



DEPARTMENT OF ECONOMICS,
MANAGEMENT AND STATISTICS
UNIVERSITY OF MILAN – BICOCCA

DEMS WORKING PAPER SERIES

**Evidence-based threat-of-audit letters:
do taxpayers respond strategically in a complex
environment?**

Carlo Fiorio, Alessandro Santoro

No. 372 – September 2017

Dipartimento di Economia, Metodi Quantitativi e Strategie di Impresa
Università degli Studi di Milano - Bicocca

<http://dems.unimib.it/>

Evidence-based threat-of-audit letters:
do taxpayers respond strategically in a complex
environment?

Carlo V. Fiorio*

University of Milan, Irvapp-FBK
and Dondena Centre

Alessandro Santoro†

University of Milan-Bicocca
and Dondena Centre

Abstract

We study the impact on tax compliance of a letter sent by the Italian Revenue Agency in 2009 to taxpayers suspected to have overreported some costs in order to decrease their taxable income without increasing the probability to be audited. We find strong and robust evidence that the letter did reduce manipulation of targeted costs, but that the impact on taxable income was much smaller. This difference is likely due to a strategic response by the taxpayers who increased other costs that they perceived were not targeted by the Revenue Agency and were able to do so without increasing the probability to be audited. This shows that strategic responses are likely to appear also in an institutionally complex environment.

JEL Numbers: H26; C55

Keywords: Tax Compliance by Businesses; Threat-of-audit letters.

*University of Milan, Department of Economics via Conservatorio 7, 20122, Milano; email: carlo.fiorio@unimi.it.

†Corresponding author. University of Milan-Bicocca, Department of Economics piazza Ateneo Nuovo, 1, 20126, Milano; email: alessandro.santoro@unimib.it.

1 Introduction

Firms can evade taxes either by underreporting revenues and/or by overreporting costs. However, the risk of being audited is likely to differ across these two reporting behaviours. This happens since tax authorities often have different information on these items. Information about revenues of a firm can be obtained from its customers, either from their accounting books if the transaction is B2B or from their consumption records or credit cards records if the transaction is B2C. On the other hand, information on some costs is more difficult to be obtained from third parties. While the amount of some costs can be inferred to a limited extent from the revenues reported by the seller of services or goods, there are cost items that do not correspond to a precise transaction and do not have an exact counterpart.

The pieces of information available to the Revenue Agency (RA) can be used to send taxpayers *evidence-based* threat-of-audit letters (TALs). These can be described as letters “including messages that convey that the tax authority has personalized information that suggests that the taxpayer has been noncompliant” (Slemrod, 2017) and also containing an explicit audit threat if the observed behaviour did not change in the future. Evidence-based TALs are different from TALs that simply state the probability of a given tax return to be examined (such as those studied by Slemrod et al., 2001; Kleven et al., 2011) because they are based on some information which is disclosed to the taxpayer. Providing evidence to the taxpayer of her own noncompliance can enhance the effectiveness of the letter, but it might also have some countereffect, since an evidence-based TAL makes the taxpayer believe that the RA possesses only partial information. This can trigger a strategic response by the taxpayer on items that she believes remain unchecked by the RA. The evidence provided by Carrillo et al. (2016) goes towards this

direction: taxpayers become more compliant with reporting items they know are targeted by the tax authorities but offset this by becoming less compliant with unobserved reports.

The same pattern is exhibited by small businesses who know that third-party information is available to the tax authorities on revenues but not on expenses (Slemrod et al., 2017).

In Italy, small businesses pay taxes on the differences between reported revenues and costs and are subject to an auditing method known as *Studi di Settore* (business sector studies, henceforth SDS). By using SDS, RA calculates the individual presumptive value of revenues on the basis of the values of some inputs and costs reported by the taxpayer. For every given revenue report, the probability of an audit increases in the difference between presumptive and reported revenues.

The RA discloses the level of presumptive revenues to the taxpayer and allows her to modify her level of inputs and costs before finally submitting the tax form. The taxpayer also knows that the probability of being audited increases the smaller is the reported as opposed to the presumptive revenues. In general, underreporting of revenues (below the presumptive level) and overreporting of some inputs for a given report of revenues increases the probability of an audit. As inputs are related to costs for accounting reasons, overreporting costs holding reported revenues constant for reducing taxable income can, in some cases, increase the level of presumptive revenues, hence increasing the probability of an audit. This is a built-in disincentive to overreport costs.

However, until 2009 not all cost reports were used to estimate presumptive revenues. In particular an item known as *residual costs* (RC) was not used to estimate presumptive revenues. Residual costs do not refer to a single

transactions and do not have exact counterparts. It is a cost category that contains a miscellaneous of costs for which third-party information is generally unavailable.

During the first year of SDS application it clearly emerged that taxpayers reacted strategically on several dimensions. For instance, Santoro and Fiorio (2011) showed that reported revenues concentrated at the presumptive level within the very first years of SDS implementation. The RA reacted in various ways including with an extensive evidence-based TAL campaign, which started in 2005 and focused on some input reports. After this TAL campaign, the RA observed a change of inputs and costs reports but also a large increase in the level of RC declared.

Figure 1 provides a visual description of the ratio of residual costs relative to taxable income over time. The median and the average incidence of RC were relatively stable in 2006 and 2007 and they both largely increased in 2008. The large increase of average incidence is also an effect of an increased frequency of very large incidence values (i.e. outliers).

This prompted the RA reaction. In 2009, all costs (including RC) were used to estimate presumptive revenues and a new evidence-based TAL addressing anomalous level of RC reported in 2008, was sent before taxpayers filled in their 2009 reports. In this paper, we study the response of firms to the evidence-based TAL addressing RC that was first introduced between 2008 and 2009. Figure 1 suggests that the letter was effective in reducing RC.

Our main research question is to investigate whether taxpayers react strategically to an evidence-based policy. Recently, Carrillo et al. (2016) show that a TAL sent to Ecuadorian firms warning them on the level of their reported revenues induced them to increase revenues but overreport costs to contain the increase in tax payment. This reaction was explained by observing

that overreporting costs is more difficult to detect than underreporting of revenues due to different availability of third-party information. Similarly, Slemrod et al. (2017), use data from Form 1099-K, an information report introduced in 2011 which provides the Internal Revenue Service with information about credit card sales. They identify a set of taxpayers especially sensitive to the new policy, who report receipts equal to or slightly exceeding the receipts reported on 1099-K. However, these taxpayers largely offset increased reported receipts with increased reported expenses, which do not face information reporting, diminishing the impact on reported taxable income.

In both of these papers taxpayers' response is simple: an increase in revenues is offset by an increase in costs. Here we focus on taxpayers' behaviour in a complex environment, such as the one created by SDS, where taxpayers have to take into account *both* the level of taxable income and the probability of receiving an audit when choosing which costs to manipulate. Our paper is the first which investigates taxpayers' *strategic* response to evidence-based policies in a complex environment.

The paper is organized as follows. Sections 2 and 3 describe the institutional background. Section 4 illustrates our dataset and our empirical strategy. Section 5 contains our main results, which we discuss by distinguishing between *direct* results on variables targeted by the TALs, i.e. residual costs and taxable income, and *indirect* results on other costs and on presumptive revenues. These results show that taxpayers' respond strategically also in a complex environment. We analyze the implications of these findings in the context of the literature in Section 6.

2 Italian Tax System and SDS

In Italy, not incorporated small businesses (sole proprietorships, joint ownerships, unlimited liability companies) are taxed on the difference between reported revenues and costs. 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 revenue below €5,000,000. SDS are 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 ¹ To describe it, we first focus on the derivation of presumptive revenues for each business and then on the characterization of the audit probability function.

The RA collects information on structural variables (e.g., size of offices and warehouses, the location, type of market and of clientele, main characteristics of customers and providers, etc.) and on accounting variables (inputs and costs).

First of all, the RA divide taxpayers into C clusters on the basis of structural and less manipulable variables. Each taxpayer is allocated to a specific cluster. Then, RA selects within each cluster $c = \{1, 2, \dots, C\}$ 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:

$$R_{c,i,t-3} = \beta'_{c,t-3} \mathbf{x}_{c,i,t-3} + \epsilon_{c,i,t-3} \quad (1)$$

where $R_{c,i,t-3}$ is the value of revenues reported by i at time $t-3$, $i \in \{1, \dots, N_c\}$, $\mathbf{x}_{c,i,t-3}$ 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 $J \times 1$

¹For a more detailed description and analysis of SDS, see Santoro and Fiorio (2011) and Santoro (2008).

vector of unknown productivity parameters for cluster c , which – once estimated by using standard regression techniques – is denoted $\widehat{\beta}_{c,t-3}$. Finally, the RA defines the $J \times R_c$ vector of productivity parameters coefficient at time t as

$$\mathbf{b}_{c,t} := \widehat{\beta}_{c,t-3}.$$

Taxpayers are provided with a freely downloadable software, called Ge.ri.co, providing the value of each element of \mathbf{b}_{ct} . Although the productivity vector is exogenous to the taxpayer, she is allowed to use it for deciding her own vector of inputs to declare, \mathbf{x}_{cit} . Until 2009, this vector usually included:

1. the value of the stock of capital goods, which is related to the cost for depreciation;
2. the number of workers employed in the business, which is related to labour costs;
3. the variation in inventories, i.e. the difference between the final and the initial value of inventories, which is related to the cost for intermediate goods;
4. the cost for services;
5. the cost for final goods used by the firm.

Since 2009, all costs, including residual costs were included, in the vector of inputs.

Hence, presumptive revenues for the taxpayer i belonging to the population of active taxpayers in cluster c and tax year t are calculated as

$$\overline{\overline{R}}_{cit} = \mathbf{b}'_{ct} \mathbf{x}_{cit}.$$

The relationship between R_{cit} and $\overline{\overline{R}}_{cit}$ defines the *congruity status* of the taxpayer. A taxpayer is said to be incongruous if $R_{cit} < \overline{\overline{R}}_{cit}$ and congruous otherwise.

A peculiarity of SDS is that incongruous taxpayer have a higher chance to be audited by the RA and this is know to taxpayers, although the exact probability function is RA's private information. Without loss of generality, we can write i 's perceived probability to be audited as $p_{cit} = p(\overline{\overline{R}}_{cit} - R_{cit})$. If a firm is incongruous, $p_{cit} > 0$. Moreover, if a firm is incongruous, the burden to prove that R_{cit} is a legitimate report is onto the taxpayer and, if occuring, the audit will be based on the amount of the incongruity. If a taxpayer is congruous her probability to be audited is instead perceived as close to zero.

Typically, taxpayers choose the vector of inputs to declare \mathbf{x}_{ci} , and, by using the provided Ge.ri.co software, they assess the corresponding level of presumptive sales ($\mathbf{b}'_c \mathbf{x}_{ci}$), 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 taxpayer's will.

3 Types of costs and threat-of-audit letters

The main reason for SDS to disclose information to the taxpayer was to establish a collaboration with her and increase tax revenues. However, the structure of SDS also introduced new opportunities of strategic behaviour by taxpayers.

The TALs policy put in place by the Italian RA since 2005 was addressed to taxpayers who allegedly manipulated capital goods (and thus depreciation costs) and variation in inventories (and thus costs for intermediate goods). No letter was ever sent for the manipulation of the number of workers, nor for the cost for services and for final goods, because labour audits is competence of the

Ministry of Labour and because budget constraints might have also played a role in the estimation of presumptive revenues. TALs on residual costs (RC) were sent only starting from the beginning of 2009. RC are a miscellaneous cost category where various types of administrative costs, not immediately directed to production, are included. Because of their nature, until 2009, these costs were not used for the estimation of presumptive revenues. Thus overreporting RC decreased taxable income without increasing the probability to be audited.

Hence not all costs that enter the determination of taxable income were relevant for SDS, as neither them nor the inputs that generated them entered the calculation of presumptive revenues. Moreover, some of them were addressed by TALs since 2005, other since 2009, others never have. Table 1 summarises these differences among deductible costs.

In this paper, we observe letters sent for tax year 2009 only. The distinguishing feature of this letter campaign was that, for the first time, TALs were sent also to taxpayers suspected to have manipulated residual costs, which do not enter SDS and the calculation of presumptive revenues.

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

- (a) the value of RC they reported in 2008 (i.e for tax year 2007) was deemed 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 believes that this anomalous value may 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 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 RC in their 2008 tax reports. The letter did not mention that a decrease in residual costs in 2009 would have, *coeteris paribus*, reduced presumptive revenues in 2009 because from that year all costs entered in the presumptive revenues calculation. We discuss the importance of this feature of the letter with respect to the strategic response put in place by taxpayers in Section 5.2.

In general, firms can receive TALs for only one anomaly, though firms can be anomalous on more than one dimension. Firms that had already been audited or that 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 2009 tax reports.

4 Data description and empirical strategy

In this paper we use a data set produced by the Italian RA for the analysis of evidence-based TALs campaigns. It is made of a random sample of firms treated with a TAL and of a random sample of firms that never received any TAL nor a tax audit, and serve as control sample. The dataset is a balanced panel over the period 2006-2009. The treated sample counts 52,782 individual units per year, which account to about 50% of firms; the control sample is made of 125,231 yearly observations, about 20% of the total population of firms that never received a TAL nor an audit over the period considered.

As the only aim of the RA was to maximise tax revenues, reducing input manipulation and increasing tax revenues, all firms identified as possibly anomalous according to RA criteria were sent the letter. This poses a challenge to our estimation of the causal effects of TALs on residual costs.

Our empirical strategy is based on a difference-in-difference (DD) approach, which relies on the crucial parallel trend assumption, i.e. the assumption that before treatment treated and control units have an equal trend, possibly allowing for a difference in levels. To assure that this assumption is valid we exploited the richness of our database.

Beside providing standard information on each firm's area of residence (divided five major areas, North-West, North-East, Center, South, Islands), size of offices and warehouse and a list of economic variables, which are declared in tax forms, it also provides detail of the cluster of activity, which we use extensively. These clusters are defined by the RA for SDS definition and provide a thin partition of each firms activity, as the number of clusters is 395, containing on average 0.25% of the total population of firms.

We drop from our treated sample all firms that received a letter for other types of anomalies and all clusters with less of 100 observations, for maintaining a reasonable degree of precision.

Hence, we run parallel trend tests by cluster of activities and we removed from the final estimation sample all those firms belonging to a cluster where the parallel trend assumption was rejected 50% of times.

Hence, we estimated by OLS the model:

$$y_{it} = \alpha + \sum_t \beta_t Year_t + \gamma Treated_i + \sum_t \delta_t (Year_t \times Treated_i) + Z'_{it} \zeta_t + \epsilon_{it}, \quad (2)$$

where y_{it} is the outcome variable, $Year_t$ is a year dummy which 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 one if individual i is treated and zero otherwise, $(Year_t \times Treated_i)$ is the year dummy interacted with the treatment dummy,

and ϵ_{it} is an error term. To control for remaining heterogeneity we also introduced a large set of interactions, included in the matrix Z_{it} , encompassing interactions between cluster and year dummies to control for cluster specific trend, between cluster and region of residence dummies, for cluster specific effect of regions, and between cluster dummies and surface size of the firms, for average size of firms by cluster. Saturating the main equation by the largest possible set of interactions of exogenous variables is methodologically similar to an exact matching procedure (Iacus et al., 2011).

Among coefficients to be estimated ($\alpha, \beta_t, \delta_t$ and ζ_t for $t = 2007, 2008, \dots, 2011$), the main interest is on δ_{2009} , which shows the effect of the TAL policy on the treated. However, also the coefficient δ_{2007} is of interest as it tests the parallel trend assumption of the outcome variable in years 2006 and 2007.

Table 2 shows some descriptive statistics for the (log-transformed) outcome variables used in the next section over the whole period. It shows that log-RC are about one third of taxable income but its standard error is larger. Variability is also large for depreciation costs and intermediate goods, which have been addressed in early TALs. The average log-presumptive revenues are about four times the size of log-taxable income but their standard errors are relatively much lower, which is a likely consequence of the way they are computed, i.e. by using average productivity prices, as described in Section 2.

5 The impact of the letter

Letters sent in 2009 to taxpayers suspected of residual costs manipulation were aimed at reducing them and, by reducing total deductible costs, increase taxable income. This is the expected direct effect of TALs on RC.

There might however be indirect effects of these letters, as taxpayers could

react strategically to contain the increase of tax payment. Given the evidence-based nature of the TAL, they might find it credible and agree on reducing RC, nonetheless reducing other costs. In the SDS-based context, taxpayers strategic responses are more complex because by increasing other costs while reducing RC, they would cause an increase of related input usage, hence an increase in presumptive revenues, prompting an increase of the probability of being audited by the RA, holding reported revenues constant.

5.1 Direct effects of TALs on RC and taxable income

In Table 3 we present the estimation of equation (2), where the outcome variable, y_{it} , is the log of residual costs. In the first column we do not control for the full set of interactions with cluster dummies, Z_{it} . In the second column we test whether introducing these controls changes our main results. It shows that on average treated firms report a higher level of residual costs than control firms. The coefficient of $(Treated \times Year\ 2007)$ is not significant in both specification, suggesting that there is no evidence to reject the parallel trend assumption in years 2006 and 2007.

The coefficient of $(Treated \times Year\ 2008)$ is positive and statistically significant providing an estimate of nearly 1 log-point in RC with respect to Year 2006, which is the reference year. This is evidence of an increase in manipulation of RC, probably caused on TALs addressing cost for depreciation and intermediate goods. One may observe that the sharp increase of RC in 2008, which prompted the RA's reaction, was itself a sign of strategic firms. As TALs addressing depreciation and intermediate deductible costs had been around for three years by 2008, strategic firms decided to reduce other deductible costs and many of them chose to reduce residual costs, which at that time did not enter SDS and left presumptive revenues unaffected.

Finally, the coefficient of ($Treated \times Year\ 2009$) is negative and statistically significant providing an estimate of 0.4 log-point in RC with respect to Year 2006, which is already suggestive of the direct effect of the letter on RC.

Table 4 present the estimation of equation (2) where the outcome variable is the log-taxable income. As opposed to previous table, controlling for the full set of interactions, which is similar to exact matching, does have an effect on coefficient estimates. The second column shows that the average log-taxable income is 0.126 log-points lower for treated firms. This table also shows that the parallel trend assumption in years 2006 and 2007 is not rejected and the direct effect of the 2009 TALs on RC was as expected by the RA, i.e. towards the increase of taxable income.

Table 5 presents the estimation of equation (2) where the dependent variable is the first difference transformation of log-RC (first column) and of log-taxable income (second column). This allows us to immediately assess the impact of the TALs on RC. The coefficient of the ($Treated \times Year\ 2009$) shows that RC decreased by 1.4 log-points after the 2009 letter, which is about half the average log-RC for treated firms in 2008.² The increase of 0.063 log-points of taxable income for treated firms amounts to about about 2% increase with respect to previous year.³

5.2 Indirect effects as evidence of strategic response

Results discussed in previous Section shows that the RC-TALs had the expected results in terms of direction of the effect, as they reduced RC and increased taxable income. Here we focus on whether there is any evidence that the increase in taxable income was reduced by the strategic behaviour of firms who received

²According to estimates in Table 3, average log-RC of treated firms in year 2008 are equal to $0.638 + 1.354 + .907 = 2.899$.

³From Table 4, the average log-taxable income of treated firms is equal to $2.735 - 0.126 + .390 - 0.005 = 2.994$.

and reacted to the TAL they received.

Table 5 shows the estimation of equation 2 for the first difference of log-costs that were discussed in Section 2 and of presumptive revenues, as computed in SDS.

Focussing on the ($Treated \times Year$ 2009) coefficients, some interesting results emerge. First, the letter on RC had no effect on depreciation, labour and intermediate goods costs, but it caused an increase of costs of services and costs of final goods consumed by the firm. Recalling Table 1, an explanation comes from noticing that no TALs on costs of services nor on costs of final goods have ever been sent and taxpayers reacted by inflating them to limit the increase of 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 did not react varying the number of workers because this is a very costly activity and a crime.

Second, presumptive revenues decrease in 2009 with respect to 2008. At first glance this result may seem contradictory since the increase in costs should have caused an increase in presumptive revenues, exactly as the decrease in costs in 2008 caused a decrease in presumptive revenues. However, we should recall here that, along with the TAL on residual costs, since 2009 the Revenue Agency inserted residual costs among the variables used to calculate presumptive revenues. Although these two policies were both aimed at reducing incentives to manipulate RC, they were not reciprocally coordinated, since the letter was written without mentioning the change in the calculation of presumptive revenues. To sum up, taxpayers react to the letter by:

- reducing residual costs to avoid the risk of being audited for the

manipulation of RC and to decrease the probability to be audited for incongruity with respect to presumptive revenues;

- increasing other costs which they perceived to be less intensively scrutinized by the Revenue Agency.

5.3 Robustness checks

Our identification strategy relies on the existence of parallel trend on years 2006 and 2007 between treated and untreated firms. Previous results are found on a sample of firms selected by analyzing them within clusters and holding only those firms that belonged to a cluster where it was very unlikely that the parallel assumption failed to hold. This was done by dropping firms belonging to clusters where the probability of rejecting the parallel trend assumption when true occurred 50% of times. In other words, by exploiting the richness of data and the data size, we assured the satisfaction of the parallel trend assumption by allowing us to increase type I error. Robustness checks are provided by making this threshold slightly more and slightly less conservative, respectively changing it to 60% and to 40%. Results are presented in tables 7, 8, 9 and 10 roughly confirming results discussed in the previous Section.

6 Concluding remarks

In his comprehensive and updated review Slemrod (2017) notes that “providing concrete evidence that the government has information indicating the presence of evasion consistently works to reduce evasion, apparently because it combines an on-our-radar message with actionable information about noncompliance”. Evidence-based threat-of-audit letters are one of the best known examples of this combination of message and action.

However, Slemrod (2017) also interprets the results obtained by Carrillo et al. (2016) as a warning that "the comprehensiveness of third-party information is crucial and (...), in the context of many developing countries where such comprehensive information is not available, one might not observe a fall in overall evasion from using an additional source of third-party information".

The case we studied in this paper is more complex than that examined by Carrillo et al. (2016). We consider a developed country, albeit with a tax evasion record exceptionally high, such as Italy. Cost reports are used by the Revenue Agency not only to determine the taxable income, but also to calculate presumptive revenues and, thus, the probability of audit. The taxpayer is aware of that and her cost reporting decision has consequences on both the taxable income and the probability to be audited.

What we observe in the paper can be summarized as follows.

Before 2009, taxpayers could overreport residual costs decreasing taxable income without increasing the probability to be audited for a given report of revenues. In 2009, the RA sends threat-of-audit letters to taxpayers suspected to have overreported residual costs and it includes them in the formula to calculate presumptive revenues. Taxpayers react to the letter by decreasing residual costs and increasing other costs at the same time. By doing so, taxpayers increase taxable income to a much smaller extent than the reduction of residual costs. Also, the combination of cost reports is such that presumptive revenues does not increase (they rather actually decrease on average) after the letter.

This response is strategic for two reasons. First, increased costs are those that taxpayers perceive not being monitored by the Revenue Agency. Second, reducing the probability to be audited while increasing costs was possible because the Revenue Agency also changed the formula for presumptive revenues.

The lessons we learnt can be summarized as follows. In a developed country where comprehensive information about revenues and costs is available, it is crucial that taxpayers perceive that this information is used and fully exploited. Also, when this happens policies have to be internally consistent and realized in way which anticipates the possibility of a strategic response by taxpayers.

7 Graphs and Tables

Table 1: Types of costs, SDS and TALs.

Cost item	Relevant for SDS	Addressed by TALs
Depreciation costs	Yes	Since 2005
Labour costs	Yes	No
Cost of intermediate goods	Yes	Since 2005
Cost for services	Yes	No
Cost for final goods	Yes	No
Residual cost	Since 2009	Since 2009

Table 2: Some descriptive statistics of the outcome variables used.

Variable	Obs	Mean	Std. Dev.	Min	Max
Log of residual costs	69,361	1.017	1.695	-6.908	8.332
Log of taxable income	69,361	3.029	0.982	-6.908	7.393
Log depreciation	64,495	1.037	1.417	-6.908	6.762
Log labour costs	25,595	2.785	2.048	-6.908	7.862
Log costs of intermediate goods	41,360	3.183	2.269	-19.854	8.387
Log costs of services	67,090	2.270	1.410	-6.908	8.283
Log costs of final goods	34,801	1.804	1.505	-6.908	6.996
Log presumptive revenues	69,287	11.290	1.120	2.996	15.520

Notes: Our calculations on RA tax data.

Figure 1: The incidence of residual costs on taxable income over time

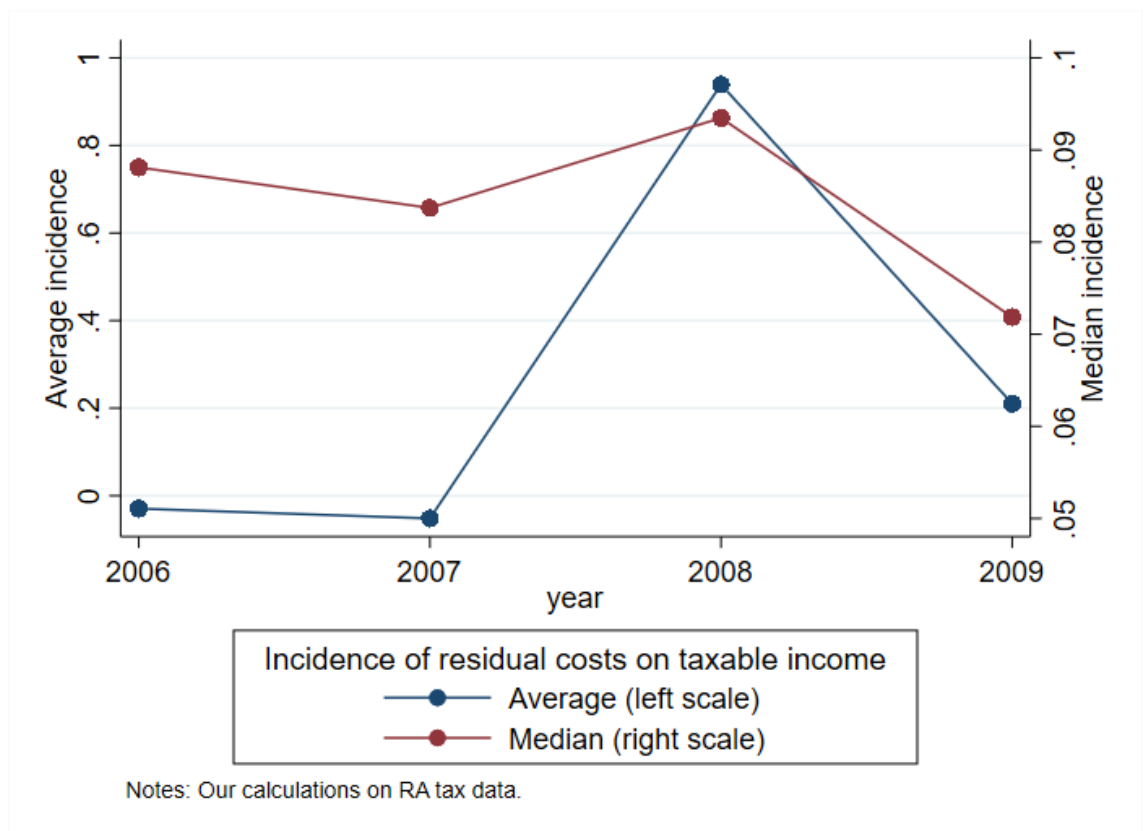


Table 3: Log of residual costs. Difference in difference model estimation selecting only clusters where the parallel trend test has p-value larger than 0.5.

Log of residual costs		
Treated	1.354***	1.267***
	(0.029)	(0.027)
Year 2007	-0.036*	0.283
	(0.020)	(0.269)
Year 2008	-0.000	0.237
	(0.020)	(0.271)
Year 2009	-0.072***	0.225
	(0.021)	(0.273)
Treated x Year 2007	0.053	0.044
	(0.039)	(0.036)
Treated x Year 2008	0.907***	0.995***
	(0.039)	(0.036)
Treated x Year 2009	-0.395***	-0.380***
	(0.040)	(0.037)
Full set of interactions	No	Yes
Constant	0.638***	0.539**
	(0.015)	(0.246)
R-squared	0.182	0.431
N	64232	63291

Notes: Our calculations on RA tax data. The full set of interactions is described in Section 4 and corresponds to matrix Z_{it} in equation (2).

Standard errors in parentheses. * for $p < .1$, ** for $p < .05$, and *** for $p < .01$.

Table 4: Log of taxable income. Difference in difference model estimation selecting only clusters where the parallel trend test has p-value larger than 0.5.

Log of taxable income		
Treated	0.142*** (0.019)	-0.126*** (0.018)
Year 2007	0.109*** (0.013)	0.186 (0.175)
Year 2008	0.156*** (0.013)	0.390** (0.175)
Year 2009	0.097*** (0.013)	0.232 (0.177)
Treated x Year 2007	-0.028 (0.025)	0.015 (0.024)
Treated x Year 2008	-0.059** (0.025)	-0.005 (0.024)
Treated x Year 2009	0.027 (0.025)	0.063*** (0.024)
Full set of interactions	No	Yes
Constant	2.901*** (0.010)	2.735*** (0.159)
R-squared	0.006	0.304
N	64232	63291

Notes: Our calculations on RA tax data. The full set of interactions is described in Section 4 and corresponds to matrix Z_{it} in equation (2).

Standard errors in parentheses. * for $p < .1$, ** for $p < .05$, and *** for $p < .01$.

Table 5: First difference of log-residual costs and log-taxable income. Difference in difference model estimation selecting only clusters where the parallel trend test has p-value larger than 0.5.

	First-difference of Log residual costs	First-difference of Log taxable income
Treated	0.090*** (0.025)	0.027** (0.013)
Year 2008	-0.288 (0.253)	0.142 (0.134)
Year 2009	-0.475* (0.258)	-0.302** (0.136)
Treated x Year 2008	0.819*** (0.034)	-0.053*** (0.018)
Treated x Year 2009	-1.446*** (0.034)	0.063*** (0.018)
Full set of interactions	Yes	Yes
Constant	0.263 (0.243)	0.014 (0.128)
R-squared	0.186	0.061
N	44510	44510

Notes: Our calculations on RA tax data. The full set of interactions is described in Section 4 and corresponds to matrix Z_{it} in equation (2).

Standard errors in parentheses. * for $p < .1$, ** for $p < .05$, and *** for $p < .01$.

Table 6: Strategic response. First difference of log-costs and log-presumptive revenues. Difference in difference model estimation selecting only clusters where the parallel trend test has p-value larger than 0.5.

	Log depreciation	Log labour costs	Log costs of intermediate goods	Log costs of services	Log costs of final goods	Log presumptive revenues
Treated	0.054*** (0.017)	0.026 (0.033)	0.045 (0.031)	0.003 (0.015)	-0.005 (0.025)	0.122*** (0.008)
Year 2008	0.159 (0.166)	0.012 (0.195)	-0.125 (0.188)	0.044 (0.144)	0.107 (0.201)	0.049 (0.076)
Year 2009	0.031 (0.169)	0.240 (0.201)	-0.184 (0.191)	-0.049 (0.147)	-0.064 (0.208)	-0.098 (0.077)
Treated x Year 2008	-0.081*** (0.023)	-0.064 (0.044)	-0.146*** (0.040)	-0.200*** (0.020)	-0.018 (0.033)	-0.032*** (0.010)
Treated x Year 2009	0.012 (0.023)	0.061 (0.044)	0.039 (0.040)	0.425*** (0.020)	0.106*** (0.033)	-0.182*** (0.010)
Full set of interactions	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-0.031 (0.164)	0.148 (0.211)	0.046 (0.183)	0.089 (0.139)	0.154 (0.237)	0.079 (0.073)
R-squared	0.049	0.143	0.184	0.073	0.100	0.148
N	40523	14588	24250	42532	20274	44449

Notes: Our calculations on RA tax data. The full set of interactions is described in Section 4 and corresponds to matrix Z_{it} in equation (2). Standard errors in parentheses. * for $p < .1$, ** for $p < .05$, and *** for $p < .01$.

Table 7: Robustness check. First difference of log-residual costs and log-taxable income. Difference in difference model estimation selecting only clusters where the parallel trend test has p-value larger than 0.6.

	First-difference of Log residual costs	First-difference of Log taxable income
Treated	0.085*** (0.028)	0.034** (0.014)
Year 2008	0.762*** (0.277)	-0.409*** (0.143)
Year 2009	0.144 (0.277)	-0.358** (0.143)
Treated x Year 2008	0.853*** (0.037)	-0.064*** (0.019)
Treated x Year 2009	-1.444*** (0.038)	0.056*** (0.019)
Full set of interactions	Yes	Yes
Constant	-0.015 (0.281)	0.315** (0.145)
R-squared	0.190	0.060
N	36649	36649

Notes: Our calculations on RA tax data. The full set of interactions is described in Section 4 and corresponds to matrix Z_{it} in equation (2). Standard errors in parentheses. * for $p < .1$, ** for $p < .05$, and *** for $p < .01$.

Table 8: Robustness check. First difference of log-residual costs and log-taxable income. Difference in difference model estimation selecting only clusters where the parallel trend test has p-value larger than 0.6.

	First-difference of Log residual costs	First-difference of Log taxable income
Treated	0.094*** (0.024)	0.024* (0.013)
Year 2008	-0.287 (0.253)	0.141 (0.133)
Year 2009	-0.473* (0.258)	-0.301** (0.135)
Treated x Year 2008	0.815*** (0.033)	-0.048*** (0.017)
Treated x Year 2009	-1.457*** (0.033)	0.063*** (0.017)
Full set of interactions	Yes	Yes
Constant	0.263 (0.243)	0.015 (0.128)
R-squared	0.186	0.061
N	48108	48108

Notes: Our calculations on RA tax data. The full set of interactions is described in Section 4 and corresponds to matrix Z_{it} in equation (2). Standard errors in parentheses. * for $p < .1$, ** for $p < .05$, and *** for $p < .01$.

Table 9: Robustness check of strategic responses. First difference of log-costs and log-presumptive revenues. Difference in difference model estimation selecting only clusters where the parallel trend test has p-value larger than 0.6.

	Log depreciation	Log labour costs	Log costs of intermediate goods	Log costs of services	Log costs of final goods	Log presumptive revenues
Treated	0.055*** (0.019)	0.045 (0.041)	0.050 (0.033)	0.011 (0.017)	-0.018 (0.029)	0.134*** (0.009)
Year 2008	0.153 (0.178)	0.303 (0.236)	-0.158 (0.197)	-0.145 (0.160)	-0.339* (0.190)	-0.059 (0.084)
Year 2009	0.053 (0.179)	0.147 (0.237)	-0.313 (0.194)	0.209 (0.160)	-0.303 (0.193)	-0.080 (0.084)
Treated x Year 2008	-0.089*** (0.025)	-0.093* (0.055)	-0.130*** (0.043)	-0.211*** (0.023)	0.025 (0.038)	-0.037*** (0.011)
Treated x Year 2009	0.011 (0.025)	0.032 (0.055)	0.024 (0.043)	0.422*** (0.023)	0.132*** (0.038)	-0.197*** (0.011)
Full set of interactions	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-0.138 (0.183)	-0.065 (0.273)	0.167 (0.201)	-0.039 (0.162)	0.356* (0.204)	0.118 (0.085)
R-squared	0.050	0.144	0.189	0.073	0.099	0.148
N	33275	10977	19562	34915	15905	36600

Notes: Our calculations on RA tax data. The full set of interactions is described in Section 4 and corresponds to matrix Z_{it} in equation (2). Standard errors in parentheses. * for $p < .1$, ** for $p < .05$, and *** for $p < .01$.

Table 10: Robustness check of strategic responses. First difference of log-costs and log-presumptive revenues. Difference in difference model estimation selecting only clusters where the parallel trend test has p-value larger than 0.4.

	Log depreciation	Log labour costs	Log costs of intermediate goods	Log costs of services	Log costs of final goods	Log presumptive revenues
Treated	0.056*** (0.016)	0.019 (0.032)	0.052* (0.029)	0.003 (0.014)	-0.020 (0.024)	0.121*** (0.007)
Year 2008	0.160 (0.167)	0.012 (0.198)	-0.127 (0.185)	0.042 (0.144)	0.108 (0.202)	0.048 (0.076)
Year 2009	0.031 (0.170)	0.236 (0.204)	-0.181 (0.187)	-0.048 (0.147)	-0.066 (0.208)	-0.098 (0.077)
Treated x Year 2008	-0.089*** (0.022)	-0.063 (0.043)	-0.139*** (0.038)	-0.192*** (0.019)	-0.014 (0.032)	-0.030*** (0.010)
Treated x Year 2009	0.010 (0.022)	0.083* (0.043)	0.026 (0.038)	0.420*** (0.019)	0.131*** (0.032)	-0.181*** (0.010)
Full set of interactions	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-0.032 (0.165)	0.149 (0.214)	0.043 (0.179)	0.089 (0.139)	0.155 (0.237)	0.080 (0.073)
R-squared	0.050	0.140	0.186	0.071	0.102	0.148
N	43754	15706	26291	45942	21662	48043

Notes: Our calculations on RA tax data. The full set of interactions is described in Section 4 and corresponds to matrix Z_{it} in equation (2). Standard errors in parentheses. * for $p < .1$, ** for $p < .05$, and *** for $p < .01$.

References

- Carrillo, D., D. Pomeranz, and Sing (2016). Dodging the taxman: Firm misreporting and limits to tax enforce. *American Economic Journal: Economic Policy*.
- Iacus, S. M., G. King, and G. Porro (2011). Causal inference without balance checking: Coarsened exact matching. *Political Analysis*.
- Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez (2011). Unwilling or unable to cheat? evidence from a tax audit experiment in denmark. *Econometrica* 79(3), 651–692.
- Santoro, A. (2008, July). Taxpayers’ choices under studi di settore: what do we know and how we can interpret it? *Giornale degli Economisti* 67(2), 161–184.
- Santoro, A. and C. V. Fiorio (2011, January). Taxpayer behavior when audit rules are known: Evidence from italy. *Public Finance Review* 39(1), 103–123.
- Slemrod, J. (2017). Tax compliance and enforcement. Technical report, University of Michigan.
- Slemrod, J., M. Blumenthal, and C. 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, J., B. Collins, J. Hoopes, D. Reck, and M. Sebastiani (2017). Does credit-card information reporting improve small-business tax compliance? *Journal of Publ* 149, 1–19.