audit risks within the gains and losses from disclosure. We decompose our results with these differences in mind.

Reporting elasticities hold a relatively more important role in determining the probability reduction effect. While firms freely respond to any drop in audit risks induced by disclosure with a full horizontal shift from $\bar{y}_C$ to $\bar{y}_L$, the incentive provision effect raises revenues up to the location of the disclosed threshold. This implies that any revenue-raising adjustment regulated by the elasticity is limited by the audit rule cutoff. On the other hand, audit incentives have a relatively larger impact on revenue gains than losses: the difference in audit risks $\Delta p = p_H - p_L$ drawing firms to the threshold from below is at least weakly larger than the audit risk drop $\Delta p_C = p_C - p_L$ that might reduce compliance upon disclosure.

Taken together, these observations imply that the revenue gains from disclosure are more likely to exceed the losses when elasticities are small in the face of the substantial bunching we observe. This is because in that case audit incentives, which hold a relatively larger role in revenue gains, would be more likely to explain the observed bunching. Kleven et al. (2011) use audit data and quasi-experimental variation to argue that the elasticity of evasion responses to tax rates among Danish PIT payers is modest. In our context, to the extent that the enforcement environment is set at the regional level as we assume in the structural model, we might draw suggestive evidence on how elastic firms are from *subregional* patterns in bunching and tax rates. Figure A20 shows the results of region fixed-effects regressions of the SeS log-revenue responses on the logarithm of the average municipal PIT surcharge rate across taxpayer groups. Tax rate coefficients are generally positive but very low in practice, even accounting for the imprecision of the estimates.[44]

The large bunching we observe in the data, coupled with the low implied elasticities, suggest that the signal provided by the tax administration through audit rule disclosure and the firms responsiveness to it might be the main drivers behind the policy effects. This echoes the conclusions in Kleven and Waseem (2013) on the effect of notches in the Pakistani PIT schedule. In that case, the authors reconcile the large observed bunching with their small estimated elasticities by pointing to the sizable distortions induced by the average tax rate jumps that the local tax law ties to specified income thresholds. In the case of disclosure, to the extent that the size of the perceived audit risk gaps generated by the policy dominate firms' responsiveness to taxes, revealing an audit rule of the kind we examine is likely to improve net compliance.

8.3. **Cost effectiveness and welfare.** We can compute the cost effectiveness of SeS disclosure with a conservative back-of-the-envelope calculation. For the tax revenue potential, we consider disclosure in the scenario with the minimum revenue effect (6.3%), where $p_C = p_H$ and $\bar{y}_C = \bar{y}_H$. We apply the average due PIT-rates to the extra reported profit that would be generated by disclosure. Specifically, we convert revenues into profits assuming that the observed profit-to-revenues ratio was not affected by the introduction of SeS.[45] We then aggregate this average effect across all SeS files involved in our structural analysis in any given year, and divide by the total value of *SOSE*'s production as an upper bound to the administrative cost of disclosure.[46] Overall, every Euro spent on SeS implementation would generate:

$$\text{(8.1)} \qquad \begin{aligned} \text{SeS effectiveness} &= \frac{\bar{\tau}\cdot\left(\min\left\{\frac{\bar{y}_D-\bar{y}_C}{\bar{y}_C}\right\}\right)\cdot\left(\frac{\bar{\pi}}{\bar{y}}\right)\cdot\bar{y}_H\cdot N}{\text{Administrative Costs}} \\ &= €64.21 \end{aligned}$$

among filing PIT-payers in any year of the tax period 2007-2010.

[44]Appendix I provides a simple microfoundation of the negative relationship between audit risks and reporting elasticities for a given bunching. In particular, we show that the level of correlation between bunching and tax rates might be directly informative for the audit risk gaps that explain the observed bunching.

[45]Table 1 reports the mean tax rates, profits, and revenues that we observe for our structural analysis. We assume the profit rate to be the same as under disclosure since our model cannot predict profit across counterfactuals. Yet, this is more conservative than dividing observed mean profits by counterfactual mean revenues, which are lower than the observed mean.

[46]This value was €12.6 million in 2010 (SOSE, 2011). We do not account for the cost of running SeS-based audits, as there is no separate, well-defined budget or auditor group dedicated solely to this end. Most importantly, we focus on the disclosure component of SeS, which does not imply the SeS policy increases the volume of audits in and of itself.

While this is desirable for a collection-maximizing administration, disclosure might have opposite welfare implications for different taxpayers. This is because revealing the threshold might raise or reduce the perceived relative risk depending on each taxpayer's counterfactual location. To account for this, we estimate the policy's marginal value of public funds (MVPF) as suggested by Hendren and Sprung-Keyser (2020).[47]

We define the MVPF as the ratio between the average taxpayers' willingness to pay for disclosure and the mean net cost of disclosure from the perspective of the tax administration. The numerator includes *i*) a positive willingness to pay by taxpayers who perceive an audit risk reduction, and *ii*) a negative willingness to pay (that is, a willingness to avoid the policy change) by taxpayers with an audit risk hike. The sign of the numerator depends in turn on *i*) the relative size of the two groups of taxpayers in the counterfactual, and *ii*) the magnitude of the involved willingness to pay, which is proportional to the perceived audit risk changes. On the other hand, the mean net cost of disclosure will be stably negative, since the potential tax gains from disclosure implied by our calculations greatly outweigh the administrative costs (as showed in 8.1). Appendix J details our estimation procedure.

Figure 9 presents the aggregate results of our MVPF analysis, which we otherwise decompose in Table A5. We display each scenario's MVPF ratio on the vertical axis, and focus on the range of counterfactual risks we deem sensible (the blue band in the graph). Our ratio estimates are generally small in absolute value (within the sensible range, 0.04 at the highest when $p_C = p_L$, and -0.042 at the lowest when $p_C = p_H$). This partly results from our policy experiment: the opposite effects of disclosure across taxpayers imply that its average welfare cost is bound to be modest.[48] In addition, estimates in the second half of the sensible range are negative. The MVPF in these cases is infinite: disclosure is raising revenues generating a positive total willingness to pay. Overall, this exercise suggests that the revenue gains from disclosure do not come at significant welfare costs.[49]

9. Disclosure effects II: natural experiment

While we estimate a positive effect of SeS disclosure on reported revenues, firms might still raise their reported costs to neutralize the tax base effects of higher revenues (Carrillo et al., 2017b). Unfortunately, the structure of SeS does not lend itself to a clear way to model the cost and profit responses of firms, since the threshold we study is in terms of revenues and

[47] For a discussion on the relationship between the MVPF and other welfare measures, we refer the reader to Hendren (2016).

[48] In Section G.1.4 in the Appendix, we extend this reasoning to the case where taxpayers perceive additional audit costs independent of intended evasion behavior. Our main conclusion should be robust to reasonable assumptions on the distribution of these audit costs around the presumed revenue threshold.

[49] Welfare comparisons across policies might require trading off the gains and losses across taxpayers below and above the threshold in the policy counterfactual, for example by assigning different welfare weights across these two groups. However, given that SeS presumed revenues are idiosyncratic to the firm in the filing year, we do not see any special justification for welfare weight assignments based on this particular policy threshold.

depends on reported costs and inputs. We thus turn to a natural experiment to assess the impact of disclosure-based policies on a broader set of compliance margins. Conveniently, the staggered introduction of the 2011 reward regime ("*regime premiale*") closely resembles the logic of our original policy exercise, where disclosure affects reporting incentives in opposite directions depending on the relative position of each taxpayer. In line with our conceptual framework, we show that taxpayers approach their presumed revenues from both sides of the threshold. Still, mean gross profits increase in response to the reform, showing that tax authorities can expand the tax base by strengthening the incentives associated to a disclosed audit rule.

9.1. **The 2011 reward regime.** Starting in 2011, the Italian government has promised stronger audit exemptions for taxpayers complying with SeS prescriptions, while threatening the others with higher chances of enforcement. Figure 10 sketches the logic of the reform. Similar to the audit risk effects of disclosure in Figure 1, the new regime influenced audit risk perceptions in opposite ways depending on the relative location of the taxpayer. Those planning on reporting more revenues than presumed while complying with several accounting indicators put forth by the tax authority would experience comparatively greater protection from enforcement for that year's report ($p_{L,reward} \leqslant p_L$). On the other hand, the reform encouraged greater scrutiny over those failing to comply ($p_{H,reward} \geqslant p_H$). The combination of these measures implies that $\Delta p_{reward} \geqslant \Delta p$, that is, taxpayers should perceive a larger audit risk discontinuity at the presumed revenue threshold after the reform.

We exploit the staggered inclusion of SeS sectors into the reward regime over the 2011-2016 (Figure A21) period to evaluate the reform's effects. We focus on businesses in 155 treated sectors across manufacturing, commerce, services, and the skilled professions, and create a balanced panel of those continuously filing for SeS over the 2007-2016 decade.[50] Since we observe each SeS sector s entering the regime in a specific tax year t, we set up an event-study design to estimate equations with the following structure:

$$y_{s,t} = \lambda_s + \gamma_t + \sum_{q=-k}^{+k'} \beta_q \cdot I\left(Q_{s,t} = q\right) + \sum_{r=2007}^{2016} \delta_r \cdot X_s \cdot I\left(t = r\right) + \varepsilon_{s,t}. \tag{9.1}$$

For any given sector-by-tax year outcome $y_{s,t}$ covered below, coefficients β_q capture the effect of including a sector into the reward regime in each period q relative to sector entry. Identification of these effects relies on a parallel path assumption. Specifically, we assume that outcomes in a sector currently under treatment would evolve in a similar fashion to those in yet-to-be-treated sectors absent the reform. We further control for sector and tax

[50]Differently from the other SeS macro-industries, the Revenue Agency has included only three out of twenty-four SeS sectors in the skilled professions by 2016, only to overhaul the SeS system for all sectors in 2018. Our results may thus not be fully representative for all professional groups. In addition, panel balancing tends to overrepresent businesses with larger size and better SeS compliance, as shown in Table A1.

year fixed effects λ and γ, respectively, and a vector of pre-treatment features summarized by X, interacted with tax year dummy variables.[51] Lastly, we weight each sector by the number of SeS files submitted at the outset of our sample period, and cluster standard errors at the sector level, following the recommendation in Bertrand et al. (2004) for treatment-level clustering.[52]

9.2. **Distribution shifts.** As outlined in Section 2, disclosure-based policies such as the reward regime might reduce perceived risks above the revealed threshold and raise them below. As a result, bunching may come from taxpayer adjustments of opposite signs, with different implications for relative compliance.

The introduction of the reward regime provides a chance to assesses whether bunchers originate both from below and above the SeS threshold. In the data, we group taxpayers by their relative distance from the presumed revenues in the year before their sector's reform. We set up six symmetric categories of filers around $\hat{y}$, based on whether they reported revenues within 5, 5 to 10, or more than 10 percentage points from what presumed just before the reform. For each of these six groups, we measure the share of files located in each one percentage point bin in every year. We then estimate (9.1) using these shares as outcomes of separate event-studies around a sector's regime entry.

Figure 11 shows the results. In each panel, we plot for each one percentage point bin the average of the six treatment coefficients β_q and the 95% confidence interval of this linear combination. In the background, a green band marks the range where each group was located the year before the introduction of the reward system.

A stark pattern emerges: whether taxpayers start out below $\hat{y}$ or not, the reform's larger risk gap at $\hat{y}$ draws a larger number of them to their threshold or just above it. In addition, the stronger drop in bin shares for bins below but closer to the threshold is, all else equal, consistent with a lower cost of achieving congruence for those having to travel a relatively shorter distance. These patterns are consistent with our theory: taxpayers facing an increase in risks below the threshold tend to raise their relative compliance, while those awarded stronger protections tend to reduce it.

9.3. **Tax base effects.** The tax base effect of disclosure in our context is ambiguous for at least three reasons. First, the opposite relative revenue shifts induced by disclosure could be revenue-reducing on the net. Second, taxpayers might comply with a presumed revenue threshold by either raising their revenues or by cutting their reported inputs and costs, since

[51]Controls include dummies for the categories of manufactures, commerce, services, and the professions as defined by the Revenue Agency; and 2007-2010 averages for a set of variables including revenues, gross profits, the incidence of employment costs on turnover, and yearly growths of employment cost rates and revenues.
[52]Weighting by the number of files submitted allows us to capture the behavior of the average taxpayer in our data.

this might reduce presumed revenues.[53] Third, taxpayers might also opt for raising their costs along with their revenues to avoid reporting larger profits. The fact that the reward regime tied its new audit benefits to input reporting compliance (the criteria of normality and coherence as seen in TableA2) might further affect the reported tax base.

To address these concerns, we exploit the adoption of the reward regime to study the period-by-period mean effect of disclosure along a number of reported margins. Figure 12 Panels A and B first show the full set of β_q coefficients from (9.1) when the outcome is mean reported revenues by sector and tax year (in logarithms and Euros, respectively). Ahead of the reform, treated sectors report slightly less revenues on average, but the path is fairly stable as we approach the reform period. After a sector's reform, reported revenues are on average 2.4% higher than in control sectors in the first year, and up by about 20.4% by year six.[54]

Next, Panels C and D study net reporting behavior in terms of gross profits. Just as for revenues, the stronger audit incentives introduced by the reward regime appear to have stimulated the emergence of a larger tax base. On average each year, firms in a treated sector report 16.2% higher gross profits than those in sectors still to treat. The pattern of coefficients is once again increasing, suggesting that familiarity with the new system improves compliance over time. Overall, our estimates imply that the reform encouraged a cumulative gain of €33,671.77 in taxable profits from the average treated business.

The profit increase we document is however smaller in magnitude than that in revenues. Figure 13, Panels A and B summarize the effect of the reward system on the difference between revenues and profits, which provides an aggregate measure of the costs reported in each SeS file. The resulting patterns are remarkably similar to those in the previous figure, with treated sectors reporting average costs from 2% to 20.7% higher than in control sectors in the first and in the last available year, respectively. Part of this increase is seemingly driven by businesses' revisions to their employment cost margins, which we directly observe in the data. Panels C and D show that taxpayers report both a larger mean cost of employment and a (weakly) higher number of employees as a result of the introduction of the reward regime.

Abstracting from the details of SeS implementation, our results might point to a desirable feature of audit rules with predicted revenue thresholds. Revenue prediction models tend to presume higher turnover for firms mobilizing larger resources. For a given reported revenue level, larger reported costs translate into higher revenue thresholds, making it harder for the firm to obtain the desired audit exemption. This design might limit the scope for cost adjustments that keep the tax base and tax liability constant when reported revenues increase.

[53] As long as the coefficients in the revenue prediction model estimated by the tax authority are smaller than 1 on the most manipulable inputs, taxpayers would find it more efficient to raise revenues than to cut costs. Adjusting either margin by the same Euro amount would result in the same income tax cost, but differ in its effect on the distance $y - \hat{y}(X)$.

[54] Since our main outcomes are specified as averages, we do not worry that their increase is driven by extensive margin responses, which we do not model in our framework.

While we show that revenue responses were indeed larger than implied cost adjustments in the case of the reward regime, this logic might extend to the counterfactual analysis of SeS disclosure in the first part of the paper.

Recent contributions on two-way fixed effects estimation have elucidated a number of potential issues in interpreting the dynamic treatment coefficients of standard event-study designs. To address these concerns, Appendix K replicates the estimation with the robust estimators in Sun and Abraham (2020) and de Chaisemartin and D'Haultfœuille (2020), obtaining a similar pattern of results to those in our baseline.

10. Conclusion

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 applying structural and quasi-experimental techniques to the case of Sector Studies (SeS), an Italian policy informing small firms of their relative audit risk around a revenue threshold.

The distribution of SeS files reveals that taxpayers are especially aware of and willing to adjust to clear audit risk signals. For the vast share of SeS filers subject to personal income taxation in 2007-2010, we find that mean reported revenues under disclosure are 6.3-7.7% higher than in several counterfactual scenarios with constant audit risks. We complement these results studying the gross profit effects of disclosure across all firm types. To do so, 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.

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. We leave the study of realized rather than threatened audits, as well as the optimal design of disclosure, to future research.

References

Aghion, P., U. Akcigit, M. Lequien, and S. Stantcheva (2017). "Tax Simplicity and Heterogeneous Learning". *NBER Working Paper No. 24049*.

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, pp. 1–38.

Arachi, G. and A. Santoro (2007). "Tax Enforcement for SMEs: Lessons from theItalian Experience?" *eJournal of Tax Research (Michigan Issue)* 5.2, pp. 225–243.

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, pp. 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, 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*.

Bertrand, M., E. Duflo, and S. Mullainathan (2004). "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics* 119.4, pp. 249–275.

Borusyak, K., X. Jaravel, and J. Spiess (2021). "Revisiting Event Study Designs: Robust and Efficient Estimation". *working paper*.

Brosio, G., A. Cassone, and R. Ricciuti (2002). "Tax Evasion across Italy: Rational Noncompliance or Inadequate Civic Concern?" *Public Choice* 112, pp. 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., E. Castro, and C. Scartascini (2017a). "Do Rewards Work? Evidence from the Randomization of Public Works". *IDB Working Paper IDB-WP-794*.

Carrillo, P. E., D. Pomeranz, and M. Singhal (2017b). "Dodging the Taxman: Firm Misreporting and Limits to Tax Enforcement". *American Economic Journal: Applied Economics* 9.2, pp. 144–64.

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, pp. 273–314.

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

Chetty, R., J. Friedman, T. Olsen, and L. Pistaferri (2011). "Adjustment Costs, Firm Responses, andMicro vs. Macro Labor Supply Elasticities:Evidence from Danish Tax Records". *The Quarterly Journal of Economics* 126.2, pp. 749–804.

Chetty, R., J. Friedman, and E. Saez (2013). “Using Differences in Knowledge Across Neighborhoods to Uncover the Impacts of the EITC on Earnings”. *American Economic Review* 103.7, pp. 2683–2721.

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, pp. 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*.

Galbiati, R. and G. Zanella (2012). “The Tax Evasion Social Multiplier: Evidence from Italy”. *Journal of Public Economics* 96, pp. 485–494.

Hendren, N. (2016). “The Policy Elasticity”. *Tax Policy and the Economy* 30.1, pp. 51–89.

Hendren, N. and B. Sprung-Keyser (2020). “A Unified Welfare Analysis of Government Policies”. *The Quarterly Journal of Economics* 135.3, pp. 1209–1318.

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.

IRS (2016). “Federal Tax Compliance Research: Tax Gap Estimates for Tax Years 2008–2010”. *Research, Analysis Statistics, Publication 1415 (Rev. 5-2016)*.

– (2019). “Federal Tax Compliance Research: Tax Gap Estimates for Tax Years 2011–2013”. *Research, Applied Analytics Statistics, Publication 1415 (Rev. 9-2019)*.

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.

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, pp. 651–692.

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, pp. 669–723.

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, pp. 881–896.

Lazear, E. (2006). "Speeding, Terrorism, and Teaching to the Test". *The Quarterly Journal of Economics* 121.3, pp. 1029–1061.

Ministry of Economy and Finance (2011). "Economia Non Osservata e Flussi Finanziari. Rapporto Finale sull'Attività del Gruppo di Lavoro sull'Economia Sommersa e i Flussi Finanziari".

– (2016). "Relazione sull'Economia Non Osservata e sull'Evasione Fiscale e Contributiva".

– (2019). "Relazione sull'Economia Non Osservata e sull'Evasione Fiscale e Contributiva".

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

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

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

– (2019). *Tax Administration 2019: Comparative Information on OECD and Other Advanced and Emerging Economies*. OECD Publishing, Paris.

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, pp. 2539–2569.

Reinganum, J. and L. Wilde (1985). "Income Tax Compliance in a Principal-Agent Framework". *Journal of Public Economics* 26.1, pp. 1–18.

Saez, E. (2010). "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy* 2, pp. 180–212.

Sánchez, I. and J. Sobel (1993). "Hierarchical Design and Enforcement of Income Tax Policies". *Journal of Public Economics* 50.3, pp. 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.*

– (2017). "Do Small Businesses Respond to a Change in Tax Audit Rules? Evidence from Italy". *Public Finance Review* 45, pp. 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, pp. 904–954.

SOSE (2011). "Bilancio dell'esercizio 2010", available at https://www.sose.it/sites/default/files/10184-1809907–SOSEBilancio2010.pdf.

Sun, L. and S. Abraham (2020). "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects". *Journal of Econometrics* forthcoming.

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

Thuronyi, V. (1996). *Presumptive Taxation*. Ed. by Victor Thuronyi. in Tax Law Design and Drafting, International Monetary Fund.

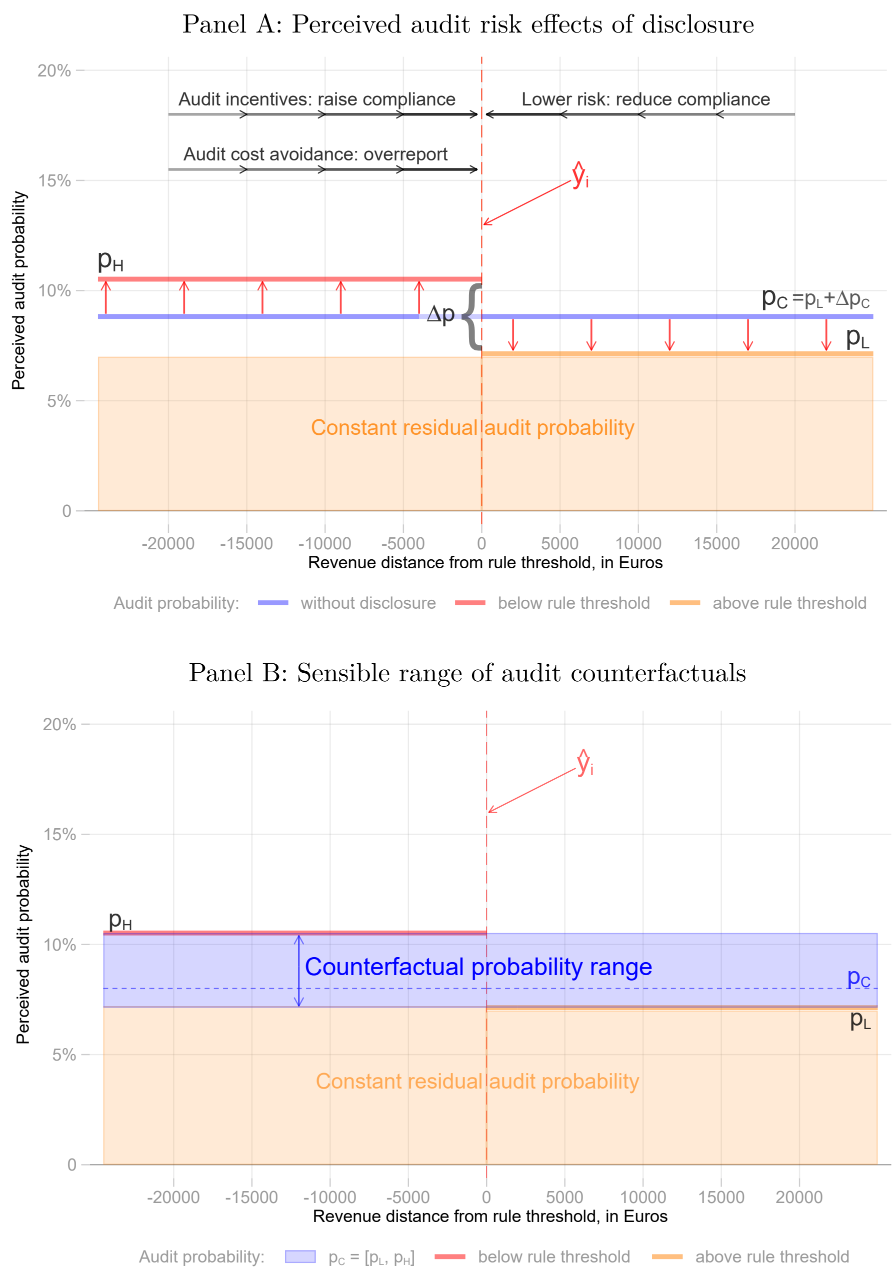
Vallanti, G. and G. Gianfreda (2020). "Informality, Regulation and Productivity: Do Small Firms Escape EPL through Shadow Employment?" *Small Business Economics*.

Wong, W. (2021). "Optimal Monitoring and Bureaucrat Adjustments". *working paper*.

Zanella, G. (2012). "Un Fatto Semplice-Semplice sull'Evasione Fiscale". *commentary on www.noisefromamerika.org, February 1, 2012.*

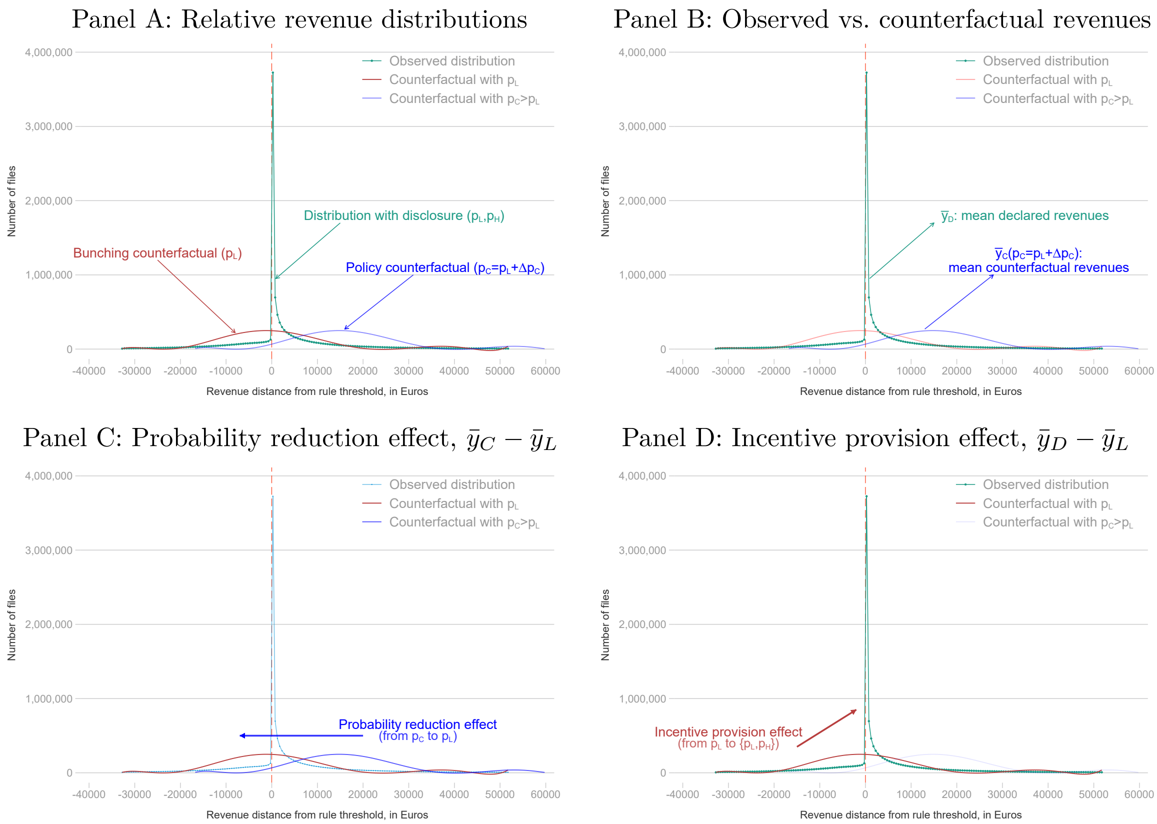
Figures

Figure 1. Perceived audit risk without and with audit rule disclosure



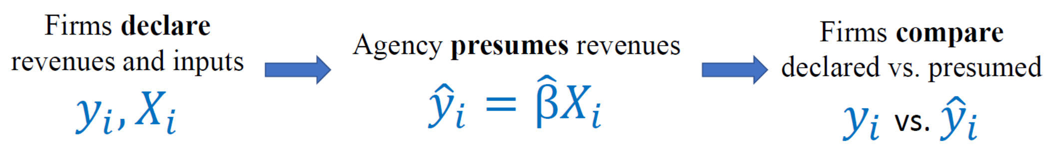
Notes: in this Figure, Panel A shows the probability of undergoing an audit as perceived by firms around a certain audit rule threshold $\hat{y}$ (the dashed red line), with and without disclosure. When the threshold is not known, taxpayers perceive a generic audit risk p_C (the blue horizontal line). Upon disclosure, taxpayers above the threshold, who are exempted by the type of audit governed by the rule, may lower their audit risk perceptions to p_L (the orange horizontal line) and reduce their compliance. Taxpayers planning to report revenues below the threshold may perceive a higher risk p_H (the red horizontal line), and raise their report to reduce the expected costs associated to audits and evasion detection. To the extent that audit risks change around the threshold as described, the difference between p_H and p_L will be weakly larger than that between p_C and p_L. We can thus define a sensible counterfactual probability range as in Panel B (the blue horizontal band), where p_C lies between the values of p_L and p_H perceived by the taxpayers.

Figure 2. Anatomy of disclosure: policy effects decomposition



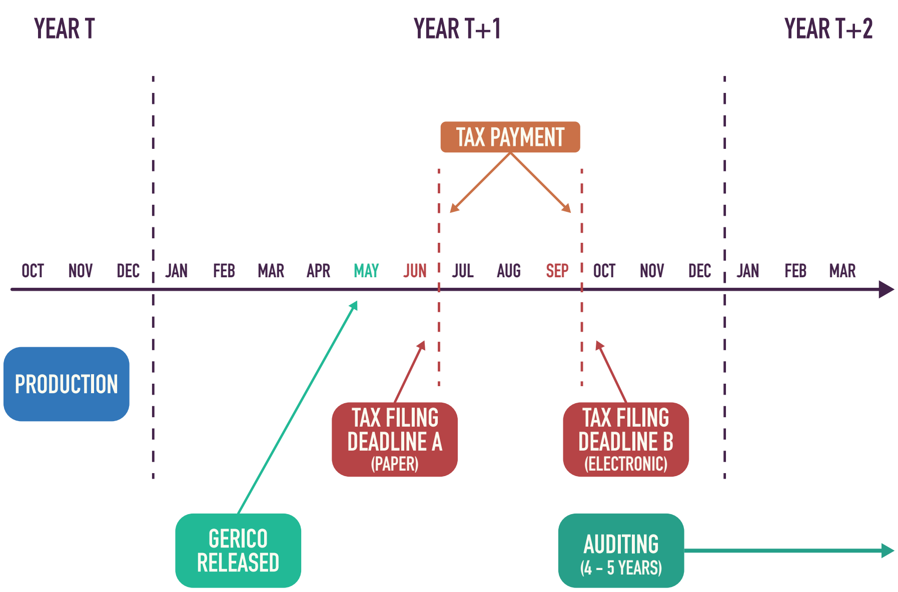
Notes: the Figure shows our conceptual decomposition of the effects of disclosure looking at three relative revenue distributions. Panel A shows an observed distribution of revenues relative to a disclosed audit rule threshold (in green), the smooth bunching counterfactual of all decisions taken when $p_C = p_L$ (in red), and a generic policy counterfactual reflecting all decisions taken when $p_C > p_L$ (in blue). Panel B highlights the two components behind the main policy effect we seek to estimate, namely the mean revenues observed under disclosure ($\bar{y}_D$) and the mean counterfactual revenues we might reconstruct based on audit risk and firm responsiveness parameters ($\bar{y}_C$). Panel C and Panel D decompose the conceptual distribution shifts that connect the policy counterfactual to the observed distribution. Panel C displays the revenue-reducing probability reduction effect, whereby the policy counterfactual shifts leftward to the location of the bunching counterfactual as taxpayers lower their audit risk perceptions from p_C to p_L. Panel D displays the revenue-increasing incentive provision effect, whereby the bunching counterfactual morphs into the observed distribution as taxpayers below the threshold face a perceived audit risk hike from p_L to p_H and the possibility of bunching to the known threshold. In all graphs, the observed distribution reflects the absolute difference between reported and presumed revenues in the 2007-2010 universe of Sector Study single-sector filers.

Figure 3. The process of Sector Studies



Notes: this Figure summarizes the implementation process of Sector Studies (SeS). Each year, Italian small businesses with turnover no greater than €5.2 million submit a dedicated SeS file to the Italian Revenue Agency. Each file contains detailed information on their past year's activity, including the revenues y_i they generated and accounting and structural information which we summarize with vectorX_i. The Agency's statistical partner company SOSE estimates sector-specific presumed revenue functions, along with a set of ancillary accounting indicators to detect anomalies in reporting. The core estimation relies on a GLS regression and is performed at least once every three years with different schedules for each sector. Regression coefficients $\widehat{\beta}$ are included in a software called *Gerico*, which is released ahead of every tax season. *Gerico* helps businesses both to produce their SeS file and to compute their presumed revenues $\hat{y}_i$ based on reported inputs X_i. Businesses can thus decide to adjust their reported revenues y_i to the amount presumed by *Gerico* $\hat{y}_i$ ahead of filing.

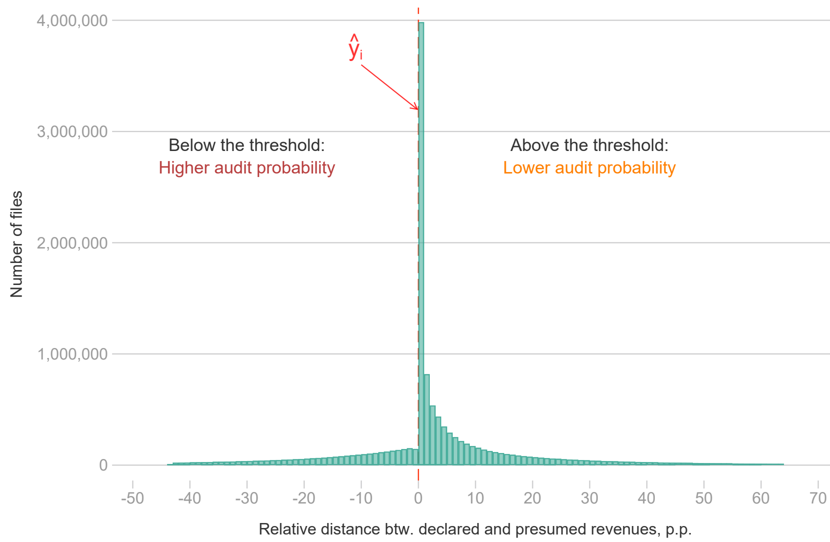
Figure 4. Timeline: taxpayer perspective



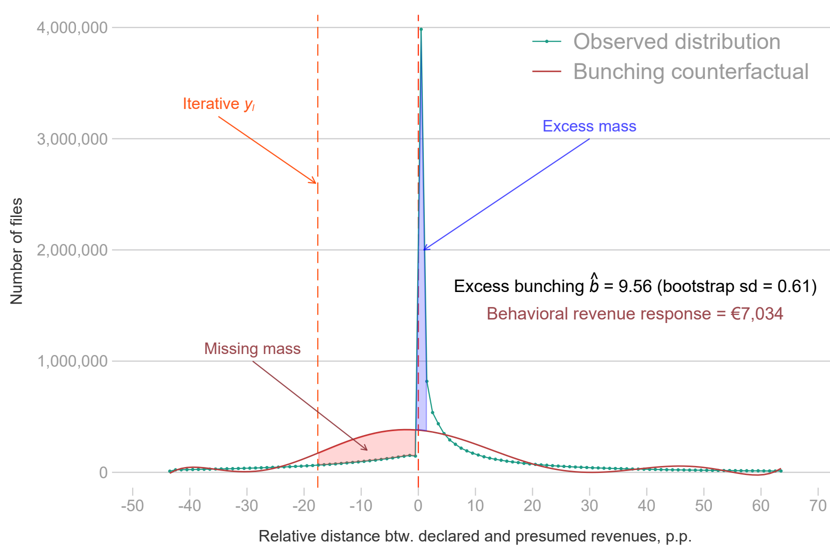
Notes: this Figure outlines the timeline of production and tax enforcement events from the perspective of taxpayers. Businesses generating revenues during year T are required to file their tax returns as well as their separate SeS file during the following year $T + 1$. SeS filing follows the tax filing 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 5. Bunching in the universe of single-sector SeS filers, 2007-2010

Panel A: Reported revenues relative to SeS threshold

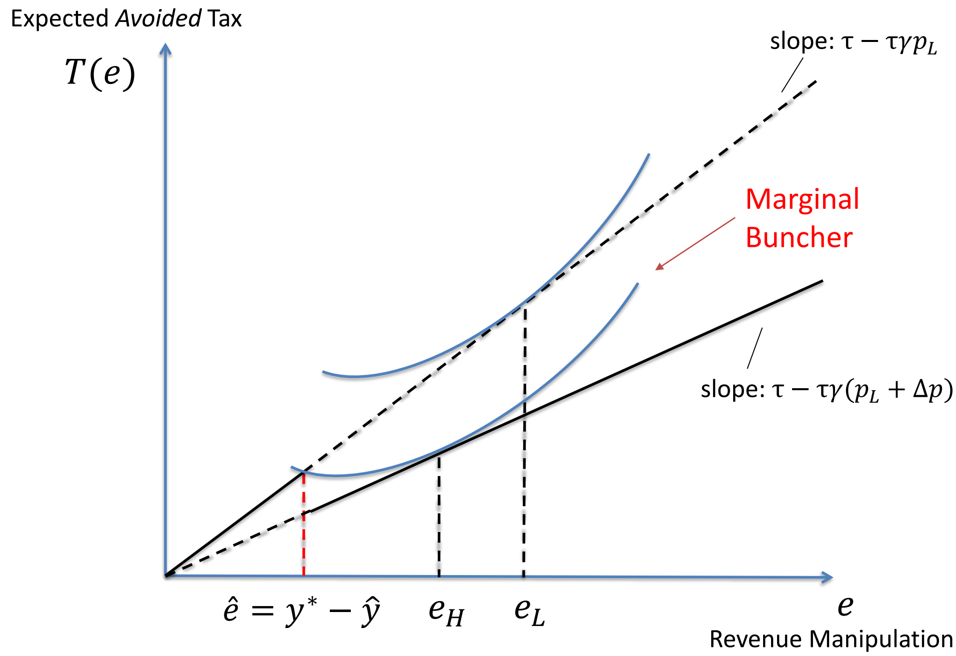


Panel B: Bunching estimates



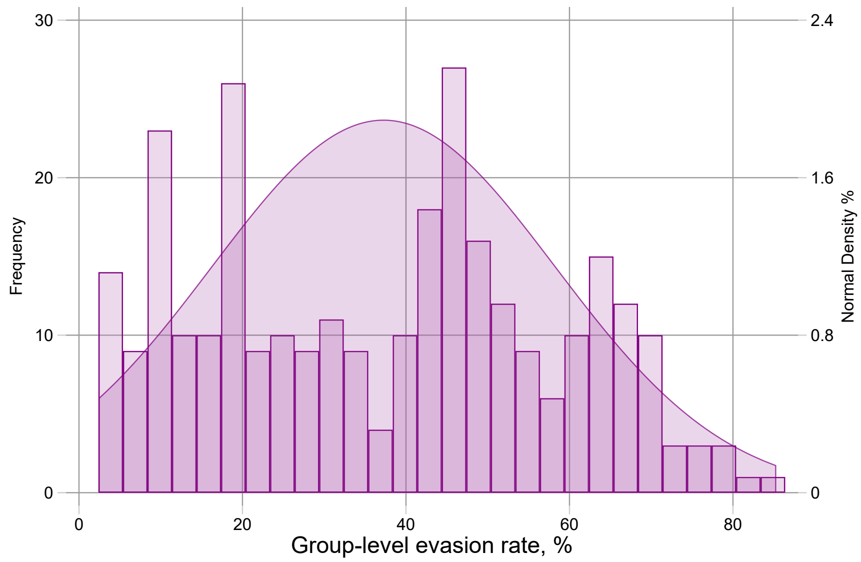
Notes: the Figure presents the distribution of $\frac{y_i - \hat{y}_i}{\hat{y}_i}$, the relative distance between reported revenues y_i and presumed revenues $\hat{y}_i$, from each SeS file in the universe of single-sector businesses in the 2007-2010 tax years. Units on the horizontal axis are percentage 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 6. Indifference condition of the SeS marginal buncher



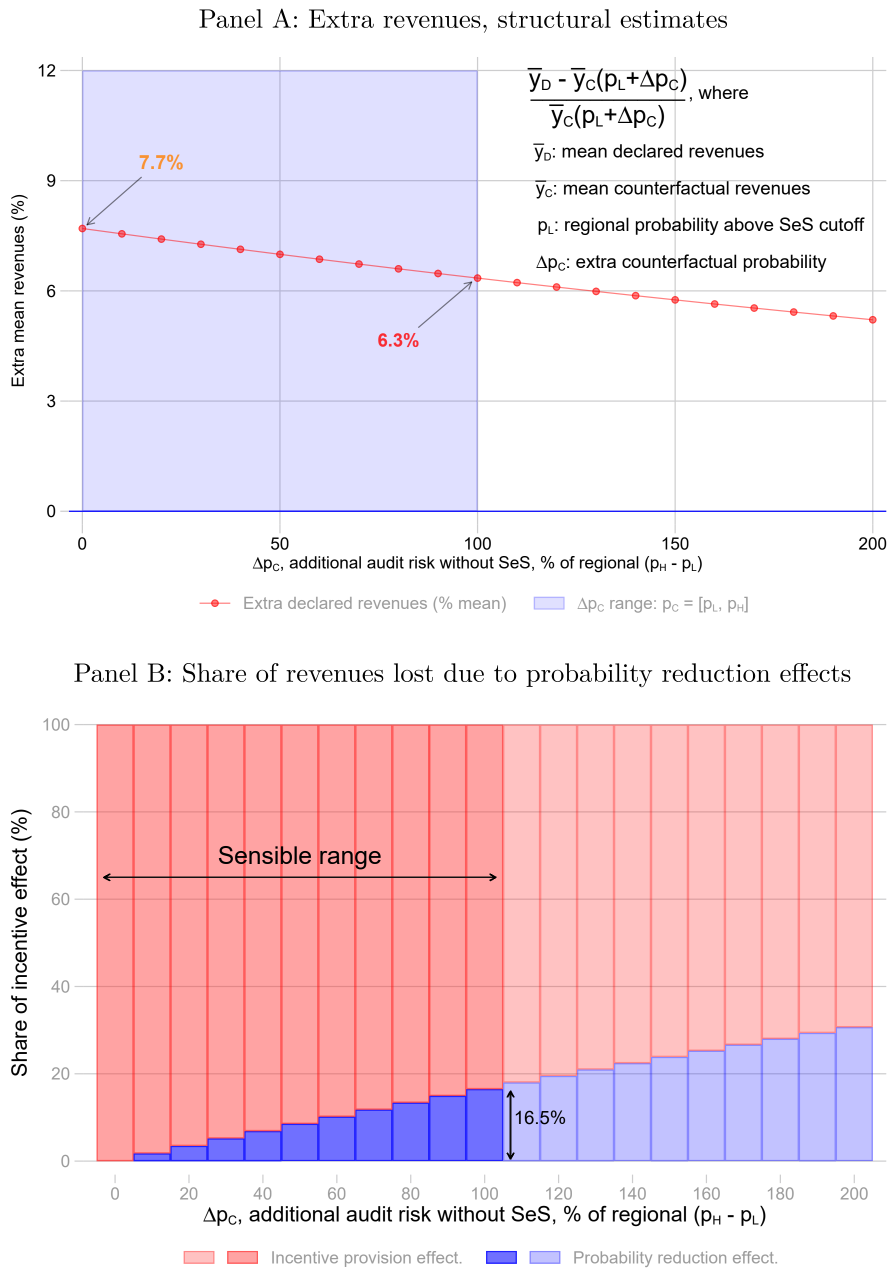
Notes: this Figure defines the marginal buncher in our discontinuous audit risk model, that is, the last taxpayer to raise reported revenues y to match the disclosed threshold $\hat{y}$. When audit risks are constant at $p_C = p_L$, each Euro of manipulated revenues $e_L = y^* - y_L$ results in an expected tax saving equal to $\tau - \tau\gamma p_L$ given a tax rate τ and an administrative penalty rate $\gamma > 1$. When disclosure reveals that audit risks are higher by Δp below a manipulation amount $\hat{e} = y^* - \hat{y}$, some taxpayers reduce their underreporting from their original interior solution e_L under risk p_L to the level consistent with reporting revenues equal to the threshold. The last taxpayer to do so is indifferent between setting their manipulation at $\hat{e}$ with p_L and at a new interior solution e_H above it with $p_H = p_L + \Delta p$. This is summarized by their indifference curve touching both the higher and lower segments of the piecewise budget constraint at these two manipulation levels, respectively.

Figure 7. Group-level evasion rates with constant audit risk $p_C = p_L$



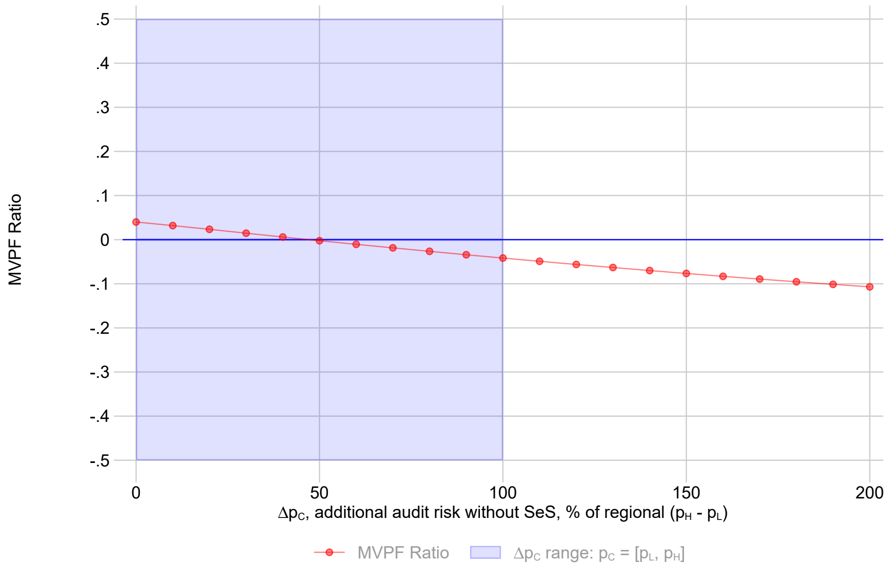
Notes: this Figure plots the distribution of evasion rates predicted by our structural model for 300 groups of PIT-payers filing for SeS over 2007-2010. Groups are each the combination of 20 regions, 5 macro-industries, and 3 presumed revenue terciles defined within each region-industry pair. Evasion rates are defined as the ratio between the average revenues underreported in each group in the model's equilibrium with constant audit risk $p_C = p_L$ and the group's median gross profits as reported in the SeS files. Equilibrium revenue evasion is defined as $e_L = k_e \left(\tau - \tau\gamma p_L\right)^{\varepsilon_e}$. In the background, we add a normal fit to smooth out the raw distribution of evasion rates.

Figure 8. Reported revenue effects of SeS disclosure across audit counterfactuals



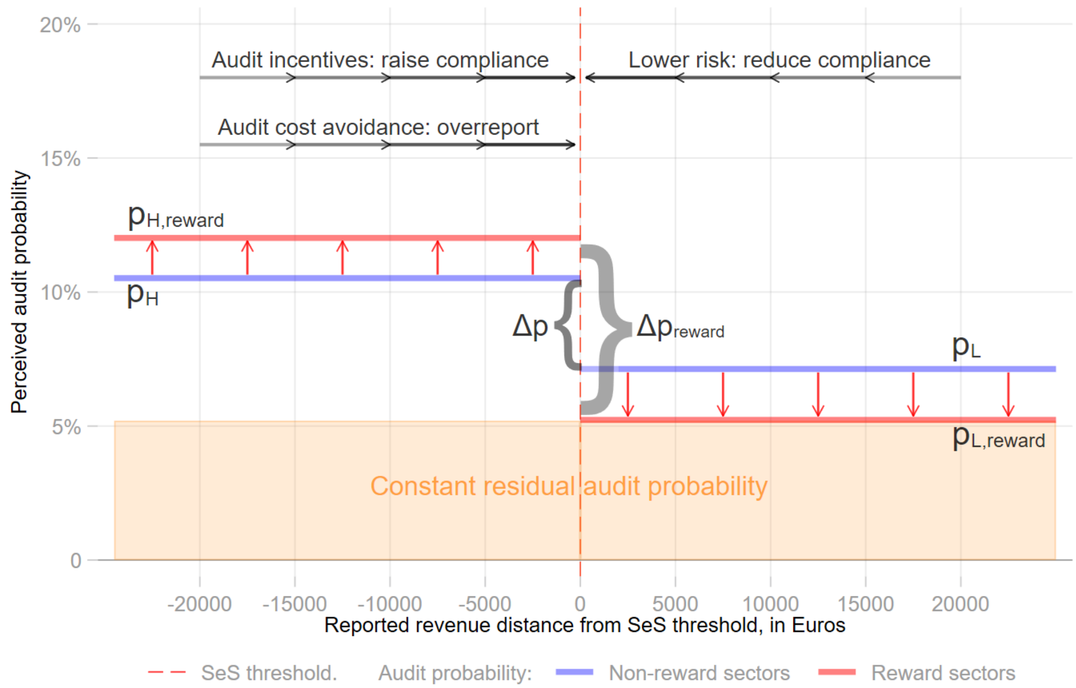
Notes: in this Figure, Panel A displays the effect of disclosing the SeS presumed revenue threshold on reported revenues relative to 21 counterfactual scenarios indexed by Δp_C. Each scenario reports the weighted average effect of disclosure across the 300 PIT-payer groups we employ in our structural analysis, with weights given by the number of SeS files in each group. In each counterfactual scenario, taxpayers perceive audit risks to be constant at $p_L + \Delta p_C$. We assign p_L to each group based on the audit risk estimate relevant to its administrative region of reference. Increments Δp_C are a percentage of the regional gap between p_L and p_H as indexed on the horizontal axis. This ensures that, within the sensible counterfactual audit range in a dark shade in both panels, the overall counterfactual risk varies exactly between p_L and p_H in each region. Panel B shows the size of the revenue losses (probability reduction effect) in percentage terms of the constant revenue gains (incentive provision effect) from disclosure. In both panels, darker colors highlight scenarios within the sensible counterfactual range.

FIGURE 9. MVPF of SeS disclosure across audit counterfactuals



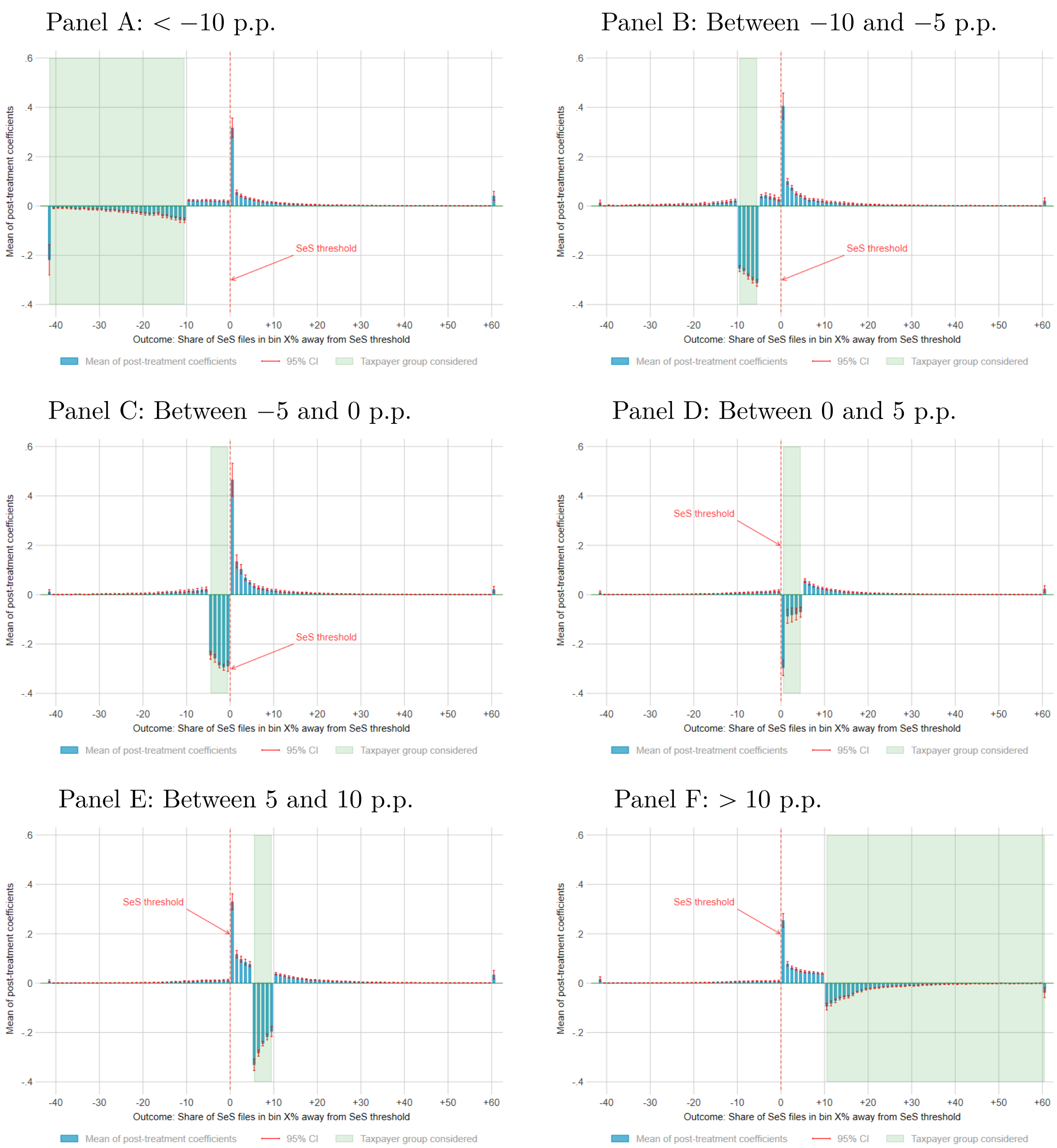
Notes: This Figure displays the estimates of the marginal value of public funds (MVPF) from disclosing the SeS presumed revenue threshold across 31 counterfactual scenarios indexed by Δp_C. We measure the MVPF on the vertical axis. We compute its numerator as the average willingness to pay for an audit risk change from p_C to p_L or p_H for taxpayers falling above or below the disclosed threshold in each of the 300 PIT-payer groups we employ in our structural analysis, with weights given by the share of SeS files in each group. We define the denominator of the MVPF as the mean administrative cost of disclosure, proxied by the total value of *SOSE*'s production in 2010 divided by the yearly number of 2007-2010 files used in the structural analysis, net of the gain in mean reported revenues we estimate in Section 8 taxed at the average PIT rate computed across all structural groups. Counterfactual scenarios are indexed on the horizontal axis. In each counterfactual scenario, taxpayers perceive audit risks to be constant at $p_L + \Delta p_C$. We assign p_L to each group based on the audit risk estimate relevant to its administrative region of reference. Increments Δp_C are a percentage of the regional gap between p_L and p_H as indexed on the horizontal axis. This ensures that, within the sensible counterfactual audit range in the dark shade, the overall counterfactual risk varies exactly between p_L and p_H in each region.

FIGURE 10. Perceived audit risk effects of the 2011 SeS reward regime



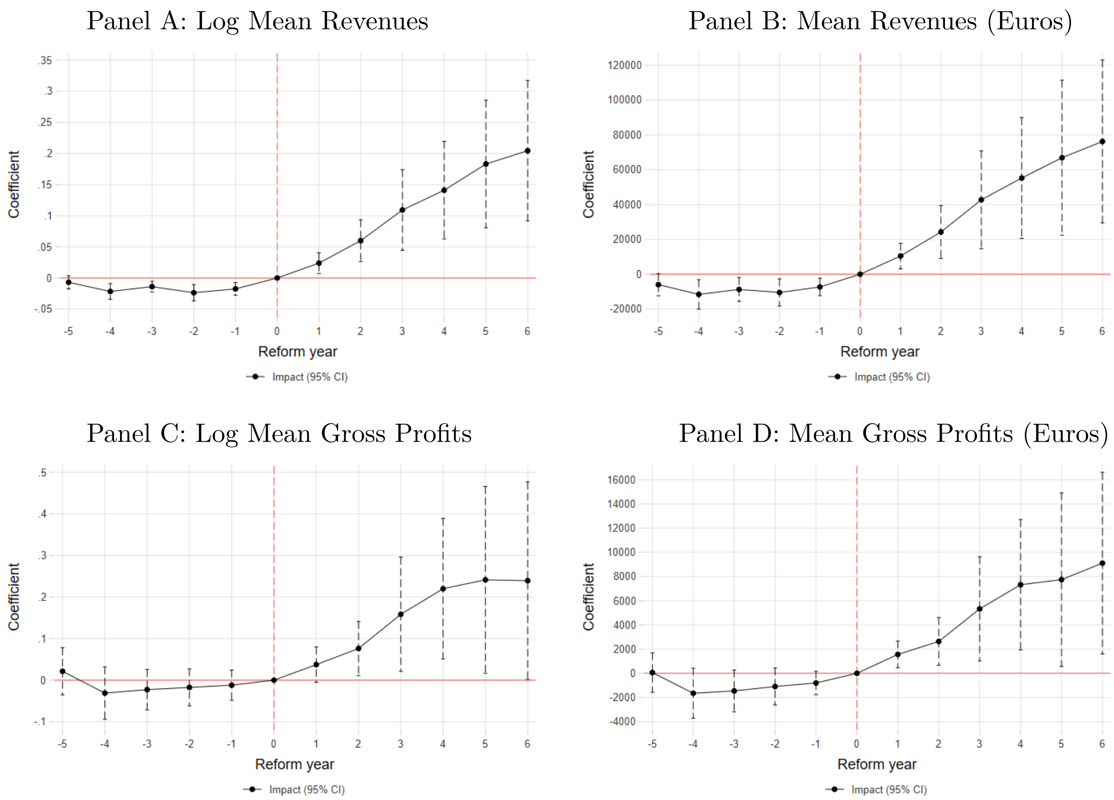
Notes: this Figure offers a rationalization of the reward regime's compliance effects on both sides of the SeS presumed revenues. The horizontal axis represents the distance in Euros from the presumed revenue amount, while the vertical axis reports the perceived probability of receiving an audit. Relative to the standard audit exemptions perceived by those declaring at or above their SeS threshold, compliance within SeS reward-regime sectors would result in a larger discrete reduction of audit risks $\Delta p_{reward} > \Delta p$. This would reinforce motives to increase (decrease) reported revenues for those below (above) the cutoff.

FIGURE 11. Reward regime-induced distribution shifts, by presumed revenues distance before the reform



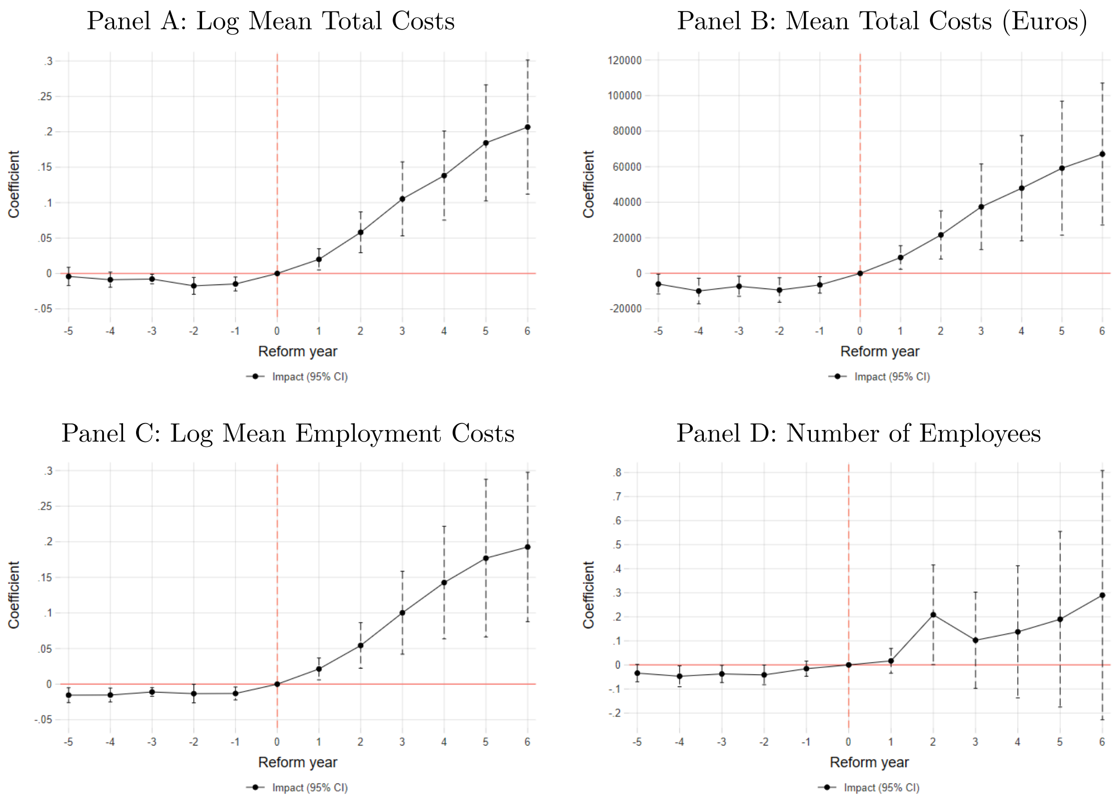
Notes: this Figure shows the effect of the reward regime on the average share of SeS files in bins of size one percentage point in presumed revenue terms. Each panel refers to one of six taxpayers' groups defined by their distance from the presumed revenue amount in the year before their sector's reform. The original location of each group is highlighted by the green band in each panel. Each bar represents the average of six group-specific post-treatment coefficients from an event-study based on the specification in (9.1). 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. Each panel represents a group of taxpayers defined as follows: taxpayers who reported revenues 10 p.p. or more below (Panel A), between 10 and 5 p.p. below (Panel B), between 5 and 0 excluded below (Panel C), between 0 and 5 p.p. above (Panel D), between 5 and 10 above (Panel E), and 10 p.p. or more above the presumed revenue amount the year before the reform (Panel F).

Figure 12. Reward regime effects on mean revenues and profits



Notes: this Figure shows the effects of the reward regime's introduction in a sector on mean reported revenues (Panels A and B) and mean gross profits (Panels C and D). Dependent variables are expressed in logarithms (left panels) or in Euro terms (right panels). 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 event-study specification in (9.1). 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 13. Reward regime effects on costs and inputs



Notes: this Figure shows the effects of the reward regime's introduction in a sector on aggregate costs and selected reported inputs. Panels A and B display the estimated effect on mean total costs defined as the difference between reported revenues and gross profits, in logarithms and Euros, respectively. Panels C and D display the estimated effect on mean employment costs in logarithm terms and on the mean number of employees, respectively. 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 event-study specification in (9.1). 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.

Tables

Table 1. Structural groups: mean statistics by industry

Sector (60 groups each)	Files (mil.)	Mean y (€,000)	Mean π (€,000)	Mean τ (PIT)
Manuf & Constr	2.18	152.3	24.9	20.4%
Wholesale	0.92	151.8	29.5	21.5%
Retail	1.82	160.8	19.1	16.3%
Professions	2.47	68.6	42.2	23.1%
Other Services	2.92	97.8	19.1	16.9%
All	10.31	126.3	26.95	19.6%

Notes: the Table reports the total number of SeS files and the mean features of the 300 groups of PIT-payers filing in 2007-2010 used in our structural analysis. Groups are a combination of 20 regions, 5 macro-sectors, and 3 presumed revenue terciles defined for each region-industry pair. Taxpayers are trimmed at each group's 5-95th percentile of the absolute difference between reported and presumed revenues. We focus on taxpayers whose location we can reconstruct up to the municipal level.

Table 2. Structural groups: median statistics by industry

Sector (60 groups each)	Files (mil.)	Median y (€,000)	Median π (€,000)	Median τ (PIT)
Manuf & Constr	2.18	71.2	23.4	21.0%
Wholesale	0.92	52.9	29.9	23.5%
Retail	1.82	82.6	14.3	14.9%
Professions	2.47	43.5	30.3	25.0%
Other Services	2.92	54.7	18.6	18.1%
All	10.31	59.8	21.7	19.5%

Notes: the Table reports the total number of SeS files and the median features of the 300 groups of PIT-payers filing in 2007-2010 used in our structural analysis. Groups are a combination of 20 regions, 5 macro-sectors, and 3 presumed revenue terciles defined for each region-industry pair. Taxpayers are trimmed at each group's 5-95th percentile of the absolute difference between reported and presumed revenues. We focus on taxpayers whose location we can reconstruct up to the municipal level.

Table 3. Structural parameters: estimates summary

Parameter	N	mean	sd	min	max	*D'Agosto et al.*
p_L	20	10.8%	3.1%	4.6%	15.2%	7.13%
p_H	20	15.6%	2.2%	10.8%	19.7%	10.52%
Δp	20	4.8%	1.9%	2.7%	8.6%	3.39%
k_e	3	35.76	8.06	29.86	44.94	
ε_e	5	2.15	0.42	1.80	2.80	

Notes: this Table reports summary information on the structural parameter estimates. We use the 2007-2010 behavioral revenue responses of 300 PIT-payer groups to fit 48 parameters. Specifically, we assign a common pair of perceived audit risks $\{p_L, p_H\}$ for groups in the same region, a common parameter ε_e for each of five industries, and a common parameter k_e for each of three presumed revenues terciles defined within each region-industry combination. Calibration of the policy parameters include average PIT rate due in each group and the lower-bound penalty rate set by law on detected evasion. Parameter k_e is in €10,000 terms. Mean perceived audit probabilities can be compared with the estimates in the last column. These estimates reflect our calculation of the implied SeS file audit risk for taxpayers reporting above and below presumed revenues in the 2007-2010 tax period based on the aggregate audit frequencies in D'Agosto et al. (2017).

Table 4. Structural groups: SeS responses and evasion by industry

Sector (60 groups each)	Bunching $\hat{b}$	$\Delta\hat{y}$ (€,000)	e^{eq} (€,000)	$g(e)$ (€,000)	$\frac{e^{eq}}{\pi^{med}}$
Manuf & Constr	10.1	5.05	8.60	0.98	39.7%
Wholesale	7.1	3.53	5.61	0.73	21.7%
Retail	10.4	5.22	8.63	0.81	58.1%
Professions	7.5	3.73	3.74	0.66	9.0%
Other Services	10.5	5.25	9.27	0.85	57.9%
All	9.11	4.55	7.17	0.81	37.3%

Notes: the Table reports the average SeS responses and equilibrium revenue evasion with $p_C = p_L$ of the 300 groups of PIT-payers filing in 2007-2010 used in our structural analysis. Groups are a combination of 20 regions, 5 macro-sectors, and 3 presumed revenue terciles defined for each region-industry pair. Taxpayers are trimmed at each group's 5-95$^{\text{th}}$ percentile of the absolute difference between reported and presumed revenues. We focus on taxpayers whose location we can reconstruct up to the municipal level. Bunching relies on bins of width €500, 7$^{\text{th}}$order polynomials, and a bunching upper bound set at the threshold bin. $\left(\frac{e^{eq}}{\pi^{med}}\right)$ is the mean across groups' ratios of the equilibrium evasion and median reported gross profits.

APPENDIX - For Online Publication

Appendix A. Condition for revenue-improving disclosure

We briefly derive Result 1 exploiting the set-up and definitions in Section 2. Audit rule disclosure is revenue-enhancing if and only if:

$$\bar{y}_D > \bar{y}_C = \bar{y}_L + \text{PRE}, \tag{A.1}$$

where we consider D as the policy counterfactual with disclosure, C the one where firms perceive a constant audit risk p_C, and L the one with the low constant risk $p_C = p_L$. The PRE is the absolute value of the loss in mean reported revenues that a drop in perceived audit risks $\Delta p_C = p_C - p_L > 0$ would bring about. We define the aggregate elasticity of declared revenues $\varepsilon^{agg}(\bar{y}_\theta) = E_{y^\star}[\frac{\partial y_\theta(y^\star)}{\partial T'_\theta(y^\star - y_\theta(y^\star))} \cdot \frac{T'_\theta(y^\star - y_\theta(y^\star))}{\bar{y}_\theta}] < 0$, or ε^{agg} for brevity, and assume it to be constant for any $\bar{y}_\theta$. We assume linearity of the expected avoided tax liability $T_\theta(e)$ in a p-constant policy counterfactual θ, so that $T'_\theta(e) = \tau - \tau\gamma p_\theta$, where τ is a flat tax rate on income and $\gamma > 1$ is a penalty rate on detected evasion. Changes in perceived audit risks affect mean reported revenues by changing how much taxes firms can expect to avoid with evasion.

We wish to characterize the relationship between the mean reported revenues $\bar{y}_C$ and $\bar{y}_L$. The latter is brought about by a change in the marginal expected avoided tax liability $T'_C(e) - T'_L(e) = -\tau\gamma\Delta p_C$. This relationship is linear, and we can write it as:

$$\bar{y}_L = \bar{y}_C + \bar{y}_C \varepsilon^{agg} \frac{\tau\gamma\Delta p_C}{\tau\left(1 - \gamma p_C\right)}. \tag{A.2}$$

Rearranging, we can see that $\frac{\bar{y}_C}{\bar{y}_L} = \frac{1}{1+\frac{\varepsilon^{agg}\gamma\Delta p_C}{1-\gamma p_C}}$. Using the relationship in (A.2), we can write an expression for the PRE in terms of either $\bar{y}_C$ or $\bar{y}_L$ as follows, simplifying for the tax rate:

$$-\bar{y}_C \varepsilon^{agg} \frac{\gamma\Delta p_C}{1 - \gamma_C} = -\bar{y}_L \frac{\varepsilon^{agg}\gamma\Delta p_C}{1 - \gamma p_C + \varepsilon^{agg}\gamma\Delta p_C}. \tag{A.3}$$

We can now rewrite the condition in (A.1) with the PRE expression in terms of $\bar{y}_L$ from (A.3) to obtain our result:

$$\frac{\bar{y}_D}{\bar{y}_L} > \frac{1}{1 + \frac{\varepsilon^{agg}\gamma\Delta p_C}{1-\gamma p_C}}. \tag{A.4}$$

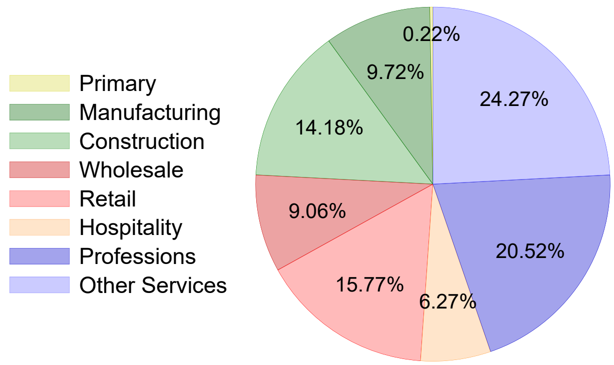
Appendix B. Details on SeS data

Figure A1. SeS dataset overview

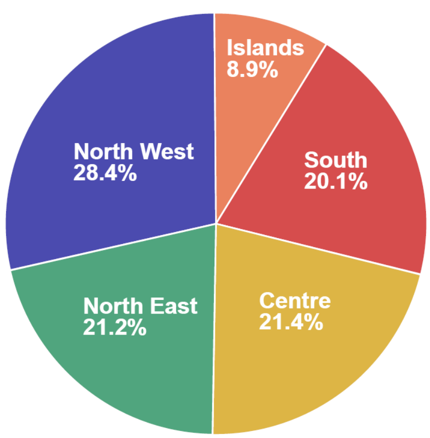
Panel A: Dataset structure

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

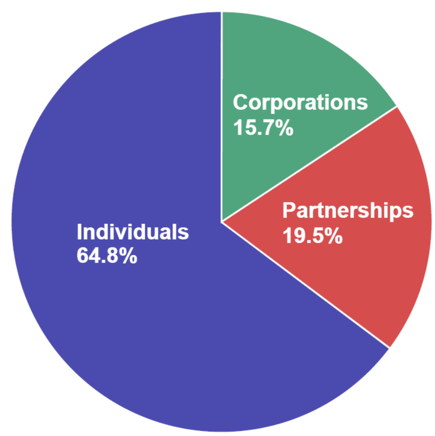
Panel B: 2007-2010 macro-sectors



Panel C: 2007-2010 geographic composition

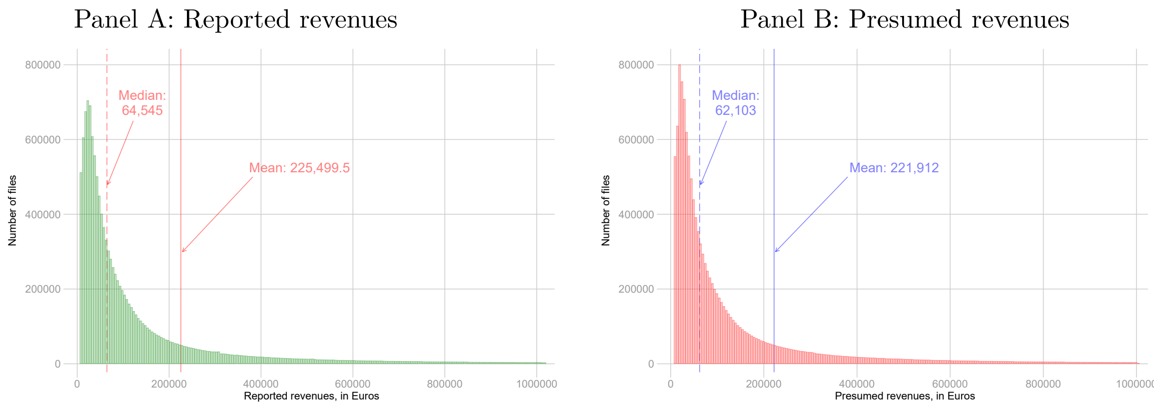


Panel D: 2007-2010 legal status



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 A2. 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^{th} and 95^{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.

Table A1. 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.

Appendix C. Details on additional data sources

C.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. Relative gap 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 et al. (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 et al. (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 et al. (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 et al. (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 A4.

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.

C.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.

Appendix D. Bootstrap Procedure for Bunching Confidence Intervals

We compute standard errors to bunching estimates using a semi-parametric bootstrap procedure. The following specification, which we used to compute the bunching counterfactual, provides the structure for our routine:

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

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 (D.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 (D.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.

Appendix E. Reporting vs. production responses

We present evidence consistent with the idea that firms respond to SeS incentives adjusting their reports rather than their production. This motivates the assumption of separability between the reporting and production margins that we introduce in Section 2.

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.[55] Specifically, we regress the bunching estimates for all provinces i on one evasion proxy j at a time according to the following model:

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

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. Figure A5 displays the standardized coefficients for all our evasion proxies. All of our estimates turn out to be positive, and most are significant and meaningful in magnitude. This result holds not just for the proxies we draw from the existing economic and administrative literature, but also for those we build from over 620,000 "whistleblower" reports submitted by consumers to the private website *evasori.info* over 2008-2011. Relying on a first principal component of the various measures does not alter the pattern of results. Finally, Figure A6 disaggregates our analysis whenever a finer 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 A7 displays a positive and significant conditional correlation between LLM bunching and the weighted average of municipal PIT surcharge rates. Municipalities can impose a surcharge rate of less than 1% on top of the national personal income tax schedule. The standardized association between bunching and these local tax rates is of the same magnitude as that between bunching and the average length of provincial tax litigation, which provides a plausible proxy for a sizable enforcement cost borne by taxpayers (Figure A8). We also find higher bunching among firms that are more exposed to the final consumer (Figure A9), among taxpayers with relatively lower turnover (Figure

[55]Figure A4, 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.

A10), and among businesses with relatively fewer reporting requirements, as in the case of individual businesses as opposed to the partnerships and corporations in our data (Figure A11). This aligns with the literature's suggestions that these features ease the concealment of true production due to the structure of VAT incentives and hurdles in the successful monitoring of smaller enterprises.

The sharp bunching observed in Figure 5, as well as the fact that knowledge of the exact location of the threshold is acquired after the end of the production period, make it unlikely that taxpayers respond to SeS by adjusting their true production. However, while a new edition of *Gerico* is released every tax season, a sector's underlying presumed revenue function is revised only once every three years. Therefore, taxpayers might learn how to fine-tune production over the course of a three-year cycle.

We assess the learning-to-adjust hypothesis in two ways. First, we estimate bunching for every sector and year, and residualize these estimates by sector and year fixed effects. Figure A12 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 A13 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.

Appendix F. Indifference condition in the structural model

Section 6 aims at defining an indifference condition which only depends on behavioral and policy parameters and revenue responses to the threshold. Here, we outline the necessary steps to obtain that expression.

We define revenue underreporting or evasion as $e = y^* - y$, that is the difference between a firm's true and reported revenues. Consider the firm revenue reporting model introduced in Section 2 and specified in Section 6, the optimality condition in (2.2) and the iso-elastic manipulation cost function $g(e)$ in (6.1). When firms perceive a constant audit risk $p_C = p_L$, their equilibrium level of reported revenues is:

$$y_L = y^* - k_e \left(\tau - \tau\gamma p_L\right)^{\varepsilon_e}, \tag{F.1}$$

or equivalently their equilibrium evasion amounts to $e_L = k_e \left(\tau - \tau\gamma p_L\right)^{\varepsilon_e}$. In the model, evasion is independent of production ability ξ due to the assumed separability between reporting and production margins. On the other hand, evasion depends on perceived audit risks, tax and penalty rates, and the parameters of the manipulation cost function. To the extent that the values of these parameters are constant in a given firm class, equilibrium evasion is also constant across firms in that class ahead of disclosure. However, the location of any firm along the reported revenue range will vary reflecting the distribution of production abilities, which we assume to be heterogeneous in each class.

When disclosure increases perceived audit risks from p_L to p_H by Δp below the threshold, some firms will bunch by raising their reported revenues to $\hat{y}$. We call marginal buncher the firm who is indifferent between reporting at two points along the reported revenue range: *i)* the threshold $\hat{y}$, and *ii)* y_H, the new interior solution when perceived risks rise to the level p_H. This firm reports $\hat{y} - \Delta\hat{y}$ when risk is constant at $p_C = p_L$ and produces true revenues $\hat{y}^* - \Delta\hat{y}^*$ with ability $\xi - \Delta\xi$.

We now write out the value that the marginal firm perceives at the two reported revenue levels $\hat{y}$ and y_H, or equivalently when its evasion amounts to $\hat{e} = \hat{y}^* - \Delta\hat{y}^* - \hat{y}$ and $e_H = y^* - y_H$, respectively. From (2.1) the value of the firm at the notch is:

$$V^N = \left(\tau - \tau\gamma p_L\right) \cdot \left(\hat{y}^* - \Delta\hat{y}^* - \hat{y}\right) - \frac{k_e}{1 + \frac{1}{\varepsilon_e}} \cdot \left(\frac{\hat{y}^* - \Delta\hat{y}^* - \hat{y}}{k_e}\right)^{1 + \frac{1}{\varepsilon_e}}, \tag{F.2}$$

On the other hand, by (F.1) revenues reported at the new interior solution will be $y_H = \hat{y}^* - \Delta\hat{y}^* - k_e \left[\tau - \tau\gamma \left(p_L + \Delta p\right)\right]^{\varepsilon_e}$. Using the corresponding expression for e_H, we can simplify the marginal buncher's value at this interior point as:

$$V^H = \frac{k_e}{1 + \varepsilon_e} \left[\tau - \tau\gamma \left(p_L + \Delta p\right)\right]^{1 + \varepsilon_e}. \tag{F.3}$$

The marginal buncher's indifference condition is defined as $V^N = V^H$. We wish to express this condition as a function of the unobservable behavioral parameters k_e, ε_e, p_L, and Δp (or p_H), the policy parameters τ and γ, as well as revenue responses $\Delta\hat{y}$. For the latter to appear in the marginal buncher's values, we use the relationship between the marginal buncher's reported and true revenues in the pre-disclosure equilibrium, that is $\hat{y} - \Delta\hat{y} = \hat{y}^* - \Delta\hat{y}^* - k_e \left(\tau - \tau\gamma p_L\right)^{\varepsilon_e}$. We can now work through $V^N - V^H = 0$ to obtain the indifference condition in (6.2):

$$\tau^e \left[-\frac{\Delta\hat{y}}{k_e} + \left(\tau^e\right)^{\varepsilon_e} \right] - \frac{\varepsilon_e}{1+\varepsilon_e} \left[-\frac{\Delta\hat{y}}{k_e} + \left(\tau^e\right)^{\varepsilon_e} \right]^{1+\frac{1}{\varepsilon_e}} - \frac{1}{1+\varepsilon_e} \left[\tau^e + \Delta\tau^e\right]^{1+\varepsilon_e} = 0, \tag{F.4}$$

where we define $\tau^e \equiv T'\left(e_L\right) = \tau - \tau\gamma p_L$ and $\Delta\tau^e \equiv T'\left(e_H\right) - T'\left(\hat{e}\right) = -\tau\gamma\Delta p$ for brevity. This is the form of the indifference condition that underlies our structural estimation.

Appendix G. Alternative models: Audit Cost and Cost Misreporting

G.1. **The role of the business cost of audits.** Our structural analysis assumes that the cost of being subject to an audit is proportional to the amount of underreporting. An alternative model could also feature a separate cost a from undergoing an audit unrelated to the administrative fines applied to any uncovered evasion. This cost can reflect the psychological or administrative burden that firms bear when engaging with the tax administration during an inspection, a tax assessment, or at any step of an ensuing tax litigation. Below, we provide our considerations on the potential role of these costs in our analysis, and explain why we assume $a = 0$ to deliver our main results. First, we discuss why we believe that these costs are not of first order importance for taxpayers' SeS behavior, providing empirical evidence to support our claim. Second, we explain why our assumption is conservative in light of the main purpose of our empirical analysis, that is to establish whether SeS disclosure is revenue-enhancing.

G.1.1. *Cost of audit as a driver of bunching.* If businesses perceive tax audits to be costly, disclosure provides a path to reduce their expected cost, as taxpayers can adjust their reports to avoid audits. In the analysis of SeS, we can however constrain the practical role of such costs in a number of cases.

First, if businesses bear a fixed cost conditional on undergoing an audit, any reporting behavior would not be influenced by this cost in equilibrium. This stems from the fact that, in the interior solution of a standard equilibrium evasion model, a fixed cost has no impact on marginal decisions. As a result, a model with $a > 0$ should not predict a significant correlation between the level of bunching in the SeS and our proxies of evasion intensity. This is contradicted by our results in Appendix E.

Second, let's consider a model where the cost of undergoing an audit depends instead on a taxpayer's level of evasion, or $\frac{\mathrm{d}a}{\mathrm{d}e} \neq 0$. If we assume that a is linearly related to evasion, the interpretation of this model is equivalent to rescaling the penalty factor γ or the perceived audit risk p in our baseline model.

Third, we can consider the case where audit costs are so large and the propensity to audit avoidance is so strong, that taxpayers are willing to report more revenues than actually generated by their business. Of course, it seems unlikely that broad groups of small firms and the self-employed are coerced into reporting more than earned, when all existing estimates point to positive tax gaps among these categories (Ministry of Economy and Finance, 2019).[56] However, if that was the case, the intensity of this revenue "overreporting" should also

[56] Galbiati and Zanella (2012) discuss this point based on the universe of audits run by the Italian Tax Police on small businesses and the self-employed for the 1987 tax year. In their records, only 10% of taxpayers report more taxable income than they should have. Since discrepancies are of small magnitude, the authors conclude that these might be reporting mistakes. While these observations come from a selected sample collected before the introduction of SeS, they suggest that the extent of overreporting behaviors is limited in practice.

correlate with underreporting on other margins or with revenue underreporting in other tax-relevant files as seen in the correlations between bunching responses and evasion in Appendix E. Once again, this seems unlikely.

We thus conclude that a model that features audit costs unrelated to administrative sanctions as the main driver of firm behavior in the SeS setting would be inadequate to explain the descriptive evidence we draw from our data.

Finally, we could think that the correct structural model of firm reporting is the following:

$$V(y^\star, y, T_\theta(y^\star - y)) = T_\theta(y^\star - y) - g(y^\star - y) - p\cdot a, \tag{G.1}$$

where $p\cdot a$ reflects the expected cost of undergoing a tax audit aside from the penalties applied to any uncovered evasion. If that is in fact the case, the results we present in our main analysis should provide an upper bound to the propensity of revenue manipulation, since we would be discounting the contribution of audit costs. Indeed, given the location of the marginal buncher (or equivalently $\Delta\hat{y}$), assuming $a = 0$ would attribute the entire bunching responses to the manipulation elasticity we intend to estimate. On the other hand, estimating one or more parameters a would require a heavier structure with the appropriate definition of additional identification restrictions and, potentially, finer empirical groups.

G.1.2. *Audit cost empirics.* To the extent that firms are sensitive to the tax enforcement environment they operate in, reporting responses in the SeS setting should correlate with evasion incentives such as tax rates and fines associated to any uncovered noncompliance, and with the size of audit costs.

Unfortunately, there is no straightforward way of measuring the overall magnitude of the costs borne by a business undergoing a tax audit, so our empirical exercises can only be suggestive. In particular, we assess the correlation between SeS bunching and measures we believe might capture a part of the firm's perceived audit costs.

We proxy the expected cost of audit with bunching at the thresholds provided by other policies that might affect these costs directly. On one hand, we exploit a 2011 reform (Law Decree 70/2011) that reduced the maximum duration of on-site tax inspections from 30 to 15 days for PIT-payers reporting less than €400,000 in the service sector or €700,000 elsewhere.[57] On the other, we measure the extent of bunching at the pre-2011 revenue threshold (€30,000) allowing SeS individuals to opt into a minimum taxpayer regime and out of the SeS system altogether. In both cases, the larger the bunching below these thresholds at the local level, the larger the benefit that businesses may perceive from a less stringent audit system. Figure A14 shows no significant correlation across provinces between SeS bunching and the intensity of

[57]These thresholds coincide with the maximum turnover allowed for firms opting into a simplified accounting regime.

responses to the incentives provided by any of these two policies. We conclude that bunching at the SeS threshold is not well captured by simple audit cost avoidance behaviors.

G.1.3. *Disclosure effects with audit costs.* Ultimately, the relative importance of audit costs to our analysis depends on their impact on the effects of disclosure that we uncover. The main goal of our analysis is indeed to assess whether disclosure can be revenue-enhancing. We address this question in Section 8. There, we decompose the effects of disclosure on the relative revenue distribution into two theoretical components: a *probability reduction effect*, whereby disclosure may lower perceived audit risks from p_C to p_L, and reported revenues shrink following the increase in the marginal benefit of evasion; and an *incentive provision effect*, whereby businesses below the presumed revenue threshold become aware of their higher audit risk $p_H > p_L$ and of the risk discontinuity at $\Delta\hat{y}$, which leads to a reported revenue increase.

Consider now the case with a fixed cost of audit $a > 0$. Fixed costs don't influence evasion decisions, and as such leave unaltered the fall in reported revenues associated to the probability reduction effect upon disclosure. On the other hand, the existence of audit costs might induce the taxpayers who perceive a new, higher threat of audit to avoid these costs by complying with SeS prescriptions. Thus, the incentive provision effect of disclosure below the presumed revenue threshold should be stronger *ceteris paribus* when $a > 0$, since taxpayers have two distinct motives to increase revenues below the SeS threshold: evasion incentives are lower, and audit cost avoidance incentives are higher.

It follows that the case with $a = 0$ rests on a conservative assumption if we aim to asses the reported revenues effect of SeS disclosure. When we set $a = 0$, we obtain an upper bound to the propensity to manipulate revenues in response to changes in evasion incentives. In this case, we thus maximize the potential reported revenues losses due to the probability reduction effect of disclosure. Hence, our results would be even stronger in the presence of $a > 0$.

Therefore, rather than simply concluding that a SeS model that purely relies on audit costs is counterfactual to our descriptive evidence, we stress that it would also lead to less conservative interpretations of any result suggesting that disclosure is revenue-enhancing.

G.1.4. *MVPF with audit costs.* In Section 8.3, we derive estimates for a MVPF ratio by applying an envelope theorem to our baseline model with $a = 0$ (see Appendix J for details). However, it is likely that our main conclusion - that the net welfare cost associated to disclosure is likely small, and that the MVPF might be infinite - would hold even with $a > 0$.

In particular, the mechanical effect of disclosure for taxpayers concerned about $a > 0$ would include this additional cost along with the penalty cost for detected evasion. As in the baseline case, however, taxpayers at or above the disclosed threshold in the policy counterfactual obtain a reduction in the expected cost of incurring in such audit-related hurdles. This can at least partially compensate for the perceived increase in audit-based costs for taxpayers below the threshold in the policy counterfactual.

Clearly, this reasoning depends on the assumption that audit risks are homogeneously distributed across the two groups of taxpayers or, more in general, along the range of relative reported revenues. While the business cost of undergoing an audit may be correlated with several observables and unobservable taxpayer characteristics (such as the local level of administrative efficiency, the level of education of the taxpayer, etc.), we do not see a particular reason why it should be strongly correlated with the relative location of the taxpayer around their own the SeS presumed revenue level. This stems from the nature of these thresholds: presumed revenues are idiosyncratic estimates of the business' revenue-generating potential in any given year. They do not necessarily reflect any single fundamental that spans across businesses, such as overall business size, profitability etc. As such, expecting the welfare costs of disclosure to be especially large requires assuming that the costs from undergoing an audit - aside from administrative penalties on detected evasion - are concentrated among businesses falling below these idiosyncratic thresholds in the policy counterfactual.

G.2. **Cost misreporting and presumed revenue manipulation.** Our main estimates of the reported revenue effect of disclosure are conservative also relative to augmenting the structural model with cost manipulation. The logic is similar to that outlined above for audit costs: an extended model would only reduce the scope for the revenue elasticity to explain bunching and thus reduce the predicted amount of revenue losses from disclosure.

Suppose firms could report several fully deductible cost items c_j different from their true production costs c_j^* in order to affect their assigned presumed revenues $\hat{y} = \phi(c)$. We can assume $\frac{\partial \phi(c)}{\partial c_j} > 0$ for at least some j items, since presumed revenue functions generally predict larger turnover for firms mobilizing larger resources.[58] If this is the case, firms would lower their reported value of these items on *Gerico*, so that the software assigns them a lower value of $\hat{y}$ and exempts them from SeS-based audits with higher likelihood for given reported revenues.

How does this situation compare to the one where firms only manipulate revenues in response to the SeS disclosure? Positive cost manipulation implies that a given level of bunching is due to two different forces. Think about two sequential moves: firms set their own threshold first, and manipulate revenues next. Therefore, for a given revenue elasticity, bunching would be higher if firms could also manipulate costs as an additional adjustment margin. To the extent that our baseline approach mutes this margin, our estimates of the revenue misreporting elasticity from the observed bunching are an upper bound of the true one, since they attribute to revenue reporting the part of adjustment due to costs.

[58]For the cost items that directly correspond to *Gerico*'s model variables, $\frac{\partial \phi(c)}{\partial c_j} = \beta_j$. This parameter is the linear regression coefficient on variable j that taxpayers might find in the Revenue Agency's yearly technical report on their sector's model.

As for the extension with audit costs a, this has implications for the estimated probability reduction effect (revenue losses), but not for the incentive provision effect (revenue gains) of disclosure. We measure the revenue gains directly in the data, computing the difference between the average revenues reported under disclosure and in the bunching counterfactual. The assumptions these values depend on are at most those behind bunching estimation, which is not affected by any structural considerations, including those on cost manipulation. On the other hand, for a given level of bunching, the estimated revenue losses would be smaller with cost manipulation: upon disclosure, a smaller revenue elasticity would produce a smaller revenue loss in response to the same audit risk drop Δp_C.

Finally note that, for a given reported revenue amount, lower costs would also correspond to a larger tax base through higher reported gross profits $\pi = y - c$, and a potentially higher tax liability $\tau \cdot \pi$. As a result, in the case of cost manipulation, policy evaluation should also account for the additional tax gains from cost underreporting, which would strengthen the case for disclosure further.

Appendix H. Structural model validation

We assess the sensibility of our structural model's predictions in light of the standard results in the evasion literature and the misreporting patterns observed in Italy.

Figure A16 breaks down our 300 structural groups into two broad categories: downstream sectors mostly exposed to the final consumer, and upstream sectors mostly transacting with other businesses along the supply chain. We define these categories based on *ISTAT* input-output tables, verifying whether each of the five macro-sectors used to group SeS filers displays below or above-median domestic value added shares from final consumption over 2010-2013. Although there exists some overlap in the two groups' evasion rate distributions, we predict the downstream sectors (retail and services) to underreport revenues with higher intensity relative to own median profits. This is in line with the common finding that transacting with the final consumer offers greater opportunities for revenue concealment given the reduced reporting incentives at both ends of the transaction (Pomeranz, 2015; Naritomi, 2019).

Historically, Italian evasion displays stark geographic patterns, with higher intensity in the South and in the Islands, and higher compliance in the Center-North (Brosio et al., 2002). A sensible model should render a similar picture. In Figure A17 we aggregate our evasion estimates by region, appropriately weighting each structural group's contribution by the number of its SeS files. We then associate the evasion predicted in each of the twenty regions to the corresponding intensity of local evasion from Galbiati and Zanella (2012). The authors access detailed data on the universe of audits carried out by the *Guardia di Finanza* (the Italian Tax Police) on a 1987 population of small businesses and professionals comparable to ours.[59] They leverage the result of these audits to compute the share of true profits originally concealed to tax authorities. The resulting scatter plot suggests that our model replicates the geographical dispersion in Italian small business evasion to a substantial extent.

We quantify and extend the latter result in Figure A18. Provided that multiple evasion estimates exist in the literature, differing by method, time period, and tax base, we gather a non-exhaustive list of six regional evasion proxies from both administrative and academic sources, including the one discussed above. After regressing each of these proxies on our predicted regional evasion rate, we plot the resulting slope coefficient along with the relevant R-squared. A one standard deviation change in the model's predicted evasion rate is associated with a change by 0.57-0.78 standard deviations in our reference proxies. This sizable association is reflected in the models' fit, with our predictions capturing between one and two thirds of the variation in the available estimates of regional evasion in Italy.

[59]Despite the long time lag between our sample period and that of Galbiati and Zanella (2012), underreporting intensity in Italy might not have abated. For example, Figure A19 shows that VAT evasion has been relatively stable around 33% of the potential base in the 1980-2014 period while the standard VAT rate increased.

Appendix I. Audit risks and elasticity: a microfoundation

In Section 8 we argue that, conditional on observing bunching at an audit rule threshold, evidence of a low reporting elasticity implies that the audit risk gap induced by disclosure is a relatively stronger driver of that bunching. We now provide a simplified model to formalize the relationship between audit risks and elasticity in reduced form. In the process, we adapt Kleven and Waseem (2013)'s logic for expressing a reduced-form elasticity in the presence of a notch.

Consider the business decision model in Section 6, which we simplify in the following two ways. First, audit risks are $p_L = 0$ when $y \geqslant \widehat{y}$ and $p \equiv p_H > 0$ when $y < \widehat{y}$, respectively. This implies that $p = \Delta p$, the audit risk gap around the disclosed threshold. Second, SeS-based audits of a firm's revenues can only claim arrears on a tax base $(\widehat{y} - y)$, rather than $(y^* - y)$.[60] That is, the tax administration does not necessarily recover the full extent of evasion, but is constrained by the presumption established by the relevant SeS prediction model.

We focus on the marginal bunchers adjusting their revenues from $\widehat{y} - \Delta\widehat{y}$ to $\widehat{y}$ upon disclosure, and compute the implied marginal tax rate t^* they face along the bunching interval. To this end, define the two tax liabilities they face respectively at their final and original location:

$$T(\widehat{y}) = \tau\widehat{y}$$

$$\begin{aligned} T(\widehat{y} - \Delta\widehat{y}) &= \underbrace{\tau(\widehat{y} - \Delta\widehat{y})}_{\text{reported taxes}} + \underbrace{\tau\gamma p(\widehat{y} - y)}_{\text{expected audit loss}} \\ &= \tau(\widehat{y} - \Delta\widehat{y}) + \tau\gamma p\Delta\widehat{y}, \end{aligned}$$

where we have used the fact that, for marginal bunchers, $\widehat{y} - y = \widehat{y} - (\widehat{y} - \Delta\widehat{y}) = \Delta\widehat{y}$. We can now define the marginal buncher's implicit marginal tax rate as the average reporting incentive they face over the bunching interval:

$$\begin{aligned} t^* &= \frac{T(\widehat{y}) - T(\widehat{y} - \Delta\widehat{y})}{\Delta\widehat{y}} \\ &= \frac{\tau\Delta\widehat{y} - \tau\gamma p\Delta\widehat{y}}{\Delta\widehat{y}} \\ &= \tau - \tau\gamma p. \end{aligned}$$

We label $t = \tau$ the pre-disclosure level of reporting incentives. Then, the variation in t^* over the bunching interval is $\Delta t^* \equiv t^* - t = -\tau\gamma p$, so that the partial derivatives of t^* and Δt^* with respect to the tax rate τ are $\frac{\partial t^*}{\partial \tau} = 1 - \gamma p$ and $\frac{\partial \Delta t^*}{\partial \tau} = -\gamma p$, respectively.

In our structural model, we consider the reporting elasticity to the marginal incentive for manipulation. Here, we approximate it through a reduced form elasticity of manipulation to

[60] Production costs are fully deductible, so we omit them from the discussion for simplicity.

the implicit tax rate t^*:

$$\varepsilon_{RF} = \frac{\Delta\widehat{y}}{\widehat{y}} / \frac{\Delta t^*}{1-t^*} = \frac{\Delta\widehat{y}}{\widehat{y}} \cdot \frac{1-t^*}{\Delta t^*}.$$

Assume ε_{RF} is independent of the tax rate τ. When we take the total derivative of this elasticity with respect to the tax rate τ and solve for p, we find:

$$0 = \frac{1}{\widehat{y}} \underbrace{\frac{d\Delta\widehat{y}}{dt^*}\frac{dt^*}{d\tau}}_{\frac{d\Delta\widehat{y}}{d\tau}} \left(\frac{1-t^*}{\Delta t^*}\right) + \frac{\Delta\widehat{y}}{\widehat{y}} \left[\frac{-\frac{dt^*}{d\tau}\Delta t^* - (1-t^*)\overbrace{\frac{d\Delta t^*}{dt^*}\frac{dt^*}{d\tau}}^{\frac{d\Delta t^*}{d\tau}}}{\left(\Delta t^*\right)^2} \right]$$

$$p = -\frac{1}{\tau\gamma}(1-\tau) + \frac{1}{\tau\gamma}\frac{\frac{\Delta\widehat{y}}{\tau}}{\frac{d\Delta\widehat{y}}{d\tau}},$$

where we assume that the audit rule threshold is independent of the tax rate. Finally, let's define $\varepsilon_{\Delta\widehat{y},\tau} = \frac{\frac{d\Delta\widehat{y}}{d\tau}}{\frac{\Delta\widehat{y}}{\tau}} \approx \frac{\partial \log \Delta\widehat{y}}{\partial \log \tau}$, the elasticity of the revenue response or length of the bunching interval to the tax rate, a quantity we can estimate in the data. We can thus rewrite the perceived audit risk as:

$$p \approx \frac{1}{\tau\gamma}\left[\frac{1}{\varepsilon_{\Delta\widehat{y},\tau}} - (1-\tau)\right]. \tag{I.1}$$

In our simplified set-up, the expression in (I.1) highlights the conditional negative relationship between audit risks (or the audit risk gap at the threshold) and the responsiveness of bunching firms to the tax rate:

$$\frac{\partial p}{\partial \varepsilon_{\Delta\widehat{y},\tau}} = -\frac{1}{\tau\gamma}\frac{1}{\left(\varepsilon_{\Delta\widehat{y},\tau}\right)^2} < 0.$$

Conditional on observing a given level of bunching, the model suggests that the smaller the revenue response to the tax component of reporting incentives, the larger must be the audit risk gap induced by disclosure that can explain the observed bunching.

Appendix J. MVPF estimation procedure

We detail the construction of the MVPF estimates presented in Section 8.3, starting from the definition of taxpayers' willingness to pay for disclosure in our structural model.

An expression for the willingness to pay comes from the application of the envelope theorem to the value function (2.1) with constant audit risks. We consider a small or marginal audit risk change in a pre-disclosure policy counterfactual, where audit risk is perceived to be constant at a level p_C. By the envelope theorem, any marginal audit risk update in our model results in a mechanical cost or gain to the taxpayer equal to the increased or decreased expected cost of detected evasion, respectively.

In particular, for taxpayers who would fall below the disclosed threshold in the policy counterfactual, audit risk weakly rises from p_C to p_H. Their negative willingness to pay amounts to:

$$WTP_{\text{Below},p_C} = -\left(p_H - p_C\right) \cdot \left(\tau \cdot \gamma \cdot e_C\right), \tag{J.1}$$

where e_C represents the level of revenues underreported in equilibrium with a constant audit risk p_C. On the other hand, for taxpayers who would fall at or above the disclosed threshold in the counterfactual distribution, audit risk weakly decreases from p_C to p_L. Thus, their positive willingness to pay amounts to:

$$WTP_{\text{Above},p_C} = \left(p_C - p_L\right) \cdot \left(\tau \cdot \gamma \cdot e_C\right). \tag{J.2}$$

We estimate our MVPF ratio in two steps. At the numerator, we compute the willingness to pay of taxpayers falling below and above the threshold across all taxpayer groups defined for our structural estimation, separately for each of several levels of counterfactual risk p_C. For each of these levels, we average the willingness to pay of all groups by their SeS file share. Notice that the share of taxpayers falling below and above the threshold in each group also depends on the selected level of risk p_C.

At the denominator, we measure the mean net cost of disclosure to the government as the difference between the administrative costs in Section 8.3, divided by the yearly number of SeS files in the structural analysis, and our estimates of the increase in mean reported revenues $\bar{y}_D - \bar{y}_C$ at each level of counterfactual audit risk, taxed at the average personal income tax rate weighted across all structural groups.

Appendix K. Event study estimates robustness

Recent contributions on two-way fixed effects estimation have elucidated a number of potential issues in interpreting the dynamic treatment coefficients of standard event-study designs. In (9.1), we set $-k$ so that the two earliest relative treatment periods in our data do not get individual dummies. This should address the concerns in Borusyak et al. (2021) that relative treatment period coefficients are not point-identified in a fully-dynamic OLS model without never-treated units. However, our results might still spuriously capture treatment effect heterogeneity across years and sectors or across treatment cohorts.

We thus re-estimate our baseline results on the tax base effects of the reward regime with the two alternative estimators proposed in de Chaisemartin and D'Haultfœuille (2020) and Sun and Abraham (2020), respectively.[61] Applying these methods comes at the cost of reducing the number of post-treatment effects we can estimate, since both require using some treatment cohort as a comparison group in any given year. Still, the pattern of our baseline results hold when we apply the alternative estimators.

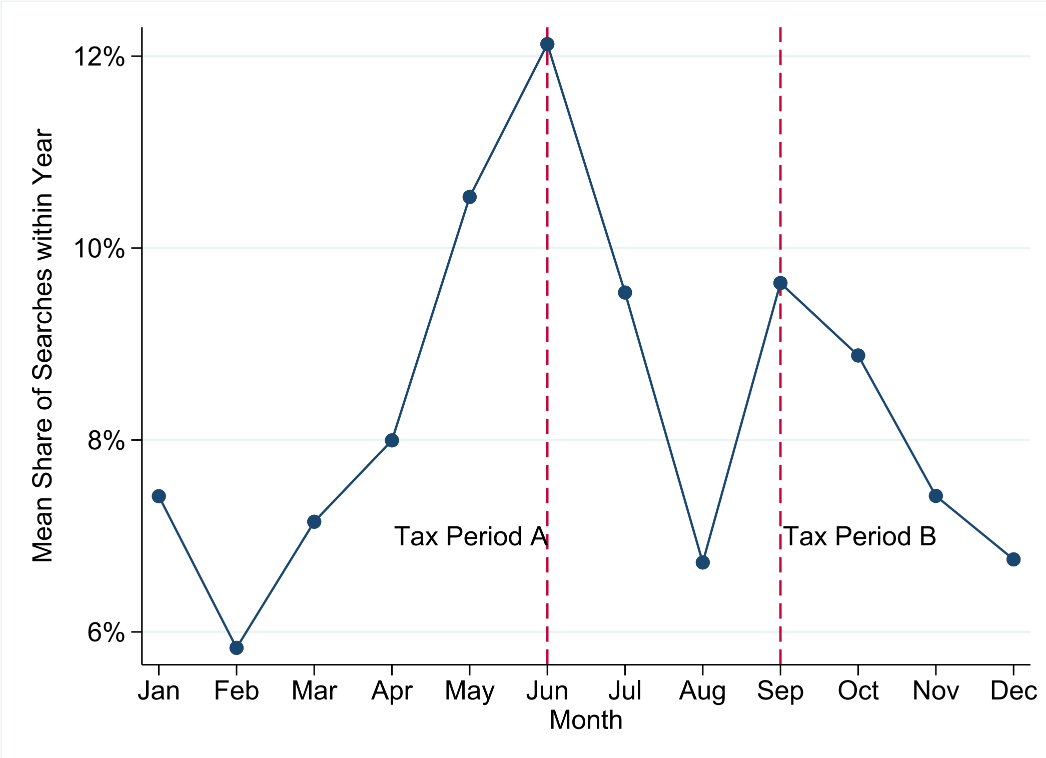
Figure A22 and Figure A23 compare the treatment effect dynamics of our baseline fixed-effects estimator with that from the interaction-weighted (IW) estimator proposed by Sun and Abraham (2020). The IW estimator ensures that each relative treatment period coefficient is a convex combination of the relevant cohort-period average treatment effects and is robust to treatment heterogeneity across cohorts. We follow the authors' recommendation using the last two reward regime cohorts as control groups, and show that the available IW point estimates of the reform impact are within the baseline confidence intervals.[62]

Figure A24 and Figure A25 implement the DID_M estimator from de Chaisemartin and D'Haultfœuille (2020), which is robust to treatment effect heterogeneity over time and across groups. Provided that this estimator compares sectors switching treatment status with those that never switch, we are only able to estimate four treatment effects in the post-reform period. Even so, we confirm that outcome differences across sectors are relatively stable in the pre-treatment period, while entry into reform is associated with a gradual but steady rise in average reporting behavior across all outcomes.

[61] The approach in Sun and Abraham (2020) might be especially useful in our setting, given that the technical criteria with which sectors are introduced to the new regime change slightly across some of our sample years. To the extent that these changes are correlated with potential outcomes, treatment heterogeneity among cohorts may emerge. The causal interpretation of our baseline results would then deserve caution.

[62] The last two cohorts of reward regime sectors are introduced in the 2014 and 2015 tax years. Only two new sectors are treated starting in 2015, so we opt for including the larger 2014 cohort among the controls. This control selection approach requires us to drop tax years starting from 2014.

Figure A3. Google searches for "gerico" spike in tax periods, 2004-2017

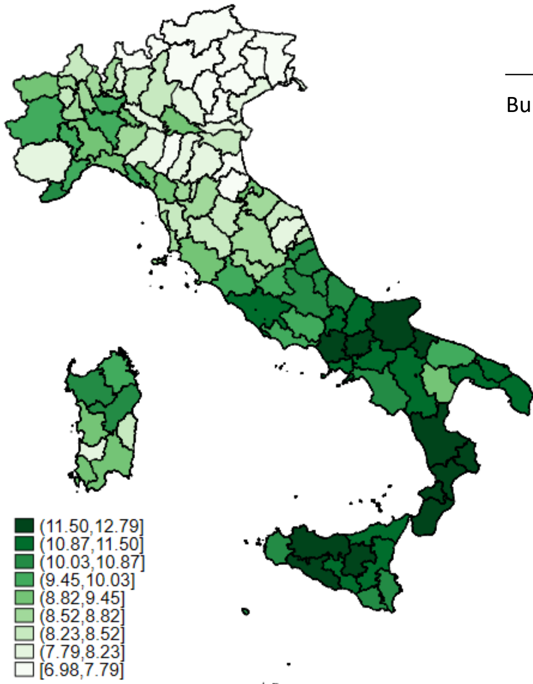


Notes: this Figure shows the month-by-month average intensity of Google searches for "*gerico*" over the 2004-2017 period in Italy. This time frame fully includes our SeS sample period, which stretches over the 2007-2016 tax years and the 2008-2017 filing years. Month-level data come from *trends.google.com*. 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.

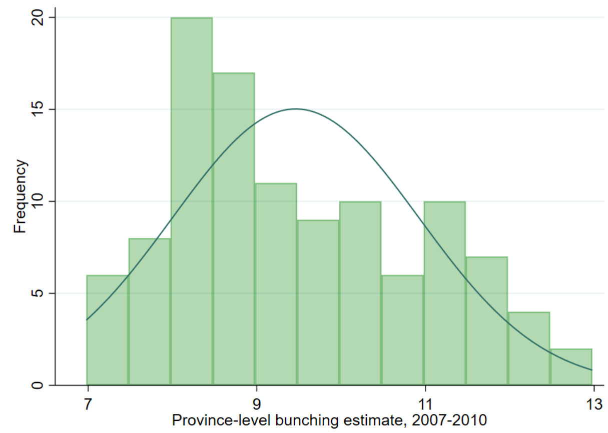
Figure A4. Local heterogeneity in bunching estimates, 2007-2010

Panel A: Bunching across 110 provinces

Structural Bunching Estimate, 2007-2010

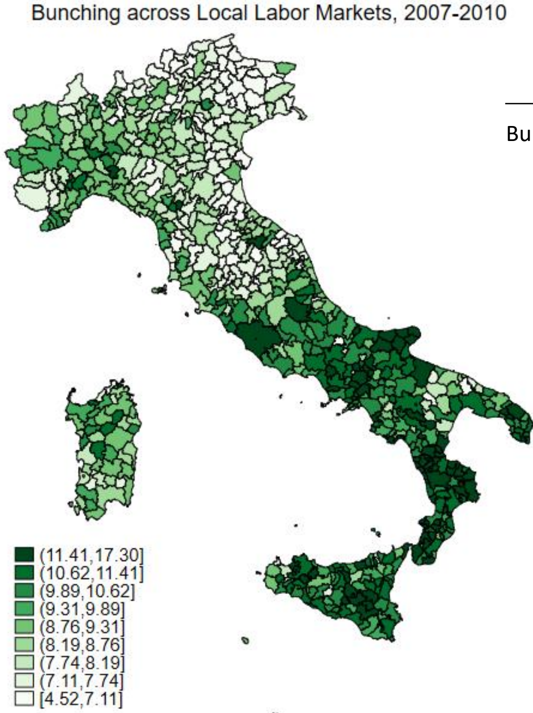


	N. Obs.	Mean	Median	SD	Min	Max
Bunching	110	9.46	9.23	1.46	6.98	12.79

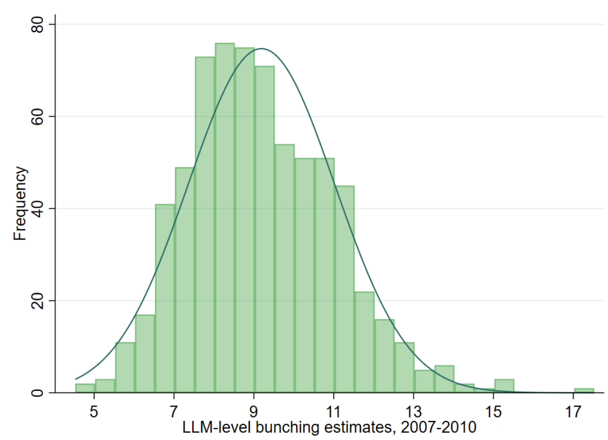


Panel B: Bunching across 686 local labor markets (2001 LLMs)

Bunching across Local Labor Markets, 2007-2010

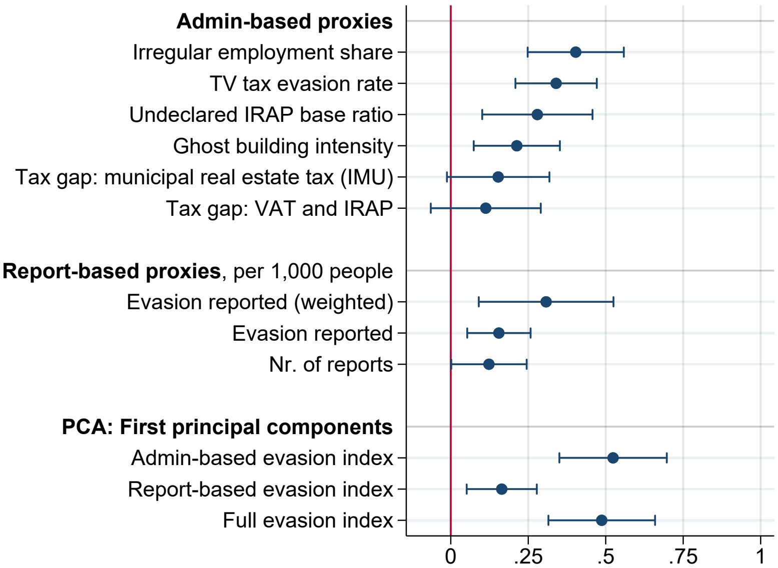


	N. Obs.	Mean	Median	SD	Min	Max
Bunching	686	9.19	8.99	1.83	4.52	17.30



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 procedure and sample definition we employ are the same as those outlined below Figure 5.

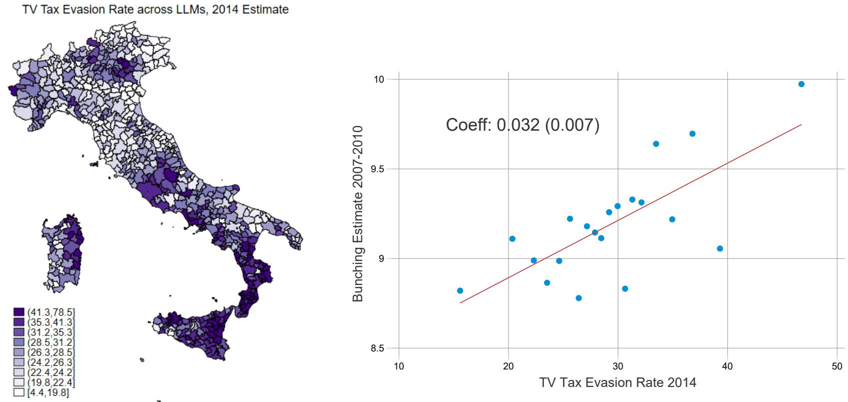
Figure A5. Provincial bunching correlates positively with local evasion



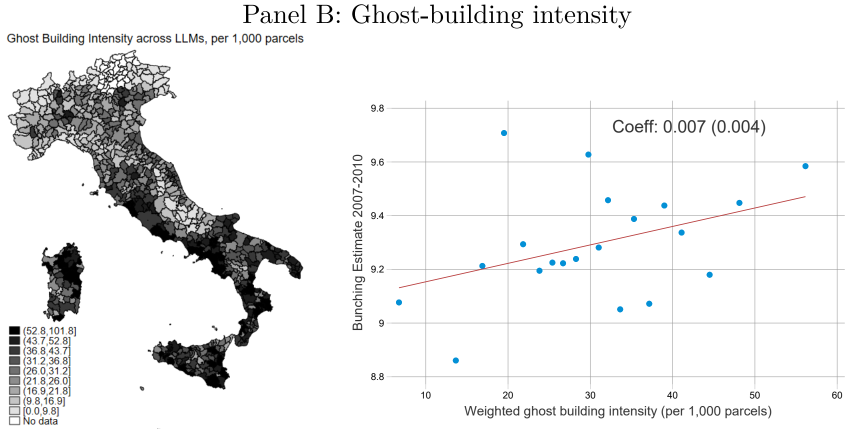
Notes: the Figure plots the standardized coefficients β and their 95% CIs from several regressions of SeS bunching on evasion proxies j across 110 provinces i according to the model: $\text{bunching}_i = \alpha + \beta \text{Evasion}_{j,i} + \gamma \log \text{VA pc}_i + \text{macroregion}_i + \varepsilon_i$. Standard errors are robust to heteroskedasticity. Bunching at the SeS presumed revenues is computed at the province level over the 2007-2010 tax period with the procedure and sample restrictions described below Figure 5. Evasion proxies and their sources are described in Appendix C. 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 A6. LLM bunching correlates positively with local evasion

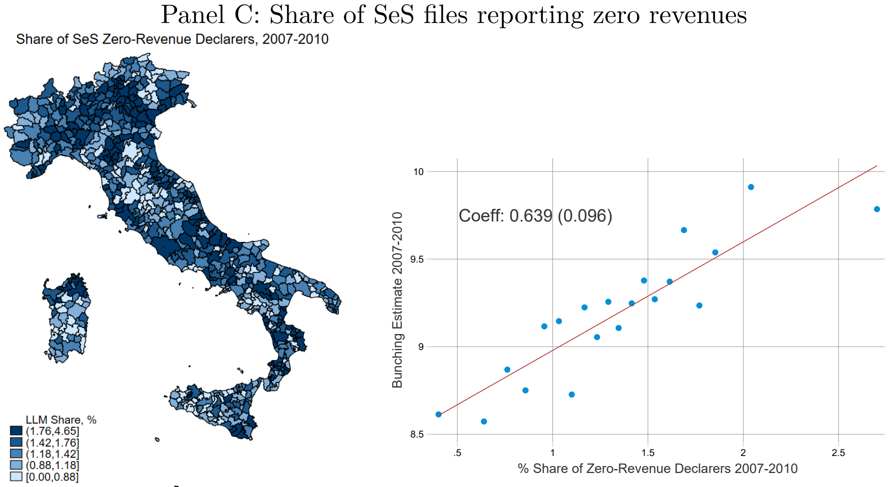
Panel A: TV tax evasion rate



Panel B: Ghost-building intensity

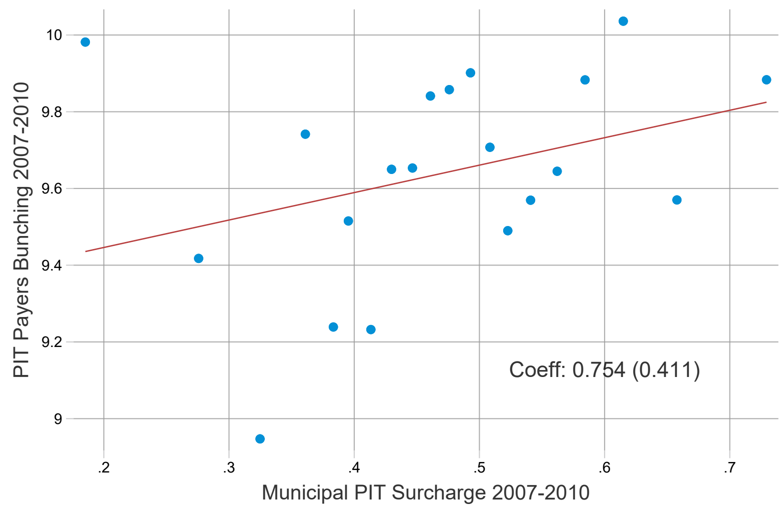


Panel C: Share of SeS files reporting zero revenues



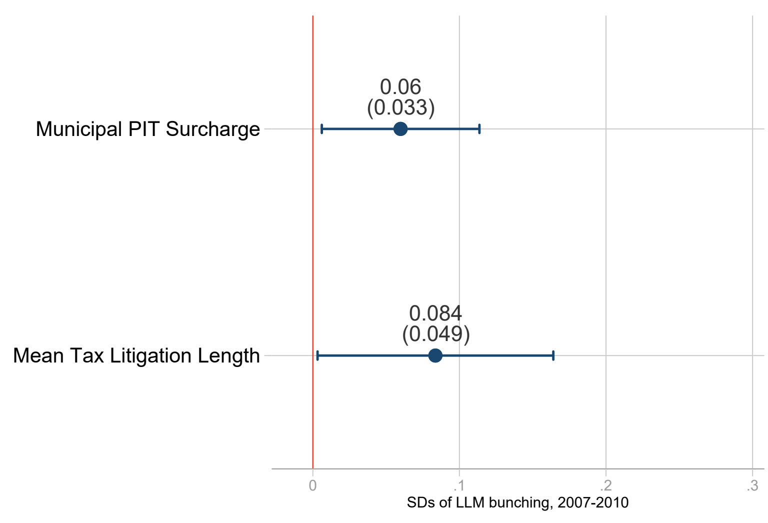
Notes: the Figure maps three LLM-level estimates of behaviors that are plausibly related to evasion or misreporting, and correlates each with local SeS bunching. The three evasion proxies are defined in Appendix C. On the right, binned scatterplots report the main slope coefficient and robust standard error from a regression of the form $\text{bunching}_i = \alpha + \beta Evasion_{j,i} + \gamma \log(\text{PIT base per taxpayer}_i) + \text{region}_i + \varepsilon_i$, 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%.

Figure A7. Bunching tracks evasion incentives: municipal taxes



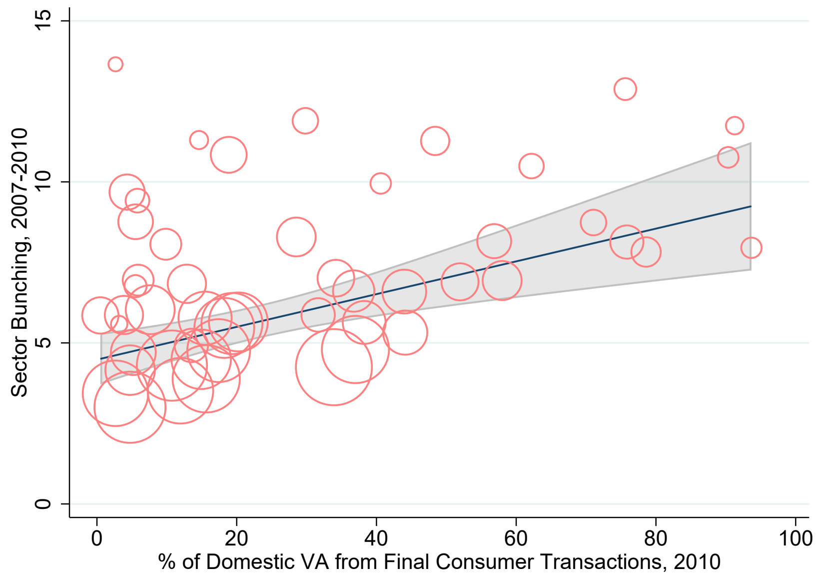
Notes: this Figure provides a binned scatterplot and the linear fit for the relation between SeS bunching among PIT-payers and the weighted average of municipal PIT surcharges for the 2007-2010 period at the LLM level. We also report the main slope coefficient and robust standard error from a regression of the form $\text{bunching}_{i,j} = \alpha + \beta(\text{PIT surcharge}_i) + \gamma \text{Litigation}_j) + + \delta \log(\text{PIT base per taxpayer}_i) + \text{region}_i + \varepsilon_{i,j}$, including the 2009-2012 mean length of litigation at the tax court of province j, 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.

Figure A8. Comparing variations: small tax vs. large administrative costs



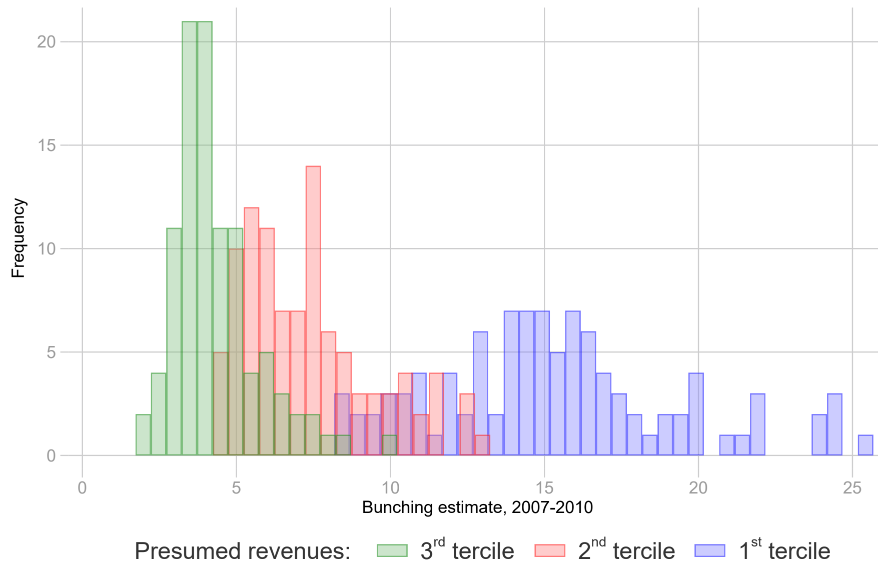
Notes: this Figure provides the standardized coefficient plot for the relation between SeS bunching among PIT-payers in 2007-2010 and two main explanatory variables: the 2007-2010 weighted average of municipal PIT surcharges at the LLM level and the 2009-2012 mean length of tax litigation at the provincial court level. We provide details for the construction of both variables in Appendix C. Municipal PIT surcharges can't exceed the national PIT schedule rates by more than 0.8%, while tax litigation lasts on average between 1 and 2 years. We use observations from 624 LLMs i in 103 provinces j in a single regression of the form: $\text{bunching}_{i,j} = \alpha + \beta(\text{PIT surcharge}_i) + \gamma(\text{Litigation}_j) + + \delta \log(\text{PIT base per taxpayer}_i) + \text{region}_i + \varepsilon_{i,j}$, including regional fixed effects and the logarithm of the average local PIT-base per individual taxpayer in the LLM. Point estimates are represented with their robust 90% CIs.

Figure A9. Bunching tracks evasion potential: downstream sectors



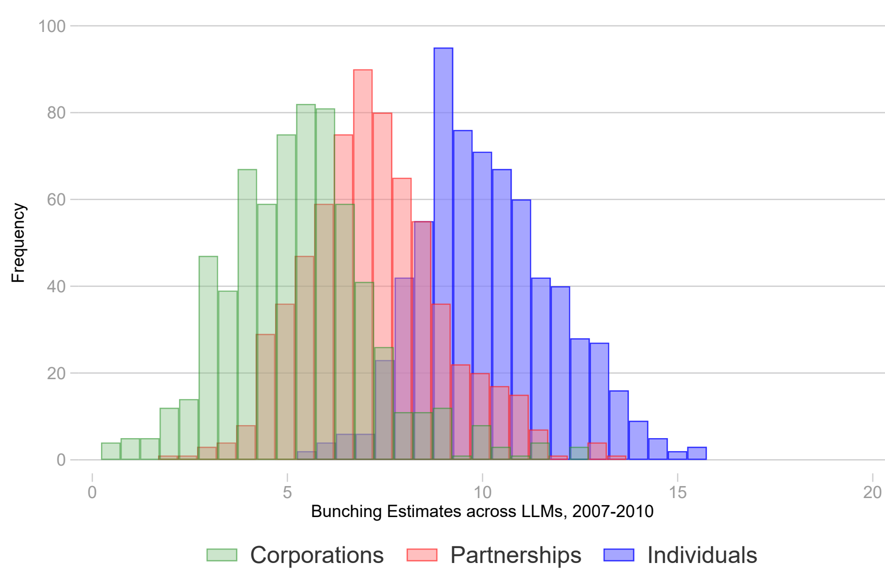
Notes: this Figure shows the sector-level scatterplot and linear fit of the relation between 2007-2010 bunching and the degree of relative exposure to the final consumer in 2010. Exposure is defined as a business sector's share of domestic value added that is determined by final consumption (see details in Appendix C). Value added is in 2010 current prices. 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).

Figure A10. Bunching reflects business size: revenue terciles



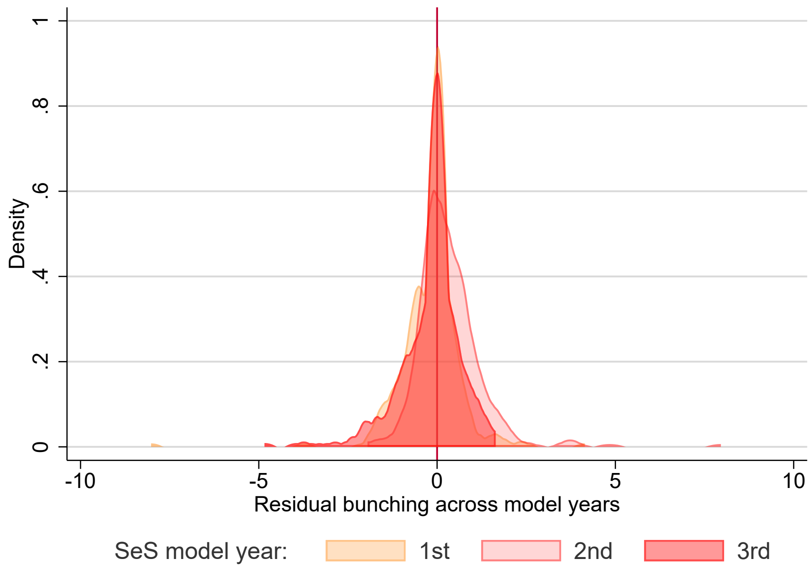
Notes: this Figure plots the 2007-2010 bunching estimates distribution for the 300 PIT-payer groups used in our structural analysis. Each group is a combination of 20 regions, 5 macro-sectors, and 3 presumed revenue terciles defined within each region-sector pair. We separate estimates based on whether they belong to groups in the first (green), second (red), and third (blue) relative presumed revenues tercile. Bunching is estimated as described below Figure 5, based on the relative distance of reported revenues from presumed revenues divided by presumed revenues.

Figure A11. Bunching tracks evasion potential: legal complexity



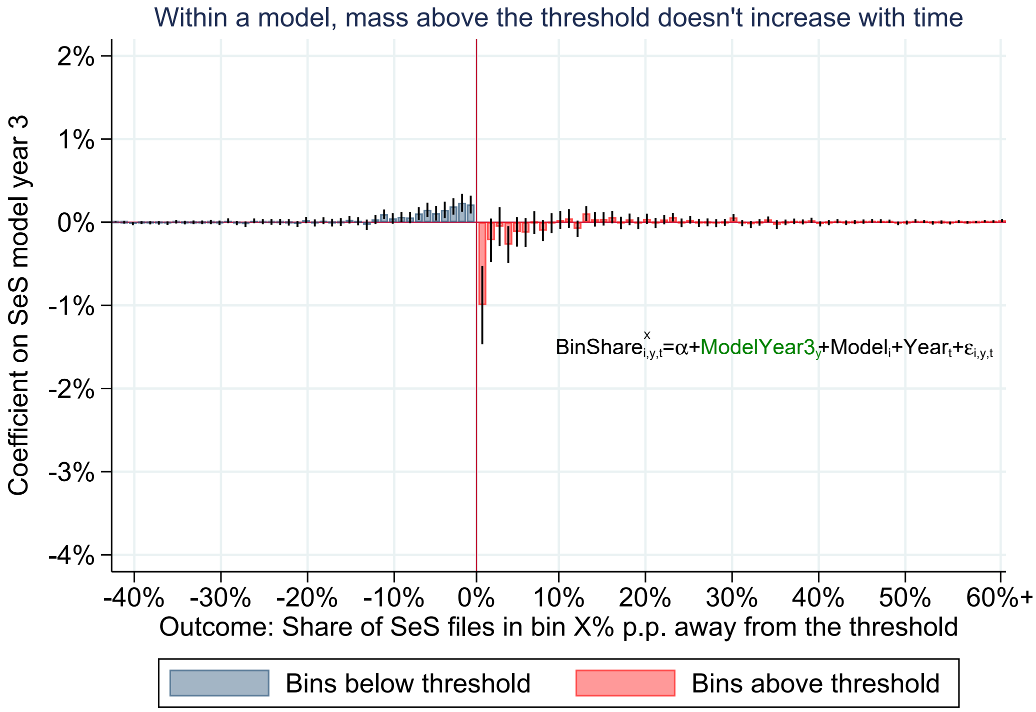
Notes: this Figure plots the distribution of 2007-2010 bunching estimates computed at the LLM-level, separately for individuals (individual businesses and self-employed individuals) in blue, partnerships in red, and corporations in green. SeS taxpayers face increasing reporting and book-keeping requirements as they formalize, with accounting complexity rising from a relatively low level in individually-owned activities to a progressively higher level among partnerships first and corporations next. For the purposes of this graph, we exclude 4% of estimates that are negative or in the 99^{th} percentile of the distribution.

Figure A12. Bunching evolution within SeS models, 2007-2010



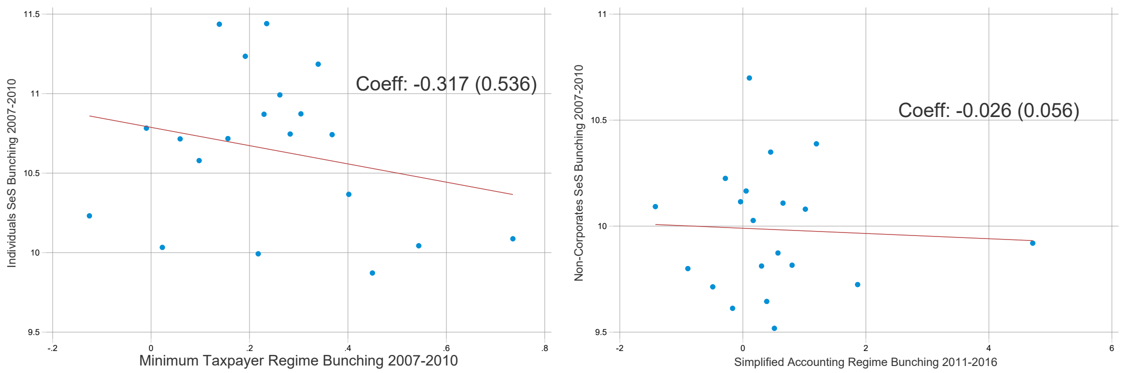
Notes: this Figure plots the distribution of 2007-2010 bunching residuals from a regression of the form: $\text{bunching}_{i,t} = \alpha + \beta_i + \gamma_t + \varepsilon_{i,t}$, where the unit of observation is a SeS model-year, and we include fixed effects for each SeS model i and calendar year t. 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 A13. Bin share effect of the last year of application of a SeS model



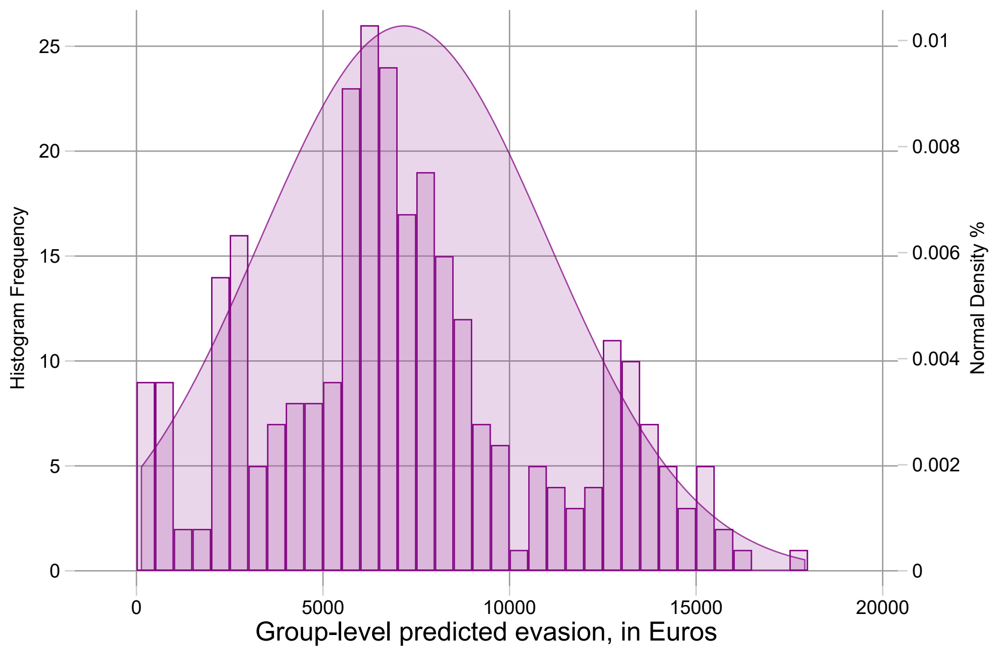
Notes: this Figure provides the coefficient plot from several regressions as the one printed above. Specifically, we observe SeS models i and consider whether they are being applied for the y^{textth} year (that is, first, second, or third year) during calendar year t. Across all SeS model-years, we 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.

FIGURE A14. Bunching at SeS presumed revenues vs. at two other thresholds



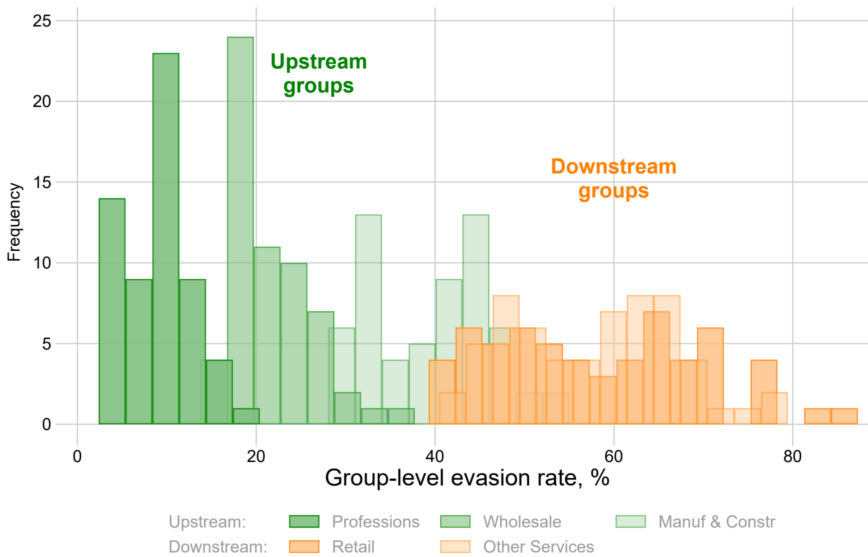
Notes: this Figure displays the (insignificant) province-level linear relationship between bunching at the SeS presumed revenues and bunching at two other thresholds around which the business cost of navigating the tax enforcement system changes discretely. On the left, we consider SeS bunching among individuals in 2007-2010 and bunching at the minimum taxpayer regime threshold (€30,000 in revenues up to 2011) below which individuals could opt out of the SeS system. On the right, we consider SeS bunching among PIT-payers (individuals and partnerships) in 2007-2010 and bunching at the simplified accounting regime threshold for service-sector businesses (€400,000 in revenues). A 2011 reform halved the maximum duration of on-site audit inspections for firms below the latter threshold. Coefficients and robust standard errors come from two regressions of the form: SeS $\text{bunching}_i = \alpha + \beta \text{Other bunching}_i^j + \gamma \log \text{VA pc}_i + \text{macroregion}_i + \varepsilon_i$ for provinces i and bunching at each alternative threshold j.

FIGURE A15. Group-level evasion levels with constant audit risk $p_C = p_L$



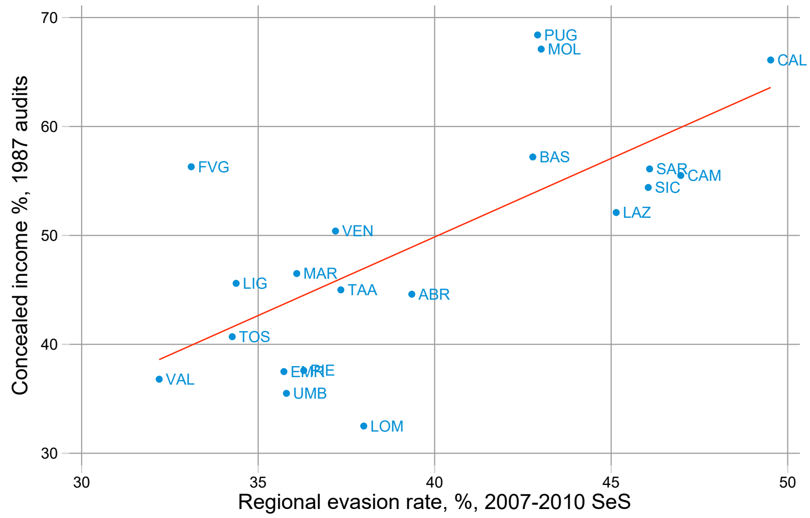
Notes: this Figure plots the distribution of average revenue evasion predicted by our structural model for 300 groups of PIT-payers filing for SeS over 2007-2010. Groups are each the combination of 20 regions, 5 macro-industries, and 3 presumed revenue terciles defined within each region-industry pair. Evasion levels are defined as the average revenues underreported in each group in the model's equilibrium with constant audit risk $p_C = p_L$. We define evasion in this equilibrium as $e_L = k_e \left(\tau - p_L \cdot \gamma \cdot \tau\right)^{\varepsilon_e}$. In the background, we add a normal fit to smooth out the raw distribution of evasion levels.

Figure A16. Validation: upstream vs. downstream evasion rates with $p_C = p_L$



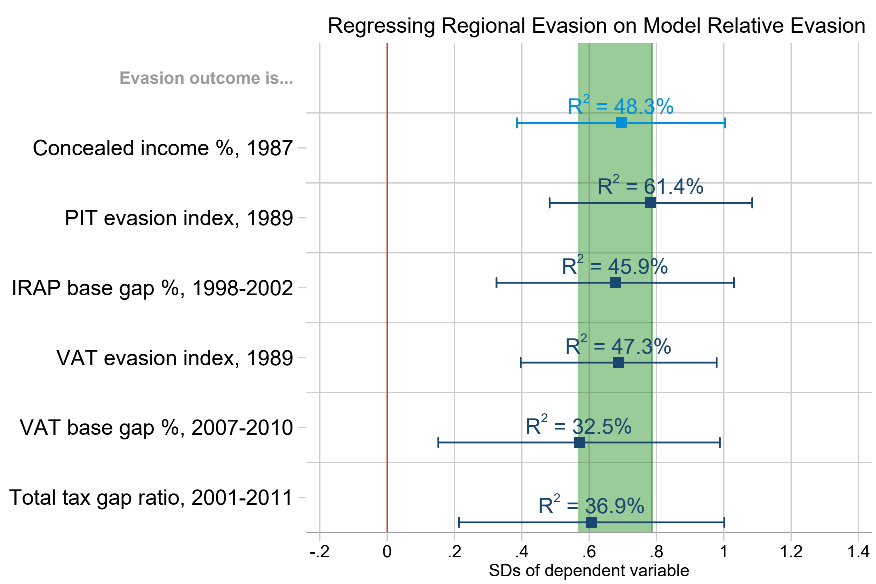
Notes: this Figure plots the distribution of evasion rates predicted by our structural model for 300 groups of PIT-payers filing for SeS over 2007-2010, split across five macro-industries and their relative exposure to the final consumer. Evasion rates are defined as in Figure 7. We label a macro-industry to be "downstream" or "upstream" along the supply chain based on whether their share of domestic value added determined by final consumption is above or below the median in *ISTAT*'s 2010-2013 input-output tables. Each macro-industry value added share is in turn the average of the domestic value added shares of the underlying ATECO two-digit sectors also appearing in the SeS.

Figure A17. Validation: audit-based vs. predicted evasion intensity



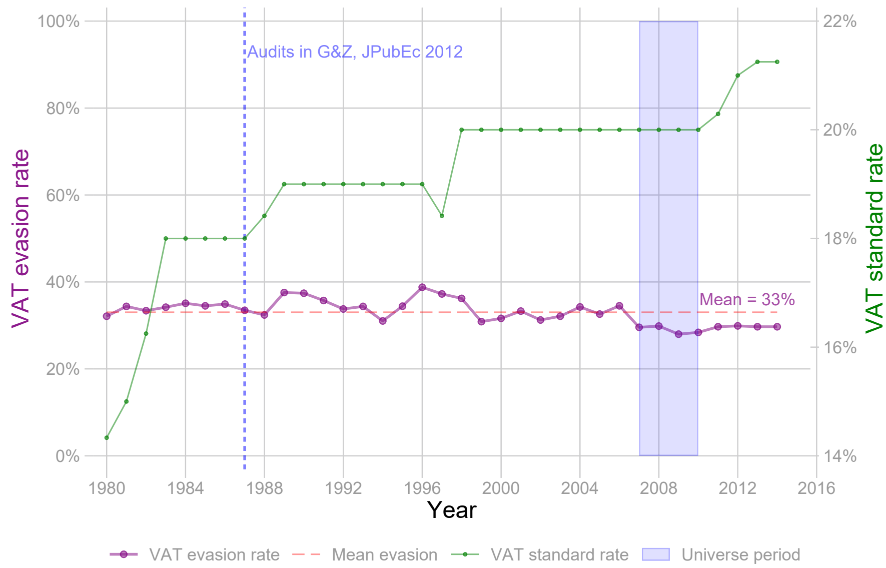
Notes: this Figure provides a scatterplot and the linear fit for the relation between twenty regional evasion rates predicted by our structural model and the share of concealed income uncovered in each region by the universe of 1987 Italian Tax Police audits of individual businesses and the self-employed (Galbiati and Zanella, 2012). Regional evasion rates are an average of the evasion rates predicted for the 300 PIT-payer groups used in our structural analysis weighted by the number of their 2007-2010 SeS files. Group-level evasion rates are the ratio of the average revenues underreported in the model's equilibrium with constant audit risk $p_C = p_L$ and the median reported gross profits observed in the data. The share of concealed income is the difference between average taxable income attested by the auditors and reported by the taxpayer as a percentage of the average attested taxable income.

Figure A18. Validation: actual vs. predicted evasion intensity



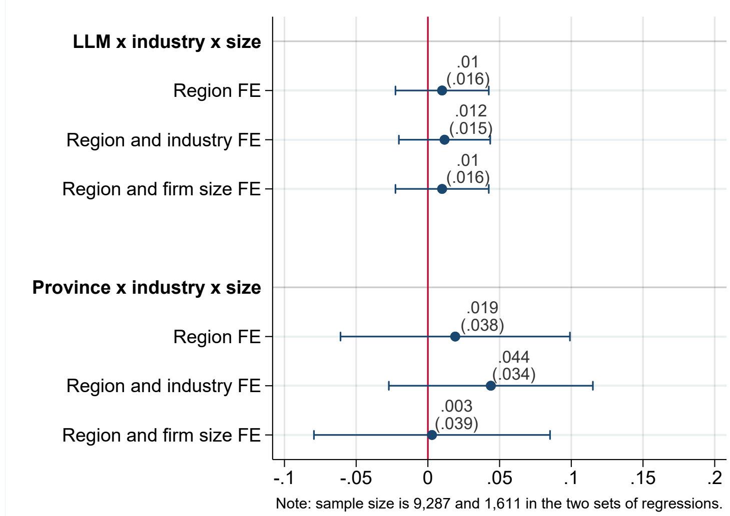
Notes: this Figure provides summary statistics from several bivariate regressions of regional evasion proxies available in the literature on the regional evasion rates predicted by our model and defined below Figure A17. Regional evasion proxies are defined in Appendix C. For each regression, we report the standardized coefficient associated to our model's regional evasion rate, its robust 95% CI, and the model's R-squared as a measure of fit. The green vertical band delimits the level of the smallest and the largest standardized coefficients.

Figure A19. Comparing VAT evasion rates across periods



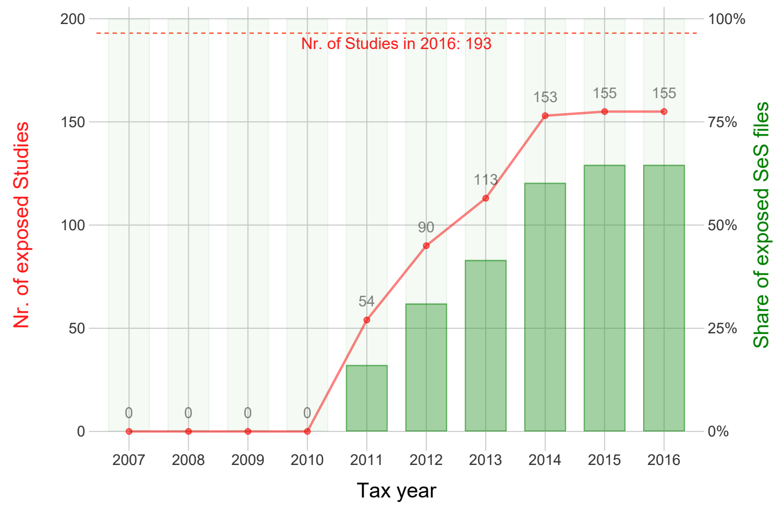
Notes: this Figure puts into perspective the time frame for our model's measure of underreporting (2007-2010) and that for concealed income in Galbiati and Zanella (2012) (1987). We do so by plotting the time series for the evasion proxy that is most consistently estimated over time, i.e. the VAT gap. Estimates come from Table 3.6 in Ministry of Economy and Finance (2011) until 2009 and Table 3.B.2 in Ministry of Economy and Finance (2016) for later years. The VAT gap is defined as the underreported VAT base as a share of the potential VAT base. The evolution of VAT evasion (in purple) appears to be relatively stable despite the long time span considered and the gradual increase in the legislated VAT standard rate (in green).

Figure A20. Subregional SeS revenue response-tax correlations, 2007-2010



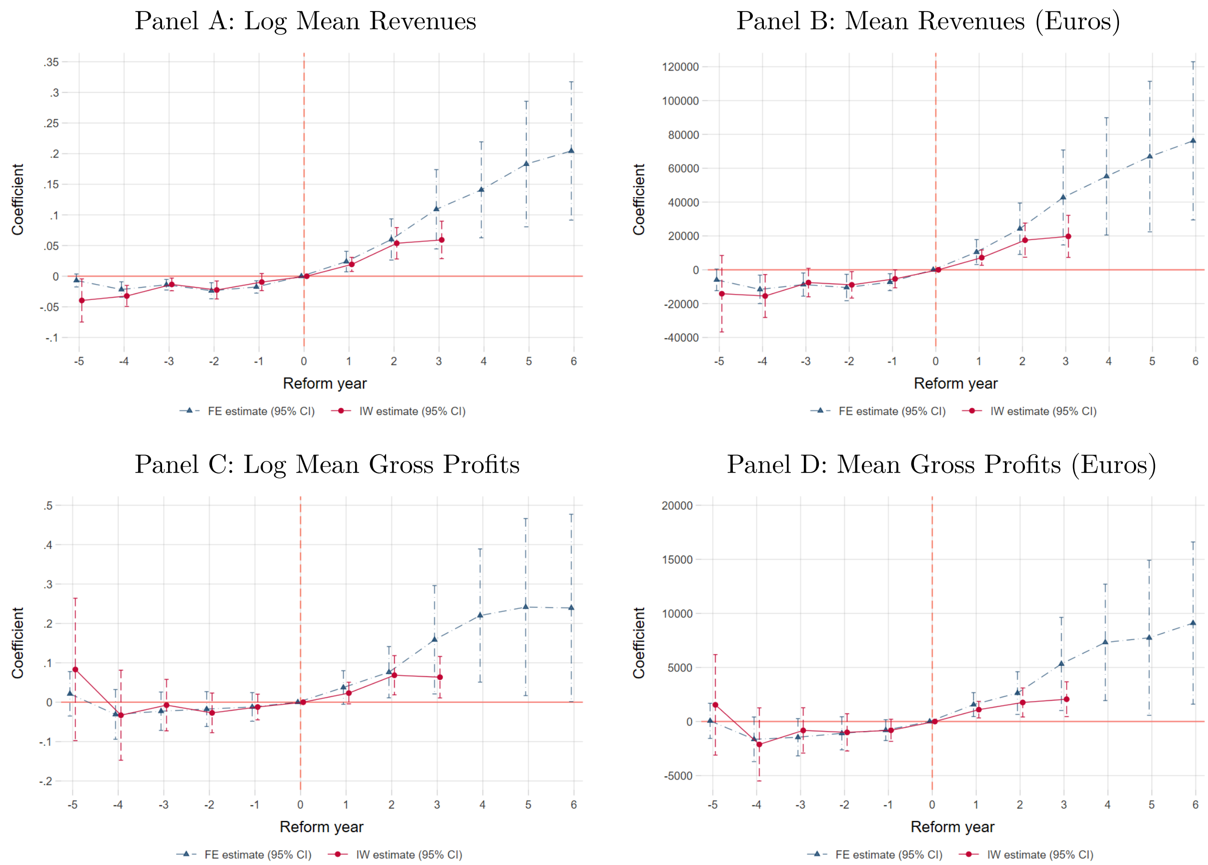
Notes: the Figure shows the main coefficient and 95% CIs from each of six regressions. We regress the 2007-2010 SeS revenue responses among PIT payers on the mean municipal PIT surcharge computed at the level of observation. We define units (in bold) at the subregional level (9,287 with local labor markets or 1,611 with provinces) by macro-sector by relative presumed revenue tercile as discussed in our structural estimation. Conditional on the displayed fixed effects choice, specifications follow: $\log{(\Delta\hat{y})_i} = \alpha + \beta \log(\text{Mean PIT Surcharge})_i + \text{region}_i + \text{sector/size}_i + \varepsilon_i$. We cluster standard errors at the regional level and trim observations at the 1^{st} and 99^{th} percentile of revenue responses.

Figure A21. 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.

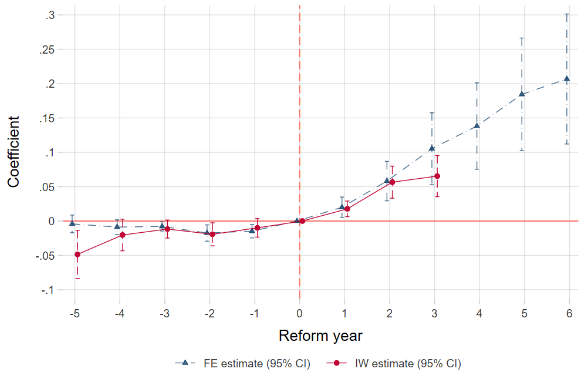
Figure A22. Robustness: Sun and Abraham (2020) reward regime effects (1)



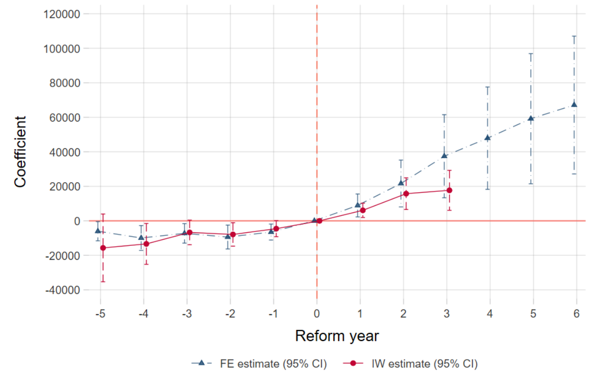
Notes: this Figure compares the effects of the reward regime's introduction in a sector from our baseline fixed-effects specification in (9.1) (FE) and that proposed by Sun and Abraham (2020) (IW). Details on sample, outcomes, and standard errors are discussed below Figure 12, while controls and weighting are as defined in the discussion of Eq. (9.1). 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. The IW estimator comes from a regression saturated in treatment cohort indicators and relative treatment period dummies. As in (9.1), we do not include dummies for the first two available pre-treatment periods and the relative period just before the reform. Following Sun and Abraham (2020), we use the last treatment cohorts as control. Given that the last cohort in the 2015 tax year consists only of two sectors, we define a broader control group starting from the 2014 tax year. This reduces the number of IW post-treatment coefficients we can estimate relative to the baseline FE specification.

Figure A23. Robustness: Sun and Abraham (2020) reward regime effects (2)

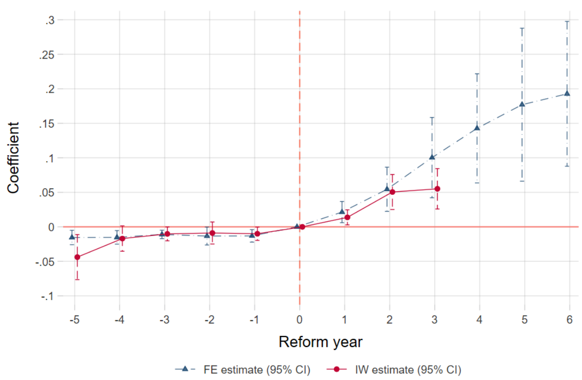
Panel A: Log Mean Total Costs



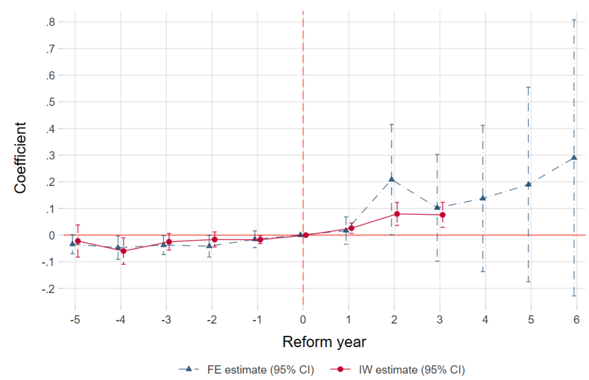
Panel B: Mean Total Costs (Euros)



Panel C: Log Mean Employment Costs

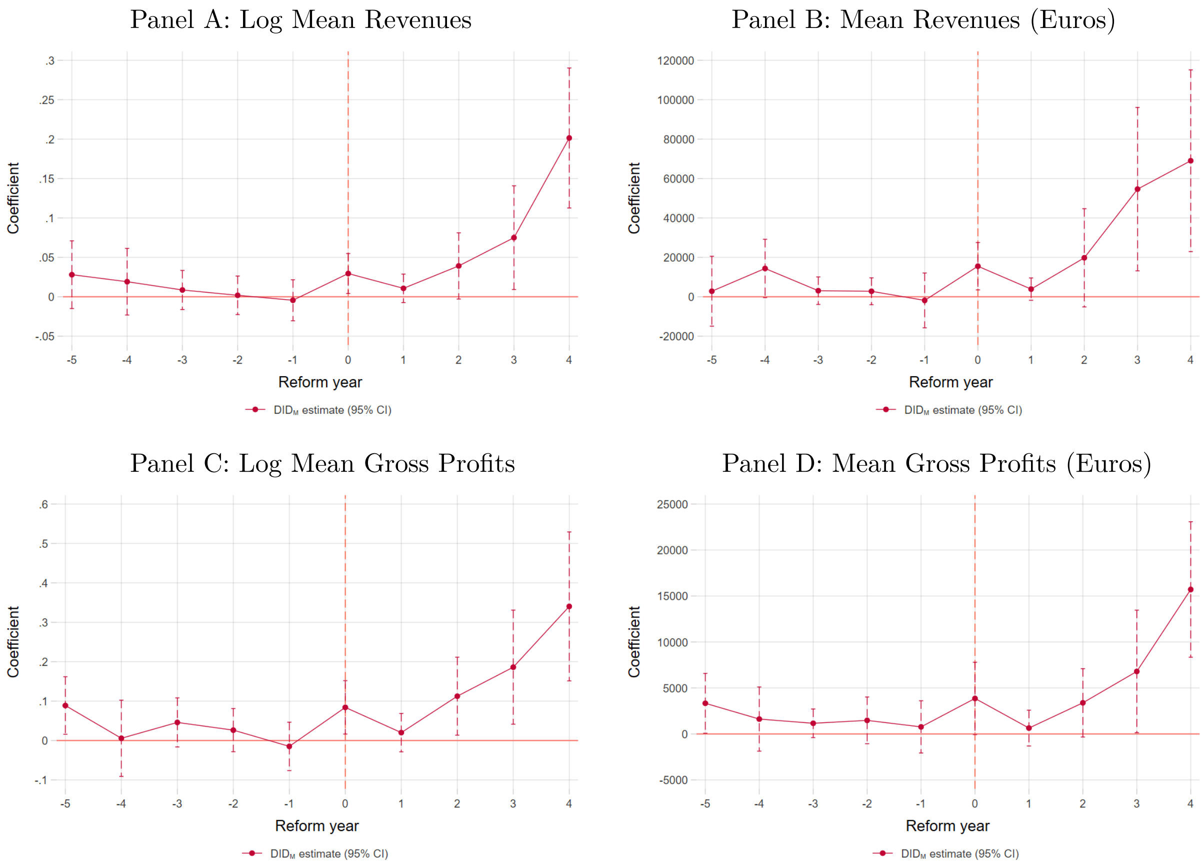


Panel D: Number of Employees



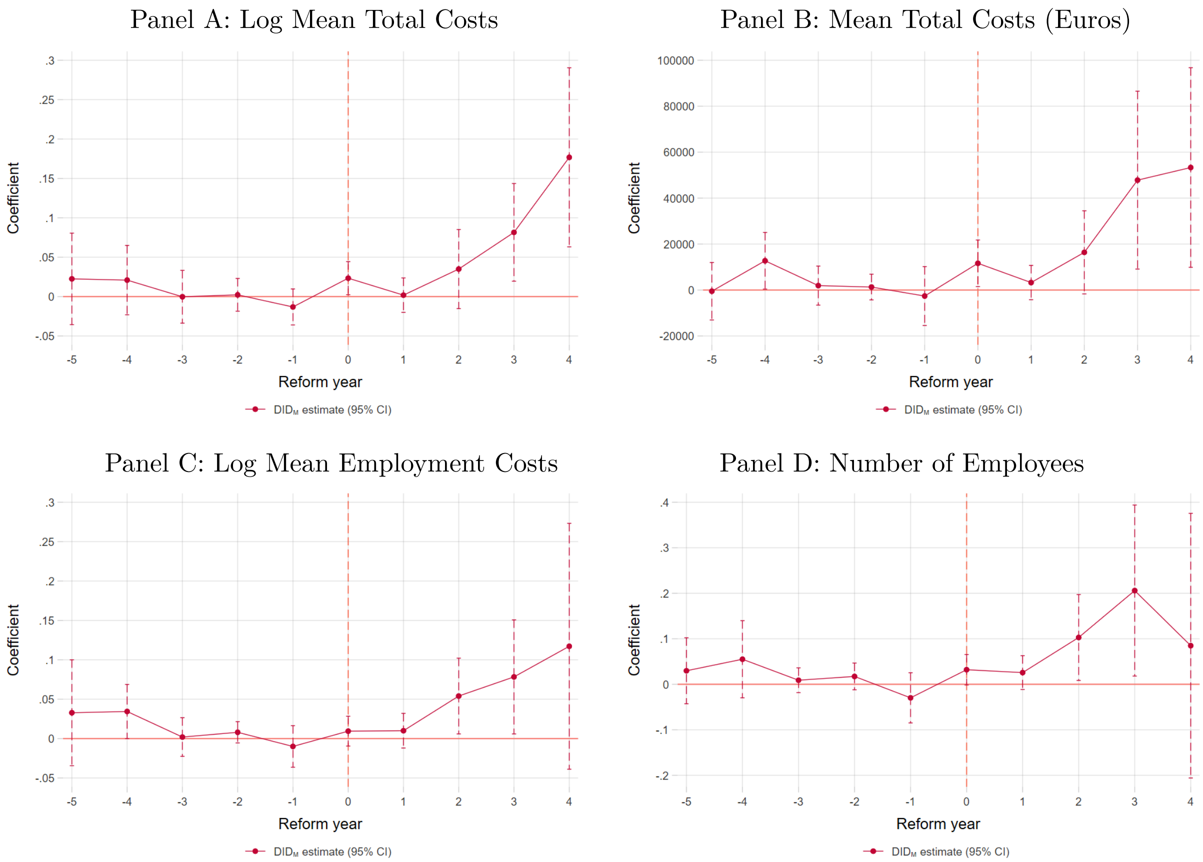
Notes: this Figure compares the effects of the reward regime's introduction in a sector from our baseline fixed-effects specification in (9.1) (FE) and that proposed by Sun and Abraham (2020) (IW). Details on sample, outcomes, and standard errors are discussed below Figure 13, while controls and weighting are as defined in the discussion of Eq. (9.1). 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. The IW estimator comes from a regression saturated in treatment cohort indicators and relative treatment period dummies. As in (9.1), we do not include dummies for the first two available pre-treatment periods and the relative period just before the reform. Following Sun and Abraham (2020), we use the last treatment cohorts as control. Given that the last cohort in the 2015 tax year consists only of two sectors, we define a broader control group starting from the 2014 tax year. This reduces the number of IW post-treatment coefficients we can estimate relative to the baseline FE specification.

Figure A24. Robustness: de Chaisemartin and D'Haultfoeuille (2020) reward regime effects (1)



Notes: this Figure shows the effects of the reward regime's introduction in a sector from the estimator proposed by de Chaisemartin and D'Haultfœuille (2020) (DID$_M$). Details on sample and outcomes are discussed below Figure 12, while controls and weighting are as defined in the discussion of Eq. (9.1). 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.

Figure A25. Robustness: de Chaisemartin and D'Haultfoeuille (2020) reward regime effects (2)



Notes: this Figure shows the effects of the reward regime's introduction in a sector from the estimator proposed by de Chaisemartin and D'Haultfœuille (2020) (DID_M). Details on sample and outcomes are discussed below Figure 13, while controls and weighting are as defined in the discussion of Eq. (9.1). 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.

Table A2. 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* $e \leq 40\% \cdot y$, $e \leq €50{,}000$	
✓	✓	✓		**1. No analytic-inductive audits** *up to any amount* **2. No synthetic audits** *up to* $\pi(s)\text{-}\pi \leq 33\% \cdot \pi(s)$ **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 A3. 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.

Table A4. *evasori.info*: summary statistics of evasion reports, 2008-2011

Data Year	Reports	Median	Mean	Provinces	Sectors
2008	$85,622$	10	193.12	109	102
2009	$99,983$	10	122.76	109	102
2010	$155,772$	10	114.23	109	104
2011	$278,956$	10	153.09	109	110
2008-2011	$\mathbf{620,333}$	**10**	**143.97**	**109**	**110**

Notes: the Table summarizes the information we draw from the citizen-supplied evasion reports submitted to *evasori.info*. This online platform was launched by an anonymous computer science professor as a citizen engagement initiative to raise awareness on the issue of tax evasion and revenue underreporting in Italy. We access all geolocated evasion reports submitted between 2008 and 2011 through a Google API set up by the website at *evasori.info/api*. Reports usually reflect instances when a business' customer is not released a receipt following a purchase transaction, but might also involve relatively large amounts as in the case of informal salaries paid to irregular workers. In th Table, median and mean refer to the median and mean Euro amount of the reported evasion instances, while province and sector refer to the number of reported provinces and business sectors we count among all reports in a given year.

Table A5. MVPF decomposition across audit counterfactuals

Mean: $\Delta p_C = \%(p_H - p_L)$	WTP_{Above} €	WTP_{Below} €	$\%Below$ %	Tax Gain €	Admin Cost €	MVPF Ratio value
0	0	-139.70	53.12	1,805.95	4.89	0.04
10	13.62	-122.56	53.12	1,774.26	4.89	0.03
20	26.55	-106.18	52.87	1,743.00	4.89	0.02
30	38.80	-90.53	52.04	1,712.17	4.89	0.01
40	50.39	-75.59	50.94	1,681.76	4.89	0.01
50	61.35	-61.35	49.85	1,651.77	4.89	-0.00
60	71.68	-47.79	49.04	1,622.21	4.89	-0.01
70	81.40	-34.89	48.16	1,593.07	4.89	-0.02
80	90.54	-22.63	47.22	1,564.34	4.89	-0.03
90	99.10	-11.01	46.15	1,536.04	4.89	-0.03
100	107.10	0	45.34	1,508.15	4.89	-0.04

Notes: the Table decomposes our estimates of the MVPF from disclosing the SeS presumed revenues threshold in 16 counterfactual scenarios indexed by Δp_C (Col. 1). MVPF estimates are displayed in the last column (Col. 7). We compute the MVPF numerator as the average willingness to pay for an audit risk change from p_C to p_L or p_H for taxpayers falling above (Col. 2) or below (Col. 3) the disclosed threshold in each of the 300 PIT-payer groups we employ in our structural analysis, with weights given by the share of SeS files in each group. For each counterfactual, we also report the weighted mean share of files falling below the threshold across all structural groups, with weights equal to each group's file share (Col. 4). We define the MVPF denominator as the mean administrative cost of disclosure, proxied by the total value of *SOSE*'s production in 2010 divided by the yearly number of 2007-2010 files used in the structural analysis (Col. 6), net of the gain in mean reported revenues we estimate in Section 8 taxed at the average PIT rate computed across all structural groups (Col. 5). In each counterfactual scenario, taxpayers perceive audit risks to be constant at $p_L + \Delta p_C$. We assign p_L to each group based on the audit risk estimate relevant to its administrative region of reference. Increments Δp_C are a percentage of the regional gap between p_L and p_H as indexed in Col. 1. This ensures that the overall counterfactual risk varies exactly between p_L and p_H in each region. In the Table, all estimates are rounded to the second decimal digit.