

Rewarding compliance: effects of the 2011 reward regime on Italian small businesses*

Enrico Di Gregorio[†] and Matteo Paradisi[‡]

May 30, 2020

Abstract

We study the effects of a unique Italian reform incentivizing voluntary tax compliance among the self-employed and small businesses. Starting in 2011, taxpayers in a growing number of sectors were promised an increase in audit exemptions upon achieving a set of desirable conditions defined by the Revenue Agency. While policy rewards might induce a tax base rise among previously non-compliant filers, curbing audit risks for broad categories of the taxpaying population might prove revenue reducing. Over the first six years of implementation, our event-study analysis of more than 9 million anonymized records reveals a substantial expansion of average declared revenues, total costs, and gross profits, with little heterogeneity across macro-regions. Although aggregate compliance does not seem to improve by policy metrics, our distributional analysis shows that large gains obtain among taxpayers appearing non-compliant in the year before their sector's reform. We also provide a dynamic perspective on bunching at salient, audit-relevant revenue thresholds generated by the system. Relative revenue reshuffling from above and below these thresholds provide evidence that bunching in our context may emerge from both desirable and adversarial updating in compliance behavior.

*We wish to acknowledge Agenzia delle Entrate for granting access to the Sector Study database and for making this project possible, as well as SOSE for their support in the interpretation of the data. We particularly thank Antonino Maggiore, Ernesto Maria Ruffini, Andrea Baldini, Danilo Ballanti, Laura Benedetti, Vieri Ceriani, Laura Limosani, Enrico Polella, and Massimo Varriale. All views expressed are solely our own. A special acknowledgment goes to Alberto Alesina, a mentor and a friend. During the development of this project, we greatly benefited from our interactions with Claudio Di Gregorio, Michele Fornino, Anders Jensen, Tommaso Nannicini, Sergio Sirabella, and Stefanie Stantcheva. All remaining mistakes are ours.

[†]Corresponding author. Kennedy School of Government and Department of Economics, Harvard University. edigregorio@g.harvard.edu

[‡]NBER and Einaudi Institute for Economics and Finance.

Premiare la “*compliance*”: gli effetti del regime premiale sulle piccole imprese italiane¹

Enrico Di Gregorio² e Matteo Paradisi³

30 Maggio, 2020

Sommario

Il presente studio analizza gli effetti di una singolare riforma volta ad incentivare la fedeltà fiscale spontanea da parte dei lavoratori autonomi e della piccola impresa interessati dagli Studi di Settore. A partire dal 2011, ai contribuenti di un numero crescente di settori è stata estesa la possibilità di maggiori esenzioni dai controlli fiscali in cambio del raggiungimento di una serie di condizioni dichiarative auspiccate da parte dell'Agenzia delle Entrate. Se da un lato i premi promessi dalla riforma avrebbero potuto indurre un'espansione delle basi imponibili riportate dai contribuenti non già allineati con i nuovi requisiti, una riduzione del rischio di controllo percepito da una larga fascia di contribuenti avrebbe potuto rivelarsi controproducente per la raccolta fiscale. Nei primi sei anni di attuazione della riforma, la nostra analisi di oltre nove milioni di posizioni anonimizzate rivela una significativa espansione dei fatturati dichiarati, dei costi aggregati, nonché dei redditi lordi, con scarsa eterogeneità fra macroregioni del Paese. Benché secondo le definizioni offerte dal sistema degli Studi di Settore la fedeltà fiscale non sia migliorata nel suo complesso, il nostro approfondimento dimostra come aumenti significativi delle basi imponibili si siano realizzati fra quei contribuenti che risultavano non congrui nell'anno precedente l'ingresso del proprio Studio nel nuovo regime premiale. L'analisi si conclude con una prospettiva inedita sulle dinamiche di “*bunching*” dei contribuenti intorno ai ricavi presunti dagli Studi, i quali generano delle discontinuità nel rischio di controllo percepito dagli utenti del sistema. Con l'avvento della riforma, molti contribuenti si riallocano verso queste soglie modificando opportunamente il proprio comportamento fiscale indipendentemente dal proprio punto di partenza. Ciò dimostra come il “*bunching*” osservato a queste soglie sia il risultato di comportamenti fiscali sia auspicabili che potenzialmente indesiderati.

¹Ringraziamo l'Agenzia delle Entrate per avere consentito l'accesso al database informativo degli Studi di Settore e per avere reso possibile questo progetto. Siamo grati a SOSE per il prezioso supporto fornito nella fase di interpretazione di questo patrimonio informativo. Siamo inoltre particolarmente riconoscenti nei confronti di Antonino Maggiore, Ernesto Maria Ruffini, Andrea Baldini, Danilo Ballanti, Laura Benedetti, Vieri Ceriani, Laura Limosani, Enrico Polella, e Massimo Varriale. Le opinioni qui espresse sono riferibili soltanto agli autori. Un ringraziamento speciale va ad Alberto Alesina, nostro mentore ed amico. Nello sviluppo del progetto abbiamo beneficiato del prezioso contributo di Claudio Di Gregorio, Michele Fornino, Anders Jensen, Tommaso Nannicini, Sergio Sirabella, e Stefanie Stantcheva. Eventuali errori sono da attribuire solo agli autori.

²Autore corrispondente. Kennedy School of Government e Department of Economics, Harvard University. edigregorio@g.harvard.edu

³NBER ed Einaudi Institute for Economics and Finance.

1 Introduction: rewarding compliance

The threat of tax enforcement fails to generate perfect tax compliance in presence of low third-party reporting. This is especially evident for small firms and the self-employed, which feature prominently in academic and practitioner discussions about “hard-to-tax” populations (Alm et al., 2004). For as recently as the 2014-2016 period, the Italian government estimates that 30.4% of all unpaid tax liabilities result from non-compliance with personal income tax obligations among these categories (Ministry of Economy and Finance, 2019). As a result, policy advice has recently shifted towards the experimentation and adoption of novel tax collection strategies that privilege voluntary tax compliance (OECD, 2017). These approaches stress the importance of reducing the pervasiveness of costly auditing for businesses, while featuring opportunities for active collaboration between taxpayers and the tax authority.

While many tax agencies moved to include voluntary compliance targets among their performance indicators, scholarly work offers no clear guidance on how to best stimulate spontaneous tax reporting.¹ Tax amnesties often fall short of generating the desired revenue effects, and their repeated use has been warned to weaken the tax morale in the long run (Andreoni et al., 1998). Tax lotteries on the other hand have offered mixed results, hinting at the importance of policy design and the nature of the rewards that the involved administration is willing to concede (the construction of sidewalks in Carrillo et al. (2017b); tax holidays in Dunning et al. (2017)). Importantly, the limited access to raffled prizes, as well as their low value, may fail to engender diffused compliance, unless lotteries are targeted to parties serving a low-cost monitoring function (Naritomi, 2019).

In this work, we study a novel policy introduced by the Italian central government in pursuit of voluntary tax compliance. Starting in 2011, small firms and the self-employed in a growing number of sectors were promised an increase in tax audit exemptions upon achieving a set of desirable conditions defined by the Revenue Agency. This reduction in audit risks sought to extend a set of pre-existing audit protections for taxpayers subject to a threshold-based enforcement system known as Sector Studies (Santoro, 2008; Santoro and Fiorio, 2011). The 2011 reform, commonly referred to as reward regime, resembles tax amnesties in that it limits the auditing powers of the tax administration over eligible taxpayers, while it differs from them by setting clear incentives ahead of reporting. It shares the latter trait with tax lotteries, but it differs from these by extending its benefits to all those who decide to comply.

We investigate the ex-ante ambiguous tax base effects that the promise of audit exemptions provided by the reward regime may generate. In particular, we conceptualize the policy as a discrete reduction in the perceived probability of audit for all potential beneficiaries. This jump creates incentives to improve compliance in order to access the featured benefits. At the same time, for some taxpayers the reduction in audit risks might cause a contraction in the reported tax base due to a fall in the expected penalty cost of evasion. To illustrate

¹For example, the tax authorities of Australia, France, New Zealand, and the United States actively target taxpayers’ compliance ahead of enforcement (OECD, 2017).

this trade-off and determine the net effect of the reform on compliance, we use more than 9 million anonymized Sector Study files covering the 2007-2016 period.

Our empirical strategy exploits the gradual introduction of the policy across sectors between 2011 and 2016 to implement an event-study methodology. We focus on sectors where the reward system was eventually introduced and use sectors that were reformed later in time as controls for those that were reformed in earlier years. Similarly to a difference-in-difference methodology, our approach relies on the assumption that the two groups of sectors would have parallel trends in the variables of interest in absence of the policy.

Our results highlight the potential of the Italian reward regime for revenue collection, as well as the limits of audit exemptions in boosting compliance. First, we show that sector exposure to the reform raise yearly business profits by 16.2% relative to the pre-reform year, with point estimates and confidence intervals growing over the six treatment years in our data. These results reflect the combination of revenue and cost responses likely induced by the peculiar accounting targets businesses must meet to access policy benefits. Importantly, we confirm the common observations that adjustments along cost margins can reduce the collection potential of successful fiscal policies (Carrillo et al., 2017a; Bachas and Soto, 2018).

Second, we present evidence of heterogeneity in reporting responses, consistent with our intuition that audit exemptions should provide opposite incentives based on a taxpayers' relative compliance ahead of the reform. Tax base gains are large among previously non-compliant filers. To the contrary, other businesses successfully inflate costs, including those related to employment, to match any additional revenues reported to meet the policy compliance requirements. As a result, average compliance doesn't seem to improve by the several metrics featured in the Sector Study system and which we present in the next Section.

Third, we document how threshold-based compliance targets can generate large discontinuities in revenue distributions. In particular, Sector Studies grant certain audit exemptions to businesses reporting revenues equal or higher than what is presumed for them by the Revenue Agency in that tax year. We use the fact that the reward regime expands the protections offered to these taxpayers to provide two sets of results. On one hand, the reform substantially increases bunching of taxpayers just above their presumed revenues in the short run. On the other, we decompose this outcome showing that taxpayers previously below the presumed revenues threshold tend to rise above it, and those previously above the threshold move in the opposite direction after the reform.

Our findings provide encouraging evidence that tax administrations can in fact rethink the philosophy guiding their tax collections efforts. To this end, approaches privileging spontaneous compliance need to put taxpayers' incentives at the core of policy design. In the case of audit exemptions, the success of any compromise between tax agencies and businesses will rest on a careful balance between the opposing pressures that determine the reported tax base.

The rest of our work is structured as follows. Section 2 presents the institutional setting and conceptualizes its implications for taxpayers' incentives. Section 3 presents the Sector

Study database we access within the limits of our special collaboration with the Italian Revenue Agency. Section 4 presents our results on average reform effects, heterogeneous responses, and distributional shifts, including our analysis of bunching as it relates to the 2011 reform. Section 5 draws out the main shortcomings and conclusions from our work, touching upon the recent introduction of an elaborate reward-based system replacing the Italian Sector Studies.

2 Sector Studies and the 2011 reward regime

A number of tax enforcement policies apply to micro firms and the self-employed in Italy. Most prominently, Sector Studies (*Studi di Settore*, henceforth SeS) were implemented as a compliance initiative starting in 1998. Each tax filing cycle, SeS require non-employee taxpayers with turnovers up to €5.2 million in the previous year to provide the tax administration with information on their revenues, inputs, and costs. The administration relies on these details to establish the degree of compliance of filing taxpayers. A publicly released software known as *Gerico* facilitates taxpayers' submission of the information required by each Study, and in turn warns them of any potential compliance issues ahead of filing. This allows businesses to update their reporting behavior in line with what the administrative database in the hands of the Revenue Agency shows to be sensitive for their sector.

Three conditions define compliance in the SeS context, and result in a set of audit exemption benefits. Congruence certifies that reported revenues are at or above the level of turnover that *Gerico* presumes from reported input patterns. Congruous taxpayers are generally shielded from SeS-based inquiries into their reported revenues. Normality and coherence establish instead that the data included in one's SeS file reflects reasonable accounting and managerial practices, respectively, by industry standards.² Specifically, *Gerico* assesses whether each taxpayer scores within acceptable ranges on a predefined set of accounting indices, granting reduced scrutiny on inputs and costs. Until 2011, being both congruous and normal resulted in partial exemptions from analytic-inductive audits, that is, the class of tax audits in the Italian enforcement system to which SeS belong.³

In December 2011, the Italian technical government adopted a set of emergency measures to consolidate the public budget. Among these, Law Decree 201/2011 strengthened the incentives for voluntary tax compliance among SeS taxpayers by gradually expanding

²Coherence and normality indicators flag abnormal behaviors undermining the reliability of the presumed revenues computed by *Gerico*. For instance, coherence may include an assessment of inventory turnover rates, while normality may examine the plausibility of residual managerial expenses as a share of total costs. The indicators also help filter out taxpayers' files in the estimation phase of each sector-specific presumed revenue function performed by SOSE.

³In an analytic-inductive tax assessment, the Agency establishes that a declaration item is unreliable and presumes a correct amount based on alternative information. This situation may arise when significant discrepancies emerge across administrative databases for the same taxpayer, or when field audits uncover the existence of unreported transactions. Similarly, a SeS file's failure to achieve congruence, coherence, or normality leads to a simple presumption of misreporting to be argued in court at the administration's discretion.

Table 1: Sector Studies compliance benefits, before and after 2011

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

Notes: the Table reports the main tax audit and assessment benefits from being congruous, coherent, and normal by the definitions provided by Sector Studies, before and after the introduction of the 2011 reward regime. 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 eligible taxpayer's file drops by one year since 2011.

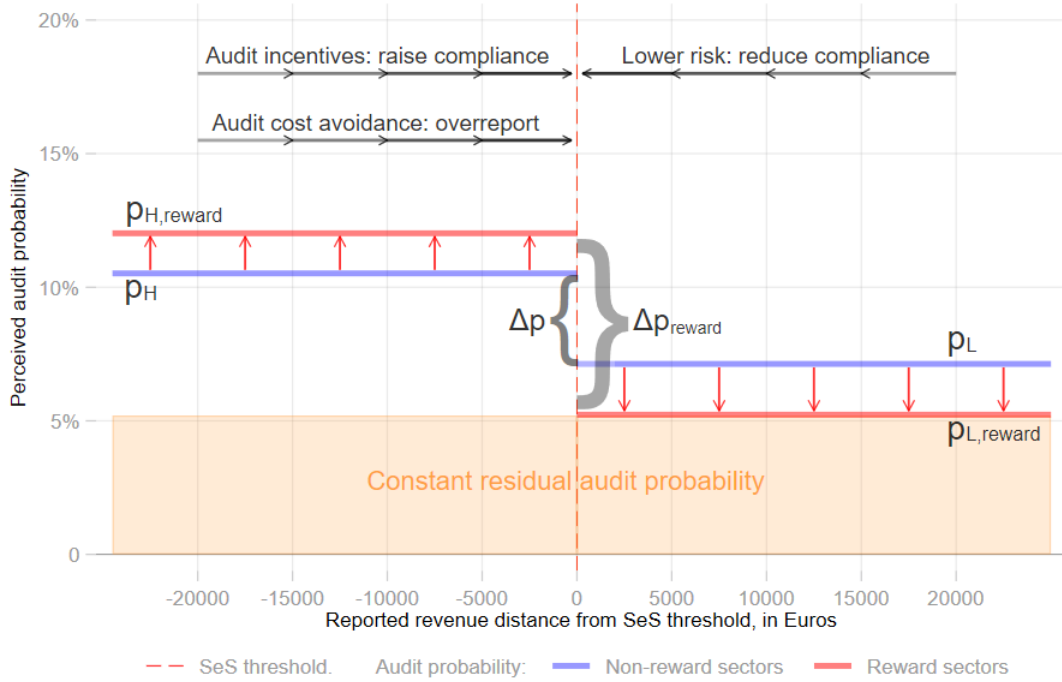
the audit protections originally tied to congruence and normality. Table 1 summarizes the policy changes induced by the reform, known as reward regime (*regime premiale*).⁴ The new system of benefits granted full exemption from analytic-inductive audits, a partial exemption from any synthetic determination of a taxpayer's income based on her private spending, and a one-year reduction of the statute of limitations for all the above classes of audit. Taxpayers would access these incentives upon jointly reaching the conditions of congruence, normality, and coherence with a truthful report. For those failing to comply, on the other hand, the tax administration would develop dedicated audit plans with the aid of modern financial investigation techniques. Protection from standard SeS-based audits remained unchanged.

2.1 Conceptual framework: reporting behavior in the reward regime

The incentives provided by the reward regime strengthen the compliance motives built into the SeS system. Figure 1 provides a simplified overview from the perspective of con-

⁴We abstract from the aspects of the reform not directly related to SeS compliance. In particular, the reform sought to reduce the administrative burden to non-corporate taxpayers. Benefits included scrapping of certain reporting requirements, automatic preparation of tax-relevant files by the Agency, and quicker VAT reimbursements. To access these rewards, taxpayers would have to open a checking account dedicated solely to their business and to share electronically all receipts pertaining to their activity with the Revenue Agency.

Figure 1: Expected effects of the 2011 reward regime



Notes: this Figure offers a rationalization of the reward regime compliance effects within Sector Studies. 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. This would reinforce motives to increase (decrease) reported revenues for those below (above) the cutoff.

gruence.⁵ SeS taxpayers perceive a discontinuous audit schedule based on whether they declare revenues below or above the presumed level computed by *Gerico*. The 2011 reform magnified the gap in expected costs associated with audits and penalties upon detection of misreporting on the two sides of the presumed revenue threshold. A relatively higher audit risk below this threshold, as well as a significant reduction above it, might convince taxpayers to improve on their compliance. Businesses who understate their true tax base may still choose to report more than their presumed revenues to limit the expected costs from enforcement. Willingness to avoid the perceived administrative costs associated with audits may also push some taxpayers into congruence. On the other hand, the protections granted to those above the threshold may come at the cost of reduced compliance, as they invite the leveling of reported tax bases onto the amounts presumed by *Gerico*. Overall, tax collection can be affected in two steps. On impact, tax revenues may rise or fall based

⁵We develop our reasoning for taxpayers who satisfy the conditions of normality and coherence. However, our results explore the effects of the regime on all compliance conditions and a selection of inputs and costs.

on which of the above compliance motives is stronger. In the medium run, selective sorting of taxpayers around SeS thresholds may affect the collection potential of tax audits. As we discuss below, we tackle the first-round effects in our results section, and leave the analysis of audit data for future research.

3 Administrative data

We work with the individual Sector Study files submitted by targeted self-employed and small businesses for the 2007-2016 tax years. The Revenue Agency, in collaboration with their statistical analysis partner SOSE, provided access to more than 26 million de-identified files. Each comes with a rich set of information, including declared and presumed revenues, gross profits, a selection of inputs and costs such as those pertaining to employment, administrative information regarding compliance with the relevant SeS (congruence, normality, coherence), as well as sector and geographic tags. While we observe the universe of declarations for the four tax years preceding the reward regime introduction (2007-2010), our data spans the remaining years with an unbalanced panel of all taxpayers continuously filing for the 2008-2010 tax period.

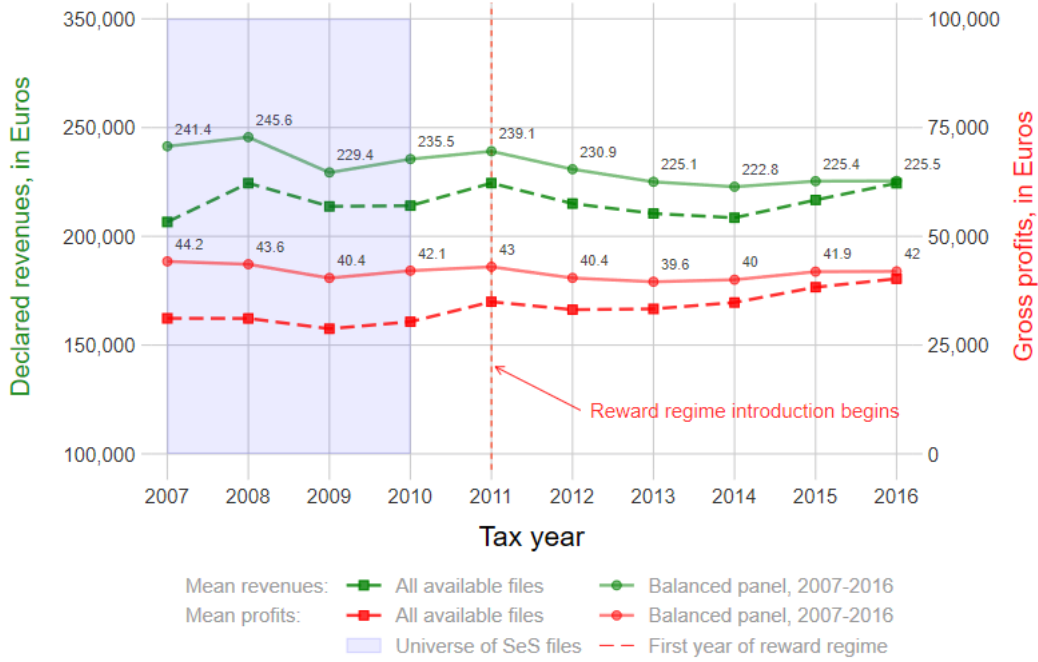
Table 2: Descriptive statistics, balanced vs. unbalanced samples, 2007-2010

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 and 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 analysis, declared revenues are winsorized at the 99th percentile of the global distribution.

We focus on the balanced panel of all single-sector filers over the 2007-2016 tax years. This should ensure data consistency and prevent any potential confounding effects of differen-

Figure 2: Mean revenues and profits, full sample and balanced panel, 2007-2016



Notes: this Figure shows the evolution of mean declared revenues and profits for two samples of taxpayers. The first is composed of all available SeS files from single-sector filers in our data, while the second includes the balanced subset of all continuous filers over 2007-2016. Labels within the graph are in thousands of Euros. We winsorize revenues at the 99th percentile of the global distribution. Declared revenues include spontaneous adjustments upon filing. Our raw data include the universe of all SeS files over 2007-2010, as well as the unbalanced panel of those continuously declaring in 2008-2010.

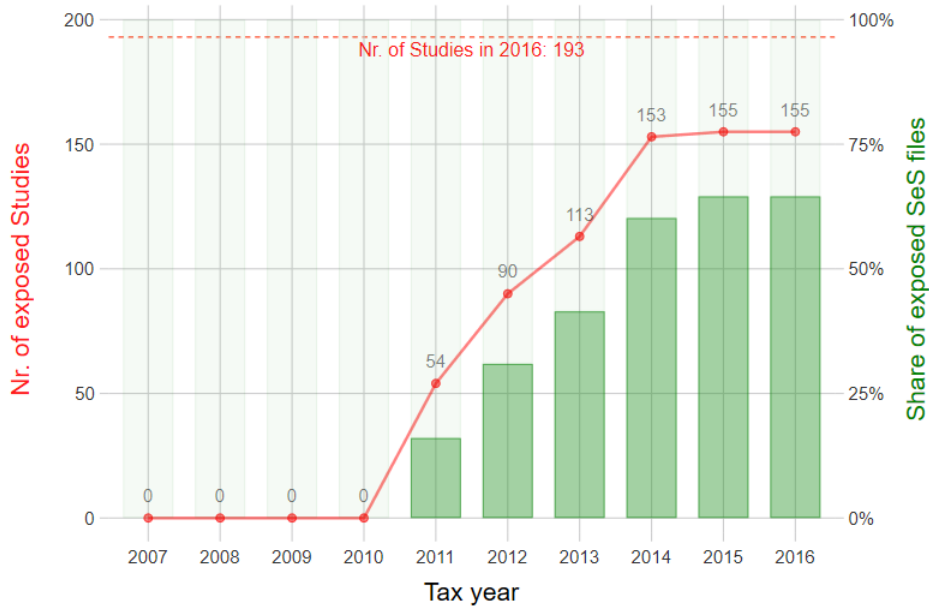
tial survival dynamics. However, our choice may bear costs in terms of representativeness relative to the general SeS population. We thus provide summary statistics to contrast the features of our balanced panel relative to the unbalanced portion of the data, as well as to the full set of files we can access. Table 2 covers the four years before the introduction of the reward regime among files in the balanced sample and the remaining part of the data. Declared revenues, gross profits, and the share of those fully compliant with their SeS are significantly and economically higher, on average, for files in our balanced panel, about 40% of all files submitted. Figure 2 further shows that balanced panel observations feature substantially larger and more profitable businesses throughout the whole period relative to the full sample.⁶ These differences, however, don't seem to bias upward our main results: the point estimates of the reward regime effects on declared revenues and profits tend to shrink towards zero when we move from the full data to our balanced panel.

⁶For reference, balanced panel files start with an average of €241,416 in revenues and €44,246 in profits - almost 17% and 42% higher than in the full sample, respectively. Gaps close by the end of the period due to the unbalanced structure of our raw data.

Access to the reward regime was granted to individual sectors starting in 2011. Each year, the Revenue Agency announced a list of admitted sectors in the early part of the tax season. Selection criteria generally involved the recent update of a number of coherence indicators and, up to 2013, having an estimated share of value added from unreported activities lower than the economy at large.⁷

To visualize the gradual pace of the reform, we define sector-year combinations as our units

Figure 3: Staggered introduction of the 2011 reward regime



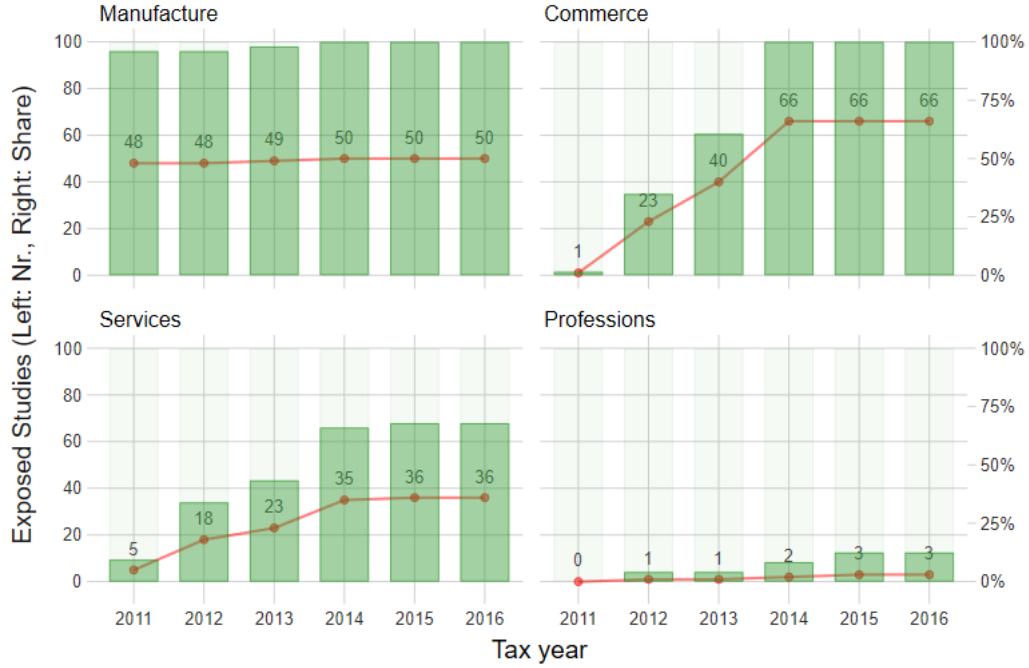
Notes: the Figure shows the staggered introduction of the 2011 reward system among existing Sector Studies. The red line displays the number of Studies with access to the regime in each year up to 2016 (scale on the left vertical axis). The dark green bars reflect the share of all files with access to the reward system 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. We code full regime access for five sectors with partial access only. The horizontal dashed line represents the total number of studies in 2016.

of analysis. Sectors within the SeS system undergo revisions of their underlying statistical model and identification code every three to four years with different schedules depending on the sector. Updated sectors change identification code and face the possibility of mergers and splits. We therefore track the evolution of SeS sectors throughout the decade, and associate each taxpayer to one of 193 SeS sectors existing in 2016, the last tax year in our data.

Figure 3 displays the staggered inclusion of these sectors into the new regime. Over the first six years of implementation, the reward regime expanded from 54 to 155 of all 2016

⁷Required indicators would generally target input efficiency, inventory management, and firm profitability. The unreported activity criterion was supposed to speed up the inclusion of low evasion-risk sectors featuring only a partial set of the required updated indicators.

Figure 4: Staggered introduction of the 2011 reward regime, by macro-sector



Notes: the Figure shows the staggered introduction of the 2011 reward regime among existing Sector Studies by macro-sector. See notes to Figure 3 for further details on its construction. In total, there were 50 sectors in manufacturing, 66 in commerce, 53 in the services, and 24 in the professions by 2016.

SeS sectors, covering 64.5% of our end-of-period files.⁸ Businesses in the manufacturing sectors were introduced most rapidly, with commerce and services following suit. On the other hand, entry for those working in the skilled professions was mostly delayed beyond the time horizon of our analysis (Figure 4).⁹

4 Empirical strategy and results

4.1 Econometric model and identification

We aim to quantify the effect of the incentives provided by the reward system on tax compliance. The main empirical challenge in setting up the analysis lies in the fact that taxpayers

⁸In a handful of sectors, the reward regime was initially introduced only for a fraction of filers. For practical purposes, Figures 3 and 4 count these sectors as fully treated.

⁹A complete overhaul of the SeS system was eventually carried out by the end of the decade. The new system known as ISA (*indicatori sintetici di affidabilità fiscale*, or indices of fiscal compliance) allowed all SeS taxpayers to access to the benefits of the reward regime. We provide a brief discussion in the conclusions.

claiming the policy benefits are not a random sample of the population, but a selected group who fulfill the prescribed requirements based on their declaration behavior. Since there exist fundamental differences in the tax declarations of the reward beneficiaries and the rest of the SeS population, the latter do not represent a reliable control group to infer what the current beneficiaries would do in absence of the policy. Realized access to the audit exemptions provided by the new system is also not a policy-relevant “treatment” per se. Indeed, since these benefits are conditioned on adopting a honest declaration behavior, declared income should look mechanically larger for beneficiaries relative to the other taxpayers.

Instead, our project asks to what extent declaration behaviors respond to the possibility, rather than the occurrence, of claiming larger benefits thanks to the reward system. The empirical counterpart of this response is often referred to as intention-to-treat in the applied microeconometrics literature. To answer our question, we must compare the declaration behavior of groups of taxpayers with different availability of the policy benefits. For this reason, our empirical strategy relies on a comparison between sectors with and without access to the new system of audit exemptions.

The gradual roll-out of the reform provides scope for a natural experiment in tax compliance. In particular, we exploit the reform schedule to implement an event study design on the 155 sectors entering the reward system between 2011 and 2016. Event studies generalize the difference-in-differences approach to settings where every unit of analysis in the sample is “treated” by the end of the sample period. Specifically, sectors where the reward system has not yet been introduced in a given year serve as controls for those where the policy has already been implemented. The former therefore provide a plausible counterfactual for how taxpayers would behave absent the possibility to claim the new benefits. The baseline specification is:

$$y_{st} = \lambda_s + \gamma_t + \sum_{q=-k}^{+k'} \beta_q \cdot I(Q_{st} = q) + \sum_{r=2007}^{2016} \delta_r \cdot X_s \cdot I(t = r) + \varepsilon_{st} \quad (1)$$

where s indexes the sector, as defined in 2016; t indexes the tax year; y_{st} is an outcome of interest such as declared revenues, income, or costs; Q_{st} measures the distance between year t and the year when the reward system was introduced in sector s so that the indicator $I(\cdot)$ is a so-called event-time dummy. β_q s are the coefficients of interest. The vector of β_q s with $q < 0$ captures the difference in declaration trends between treated and control sectors before the reward system is introduced. These coefficients provide a useful placebo test to check that the sectors where the reward system was phased-in with a lag behave similarly to sectors that experienced an early introduction. The vector of β_q s with $q > 0$ instead captures the effects of the reward system on declaration behavior in the tax years following the introduction of the policy in a given sector. To summarize our results in a single statistic we also compute the average of the coefficients β_q s with $q > 0$.¹⁰ Variables

¹⁰ As suggested by [Borusyak and Jaravel \(2017\)](#), we use the average of β_q s rather than the treatment

λ_s and γ_t capture sector and year fixed-effects, respectively.¹¹

The empirical strategy relies on the assumption that treated and control sectors would follow parallel trends in declaration behavior absent the policy. Importantly, we do not require that the two groups have the same starting levels of y , but only the same growth in y over time. We estimate the specification in (1) with a standard OLS, weighting each sector by its number of SeS files submitted in 2007. We cluster standard errors at the sector level as suggested by [Bertrand et al. \(2004\)](#).

We augment our specification with a vector of baseline covariates X_s interacted with the tax year dummies. Controls include dummies for the four SeS macro-industries defined by the tax administration (manufactures, commerce, services, and the professions); and 2007-2010 averages for a set of variables including revenues, profits, incidence of employment costs on turnover, yearly growths of employment cost rates and revenues.

Our analysis tackles the effects of the reward regime on several outcome variables. In the first part, we focus on main reporting aggregates such as revenues, profits and costs. Since different SeS may require taxpayers to report different sets of cost items, we derive a measure of total costs from the difference between revenues and profits in each file. We then study a number of metrics defined by SeS to measure the degree of taxpayer compliance. Our data provide information on congruence, coherence, and normality. In addition, we devote attention to the distribution of reported revenues relative to presumed revenues, by defining an index of “bunching” at the presumed revenues amount implied for each file by *Gerico*. Bunching allows us to assess the extent to which taxpayers declare the amount of revenues presumed by their SeS, or just above it, in order to reduce the audit risk. Specifically, bunching estimates measure the excess mass of files with turnover equal or close to their own presumed revenues relative to a counterfactual scenario. We define this counterfactual as the distribution of relative turnover that would arise if taxpayers faced a constant audit probability equal to the audit probability perceived by congruous taxpayers above the level of presumed revenues. [Appendix A](#) and [Di Gregorio and Paradisi \(in preparation\)](#) provide more details about the bunching definition and the construction of the counterfactual.

4.2 Main declaration effects of the 2011 reward regime

We start by assessing whether sector-wide exposure to the reward regime affects key reporting aggregates. We first estimate Equation (1) using the log and level of total declared revenues as an outcome.¹² [Figure 5](#) plots all event-time coefficients relative to the last year

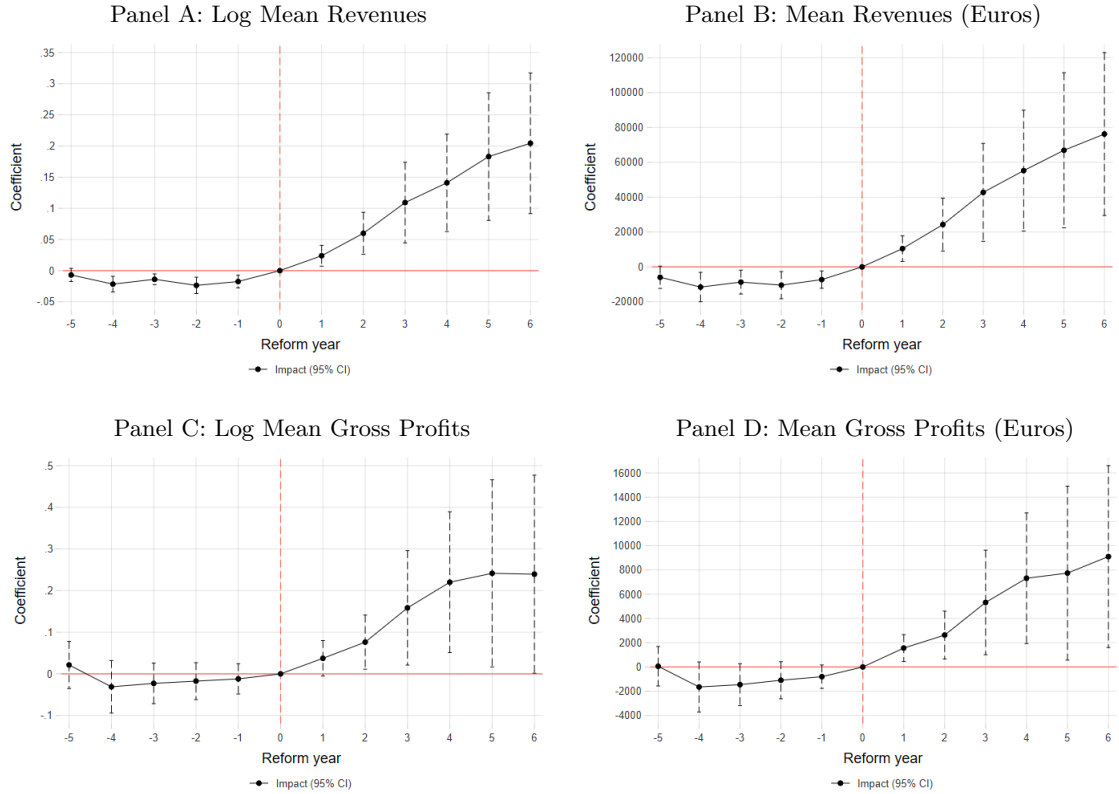
coefficient from a more classical and parsimonious specification with a single variable equal to 1 in the years after the introduction of the reward regime. This avoids the negative weighting of later periods’ effects.

¹¹In order to estimate both year and treatment time effects, we exclude the first two out of the fourteen available pre- and post-treatment time dummies. Since we also omit the treatment dummy associated with the year before each sector’s reform, the coefficients can be interpreted as treatment effects relative to the outcome level in that year.

¹²Total revenues include two components: revenues declared as part of the standard SeS filing process, matching what can be desumed from receipts, books, or the income statement for firms with such reporting requirements; and spontaneous adjustments upon filing, as foreseen for taxpayers willing to raise their

ahead of regime entry, along with 95% confidence intervals. While most pre-treatment dummies return slightly negative coefficients, there is no stable growth trend ahead of treatment. On the other hand entry into the reform is associated with 2.4% higher revenues in the first year, growing to 20.4% by year six, with all effects achieving 1% significance. Uncertainty surrounding the point estimates increases with treatment time, likely because only a subset of all sectors achieve six years of reward regime exposure by 2016.

Figure 5: Reward system effects on revenues and gross profits



Notes: the Figure shows the effects of the reward system introduction in a sector on log mean revenues (Panel A), on mean revenues in euros (Panel B), on log mean gross profits (Panel C), and on mean gross profits in Euros (Panel D). Whiskers represent 95 percent confidence intervals. Effects are relative to the year before the advent of the reform in each sector. Estimates are based on our event-study specification in (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. Declared revenues are winsorized at the 99th percentile.

To quantify the magnitude of these effects, Table 3 provides estimates of the reported revenues to the level presumed by *Gerico*.

enues cumulated by the end of the sample period. On average, reward regime taxpayers boost their reported revenues by €275,828.7 throughout six tax years relative to non-reward regime taxpayers, consistent with average revenues increasing yearly by €45,971.46.¹³

Table 3: Average reward system absolute effects, 2011-2016

Variable	Statistic	Mean of Post-Treatment Coefficients			
		Estimate	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Revenues	<i>mean</i>	45.97	14.85	16.63	75.31
	<i>25th percentile</i>	4.03	2.16	-0.23	8.28
	<i>50th percentile</i>	17.23	5.84	5.70	28.76
	<i>75th percentile</i>	63.09	22.21	19.21	106.97
	<i>Cumulative (years 1 to 6)</i>	275.83	89.12	99.78	451.88
Gross profits	<i>mean</i>	5.61	2.29	1.09	10.13
	<i>25th percentile</i>	1.89	0.93	0.05	3.73
	<i>50th percentile</i>	1.60	0.78	0.05	3.15
	<i>75th percentile</i>	3.75	1.60	0.59	6.90
	<i>Cumulative (years 1 to 6)</i>	33.67	13.72	6.56	60.79

Notes: this Table shows the average of six post-treatment coefficients from our event-study analysis of the effect of the staggered introduction of the 2011 reward regime across 155 sectors in the Sector Study system. Each outcome is evaluated at the sector-year average, 25th, 50th, 75th percentile, respectively. The last line cumulates the coefficients from the mean variable analysis. Details on the full specification are provided under Equation (1) in the text. Standard errors rely on 2016 Sector Study clustering. Units are thousands of Euros.

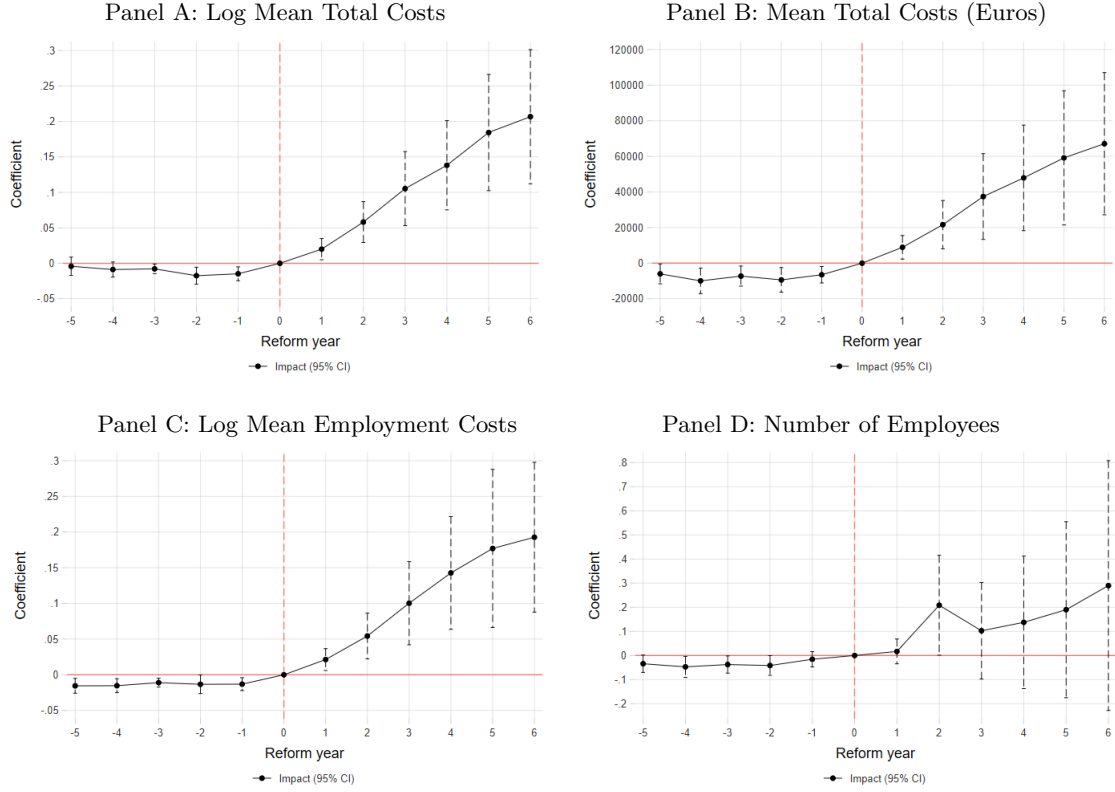
Revenue-raising policies can still be self-defeating if they trigger an offsetting adjustment in costs (Carrillo et al., 2017a; Bachas and Soto, 2018). We thus study the effects of the reform on a sector-average index of total reported costs and their employment component. Given the diverse nature of cost items reported in each SeS, we first derive our total cost outcome by computing the sector-year average difference between declared revenues and gross profits. Figure 6, Panel A shows that the cost patterns induced by the reform seem to follow those found for reported revenues, with coefficients ranging from 2% to 20.7%. About one fifth of the yearly effects appear to be explained by a raise in mean employment costs, with the number of employees at the filers' premises being positively but weakly affected (Figure 6, Panel C and D).

However, tax bases may ultimately be expanding. Despite the contemporaneous rise in revenues and costs, Figure 5, Panel C and D reveals that gross profits become higher than in the counterfactual scenario by a yearly average rate of 16.2%. Point estimates are signif-

¹³We obtain these estimates by linear combination of the six post-treatment coefficients, with weights equal to 1 and 1/6, respectively.

icant and increasing over the course of the treatment period.¹⁴

Figure 6: Reward system effects on reported costs



Notes: the Figure shows the effects of the reward system introduction in a sector on log mean total costs (Panel A), on mean total costs in euros (Panel B), on log mean employment costs (Panel C), and on the number of employees (Panel D). Whiskers represent 95 percent confidence intervals. Estimates are based on our event-study specification in (1). Standard errors are clustered at the sector level. Total costs are derived as the difference between declared revenues and profits. See notes to Figure 5 for a detailed explanation of the regression sample. Number of sector-years: 1550.

These results suggest that, on impact, exposure to the reward regime led to economically meaningful improvements in the tax base. By Table 3 estimates, the reform encouraged a cumulative gain of €33,671.77 in taxable profits from the average treated business.¹⁵ Extrapolating the reform impacts to the full treated population comes at the cost of additional

¹⁴Individual treatment effects are significant at the 5%-10% level when we take the log of mean profits, and 1%-5% significant when we use a specification in levels.

¹⁵Ultimately, net profits are taxed as either personal income for individuals and partnerships, or as business profits for the small corporations in our sample. In the pre-reform period, these three groups of taxpayers contributed 64.8%, 19.5%, and 15.7% of all SeS files, respectively.

assumptions. In particular, this requires moving from a partial to a general equilibrium setting conditional on the stability of our controls. Provided we accept these strong assumptions, by 2016 the reward regime would have contributed to a cumulative expansion of gross profits by €18.1 billion among single-sector taxpayers continuously filing over the previous decade.¹⁶

Table 4: Average reward regime relative effects, 2011-2016

Variable	Statistic	Mean of Post-Treatment Coefficients			
		Estimate	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Revenues (logged)	<i>mean</i>	0.120	0.034	0.053	0.188
	<i>25th percentile</i>	0.098	0.047	0.005	0.190
	<i>50th percentile</i>	0.110	0.046	0.019	0.201
	<i>75th percentile</i>	0.141	0.051	0.040	0.241
Gross profits (logged)	<i>mean</i>	0.162	0.072	0.019	0.305
	<i>25th percentile</i>	0.210	0.092	0.029	0.392
	<i>50th percentile</i>	0.067	0.040	-0.011	0.146
	<i>75th percentile</i>	0.077	0.042	-0.005	0.160

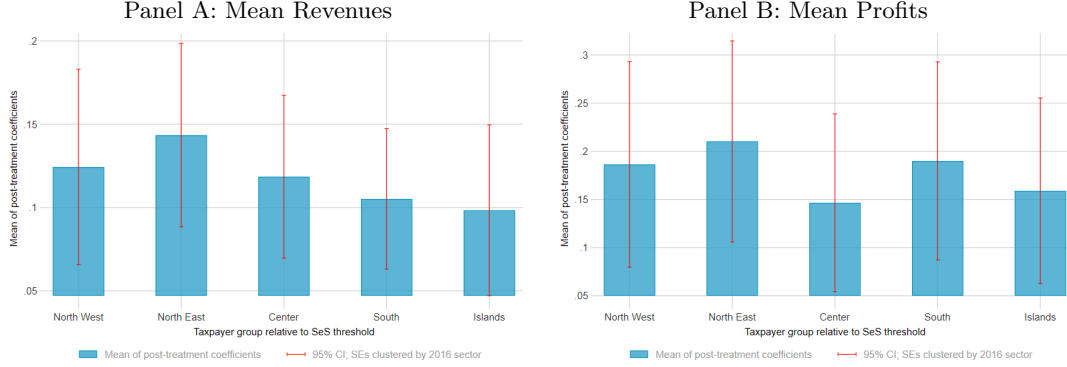
Notes: this table shows the average of six post-treatment coefficients from our event-study analysis of the effect of the staggered introduction of the 2011 reward regime across 155 sectors in the Sector Study system. Each outcome is evaluated at the sector-year average, 25th, 50th, 75th percentile, respectively, and logged. Details on the full specification are provided under Equation (1) in the text. Standard errors rely on 2016 Sector Study clustering.

Positive effects seemingly arise throughout the revenues and profits distribution. Tables 3 and 4 illustrate this point. Looking at the average of post-treatment coefficients for sector-year median revenues and profits, as well as their 25th and 75th percentiles, point estimates are larger than zero, with relative growth increasing as we move towards the top quartile of the turnover distribution and the bottom quartile of the profit distribution, respectively. Italy's heterogeneous local business conditions might explain the aggregate results. To assess this, we replicate our analysis by splitting taxpayers into five NUTS-1 macro-regions before aggregation at usual sector-year level, and allow for heterogeneous reform impacts by interacting our full set of event-time dummies with macro-regional indicators. Although the reform seems to have contributed, on average, to raise the revenues and profits of North-Eastern sectors to a greater extent, 95% confidence intervals cannot rule out homogeneous effects across macro-regions (Figure 7). Hence, while relatively more profitable taxpayers reside in the North, and expected underreporting may be more intense in the South

¹⁶We arrive at this figure in steps. First, we compute the number of new files treated in each year. Next, we associate each of the six event-time coefficients to the appropriate share of treated files in each year, since different sectors have different exposure length at each point in time. Within each tax year, we then sum the products of the number of treated taxpayers across groups with different exposure and the relevant treatment effect. Finally, we sum up the year-specific increases in the tax base.

(Ministry of Economy and Finance (2011) and Ministry of Economy and Finance (2017), respectively), on the net the promise of new audit exemptions stimulated comparable earnings responses across the board.

Figure 7: Macro-regional effects of the reward system, 2011-2016



Notes: the Figure shows the heterogeneous effect of the reward system across macro-regions on mean revenues (Panel A) and mean profits (Panel B). Bars display for each macro region the average of six group-specific post-treatment coefficients from an event-study based on an extension of the specification in (1) where each event-time dummy is interacted with a macro-region dummy in a triple-differences design. Whiskers represent 95 percent confidence intervals of these combinations of coefficients. See notes to Figure 5 for a detailed explanation of the regression sample.

4.3 Heterogeneous compliance and distributional shifts

We now examine the sources of our aggregate reporting results. The SeS system provides a number of metrics to assess the relative degree of taxpayer compliance. We first show which categories of taxpayers strive to comply with the new requirements of the reward regime. Next, we link these compliance responses to heterogeneous reporting behaviors around the presumed revenue amounts implied by *Gerico*. Finally, we discuss how the regime impacts the distribution of reported revenues relative to these presumed revenue cutoffs. We present evidence on both the dynamics of bunching at these thresholds, and the relative reallocation of taxpayers from both sides of the thresholds in response to the new incentives.

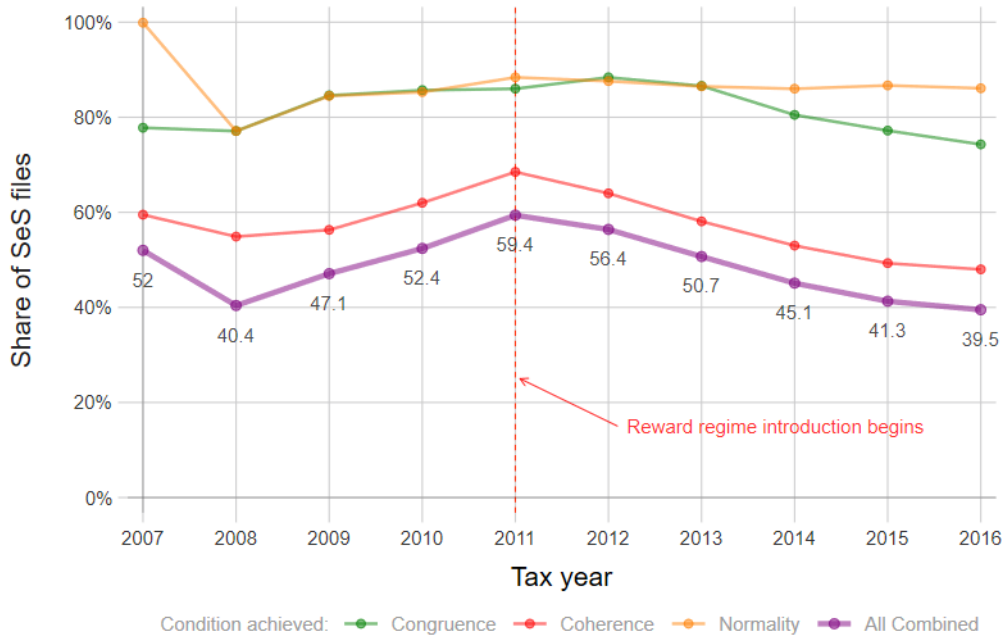
4.3.1 Sector Study compliance effects

This section assesses whether exposure to the new regime improved SeS compliance in the aggregate or if we can single out groups of the taxpaying population that were most receptive.

The SeS system defines congruence, normality, and coherence as compliance conditions. Figure 8 shows that before the reward regime reform about 80% or more of our sample either reported revenues at or above the presumed level defined by *Gerico* (congruence),

or scored within acceptable ranges on a number of accounting regularity indicators (normality). The combination of these two conditions granted access to the kind of audit exemptions external to the SeS system described in Section 2. The reward regime, however, tied these and further exemptions to additionally achieving coherence - a condition that 52.4% of SeS taxpayers in the balanced panel were meeting by 2010.¹⁷ In practice, this means that broader rewards would be granted to a more selected subset of beneficiaries.

Figure 8: Compliance evolution in the balanced panel, 2007-2016



Notes: the Figure shows the share of all files from single-sector, continuous filers over 2007-2016 achieving the access conditions for the reward regime (congruence, coherence, and normality) as defined by their respective Sector Study. Each year features 1,412,980 files. Staggered regime introduction starts in 2011.

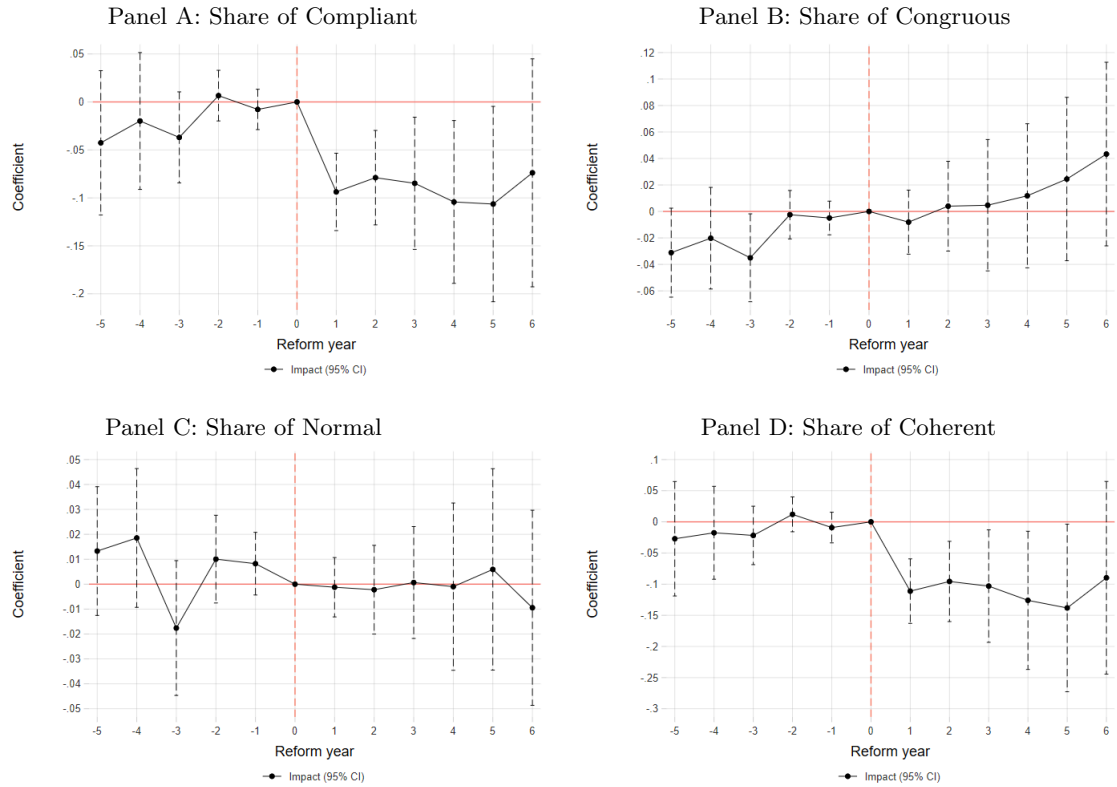
With this background in mind, we ask whether the reform affected SeS compliance across the board. Figure 9, Panel A replicates our event-study analysis using as outcomes the share of files simultaneously congruous, normal, and coherent in each sector-year. Surprisingly, this share seems negatively affected by the reform.¹⁸ The resulting 9% mean drop is entirely

¹⁷Figure A.1 suggests this share might have been even lower among sectors entering the reward regime in the following years.

¹⁸A potential explanation could be that each sector's reform happens at the same time as one of the regular updates of its SeS statistical model. Average compliance could be affected if updates adopt progressively stricter compliance criteria. The drop in overall compliance is however robust to controlling for SeS update years and their interaction with the tax year of each sector's reward regime reform.

explained by an instantaneous reduction in overall coherence (Panel D), while congruence and normality display no significant trend (Panel B and C, respectively). On the net, these results imply that the expansion of the regime did not result in an immoderate diffusion of audit protections among treated sectors relative to what would have happened absent the reform. The evidence presented is therefore reassuring from the revenue collection perspective of the Revenue Agency: the new reward system expanded the reported tax base without resulting in significantly broader audit protection.

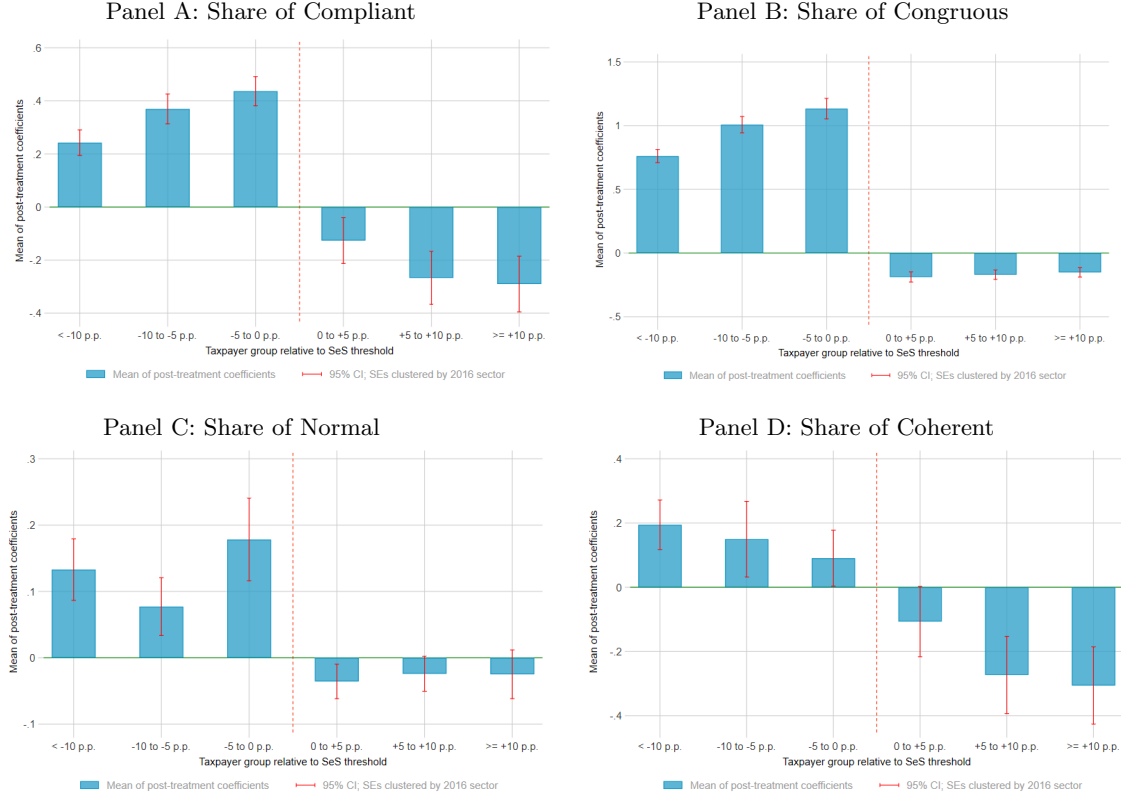
Figure 9: Aggregate reward regime effects on Sector Study compliance



Notes: the Figure shows the effects of the reward system introduction in a sector on the share of taxpayers simultaneously congruous, normal and coherent (Panel A), on the share of congruous taxpayers (Panel B), on the share of normal taxpayers (Panel C), and on the share of coherent taxpayers (Panel D). Whiskers represent 95 percent confidence intervals. Estimates are based on our event-study specification in (1). Standard errors are clustered at the sector level. See notes to Figure 5 for a detailed explanation of the regression sample. Number of sector-years: 1550.

We then turn to investigate whether compliance responses are homogeneous for taxpayers with different initial SeS conditions. In particular, we start from studying whether the reward system successfully mobilized previously non-compliant SeS filers. Each file in our

Figure 10: Heterogeneity in compliance responses by SeS presumed revenue distance



Notes: the Figure shows the effects of the reward system introduction in a sector on the share of taxpayers simultaneously congruous, normal and coherent (Panel A), on the share of congruous taxpayers (Panel B), on the share of normal taxpayers (Panel C), and on the share of coherent taxpayers (Panel D). Bars display for each of six groups the average of six group-specific post-treatment coefficients from an event-study based on the specification in (1). Whiskers represent 95 percent confidence intervals of these combinations of coefficients. Different groups are defined based on the revenues declared the year before the reward system is introduced in a given sector, with bins based on the distance from the presumed revenue amount as indicated on the horizontal axis. The vertical dashed line represents the presumed revenue amount. See notes to Figure 5 for a detailed explanation of the regression sample. Number of sector-years: 1550.

data is associated to an idiosyncratic presumed revenue amount, which we use to quantify an individual degree of compliance along the congruence dimension. Specifically, we group taxpayers according to the percentage points distance between their files' declared and presumed revenues in the year before entry in the new system of incentives. The result is a set of six groups formed symmetrically below and above the SeS presumed revenues threshold, which we then collapse by sector and year to perform our analysis.¹⁹

¹⁹Namely, we group together files below and above their SeS threshold by more than 10 p.p., between 5 and 10 p.p., and between 0 and 5 p.p., respectively. Given that congruence is also achieved by declaring

The four panels of Figure 10 reveal substantial heterogeneity in compliance responses based on a taxpayer’s relative position just before the advent of the reward regime. Compared to the baseline scenario, reward-sector groups previously below the threshold significantly improve their stances, with higher likelihood of being congruous, normal, and coherent, both individually and simultaneously. Hence, the new audit exemptions proved especially effective among previously unprotected taxpayers.

The fact that groups above the threshold seem to be moving in the opposite direction, though often with smaller magnitudes, is also enough to explain the aggregate results from Figure 9. As shown in Table A.1, the vast majority of the files we employ in our analysis report revenues above the corresponding presumed amounts. Therefore, a reduction in compliance in that revenue range has prevented a generalized rise in compliance among reward sectors at large.

4.3.2 Group heterogeneity in reporting responses

As much as for SeS compliance, gross profit responses are heterogeneous on the two sides of the SeS presumed revenue thresholds. The top panels in Figure 11 document that changes in filing behavior lead to significant increases in average profits, both in relative and in absolute terms, mostly among previously non-congruous taxpayers. The point estimates of these expansions decrease as we approach the cutoff from below and tend to lose significance above it.

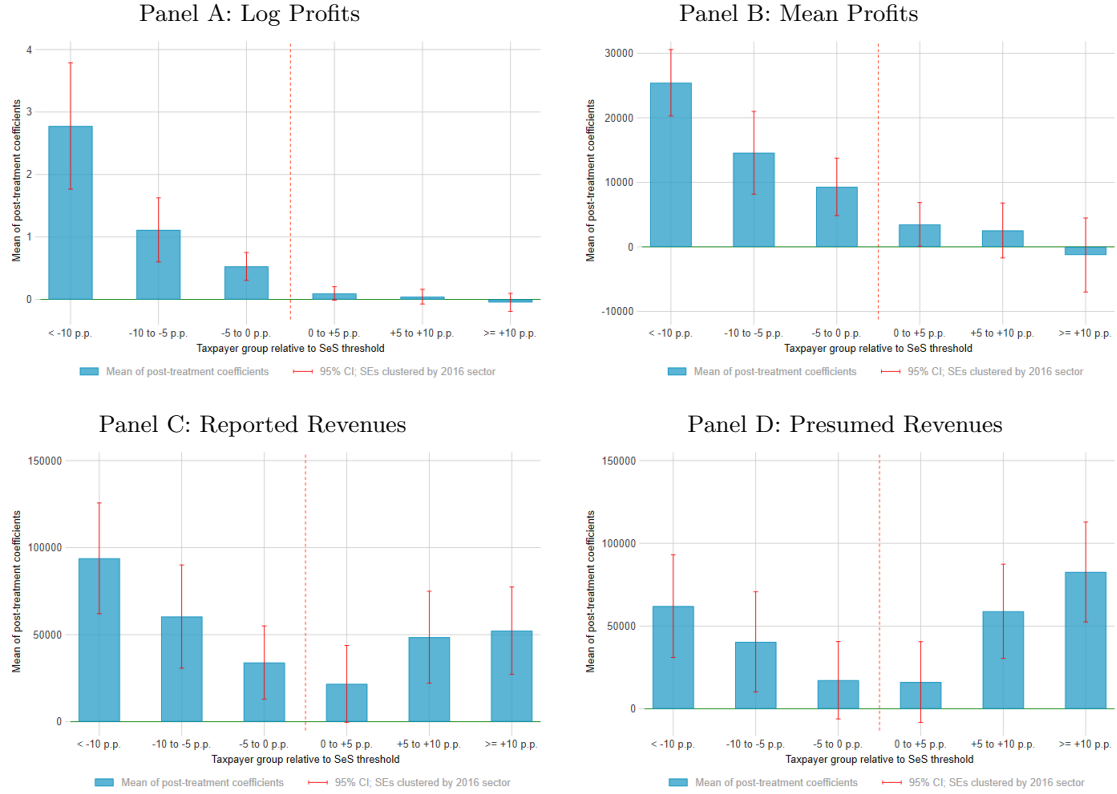
These results may be understood in light of the simultaneous adjustments operated by businesses along their revenue and cost margins. Panel C shows that almost all groups increase their declared revenues. However, input and cost reporting behavior as proxied by changes in mean presumed revenues offset any revenue gain for taxpayers who were congruous before the reform (Panel D). The fact that presumed revenues grow faster than reported revenues for these groups is behind the decrease in congruence rates observed in Figure 10, Panel B. Additional results elucidate the sources of the observed presumed revenue changes. Figure 12 suggests that the reward regime was especially effective in raising reported employment costs (Panel A) and reported employment (Panel B) among businesses more than 5 p.p. above their SeS thresholds. In particular, the average coefficients on the two top groups imply that 0.31 and 0.63 additional employees are reported at reward-sector firms, relative to a pre-reform mean of 2.3 and 1.45, respectively. This evidence suggests that the policy could have induced some groups of taxpayers to regularize irregular workers. However, more detailed data on work contracts is needed to support this conjecture.

We present some of the accounting implications of these dynamics in Panel C and D of Figure 12. While median profit-to-turnover ratios (Panel C) grows substantially among taxpayers far below the SeS threshold, employment cost-to-turnover ratios (Panel D) rise

revenues exactly at the presumed revenue level, we include such taxpayers in the group reporting within 5 p.p. above their presumed revenues.

significantly only among taxpayers further above the cutoff.

Figure 11: Profit and revenue responses by SeS presumed revenue distance



Notes: the Figure shows the effects of the reward system introduction in a sector on log mean profits (Panel A), on mean profits (Panel B), on mean reported revenues (Panel C), and on mean presumed revenues (Panel D). Bars display for each of six groups the average of six group-specific post-treatment coefficients from an event-study based on the specification in (1). Whiskers represent 95 percent confidence intervals of these combinations of coefficients. Different groups are defined based on the revenues declared the year before the reward system is introduced in a given sector, with bins based on the distance from the presumed revenue amount as indicated on the horizontal axis. The vertical dashed line represents the presumed revenue amount. See notes to Figure 5 for a detailed explanation of the regression sample. Number of sector-years: 1550.

4.3.3 Distributional shifts around Sector Study presumed revenue thresholds

Thanks to the level of detail in our data, we can further document how the new regime's advent reshaped the distribution of declared revenues relative to the level of presumed revenues.

All else equal, taxpayers benefit from being congruous, i.e. reporting at or above their presumed revenue amount, a condition which may discretely reduce their risk of undergoing

Figure 12: Input and Profit responses by SeS presumed revenue distance



Notes: the Figure shows the effects of the reward system introduction in a sector on mean employment costs (Panel A), on mean number of employees (Panel B), on profit rate (Panel C), and on the employment cost rate (Panel D). Profit and employment cost rates are computed using turnover in 2007 as a base. Bars display for each of six groups the average of six group-specific post-treatment coefficients from an event-study based on the specification in (1). Whiskers represent 95 percent confidence intervals of these combinations of coefficients. Different groups are defined based on the revenues declared the year before the reward system is introduced in a given sector, with bins based on the distance from the presumed revenue amount as indicated on the horizontal axis. The vertical dashed line represents the presumed revenue amount. See notes to Figure 5 for a detailed explanation of the regression sample. Number of sector-years: 1550.

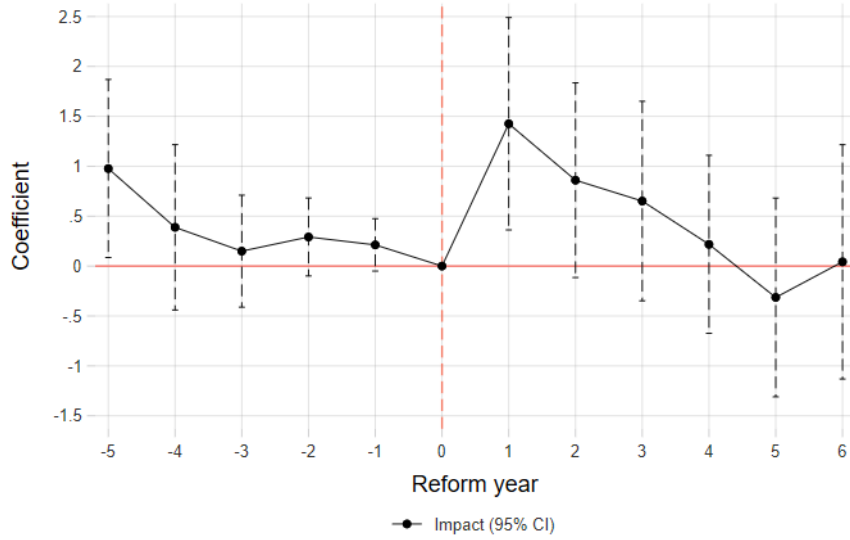
a tax assessment. Relative to a scenario without such discontinuity in incentives, the SeS system provides reason for taxpayers to bunch above but near the idiosyncratic revenue thresholds generated by *Gerico*.

Figure A.2 presents evidence of such bunching. Among sectors entering the reward regime by 2016, bunching estimates are on average between 4.05 and 5.87 between 2007 and 2010. Hence, even before the reform, the excess mass of taxpayers at the presumed revenue amount was about 4-6 times larger than in the counterfactual distribution that would arise if all taxpayers perceived the fiscal incentives available to congruous filers. This provides conspicuous evidence that the audit exemptions tied to congruence were already driving reporting

decisions in the pre-reform SeS system.

The reward regime seeks to update these exemptions by tightening access requirements on one hand, and expanding the relevant audit benefits on the other. If taxpayers perceive the resulting increase in the value of congruence, we should observe a positive shock to bunching estimates among treated sectors, with relatively more files to be found above but near the presumed revenue threshold.

Figure 13: Reward regime effects on bunching at Sector Study presumed revenues thresholds



Notes: the Figure shows the effects of the reward system introduction in a sector on the level of bunching at the Sector Study presumed revenue amount. Bunching is computed as the excess amount of Sector Study files between 0 and 1 percentage point of distance from the presumed revenue amount relative to a 7th-order polynomial counterfactual. Whiskers represent 95 percent confidence intervals. Estimates are based on our event-study specification in (1). Standard errors are clustered at the sector level. See notes to Figure 6 for a detailed explanation of the regression sample. For each bunching estimation exercise, we trim observations at the 5th and 95th percentile of presumed revenue distance. Number of sector-years: 1550.

We test this conjecture in Figure 13. We re-estimate Equation (1) using our sector-year bunching measure as a dependent variable. While pre-reform differences in bunching are insignificant, the first year of application of the reform raises bunching by about 1.43 units. Effects are positive for at least the following three tax years, although large standard errors prevent us from rejecting the hypothesis that meaningful impacts are limited to the first period.

Changes in the excess mass at the SeS presumed revenue thresholds may come from different sources. Some taxpayers previously reporting below the cutoff might be encouraged to move up along the relative revenue range. This could be beneficial to both those trying to

reduce the expected penalty from detection of underreported tax bases, and those trying to avoid audits they perceive as costly regardless of intended evasion behavior. On the other hand, some taxpayers otherwise reporting above the cutoff might be willing to reduce their relative compliance. Indeed, the promise of reduced audit risks might induce previously congruous taxpayers to shrink their (positive) revenue distance from the threshold. Either way, both types of behavior would contribute to observing more files near but above the SeS thresholds.

To assess the plausibility of these distributional shifts, we first compute the share of files at each percentage point of distance between reported and presumed revenues in every sector-year. Next, we re-estimate our main equation using these shares as outcomes to assess the effect of the reform at each point of the relative revenues distribution. In Figure 14 we then report the average of the six treatment-period coefficients for each one percentage point bin regression, displaying the impact of the new audit exemptions on both sides of the threshold.²⁰ We present our analysis separately by six groups of taxpayers based on their relative distance from the cutoff in the year before their sector’s reform.²¹

Panels A, B, and C illustrate the relative reshuffling of taxpayers previously at various distances below the threshold. The green band highlights the bins from which taxpayers would be moving based on their pre-reform location. In all three cases, a general movement away from the green band is observed, with stronger reductions in bin shares for bins closer to the threshold. All else equal, this is consistent with a lower cost of achieving congruence for taxpayers required to travel a shorter distance to obtain the new audit exemptions.

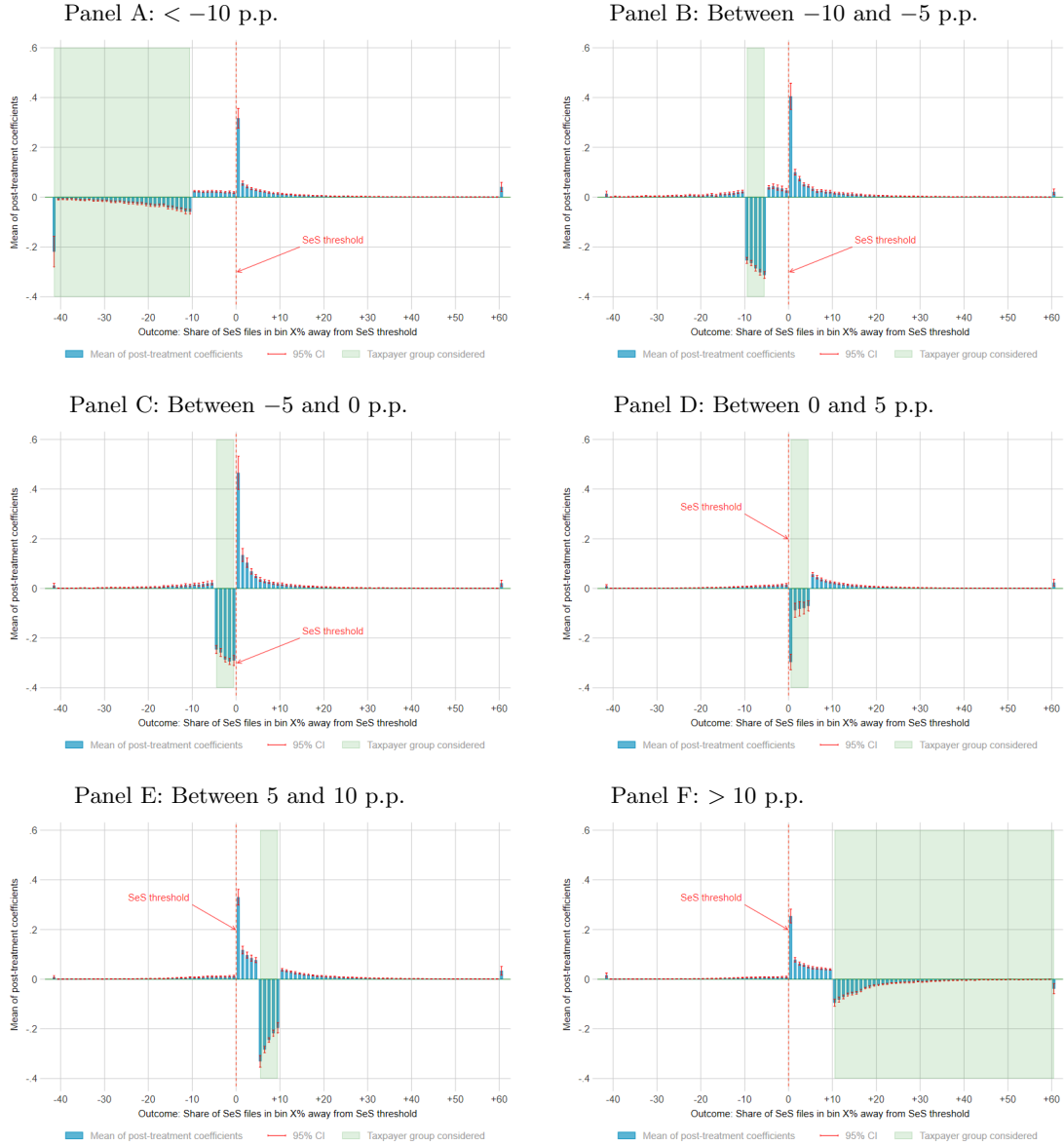
As expected from the discontinuous structure of the incentives, redistribution happens largely towards the bins just above the presumed revenue threshold. All three groups transfer a sizable amount of mass already within one percentage point of the threshold, and progressively smaller amounts thereafter. Once again, the closer the starting point, the larger the effect: the farthest, the intermediate, and the most proximate groups below the threshold increase the share of files in the first bin above the cutoff by 31.7, 40.5, and 46.5 p.p. more, respectively, than in the counterfactual scenario.

Incentives might work in the opposite way for those starting above the threshold. As discussed before, we might find relative revenue-reducing effects among those whose perceived audit risk is diminished by the new audit exemptions. Panels D, E, and F address this point. The two groups farthest above the threshold shift towards the cutoff, with taxpayers becoming more likely to be just barely congruous, and with increasing intensity for lower bins. The share gains for the bin at the threshold among these two groups are 33 and 25.4 p.p., respectively (Panels E and F). On the other hand, the taxpayer group closest to the threshold loses density near the cutoff. This loss is in favor of bins further up along the relative revenue range rather than outside of the congruence zone (Panel D).

²⁰The first and last bars relate to the residual bottom and top threshold-distance categories. In particular, they cumulate the effect of the reform on all files up to -41 p.p. and at or above +60 p.p. of the presumed revenue threshold ahead of the reform, respectively.

²¹As a reference, Figure A.3 in the Appendix plots the relative distribution of files within each of the six groups in the year before the reform.

Figure 14: Distribution shifts, by SeS threshold distance before reform



Notes: the Figure shows the effect of the reward regime on the average share of Sector Study 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 (1). Whiskers represent 95 percent confidence intervals of these combinations of coefficients. Standard errors are clustered at the sector level. See notes to Figure 5 for a detailed explanation of the regression sample. Number of sector-years: 1550. 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).

These shifts provide important caveats for thinking about the potential role of audit exemptions in encouraging tax compliance. While some taxpayers may be nudged into reporting larger tax bases than otherwise, others may react in undesirable ways to the laxer enforcement regime they now face. The direct effects of the reform in the aggregate will reflect the combination of these two opposing forces. For the medium run, however, additional considerations must be made. The tax base results reported so far potentially mask relevant implications for the quality of the ensuing tax audits. Indeed, the selection process with which taxpayers sort themselves on either side of the threshold might affect the revenue collection potential of newly admissible audits. As a result, tax agencies might need to accompany the promise of selective audit breaks with new ways to limit the extent of tax liability reduction that strategic taxpayers are willing to achieve.

5 Conclusions

Main Takeaways. In the Italian Sector Study system, small businesses and the self-employed are promised exemptions from selected types of tax audits conditional on achieving well-defined accounting targets. We investigate whether reinforcing these audit exemptions generates positive returns to the tax administration. Studying the staggered introduction of a new reward regime between 2011 and 2016, we present evidence that the average firm expands its tax base in the short run to benefit from stronger audit protections.

Our analysis features several interesting results. First, despite the rise in mean gross profits, the average sector share of taxpayers eligible to the new rewards does not increase across the board. The direct implication is that, in the case of the reform we study, the promise of stronger audit exemptions did not translate into a broader exempted population. This fact might ease concerns that this type of reforms must necessarily trade off a short-lived increase in tax revenues with a drastic reduction in the inspection powers of the tax authorities.

Second, tax base responses are heterogeneous by pre-reform compliance status. The taxpayers with the largest increases in gross profits are those that the Revenue Agency believed to be reporting less revenues than expected ahead of the reform. Since reporting less than what presumed by the administration would preclude access to the reform benefits, our evidence suggests that the new audit exemptions proved especially effective in motivating previously non-compliant taxpayers to improve their position.

Third, reporting flexibility might limit the collection potential of audit exemptions. In particular, the effective audit risk reduction among previously compliant firms attenuates their incentives for truthful reporting. Although our data doesn't allow us to establish the accuracy and completeness of reported items, we provide evidence that groups of taxpayers declaring more than 5 p.p. above their presumed revenues have increased their employment costs more than other groups and leveled down their revenues towards their compliance targets. While the reform did not significantly reduce these groups' mean profits, we question whether manipulation of reported items has prevented a potential tax base expansion also among these categories.

Data Limitations. Our work comes with two main shortcomings. First, most professionals’ sectors in our data were not granted access to the new regime within our sample period, which raises questions about the validity of our results for filers excluded from the reform. Second, the Sector Study database does not include information on which businesses undergo tax audits, nor the outcome of any resulting tax assessment. This may be especially problematic when attempting to derive the overall effects of extending audit exemptions. Depending on the identity of the novel beneficiaries, these exemptions may come at the cost of a more adversely selected audit pool, with the tax administration left to investigate unprotected firms with truthful reports. Only detailed audit data can reveal whether the magnitudes of the potential reduction in audit quality outweighs the tax base gains we document in our work. We hope to tackle this possibility in future work.

ISA: lessons for the evolution of Sector Studies. Starting in the 2018 tax year, Sector Studies evolved to fully incorporate and expand the 2011 reward regime. The new system, known as indices of fiscal compliance (*indicatori sintetici di affidabilità fiscale*, or ISA), combined the original set of compliance indicators into a synthetic score ranging from 1 to 10. Taxpayers enjoy increasing audit exemptions at higher levels of compliance based on their average performance on a broad set of accounting indices. This innovation might partly decrease the risk of adversarial reporting behavior we have shown by smoothing out the set of fiscal targets taxpayers aim for and by reducing the importance of adjusting individual items upon filing.

While our results suggest that the introduction of a broader reward system could lead to the emergence of a larger tax base, the actual effects will depend on the interplay between the specific audit exemptions provided, the enforcement strategies pursued by the tax administration, and the elasticity of each reporting item to the prevailing audit risk.

References

- J. Alm, J. Martinez-Vazquez, and S. Wallace. *Taxing the Hard-to-tax: Lessons from Theory and Practice*, volume 268. Elsevier, Amsterdam, 2004.
- J. Andreoni, B. Erard, and J. Feinstein. Tax compliance. *Journal of Economic Literature*, 36, 2: 818-60, 1998.
- P. Bachas and M. Soto. Not(ch) your average tax system: Corporate taxation under weak enforcement. *Policy Research Working Paper Series 8524, The World Bank.*, 2018. ISSN 1813-9450. doi: 10.1596/1813-9450-8524.
- M. Bertrand, E. Duflo, and S. Mullainathan. How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics*, 119(4):249–275, 2004.
- K. Borusyak and X. Jaravel. Revisiting event study designs, with an application to the estimation of the marginal propensity to consume. *working paper*, 2017.

- P. Carrillo, D. Pomeranz, and M. Singhal. Dodging the taxman: Firm misreporting and limits to tax enforcement. *American Economic Journal: Applied Economics*, 9(2):144–64, April 2017a.
- P. E. Carrillo, E. Castro, and C. Scartascini. Do rewards work? evidence from the randomization of public works. *IDB Working Paper IDB-WP-794*, 2017b. doi: 10.18235/0000673.
- E. Di Gregorio and M. Paradisi. Audit rule disclosure and tax compliance. *Working paper*, in preparation.
- T. Dunning, F. Monestier, R. Pineiro, F. Rosenblatt, and G. Tunon. Is paying taxes habit forming? theory and evidence from uruguay. *Working paper*, 2017.
- H. J. Kleven and M. Waseem. Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from pakistan. *Quarterly Journal of Economics*, 128: 669–723, 2013.
- Ministry of Economy and Finance. Economia non osservata e flussi finanziari. rapporto finale sull’attività del gruppo di lavoro sull’economia sommersa e i flussi finanziari. Technical report, 2011.
- Ministry of Economy and Finance. Relazione sull’economia non osservata e sull’evasione fiscale e contributiva. Technical report, 2017.
- Ministry of Economy and Finance. Relazione sull’economia non osservata e sull’evasione fiscale e contributiva. Technical report, 2019.
- J. Naritomi. Consumers as tax auditors. *American Economic Review*, 109:3031–3072, 2019. ISSN 0002-8282. doi: 10.1257/aer.20160658.
- OECD. Italy’s tax administration: A review of institutional and governance aspects. Technical report, 2017.
- A. Santoro. 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., 2008.
- A. Santoro and C. Fiorio. Taxpayer behavior when audit rules are known: Evidence from italy. *Public Finance Review*, 39:103–123, 2011.

APPENDIX

A Estimation of bunching at the SeS presumed revenues

Our empirical analysis employs a measure of bunching as one of the outcomes of interest. This variable helps us to quantify the extent to which taxpayers move toward the presumed revenues amount or just above it after the reward regime is introduced in their sector.

Our measure relies on the counterfactual revenue distribution that would arise if there was no discontinuity in audit enforcement at the presumed revenues threshold. Specifically, this is the distribution that would arise if the audit probability perceived by taxpayers was constant and equal to the one perceived above the level of presumed revenues in the presence of Sector Studies.

We build the counterfactual following the standard procedure in the literature, which exploits information on the observed distribution of SeS files above their presumed revenues to extrapolate the entire counterfactual. In line with [Kleven and Waseem \(2013\)](#), we estimate a flexible polynomial excluding an area $[y_L, y_U]$ around the presumed revenue amount \hat{y} . Our specification reads:

$$c_j = \sum_{i=1}^K \beta_i (y_j)^i + \sum_{k=y_L}^{y_U} \gamma_k \mathbb{1}(y_j = k) + \varepsilon_j$$

where c encodes the count of SeS files in any given bin j around the presumed revenue threshold and y reflects the bin's position in percentage points of distance from the threshold. i is the polynomial degree. We use a 7-th degree polynomial for our exercise.

In the equation, the second block of terms includes a vector of dummies for bins in the relative revenue range $[y_L, y_U]$ to ensure that the excess mass at \hat{y} does not affect the estimation of the smooth counterfactual. We set y_U visually at the first bin above \hat{y} , while we use an iterative procedure to determine y_L . Specifically, we equate the missing mass below \hat{y} , computed as the cumulative difference between the counterfactual and the observed distribution below the threshold, to the level of excess mass in the first bin just above \hat{y} . In our case, excess mass refers to the number of files observed above the counterfactual number at the first bin above the threshold. After estimating the equation above, we compute our excess bunching measure as the ratio between the excess number of files observed at \hat{y} and the counterfactual number of files at the same point.

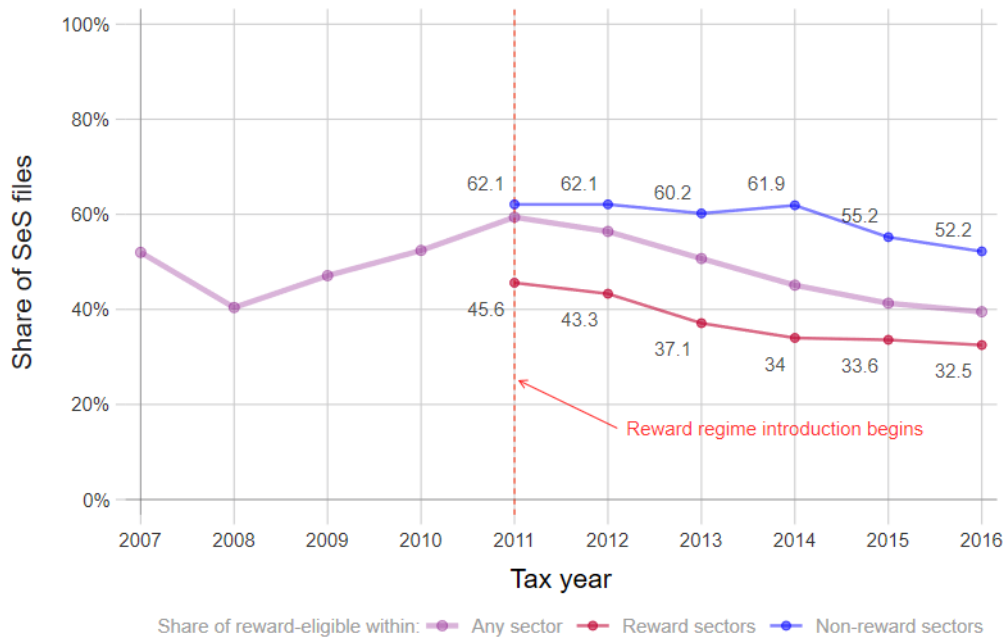
B Additional tables and figures

Table A.1: Group size by SeS presumed revenue distance

Group of taxpayers (with declared revenues relative to SeS presumed revenues)		Nr. of Files	% of All Files
1	Up to -10 p.p.	868,580	9.60
2	-10 to -5 p.p.	274,460	3.03
3	-5 to 0 (excluded) p.p.	383,980	4.24
4	0 (included) to +5 p.p.	4,081,820	45.09
5	+5 to +10 p.p.	1,187,960	13.12
6	+10 p.p. or more	2,255,600	24.92
Total		9,052,400	100

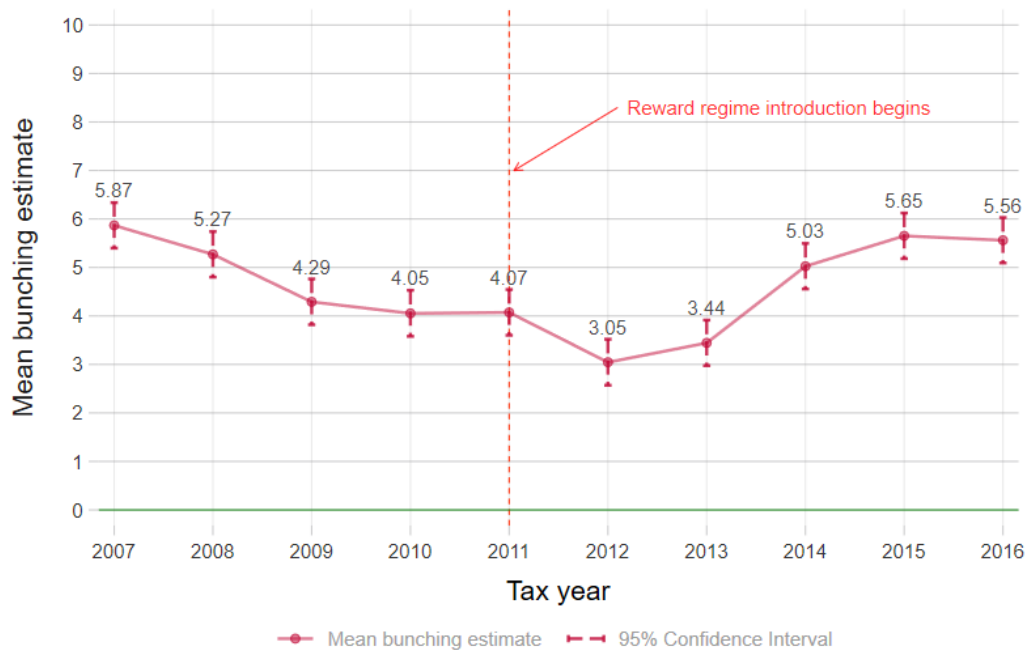
Notes: this Table reports the breakdown of all SeS files from single-sector taxpayers in the 2007-2016 balanced panel based on the relative distance of each file's reported revenue from its associated presumed revenue level as computed by Gerico the year before the taxpayer's sector entry into the reward regime. Only files from the 155 sectors entering the 2011 reward regime by 2016 are therefore considered, as in the rest of the empirical analysis. Group 4 includes revenues at the threshold.

Figure A.1: Compliance evolution inside and outside of the reward regime



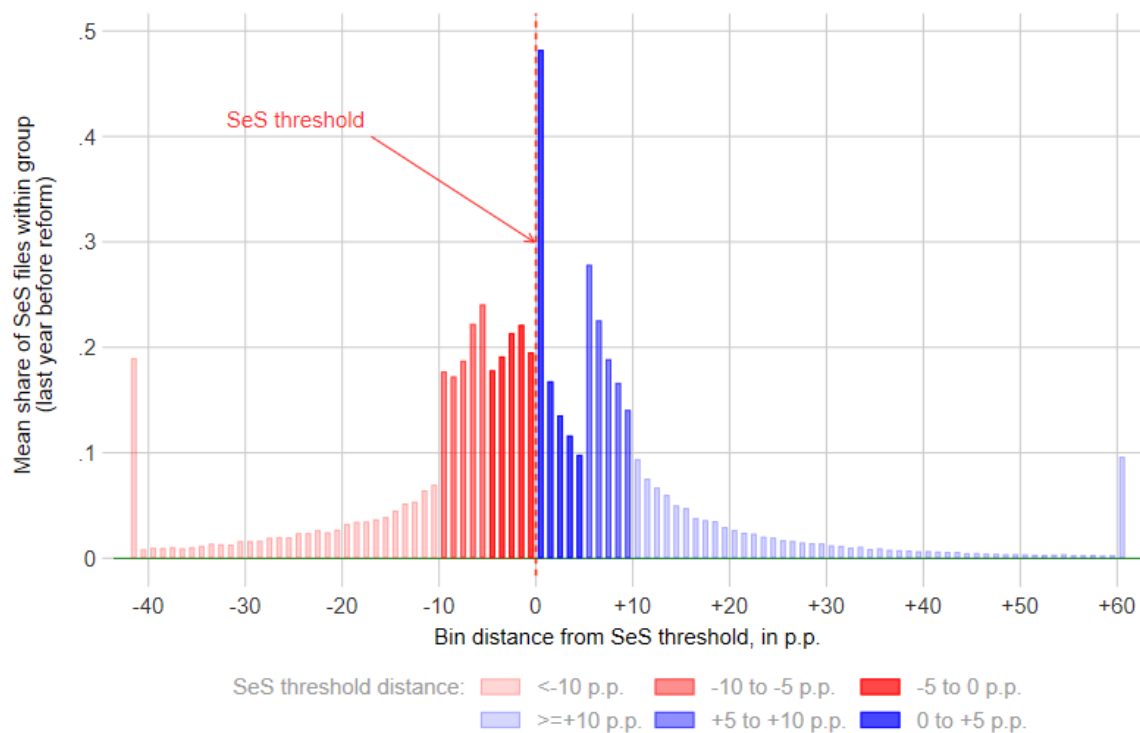
Notes: the Figure shows the share of all files from single-sector, continuous filers over 2007-2016 achieving all the access conditions for the reward regime (congruence, coherence, and normality) as defined by their respective Sector Study. After 2011, when the staggered reward system introduction started, the Figure also shows the compliance measure for sectors with and without reward system access. Each year features 1,412,980 files.

Figure A.2: Mean bunching at Sector Study presumed revenues thresholds



Notes: the Figure plots the evolution of mean bunching estimates across sectors ever accessing the reward regime in the 2011-2016 period. For each sector-year, we estimate bunching within 1 percentage point of distance above SeS presumed revenues. Data are from each SeS file from single-sector taxpayers in the 2007-2016 balanced panel. Samples are trimmed at the 5th and 95th percentiles of SeS threshold distance ahead of bunching estimation. Means and confidence intervals are from a regression of sector-year bunching estimates on tax year dummies with no model constant.

Figure A.3: Bin file share in the year before the reform, by SeS threshold distance



Notes: the Figure shows six distinct distributions for the relative allocation of Sector Study files in the year before the advent of the reward regime. Each distribution is associated to one of six groups of taxpayers based on their reported revenues' relative distance from their Sector Study presumed revenue amount in the year before their sector's reform. Bar height represents the group-specific average share of files at each percentage point of presumed revenue distance across all treated sectors in the year before the reform. The first and last bars display the average share of all files with revenues up to -41 percentage points and at or above $+60$ percentage points of distance, respectively. The starting sample consists of all Sector Study files submitted by single-sector taxpayers continuously filing over the 2007-2016 period.