Routledge
Taylor & Francis Group

Economic and Political Studies

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/reps20

Tax enforcement activities: Evidence on the impact of a threat-of-audit letter

Carlo V. Fiorio & Alessandro Santoro

To cite this article: Carlo V. Fiorio & Alessandro Santoro (2023): Tax enforcement activities: Evidence on the impact of a threat-of-audit letter, Economic and Political Studies, DOI: 10.1080/20954816.2023.2192559

To link to this article: https://doi.org/10.1080/20954816.2023.2192559

Published online: 18 Apr 2023.

Submit your article to this journal

Article views: 16

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=reps20

ECONOMIC AND POLITICAL STUDIES
https://doi.org/10.1080/20954816.2023.2192559

RESEARCH ARTICLE

Tax enforcement activities: Evidence on the impact of a threat-of-audit letter

Carlo V. Fiorio[a,b,c] and Alessandro Santoro[d,e]

[a]Department of Economics, Management and Quantitative Methods (DEMM), University of Milan, Milan, Italy; [b]The Research Institute for the Evaluation of Public Policies, Bruno Kessler Foundation (FBK-IRVAPP), Trento, Italy; [c]Dondena Centre, Milan, Italy; [d]Department of Economics, Management and Statistics (DEMS), University of Milan-Bicocca, Milan, Italy; [e]Centre for European Studies (CefES), University of Milan-Bicocca, Milan, Italy

ABSTRACT

As part of their ordinary enforcement activities, tax authorities around the world use letters and other forms of personalised communications to elicit tax compliance at a low cost. The literature shows that the impact of a threat-of-audit letter (TAL) depends critically on the credibility of the letter which, in turn, depends on the information disclosed to the taxpayer. We study a TAL that is credible because it also uses third-party information to target a form of cost manipulation but, at the same time, incentivises alternative forms of cost manipulation by not targeting them. We find that the direct impact on the targeted cost manipulation is stronger than the strategic increase in other costs, and therefore the overall effect of the TAL on taxable income is positive. Our approach can be applied to any letter or enforcement action that reveals information to taxpayers and thus prompts their strategic responses. By observing the taxpayers' responses along all relevant dimensions, the tax authority can evaluate the impact of the strategy and also obtain valuable information to target future enforcement activities.

ARTICLE HISTORY
Received 5 November 2021
Accepted 19 October 2022

KEYWORDS
Tax compliance by businesses; threat-of-audit letter; difference-in-differences; small business taxation

Introduction

Tax authorities worldwide often use personalised communications, such as letters, to encourage tax compliance at a relatively low cost. According to Slemrod (2019), there are four categories of intervention letters, but this paper focusses on two of them. Specifically, we examine threat-of-audit letters (TALs), which warn that the tax return will be closely scrutinised or audited with some probabilities, and third-party letters (TPLs), which notify taxpayers that the tax authority has received third-party information indicating possible non-compliance. TALs conveying basic threat messages, such as those studied by Slemrod, Blumenthal, and Christian (2001) or by Kleven et al. (2011), have an impact only on subpopulations whose

CONTACT Alessandro Santoro, alessandro.santoro@unimib.it Department of Economics, Management and Statistics (DEMS) and Centre for European Studies (CefES), University of Milan-Bicocca, Milan, Italy

© 2023 Economic and Political Studies

features (the absence of a tax consultant and the presence of a third party, as studied respectively in these two works) would support compliance even without a letter. TALs that try to reinforce the basic threat messages by reporting also implausibly high values of audit probability (Bérgolo et al. 2023) do not have any impact. TALs reporting some additional information on the third-party source of the information (Meiselman 2018), on the payment process (Ortega and Scartascini 2020), or on the expectations the administration has expressed in a very simple way (De Neve et al. 2021) have a positive impact on average. Hence, additional information to the taxpayer, on top of the basic threat messages, emerges as the key to the credibility of a TAL.

However, the literature on enforcement (Carrillo, Pomeranz, and Singhal 2017; Slemrod et al. 2017) shows that there are cases when the credibility of enforcement activities based on revealed information does come at a cost. By disclosing information, the tax authority also reveals – or draws attention to – the information that it does not possess, and, by doing so, it may trigger strategic responses from taxpayers.

In this paper, we study the impact of a TAL that, along with the threat message, conveys some relevant additional information to the taxpayer and hence leads to the taxpayer's strategic responses. Hence, the TAL we study is unique in the literature, as it merges some features of TALs and of TPLs.

The letter is sent by the Italian revenue authority to small businesses that are allegedly manipulating some of their cost reports, namely residual costs (RCs). The letter is based on some evidence, a mixture of third-party and own-reported information, indicating that these reports are 'anomalous'. Different from other TALs that have been studied in the literature, this evidence is disclosed to the taxpayers. By addressing RCs, the letter aims to indirectly increase taxable income. However, due to the peculiar nature of the letter and our access to a unique administrative dataset measuring a large set of variables, we are able to disentangle the direct effect of the letter, i.e. its impact on the behaviour which is targeted by the letter, from the indirect effect, i.e. the impact the letter has on other (non-targeted) forms of cost manipulation, allowing us to identify any strategic behaviour of the taxpayers. The overall effect on taxable income is the net sum of these two, and it ultimately provides evidence of the impact of the letter on tax revenues.

To address these questions empirically, we have to address an identification challenge. As the only aim of the Revenue Agency (RA) is to maximise tax revenues, all firms identified as possibly anomalous according to the RA criteria are sent the letter, without any randomisation strategy. This poses a challenge to our estimation of the causal effects of the TAL, which is based on a difference-in-differences (DD) approach, as the parallel trend assumption might not apply to all firms. Hence, we exploit the richness of our database, where businesses are clusterised according to their structural features, then run parallel trend tests between treated and control units by single clusters of activities, and finally select only firms belonging to a cluster where the parallel trend assumption is not rejected.

We find that the letter has the expected direct impact, i.e. it decreases the RCs targeted by the letter and increases taxable income, but it also triggers some alternative forms of cost manipulation, reducing the overall effect of the TAL.

Indirect effects are relevant also because they provide useful information to the tax authority. In a recent contribution, Advani, Elming, and Shaw (2021) argue in favour of audits because they provide information to the tax authority on the heterogeneity of responses across income types, which can be used to better target enforcement activities, whereas TALs reveal information to taxpayers. Here, we show that by disentangling direct from indirect effects, TALs can also provide information on the strategic behaviour of taxpayers, which could be used to better sharpen future enforcement activities.

This paper is organised as follows. We first review the literature and clarify how the credibility of intervention letters helps explain their impact. Next, we illustrate the functioning of the Italian SDS (literally *studi di settore*, or business sector studies), and provide details on the TAL sent in 2009 by the Italian tax authority to taxpayers alleged to manipulate their RCs. In the following section, we describe the data and the empirical strategy. We then examine the direct impact of the TAL on RCs as well as its impact on taxable income, and provide evidence on the indirect impact of alternative cost manipulations. The last section concludes.

Review of the literature on intervention letters

A TAL states that the tax return will be closely examined or that it will be examined with some stated probabilities, and the reaction depends on how the message changes the targeted taxpayer's perception of the probability of an audit (Slemrod 2019).

In Slemrod, Blumenthal, and Christian (2001), high-income taxpayers with self-employment or farm income receiving the TAL on average report lower income, a seemingly perverse response from the perspective of the deterrence model. The authors argue that more sophisticated taxpayers (and their tax consultants) understand an audit to be a negotiation, and view reported taxable income as the opening (low) bid in a negotiation that does not necessarily result in finding and penalising all non-compliance. Therefore, a TAL is not credible when the taxpayer knows (or believes) that the nature of the audit process makes the threat less worrying than it may seem at first glance.

In Kleven et al. (2011), the threat-of-audit experiment is conducted only on a sample of employees, i.e. it excludes any self-employed individuals. Although the TAL does not explicitly mention third-party information, employees do know that this type of information is available to the RA. This knowledge, coupled with the fact that, in this case, taxpayers are not aware that they are part of an experiment, makes the TAL studied by Kleven et al. (2011) more credible which, in turn, explains why it has a positive, albeit small, effect on compliance.

The TALs studied by Meiselman (2018) are sent by the City of Detroit on the basis of the information provided by the Internal Revenue Service (IRS), and the taxpayers are aware of this. The rationale for the TALs is that a taxpayer will think punishment is more likely if the tax authority reveals that it has relevant information. Other treatments studied by Meiselman (2018) include the 'penalty salience', where the letter states the otherwise unknown statutory penalty for the misdemeanour of failing to file a tax return, and the 'civic pride', where the letter proclaims the importance of

tax collection to the resurgence of Detroit. The treatment effect of the TAL is found to be positive and significant, albeit half of that of the penalty salience treatment. No effect is found for the civic pride treatment. A possible explanation is as follows. The penalty salience treatment is very credible because it is based on the statement of the law. The impact of the TAL is positive because it is based on disclosed third-party information. Finally, the civic pride treatment does not have any sound foundation, at least in a deterrence context.

Bérgolo et al. (2023) examine the impact of a TAL which is sent to a separate sample of firms pre-selected by the revenue authority for auditing. They divide selected firms into two groups: one receives a letter declaring that if their anomaly is not corrected they will receive an audit with a 25% probability, and the other with a 50% probability. However, 25% and 50% probabilities are exceptionally high, and cannot be considered realistic by any taxpayer who has a minimum experience with the revenue authority. The TAL informs firms of the exact audit probability assigned to them, but Bérgolo et al. (2023) find no systematic difference between the two groups in post-treatment VAT payments.

The TAL studied by Ortega and Scartascini (2020) reveals the account balance, the type of tax, and the year or month in which the tax was not paid. It also includes information on the methods of payment and the cost that the taxpayer is incurring by not paying (e.g. interest and penalties, potential legal action, and possible effects on the credit history). This TAL is reported to have a positive impact on subsequent tax compliance, albeit smaller than that of in-person visits conveying exactly the same information.

De Neve et al. (2021) focus on simplifying communications sent with the purpose of more clearly expressing what the tax administration expects from taxpayers. Simplification includes shortening letters while retaining information concerning fines and follow-up enforcement, including deadlines. The impact of this treatment is positive, and it is reinforced by a final sentence threatening to conduct an audit of uncompliant taxpayers.

To sum up, TALs conveying basic threat messages, such as those studied by Slemrod, Blumenthal, and Christian (2001) or by Kleven et al. (2011), only have an impact on subpopulations whose personal features (e.g. the absence of a tax consultant, and the presence of a third party) will induce compliance even without a letter. TALs that try to reinforce the basic threat messages by reporting also implausibly high values of audit probability (Bérgolo et al. 2023) do not have any impact. TALs reporting some additional information of the third-party source (Meiselman 2018), on the payment process (Ortega and Scartascini 2020), or on the deadlines (De Neve et al. 2021) have a positive impact on average. Additional (to the basic threat messages) information thus emerges as the key to the credibility of a TAL.

TPLs include messages conveying that the tax authority possesses third-party provided information suggesting that the taxpayer is, or has been, noncompliant (Slemrod 2019), and therefore they are intrinsically credible.

The paper by Carrillo, Pomeranz, and Singhal (2017) is a prominent example of a TPL. It studies the impact of a letter sent by the Ecuadorian tax authority to taxpayers whose sales reports on VAT forms do not match the purchase reports issued

by their customers, offering them the opportunity to file an amended return. It finds that when firms are notified by the tax authority about detected revenue discrepancies on previously filed corporate income tax returns, they increase reported revenues, matching the third-party estimate when provided. In such a case, the credibility of the letter is clearly perceived by the taxpayers, who also perceive that the tax authority is explicitly looking at their behaviour. However, firms also increase reported costs by 96 cents per dollar of revenue adjustment, resulting in minor increases in total tax collection. These unintended consequences of the letter are explained by the fact that cost manipulations are not monitored by the tax authority, which is referred to by the authors as an institutional weakness of the Ecuadorian tax system.

Slemrod et al. (2017) do not study a letter, but find a similar strategic response to the introduction of Form 1099-K, an information report introduced in 2011 which provides the IRS with information about electronic sales (e.g. credit card sales). In this case, taxpayers largely offset increased reported receipts with increased reported expenses, which do not face information reporting, thus diminishing the impact on reported net taxable income.

A TAL that reveals information to taxpayers in a way such that some additional (and not originally targeted) form of tax non-compliance may emerge, as a TPL may do, has not been examined in the existing literature.

Institutional background

The Italian tax system and SDS

Small businesses that are not incorporated, such as sole proprietorships, joint ownerships, and unlimited liability companies, are taxed on their profits, which are calculated by subtracting their business expenses from their reported revenues. Since 1998, Italy has adopted SDS to audit businesses (small businesses, corporations and professionals) conducting an economic activity on a small scale, i.e. reporting an annual value of revenues below €5,000,000. SDS is used by the RA to compute each firm's presumptive revenues, which are then used to determine the audit probability function by comparing presumptive with reported revenues. To describe it, we first focus on the derivation of presumptive revenues for each business and then on the characterisation of the audit probability function.

The RA collects information on structural variables (e.g. the size of offices and warehouses, location, the type of market and of clientele, and main characteristics of customers and providers, among other things) and on accounting variables (e.g. inputs and costs). First of all, the RA divides taxpayers into C clusters on the basis of structural and less manipulable variables. Each taxpayer is allocated to a specific cluster. Then, the RA selects within each cluster $c = \{1, 2, \ldots, C\}$ the group of taxpayers that it believes to be normal, $N_c \subseteq I_c$, in year t, where I_c is the subgroup of the total population I belonging to cluster c, where $\cup I_c = I$. Hence, it estimates c relationships as follows:

$$R_{c,i,t-3} = \beta'_{c,t-3}\boldsymbol{x}_{c,i,t-3} + \epsilon_{c,i,t-3} \tag{1}$$

where $R_{c,i,t-3}$ is the value of revenues reported by i at time $t-3$, $i \in \{1, \ldots, N_c\}$. In each cluster there are J relevant inputs, $\boldsymbol{x}_{c,i,t-3}$ is the $J \times N_c$ matrix of inputs at time $t-3$, and $\epsilon_{c,i,t-3}$ is an idiosyncratic error of i, belonging to cluster c in period $t-3$, respectively. $\beta_{c,t-3}$ is the $J \times 1$ vector of unknown productivity parameters for cluster c, which, once estimated by using standard regression techniques, is denoted as $\hat{\beta}_{c,t-3}$. Finally, the RA defines the $J \times R_c$ vector of productivity parameter coefficients at time t as

$$\boldsymbol{b}_{c,t} := \hat{\beta}_{c,t-3}$$

There are two institutional features that are particularly important, and that we shall recall when interpreting results in the subsection 'Indirect effects as evidence of strategic responses'. First, the entire process is completely transparent and relevant parameters for the calculation of presumptive revenues are known to taxpayers when they issue their tax reports. In particular, taxpayers are provided with a freely downloadable software, called Ge.ri.co, which shows the value of each element of $\boldsymbol{b}_{c,t}$. Although the productivity vector is exogenous to the taxpayers, they are allowed to use it for deciding their own vector of inputs to declare, $\boldsymbol{x}_{c,i,t}$. Second, in general, the components of the productivity vector are positive, so increasing the reported value of a cost or an input that enters the calculation also increases the presumptive revenues. Hence, presumptive revenues for the taxpayer i belonging to the population of active taxpayers in cluster c and tax year t are calculated as

$$\bar{\bar{R}}_{c,i,t} = b'_{c,t}\boldsymbol{x}_{c,i,t}$$

The relationship between $R_{c,i,t}$ and $\bar{\bar{R}}_{c,i,t}$ defines the congruity status of a taxpayer. A taxpayer is said to be incongruous if $R_{c,i,t} < \bar{\bar{R}}_{c,i,t}$ and congruous otherwise. A peculiarity of SDS is that incongruous taxpayers have a higher chance to be audited by the RA and this is known to taxpayers, although the exact probability function is the RA's private information. Without loss of generality, we can write i's perceived probability to be audited as $p_{c,i,t} = p\,(\bar{\bar{R}}_{c,i,t} - R_{c,i,t})$. If a firm is incongruous, $p_{c,i,t} > 0$. Moreover, if a firm is incongruous, the burden to prove that $R_{c,i,t}$ is legitimate is onto the taxpayer and, if occurring, the audit will be based on the amount of the incongruity. If a taxpayer is congruous, the probability of an audit is instead perceived as close to zero.

Typically, taxpayers choose the vector of inputs to declare $\boldsymbol{x}_{c,i}$, and, by using the provided Ge.ri.co software, they assess the corresponding level of presumptive sales ($b'_c\boldsymbol{x}_{c,i}$), which they need to declare to be congruous. At this stage, taxpayers can go back defining a different level of inputs (and costs) to declare and assess how much the presumptive level of sales would change and, originally, this procedure could go on at the taxpayers' will.

The TAL on RCs

In Italy, the RA uses different types of intervention letters to enhance tax compliance. Some of them are strictly related to the implementation of SDS.

In principle, input values and costs that enter the calculation of the presumptive revenues cannot be freely inflated, because, as noted above, by doing so presumptive revenues and thus, ceteris paribus, the probability of being audited, also increase. We shall see below that this does not actually prevent all forms of manipulation, but renders them more sophisticated.

Here we focus on the incentives created by SDS to manipulate costs that do not enter the calculation of presumptive revenues. The idea here is that if a taxpayer cannot overstate a cost that enters the calculation because this would increase the value of presumptive revenues, then he/she can alternatively overstate costs that do not enter this calculation. The ideal candidate for this is a cost category known as RCs. RCs represent a miscellaneous cost category where various types of administrative costs, not immediately directed to production, are included. Because of their nature, until 2010, these costs were not used for the estimation of presumptive revenues and were easy to manipulate.

In May 2009, i.e. some months before issuing their tax reports, taxpayers received a letter from the RA informing them that:

a. the RA deemed the value of RCs they reported in 2008 (i.e. for the tax year 2007) to be 'excessive', taking into account the reported value of revenues and comparing both with values reported by taxpayers belonging to the same cluster;
b. the RA believed that this anomalous value might be due to a strategy of 'false' communication of data by taxpayers subject to SDS;
c. if this anomalous report was repeated in 2009 (i.e. in their tax report for the tax year 2008, to be handed in soon), the taxpayer would certainly be included in a list of taxpayers to be audited.

The letter was originally sent to all taxpayers who, according to the information available to the RA, allegedly overstated RCs in their 2008 tax reports.

In general, firms might receive TALs for only one anomaly, though they could be anomalous in more than one dimension. Firms that had already been audited or received a TAL in previous years were excluded from this letter campaign regardless of their behaviour. The intent of the timing of the letter was clearly to induce recipients to correct spontaneously their reporting behaviour in the 2009 tax reports.

Data description and the empirical strategy

In this paper, we use a unique dataset produced for us by the Italian RA for the analysis of evidence-based TAL campaigns. The dataset is formed by the following three steps: (a) selecting businesses that presented a tax declaration in each tax year between 2006 and 2009, (b) distinguishing between businesses that received the TAL on RCs and those that did not receive any letter, and (c) taking all the observations from the first group and a random sample from the second group.

The dataset is a balanced panel over the period 2006–2009. The treated sample contains 52,782 individual units per year, which account for about 50% of all firms that received the TAL; the control sample is made of 125,231 yearly observations, about 20% of the total firms that received neither a TAL nor an audit over the period considered.

As the only aim of the RA was to maximise tax revenues, reduce input manipulations and increase tax revenues, all firms identified as possibly anomalous according to the RA criteria were sent the letter. This poses a challenge to our estimation of the causal effects of TALs on RCs, as treated firms may have some peculiar characteristics that make them different from controls. This means that not all control units are reliable counterfactuals for treated firms, and we need to find a way to ensure common support for the causal effect estimation of the TALs.

By exploiting the richness of our database, we apply a DD identification strategy, relying on the crucial parallel trend assumption that before treatment, treated and control units have an equal trend, possibly allowing for a difference in levels.

Besides providing standard information on each firm's area of residence (divided into five major areas: North-West, North-East, Centre, South, Islands), the size of its offices and warehouses, and a list of economic variables, which are declared in tax forms, the dataset also provides details of the clusters of each firm's economic activity, which we use extensively. These clusters are defined by the RA to provide a thin partition of each firm's activity, as the number of clusters is 395, accounting for on average 0.25% of the total population of firms.

Lacking an ideal setting, to conduct a reliable DD analysis we perform a set of sample selections. First of all, we drop from our treated sample all firms that received a letter for anomalies different from the one analysed here and all clusters with less than 100 observations, for maintaining a reasonable degree of precision. As the anomaly appeared sometime in 2008, we drop the year 2008 from all the empirical analyses.

Hence, we run parallel trend tests between treated and control units by single clusters of activities and we select only firms belonging to a cluster where the parallel trend assumption is not rejected τ times, where $\tau = \{0.1, 0.2, 0.3, 0.4\}$. Not rejecting the null hypothesis of the equal trend does not mean that the null is true; however, our strategy is motivated to find traces of taxpayers' strategic behaviour by performing a set of sensitivity tests and checking the robustness of the conclusions provided. Nevertheless, selecting increasing levels of τ implies that the estimation sample size decreases, introducing a trade-off between the precision of estimates and the compliance of the parallel trend assumption.

The empirical models we estimate by OLS are standard DD models as such:

$$y_{i,t} = \alpha + \sum_t \beta_t Year_t + \sum_t \delta_t (Year_t \times Treated_i) + \xi_i + \epsilon_{i,t} \tag{2}$$

where $y_{i,t}$ is the (log-) outcome variable, $Year_t$ is a year dummy that controls for common time trends and institutional changes that affect all taxpayers, such as the use of all costs in the calculation of presumptive revenues for 2009. $Treated_i$ is equal to one if firm i is treated and zero otherwise, and $\epsilon_{i,t}$ is an error term. To control for time-invariant observed and unobserved heterogeneity, we estimate all regressions using a business fixed-effect (ξ_i); and to account for the likely correlation of error

terms between firms belonging to the same cluster of economic activities, we cluster standard errors by the cluster of economic activities. The estimated coefficient of the interaction variable ($Year_t \times Treated_i$) is our key coefficient. In particular, δ_{2007} will provide the test of the parallel trend assumption between years 2006 and 2007, with the former as the reference year. If statistically significant, it rejects the assumption that the two-year trend of the outcome variable is the same for treated and control units. The estimation of δ_{2009} provides the effect of the TAL one year after it was sent. These models are estimated for different values of τ over the period 2006–2009, removing the year 2008.

Table 1 shows some descriptive statistics for the (log-transformed) outcome variables used in the next section over the three years considered. In Panel (a) we show statistics for all outcome variables considered, whereas in Panel (b) we show the same statistics for the sample in which we drop clusters where the p-value of the null hypothesis of the parallel trend is smaller than 0.4. It shows that average (log-) RCs are about 1/3 of the taxable income but their variability is large. The variability is also large for depreciation and costs of intermediate goods, which have been addressed in early TALs. The average (log-) presumptive revenues are about four times the size of (log-) taxable income and present relatively small variability, which is a likely consequence of the way they are computed, i.e. by using average productivity prices, as described in the section 'The Italian tax system and SDS'.

The impact of the TAL

Letters were sent to taxpayers suspected of manipulating their RCs in 2008 before the submission of the 2009 tax return. The purpose of these letters was to reduce the RCs and increase taxable income by reducing total deductible costs. This reduction in RCs is the expected direct effect of TALs.

Table 1. Some descriptive statistics of the outcome variables.

Variable	N	Mean	Std. dev.	Min	Max
(a) Full sample					
Residual costs	275,578	0.962	1.770	−6.908	8.332
Taxable income	269,645	3.056	1.018	−6.908	10.001
Depreciation	263,185	1.211	1.554	−6.908	7.689
Labour costs	144,525	3.279	1.811	−6.908	8.283
Costs of final goods	190,458	3.405	2.226	−20.653	9.208
Costs of services	284,028	2.492	1.648	−6.908	8.546
Costs of intermediate goods	142,585	2.141	1.525	−6.908	7.860
Presumptive revenues	292,071	11.426	1.377	0.000	17.126
(b) Sample with $\tau = 0.4$					
Residual costs	90,324	0.835	1.756	−6.908	8.096
Taxable income	168,030	3.057	1.028	−6.908	8.981
Depreciation	121,537	1.196	1.548	−6.908	7.689
Labour costs	81,125	3.294	1.812	−6.908	8.192
Costs of final goods	114,282	3.384	2.214	−19.737	9.208
Costs of services	148,758	2.433	1.615	−6.908	8.499
Costs of intermediate goods	102,837	2.142	1.498	−6.908	7.813
Presumptive revenues	26,978	11.499	1.342	2.398	15.520

Notes: Our calculations on the RA tax data. All outcome variables are log-transformed. The year 2008 is dropped from the analysis.

There may, however, be indirect effects of the letter, as taxpayers could react strategically to contain the increase in the tax payment. Given the evidence-based nature of the TAL, they may find it credible and agree on reducing RCs, nonetheless increasing other costs. Note, however, that the latter would in turn increase presumptive revenues, prompting an increase in the probability of being audited by the RA, which holds reported revenues constant.

Direct effects of the TAL on RCs and taxable income

Figures 1 and 2 show the trend of (log-) RCs and of (log-) taxable income for $\tau = 0.4$, with the 95% confidence intervals depicted as shaded areas around the pointwise coefficient estimates. The visual inspection clearly suggests that for both variables the trend between control and treated units was very similar before the TAL was sent. After 2008, RCs largely decreased for treated firms whereas they stayed rather stable for control units. On the other hand, taxable income increased between 2007 and 2009 for treated units, whereas they remained about stable for controls.

In Table 2 we present the estimation of Equation (2), where the outcome variable, $y_{i,t}$, is the log of RCs. We present results for different levels of τ to assess the robustness of the results. It shows that, on average, treated firms report a higher level of RCs than control firms. The coefficient of $Treated \times Year_{2007}$ is never significant, suggesting that there is no evidence to reject the parallel trend assumption in the years 2006 and 2007. The coefficient of $Treated \times Year_{2009}$ is negative and statistically significant providing a rather stable estimate of 0.42–0.48 log-point in RCs with respect to the year 2006, which is our reference year, suggesting a direct effect of the TAL on RCs.

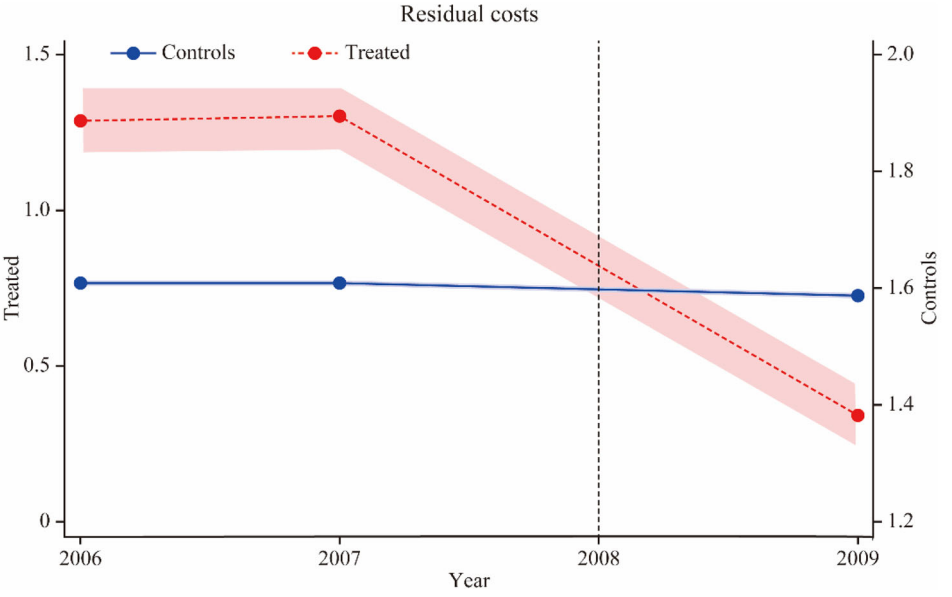

Figure 1. Trends of RCs, $\tau = 0.4$.
Notes: All variables are log-transformed. Shaded bands show the 95% confidence intervals.

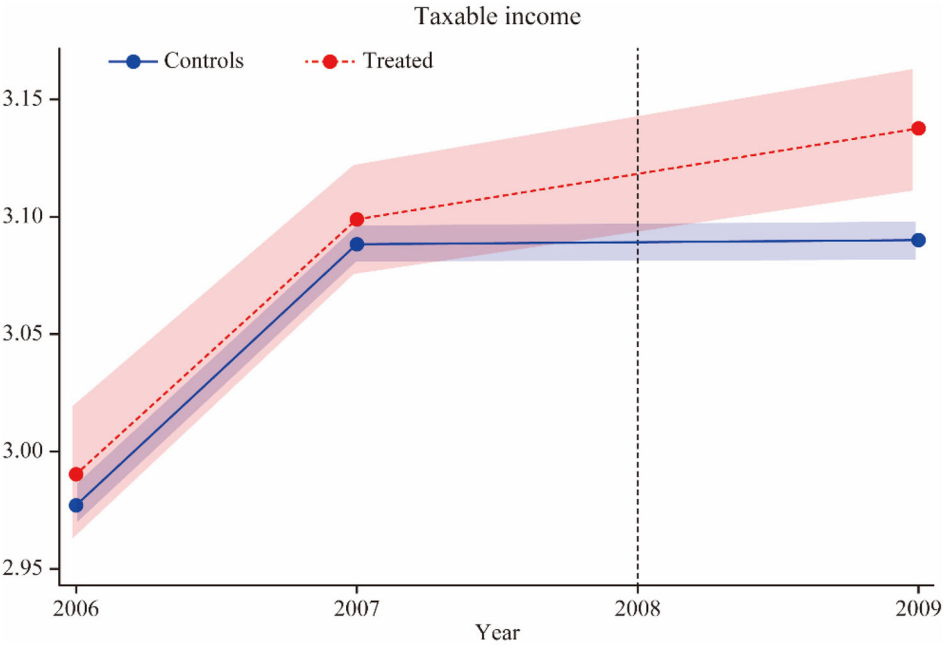

Figure 2. Trends of taxable income, $\tau = 0.4$.
Notes: All variables are log-transformed. Shaded bands show the 95% confidence intervals.

Table 2. Difference-in-differences model estimation: Log of RCs.

	$\tau = 0.3$	$\tau = 0.4$	$\tau = 0.5$
$Year_{2007}$	0.010	0.010	0.026**
	(0.009)	(0.009)	(0.012)
$Year_{2009}$	−0.038***	−0.040***	−0.025**
	(0.009)	(0.009)	(0.012)
$Treated \times Year_{2007}$	0.044	0.036	0.016
	(0.029)	(0.030)	(0.041)
$Treated \times Year_{2009}$	−0.415***	−0.427***	−0.483***
	(0.030)	(0.030)	(0.041)
Constant	0.867***	0.857***	0.827***
	(0.006)	(0.006)	(0.008)
R-squared	0.795	0.794	0.787
N	92,423	90,324	53,812

Notes: Our calculations on the RA tax data. Standard errors in parentheses. **p-value < 0.05 and ***p-value < 0.01.

Table 3 presents the estimation of Equation (2) where the outcome variable is the (log-) taxable income, for different values of τ. This table also shows that the parallel trend assumption in the years 2006 and 2007 is not rejected and the direct effect of the 2009 TAL on RCs is as expected by the RA, i.e. producing an increase in taxable income. The effects of the TAL on taxable income are estimated between 3.1 and 3.8 percentage points.

Indirect effects as evidence of strategic responses

Although TALs reduce the manipulation of RCs and taxable income, some alternative manipulations may have been introduced. To examine this issue, we need to complete

Table 3. Difference-in-differences model estimation: Log of taxable income.

	$\tau = 0.3$	$\tau = 0.4$	$\tau = 0.5$
$Year_{2007}$	0.151***	0.149***	0.151***
	(0.003)	(0.004)	(0.004)
$Year_{2009}$	0.143***	0.147***	0.144***
	(0.003)	(0.004)	(0.004)
$Treated \times Year_{2007}$	–0.007	–0.001	0.005
	(0.010)	(0.012)	(0.013)
$Treated \times Year_{2009}$	0.031***	0.031**	0.038***
	(0.010)	(0.012)	(0.013)
Constant	2.955***	2.953***	2.944***
	(0.002)	(0.002)	(0.003)
R-squared	0.821	0.820	0.823
N	211,163	168,030	127,922

Notes: Our calculations on the RA tax data. Standard errors in parentheses. **p-value < 0.05 and ***p-value < 0.01.

Table 4. Types of costs, SDS-relevance and TALs, until 2010.

Cost item	Relevant to SDS	Addressed by TALs
Depreciation	Yes	Yes
Labour costs	Yes	No
Costs of intermediate goods	Yes	Yes
Costs of services	Yes	No
Costs of final goods	Yes	No

the analysis of the institutional context. Under SDS, inflating costs that directly or indirectly enter (i.e. via input values) the calculation of presumptive revenues increase the latter, because productivity parameters are generally positive, which, ceteris paribus, increases the probability of an audit. On the other hand, underreporting these costs, which are generally deductible from the tax base, increases taxable income. These institutional arrangements seem to design a proper system of counter incentives to the manipulation of costs different from RCs.

However, it should be recalled that productivity parameters are known to taxpayers. Therefore, taxpayers could select the inputs and costs whose impact on presumptive revenues is small, or that can be counterbalanced by alternative manipulations. Evidence of these manipulations is abundant in the history of the implementation of SDS. For instance, Santoro and Fiorio (2011) show that reported revenues were concentrated at the presumptive level within the very first years of the SDS implementation. Given that taxpayers could use the software Ge.ri.co for an unlimited number of attempts before issuing their final report, they are able to optimise the balance between presumptive revenues (and thus the probability of being audited) and taxable income.

The Italian tax authority reacts to this evidence of manipulation by sending more TALs. Their structure and wording are very similar to the RC-TALs described above, but, in this case, it is purported to disincentive the manipulation of some of the costs that are included in the calculation of presumptive revenues, as shown in Table 4.

In Figure 3, we show the trend of all remaining variables, for $\tau = 0.4$, and once again we visually confirm the parallel trend assumption, which is crucial for our identification strategy. In Table 5 we provide the corresponding estimations, confirming

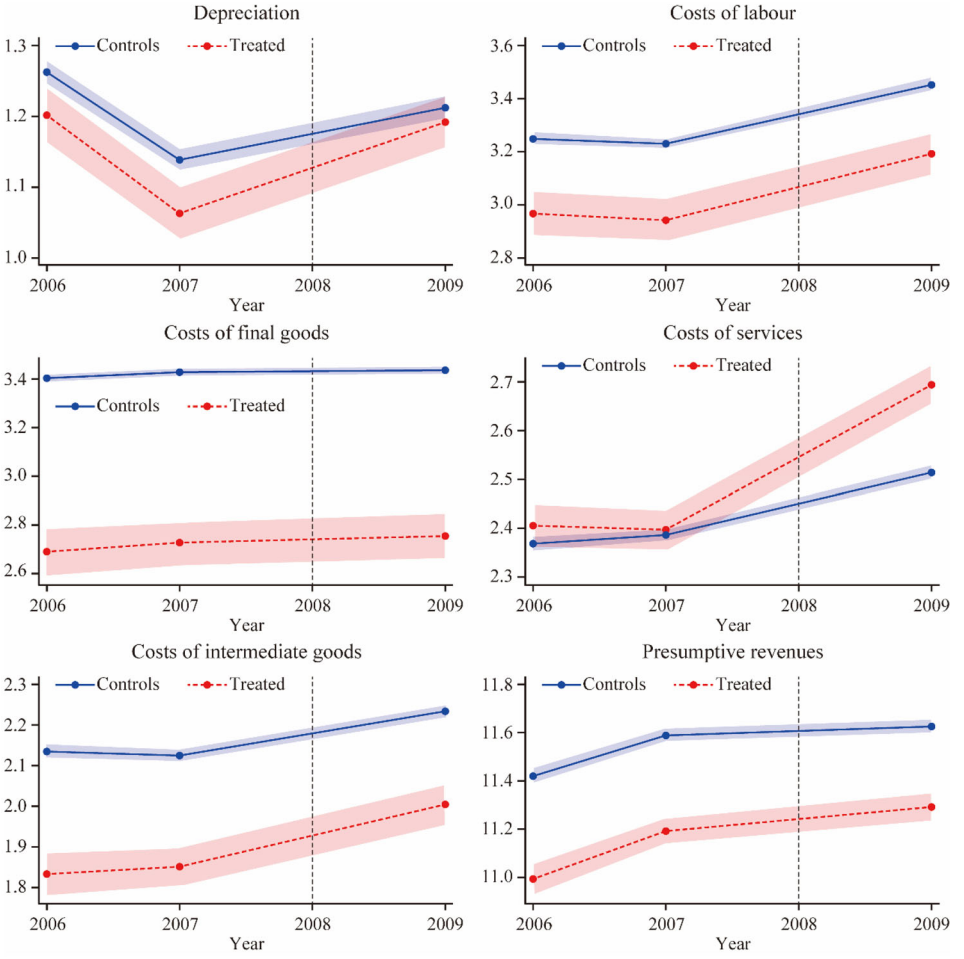

Figure 3. Trends for remaining outcome variables, $\tau = 0.4$.
Notes: All variables are log-transformed. Shaded bands show the 95% confidence intervals.

Table 5. Strategic responses, $\tau = 0.4$.

	Depreciation $\tau = 0.4$	Labour costs $\tau = 0.4$	Costs of final goods $\tau = 0.4$	Costs of services $\tau = 0.4$	Costs of intermediate goods $\tau = 0.4$	Presumptive revenues $\tau = 0.4$
$Year_{2007}$	−0.121***	0.030***	0.059***	0.024***	−0.009*	0.211***
	(0.005)	(0.006)	(0.007)	(0.004)	(0.005)	(0.007)
$Year_{2009}$	−0.074***	0.195***	0.058***	0.151***	0.080***	0.245***
	(0.005)	(0.006)	(0.007)	(0.004)	(0.006)	(0.007)
$Treated \times Year_{2007}$	0.016	−0.003	−0.006	0.004	−0.021	0.021
	(0.016)	(0.024)	(0.030)	(0.015)	(0.019)	(0.019)
$Treated \times Year_{2009}$	0.081***	0.025	−0.033	0.191***	0.034*	0.086***
	(0.016)	(0.024)	(0.030)	(0.014)	(0.020)	(0.019)
Constant	1.260***	3.216***	3.344***	2.366***	2.117***	11.335***
	(0.004)	(0.004)	(0.005)	(0.003)	(0.004)	(0.005)
R-squared	0.874	0.911	0.902	0.902	0.881	0.937
N	121,537	81,125	114,282	148,758	102,837	26,978

Notes: Our calculations on the RA tax data. Standard errors in parentheses. *$p < 0.1$, and ***$p < 0.01$. First difference of (log-) costs and (log-) presumptive revenues. Difference-in-differences model estimation removing firms belonging to a cluster where the parallel trend assumption is rejected. $\tau = 0.4$.

Table 6. Strategic responses, $\tau = 0.3$.

	Depreciation $\tau = 0.3$	Labour costs $\tau = 0.3$	Costs of final goods $\tau = 0.3$	Costs of services $\tau = 0.3$	Costs of intermediate goods $\tau = 0.3$	Presumptive revenues $\tau = 0.3$
$Year_{2007}$	–0.123***	0.028***	0.057***	0.022***	–0.011**	0.239***
	(0.005)	(0.005)	(0.006)	(0.004)	(0.005)	(0.005)
$Year_{2009}$	–0.070***	0.193***	0.061***	0.148***	0.078***	0.257***
	(0.005)	(0.005)	(0.006)	(0.004)	(0.005)	(0.005)
$Treated \times Year_{2007}$	0.017	–0.006	0.009	0.007	–0.018	0.012
	(0.015)	(0.021)	(0.026)	(0.014)	(0.018)	(0.016)
$Treated \times Year_{2009}$	0.073***	0.015	–0.026	0.187***	0.025	0.104***
	(0.015)	(0.021)	(0.026)	(0.014)	(0.018)	(0.016)
Constant	1.308***	3.192***	3.363***	2.369***	2.124***	11.256***
	(0.003)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)
R-squared	0.876	0.912	0.901	0.904	0.883	0.906
N	145,751	107,932	139,866	164,173	117,024	72,141

Notes: Our calculations on the RA tax data. Standard errors in parentheses. $^{**}p < 0.05$, and $^{***}p < 0.01$. First difference of (log-) costs and (log-) presumptive revenues. Difference-in-differences model estimation removing firms belonging to a cluster where the parallel trend assumption is rejected. $\tau = 0.3$.

Table 7. Strategic responses, $\tau = 0.5$.

	Depreciation $\tau = 0.5$	Labour costs $\tau = 0.5$	Cost of final goods $\tau = 0.5$	Costs of services $\tau = 0.5$	Cost of intermediate goods $\tau = 0.5$	Presumptive revenues $\tau = 0.5$
$Year_{2007}$	–0.121***	0.033***	0.057***	0.027***	–0.009	0.228***
	(0.005)	(0.007)	(0.007)	(0.005)	(0.006)	(0.008)
$Year_{2009}$	–0.074***	0.196***	0.059***	0.153***	0.082***	0.260***
	(0.005)	(0.007)	(0.007)	(0.005)	(0.006)	(0.008)
$Treated \times Year_{2007}$	0.012	–0.007	–0.009	0.000	–0.013	0.024
	(0.016)	(0.027)	(0.030)	(0.015)	(0.021)	(0.023)
$Treated \times Year_{2009}$	0.074***	0.035	–0.048	0.186***	0.038*	0.108***
	(0.016)	(0.028)	(0.030)	(0.015)	(0.021)	(0.023)
Constant	1.257***	3.197***	3.342***	2.378***	2.138***	11.231***
	(0.004)	(0.005)	(0.005)	(0.003)	(0.004)	(0.006)
R-squared	0.874	0.910	0.903	0.906	0.882	0.929
N	119,232	65,787	110,643	118,409	87,077	21,524

Notes: Our calculations on the RA tax data. Standard errors in parentheses. $^{*}p < 0.1$, and $^{***}p < 0.01$. First difference of (log-) costs and (log-) presumptive revenues. Difference-in-differences model estimation removing firms belonging to a cluster where the parallel trend assumption is rejected. $\tau = 0.5$.

that the parallel trend assumption is never rejected for all variables when $\tau = 0.4$. In Tables 6 and 7, we provide some robustness checks of these results when $\tau = 0.3$ and $\tau = 0.5$, respectively. Focussing on the coefficients of $Treated \times Year_{2009}$, some interesting results emerge. First, the letter on RCs has no effect on labour and final goods costs, but it causes an increase in the costs of services and intermediate goods consumed by the firm. Recalling Table 4, an explanation comes from the fact that no TALs on the costs of services or on the costs of final goods have ever been sent, and taxpayers react by inflating them to limit the increase in taxable income without increasing the probability of an audit. These costs have been perceived as freely manipulable, though being related to inputs that enter SDS and presumptive revenue calculations. Note also that, unsurprisingly, strategic taxpayers do not react by varying the number of workers because this is a very costly activity and a crime.

It is also found that presumptive revenues also increased in 2009 with respect to 2006. This result indicates that, although the TAL does prompt a strategic response, this response does not go unnoticed, and the RA could, in theory, react by auditing businesses that are inflating alternative costs.

Concluding remarks

Recent literature has emphasised that audits do have a positive impact on subsequent compliance although the magnitude and duration of the effect depend on the type of income and taxpayers targeted as well as on the outcome of the audit. Audits tend to increase future compliance when taxpayers are found to have made errors (Advani, Elming, and Shaw 2021; Mazzolini, Pagani, and Santoro 2022). Also, audited taxpayers reporting self-employment income tend to increase future reporting more but for a shorter period with respect to other taxpayers (Advani, Elming, and Shaw 2021). In general, the average net impact of audits is estimated to be higher than the administrative costs of the audits and this, in turn, justifies the increase in the number of audits from a revenue-maximising perspective. Although one should take also compliance costs into account from a welfarist perspective, the message coming from these studies is that audits are a valuable enforcement tool.

The evidence on intervention letters is somewhat more mixed. When examining the literature, we argue that when the letter conveys only the threat of the audit without reporting the source of the information that motivates such a threat or without reinforcing the message with some additional information that is valuable for the taxpayer, it tends to have a limited impact.

Along these lines, the possibility to replace costly audits with low-cost intervention letters has been explicitly questioned by Advani, Elming, and Shaw (2021). They look, in particular, at the information that the enforcement provides to the tax authority about the magnitude of taxpayers' income at a point in time. This snapshot can be used by the tax authority in the years following the audit because, for income that is more stable, large deviations in the short term are implausible. In turn, taxpayers having stable income perceive that the tax authority has gained this piece of information, and they modify their behaviour accordingly. On the contrary, Advani, Elming, and Shaw (2021) stress that the letters do not provide such a snapshot and therefore cannot induce a behavioural change.

However, this criticism applies to basic TALs. In this paper, we study a TAL that along with its direct effect on the targeted behaviour (the manipulation of RCs), it implicitly conveys the information that alternative forms of manipulation are not under examination. This prompts an undesired secondary effect, namely the manipulations of depreciation, service costs and intermediate goods costs.

In our case, after carefully defining the dataset to ensure the full respect of the parallel trend condition, we show that the direct effect overweighs the indirect effect, thus generating an increase in taxable income. Therefore, in the case we study, the letter on the whole increases revenues at a low cost.

Nevertheless, the evidence of cost manipulations within SDS is present even in this, overall satisfactory, policy. This reinforces the evidence already existing on this matter (Santoro and Fiorio 2011).

This evidence has contributed to motivating a policy change from 2019 onwards. SDS has been replaced by a new approach based on a compliance score known as ISA (*Indicatori Sintetici di Affidabilità*, or synthetic compliance indicators). There are two basic differences between SDS and ISA. First, the score in ISA depends not only on the level of turnover reported but also on the value of taxable income. This clearly limits the scope of cost manipulations of the type we examine in this paper. Second, ISA also provides, along with a threat of auditing for low scores, rewards for high scores. Aggregate results seem to suggest that this change of approach has increased average tax compliance, but detailed analyses are still ongoing.

Acknowledgements

The authors would like to acknowledge the Revenue Agency of Italy for providing the dataset used in this study.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Alessandro Santoro http://orcid.org/0000-0002-4646-3583

References

Advani, Arun, William Elming, and Jonathan Shaw. 2021. "The Dynamic Effects of Tax Audits." *The Review of Economics and Statistics* 2021 (September): 1–45.

Bérgolo, Marcelo L., Rodrigo Ceni, Guillermo Cruces, Matias Giaccobasso, and Ricardo Perez-Truglia. 2023. "Tax Audits as Scarecrows: Evidence from a Large-scale Field Experiment." *American Economic Journal: Economic Policy* 15 (1): 110–153.

Carrillo, Paul, Dina Pomeranz, and Monica Singhal. 2017. "Dodging the Taxman: Firm Misreporting and Limits to Tax Enforcement." *American Economic Journal: Applied Economics* 9 (2): 144–164.

De Neve, Jan-Emmanuel, Clement Imbert, Johannes Spinnewijn, Teodora Tsankova, and Maarten Luts. 2021. "How to Improve Tax Compliance? Evidence from Population-wide Experiments in Belgium." *Journal of Political Economy* 129 (5): 1425–1463.

Kleven, Henrik Jacobsen, Martin B. Knudsen, Claus Thustrup Kreiner, Søren Pedersen, and Emmanuel Saez. 2011. "Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark." *Econometrica* 79 (3): 651–692.

Mazzolini, Gabriele, Laura Pagani, and Alessandro Santoro. 2022. "The Deterrence Effect of Real-world Operational Tax Audits on Self-employed Taxpayers: Evidence from Italy." *International Tax and Public Finance* 2022 (29): 1014–1046.

Meiselman, Ben S. 2018. "Ghostbusting in Detroit: Evidence on Nonfilers from a Controlled Field Experiment." *Journal of Public Economics* 158: 180–193.

Ortega, Daniel, and Carlos Scartascini. 2020. "Don't Blame the Messenger. The Delivery Method of a Message Matters." *Journal of Economic Behavior & Organization* 170: 286–300.

Santoro, Alessandro, and Carlo V. Fiorio. 2011. "Taxpayer Behavior When Audit Rules Are Known: Evidence from Italy." *Public Finance Review* 39 (1): 103–123.

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

Slemrod, Joel, Marsha Blumenthal, and Charles Christian. 2001. "Taxpayer Response to an Increased Probability of Audit: Evidence from a Controlled Experiment in Minnesota." *Journal of Public Economics* 79 (3): 455–483.

Slemrod, Joel, Brett Collins, Jeffrey Hoopes, Daniel Reck, and Michael Sebastiani. 2017. "Does Credit-card Information Reporting Improve Small-business Tax Compliance?" *Journal of Public Economics* 149: 1–19.