



Contents lists available at ScienceDirect

## Journal of Economic Behavior and Organization

journal homepage: [www.elsevier.com/locate/jebo](http://www.elsevier.com/locate/jebo)Compliance spillovers across taxes: The role of penalties and detection<sup>☆</sup>Andrea Lopez-Luzuriaga<sup>a</sup>, Carlos Scartascini<sup>b,c,\*</sup><sup>a</sup> George Washington University and Inter-American Development Bank, United States<sup>b</sup> Inter-American Development Bank, 1300 New York Ave, NW, United States<sup>c</sup> Washington, DC 20577, United States

## ARTICLE INFO

## Article history:

Received 4 December 2018

Revised 19 June 2019

Accepted 20 June 2019

Available online 11 July 2019

## JEL classification:

H26

C93

D03

H41

## Keywords:

Tax compliance

Spillovers

Evasion

Property tax

Sales tax

Randomized field experiment

Behavioral economics

## ABSTRACT

When the tax authority increases the enforcement for one tax, what happens to the level of compliance in other taxes (spillover effect)? In this paper, we present a simple analytical model that shows that the sign of the spillover depends on how taxpayers update their beliefs about penalties and detection probabilities for one tax after observing the deterrence actions the tax agency takes for another tax. As a result, when spillovers are present, penalties and detection may not necessarily be interchangeable policy tools. We evaluate the sign of the spillover in the context of a randomized field experiment in a municipality in Argentina in a sample of about 700 taxpayers who are liable for both the property and gross-sales taxes. The evidence from the intervention indicates that the spillover from a message that increases the salience of penalties and enforcement for the property tax on the declaration in the gross-sales tax is positive. Those in the treatment group increase their reported tax by two percentage points more than the control group. This result has ample implications for researchers bringing interventions to the field and for governments' enforcement strategies.

© 2019 The Authors. Published by Elsevier B.V.  
This is an open access article under the CC BY-NC-ND license.  
(<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

## 1. Introduction

Empirical studies evaluating the direct effect of enforcement on tax compliance have blossomed in the last few years (Hallsworth, 2014; Mascagni, 2018; Slemrod, 2016). However, there is little evidence of the effect of enforcing one tax on the behavior of taxpayers in other taxes (spillover effect) even though it can determine the overall success of an enforcement strategy. Should we expect positive, neutral, or negative spillovers? We explore the answer to this question by using a very simple and easily generalized analytical model à la Allingham–Sandmo that approximates the setting in which we work.

<sup>☆</sup> We would like to thank the former Mayor and staff from the Municipality of Junín for providing the data, Lucio Castro and the team at CIPPEC for helping with the original data collection and intervention, and the Institutional Capacity Strengthening Fund (ICSF) of the Inter-American Development Bank for its financial support, funded by the Government of the People's Republic of China. We would also like to thank the editor, an associate editor, an anonymous reviewer, the discussant and participants at the 2018 IRS-TPC Conference, the 2018 Advances with Field Experiments Conference, and the Development Tea Workshop at George Washington University for their comments and suggestions. The opinions presented herein are those of the authors and thus do not necessarily represent the official position of the institutions they belong to.

\* Corresponding author.

E-mail address: [carlossc@iadb.org](mailto:carlossc@iadb.org) (C. Scartascini).

In the model, taxpayers face sequential decisions about whether to pay a tax that has neither reporting nor informational asymmetries (property tax), and then how much sales to declare in a self-reporting tax where there are informational asymmetries (gross-sales tax). In that simple setting, the comparative statics are straightforward. If there is an increase in penalties (or perceived penalties), which tend to be uniform across taxes, the spillover will be positive. If there is an increase in the perceived probability of detection in one tax, the effect on the declaration of other taxes depends on how taxpayers update their beliefs about overall detection probabilities. That is, if taxpayers extrapolate the higher detection in one tax to the other taxes they owe, spillovers will be positive. However, taxpayers could also assume that, given the limited resources of the tax administration, higher detection in one tax might imply lower enforcement in other taxes, which will generate negative spillovers (Advani et al., 2017; DeBacker et al., 2015; Maciejovsky et al., 2007). Adding cash constraints or an overall budget constraint for each taxpayer could reinforce these negative effects. Consequently, while interventions affecting either the penalties or the probability of detection will have positive direct effects, the spillover effects are independent neither of whether the tax authority signals higher penalties or detection nor the assumptions taxpayers make about the enforcement capacity of the government.

In the empirical section, using data that combines a randomized field experiment for the property tax with administrative data on gross-sales tax declarations in one municipality of Argentina, the results suggest that taxpayers who received a message explaining the consequences of not paying the property tax decided to declare a higher gross-sales tax. The group that received the deterrence message with their property tax bill increased their gross-sales tax payment, on average, by two percentage points more than the control group (which translated into an increase of about 3.4 percentage points in their declared sales), which is statistically significant at the 5% level. The results are consistent with the model, particularly because the deterrence message in the treatment was mostly focused on increasing the salience of the penalty, which is the same across taxes in this city. The positive spillover result suggests that taxpayers did not believe that a higher enforcement in one tax might indicate lower enforcement in other taxes they have to pay.

The suggestive evidence that a spillover effect across taxes can exist has several important implications. First, the results and the analytical argument seem to indicate that penalties and increased detection are not necessarily interchangeable policy tools once we consider all the taxes an individual is liable for (the full tax portfolio). While an increase of the penalties will have an unmistakable positive effect, an increase on the probability of detection could have a positive or negative effect depending on the assumptions the taxpayer holds about the tax control process from the tax authority. Second, researchers should consider the spillover effect when designing an intervention. Otherwise, they risk losing from other taxes what they may gain from the tax under treatment. This puts an additional burden on the design stage of the intervention because choosing penalties or detection probability is not trivial. Additionally, when manipulating enforcement, the intervention should explicitly consider how people might update detection probabilities across taxes. Third, given that there are spillover effects, tax authorities should design deterrence strategies taking into account the full tax portfolio for any given taxpayer. Therefore, the most efficient strategy is not the one that maximizes the direct payoff but the one that maximizes tax collection across the full portfolio.

The rest of the paper is organized as follows. Section 2 presents a literature review. Section 3 presents the model. Section 4 presents an overview of the original intervention and describes the property and gross-sales taxes. Sections 5 and 6 present our empirical strategy and results. Section 7 concludes.

## 2. Literature review

There is now ample empirical literature showing that taxpayers who receive a deterrence message from the tax authority tend to react by increasing tax compliance (Brockmeyer et al., 2016; Chirico et al., 2016; Doerrenberg and Schmitz, 2017; Fellner et al., 2013; Kleven et al., 2011; Meiselman, 2018; Slemrod et al., 2001). In addition, it has been documented that an increase in monitoring has a positive effect on compliance (LaLumia and Sallee, 2013; Naritomi, 2019). There is also literature supporting the idea that individuals might exhibit sub-optimal behaviors when dealing with taxes (Abeler and Jäger, 2015; Chetty et al., 2009). In fact, when taxpayers have limited attention, messages that raise the salience of fines and legal action can increase compliance (Bernheim and Rangel, 2007; 2009; Castro and Scartascini, 2015). Hence, it is expected that if a taxpayer received a message that underlined the probability of being penalized and explained the calculation of the fine, she would increase her level of compliance.

Within this broad and rapidly expanding empirical literature, studies looking at spillovers are still scarce. We define the spillover as the indirect effect of the interventions across individuals or across margins for the same individual. Only a few studies explore the presence of spillover effects of tax enforcement across individuals. All of these studies analyze the behavior of individuals who themselves have not been subject to any enforcement but are related to someone who has. Rincke and Traxler (2011) analyze the effect of licensing inspections on the payment of TV license fees. They take advantage of the fact that inspections are not directly observable for untreated households and look at the spillover effect on their compliance generated by informal communications among neighbors. They adopt an instrumental variable approach using the intensity of winter as an instrument, because inspectors are paid a fixed fee per visit. Pomeranz (2015) shows that deterrence letters sent to taxpayers have spillover effects up the value-added-tax chain by generating a paper trail of the transaction. Drago et al. (2015) show a substantial spillover effect from treated to untreated individuals with results from a field experiment that varied the content of mailings sent to potential evaders of TV license fees. This result has important implications for deterrence policies, given that different individuals generate different spillovers according to the network

they belong to. Similarly, [Boning et al. \(2018\)](#) also show spillover effects in enforcement that are transmitted through tax-preparer networks, geographic neighborhoods, and parent-subsidiary relationships. Finally, [Carrillo et al. \(2017b\)](#) find evidence of spillovers across individuals in a setting of positive incentives instead of deterrence. In the context of a program that rewarded individuals who had complied by providing them with the construction of a new sidewalk, they find an increase in compliance by the neighbors of the winners. Interestingly, the results are heterogeneous regarding the salience of the sidewalk. This literature provides evidence that spillovers across taxpayers can exist for both deterrence and positive incentives, but the sign and size of the spillover are not independent of the design of the intervention.

There is also some evidence about how taxpayers behave across different margins of the same tax. [Carrillo et al. \(2017a\)](#) and [Slemrod et al. \(2017\)](#) make the case that when the tax authority signals having third-party information on transactions, taxpayers tend to increase their reported revenues, but these taxpayers largely offset increased reported revenues with increased reported expenses. The same phenomenon of compensating higher taxes in one margin by decreasing their reporting in another is reported by [Boning et al. \(2018\)](#). In this case, subsidiaries of treated firms remitted less tax, which is consistent either with a cash-flow effect or substitution of noncompliance to a seemingly less monitored report. There are also a couple of studies that look at the effect of enforcement for the same individual and the same tax over different periods of time ([Advani et al., 2017](#); [Kleven et al., 2011](#)). [Kleven et al. \(2011\)](#) select a sample of 40,000 income tax filers in Denmark, half of whom were audited. The following year, they randomly sent a threat-to-audit letter to taxpayers who had previously been audited and taxpayers who had not been. They find that the audit and the threat of an audit decrease evasion on self-reported income. [Advani et al. \(2017\)](#) find a similar result when studying the random audit program in the United Kingdom over five years. They find that the audit is more effective and more lasting on sources of income that are self-reported and less volatile over time.

In summary, the evidence so far indicates that for an individual taxpayer: (i) deterrence messages that increase the salience of penalties and the stringency of enforcement in one tax increase compliance with that tax, and (ii) spillovers can be positive or negative. The sign of the spillover seems to be correlated to the taxpayer evaluation of the ability of the tax agency to enforce across individuals, other taxes, or other margins of the same tax.

Our research is different from that previously described because we look at the effect of an intervention on the same individual but across different taxes. To the best of our knowledge, the only field experiment that shows some evidence regarding spillovers for an individual taxpayer across taxes is [Ortega and Scartascini \(2015\)](#). Taxpayers who received a notice from the tax authority regarding their owed taxes for the income tax, wealth tax or VAT tended to show a higher probability of canceling debts in other taxes too. Our study differs from theirs in two ways: their focus is on tax delinquencies while ours is focused on current payments and declarations (a setting which is more akin to the one the literature tends to deal with), and their focus is on the direct effect of the intervention, so they do not explore the mechanisms behind the results. Because we care about the mechanisms, we develop a simple analytical model that could serve as a building block for evaluating spillovers in the broader literature.<sup>1</sup>

### 3. A simple analytical model

We analyze the effect of an intervention designed to test the determinants of compliance with the property tax on the gross-sales tax declarations. We develop a very simple analytical model à la ([Allingham and Sandmo, 1972](#)) that approximates the setting in which we work to understand the conditions under which spillovers could be positive or negative. Within our model, taxpayers face sequential decisions about whether to pay a tax that has no reporting and informational asymmetry (property tax) and then how much sales to declare in a self-reporting tax where there are informational asymmetries (gross-sales tax). We develop this specific model in order to generate testable implications for our empirical work, but the model could easily be generalized to a context with any two uncorrelated taxes (the results would be the same), or two correlated taxes (the results would tend to be stronger *ceteris paribus*).<sup>2</sup>

In the setting in which we work, before receiving the message on the property tax bill, the individual has prior beliefs regarding both the probability that the penalties for not paying the tax will be enforced and about the amount of the penalty. The penalty is determined by law, but the taxpayer could have imperfect knowledge about how it is calculated. Upon receiving the property tax bill, the taxpayer updates either or both of those beliefs. A few days after receiving the bill, the taxpayer decides whether or not to pay it and a few weeks later, she decides how much sales to declare and whether or not to pay the gross-sales tax. We analyze the decisions of the individual in the same sequence in which she faces them. In the first stage, she decides whether or not to pay the property tax, and in the second, how much sales to declare. We are assuming a risk-averse individual who is not credit constrained. She has some wealth and enough money to pay both taxes, and her business is producing some profit. If the individual is credit-constrained, there will be an additional channel that allows the enforcement of one tax to affect the other.

<sup>1</sup> Our work also contributes to the literature on risk perception. [Bérgolo et al. \(2018\)](#) present evidence that taxpayers overestimate the probability of being audited on the reported income tax. We present a possible explanation for this phenomenon.

<sup>2</sup> Adding a binding budget constraint or evaluating the effect of spillovers across taxes enforced by different levels of government could also be useful ideas for expanding the model.

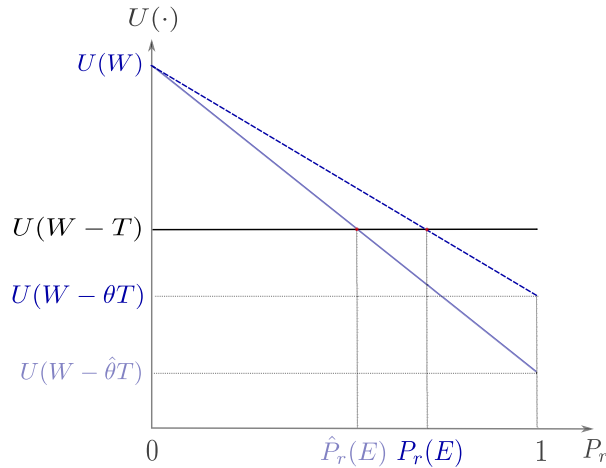


Fig. 1. Whether or not to pay the property tax.

3.1. First stage

The individual decides whether or not to pay the property tax. She has an initial level of wealth  $W$  and has to pay a tax of amount  $T$ . The utility when paying the tax is  $U(W - T)$ , where  $U(\cdot)$  is increasing and concave. If she decides not to pay, her expected utility is  $P_r(E)U(W - \theta T) + (1 - P_r(E))U(W)$ , where  $P_r(E)$  is the (perceived) probability that the city government enforces the penalties of not paying the property tax, and is a function of the overall perception regarding the enforcement capacity of the tax authority ( $E$ ). For instance, if the government increased its personnel or received more funding for tax control,  $E$  would increase and so would  $P_r$ , with  $\frac{\partial P_r}{\partial E} > 0$ .

If the government enforces the payment of the fine, the individual has to pay a fine  $\theta$  in addition to the billed tax, where  $\theta > 1$ ; for instance, for a penalty of 5%,  $\theta$  would be 1.05. The solution can easily be interpreted according to 1. She pays the tax as long as the expected utility of paying is higher than the expected utility of not paying. For a perceived fine of size  $\theta$ , the taxpayer will pay the tax if she assumes that the probability of enforcement is equal to or higher than  $P_r$ . An increase in the perceived amount of the fine will make the option of paying more attractive. Take note of how for a higher fine ( $\theta < \hat{\theta}$ ) the utility of not paying is lower for any probability of enforcement. Now, if the perceived fine goes up to  $\hat{\theta}$  ( $> \theta$ ), then taxpayers with perceived probabilities between  $\hat{P}_r$  and  $P_r$  will also decide to pay the tax. Therefore, if the tax authority is able to affect the perceived fines or the perceived probability of enforcement, it can increase tax compliance.

3.2. Second stage

In the second stage we use a traditional Allingham-Sandmo (A-S) model with a risk-averse individual with an increasing concave utility function. The individual maximizes the expected utility by choosing how much income to report. For simplicity, we assume that the only cost for the business is the tax. The individual's true sales are  $y$  and the reported sales are  $\tilde{y}$ . The reported sales are taxed at a rate  $t$ . The probability of being caught under-reporting sales is  $P_s$ , which is a function of the overall perception of the city government's enforcement capacity ( $E$ ), and a function of the enforcement in other taxes ( $P_r$  in this case). The reason is quite simple: resources are limited, so higher enforcement in one tax might imply lower enforcement in another. Assuming fixed overall resources for the tax authority is relatively standard (see Ortega and Scartascini, 2015 for a discussion). If caught cheating, the taxpayer has to pay the tax  $t$  plus a penalty  $\theta$ . The individual maximization problem can be written as:

$$\max_{\tilde{y}} : (1 - P_s(E, P_r(E)))U(y - t\tilde{y}) + P_s(E, P_r(E))U(y - t\tilde{y} - \theta t(y - \tilde{y})) \tag{1}$$

For notation convenience  $X = y - t\tilde{y}$  and  $\hat{X} = y - t\tilde{y} - \theta t(y - \tilde{y})$ . Let us denote  $V$  as the expected utility function. The first order conditions ( $\frac{\partial V}{\partial \tilde{y}} = V' = 0$ ) is:

$$-t(1 - P_s(E, P_r(E)))U'(X) + tP_s(E, P_r(E))U'(\hat{X})(\theta - 1) = 0 \tag{2}$$

Since the utility function is concave, the second order conditions are satisfied. In this simple setting, comparative statics are straightforward. Differentiating the first order conditions with respect to  $\theta$ , we find that if there is an increase in penalties (or of perceived penalties), which tend to be uniform across taxes, then the spillover is positive  $\frac{\partial \tilde{y}}{\partial \theta} > 0$ . Repeating the exercise for  $P_r(E)$ , we find that the effect of an increase in the perceived probability of detection in one tax upon other taxes strongly depends on the assumptions about how taxpayers update their beliefs regarding overall enforcement sign

$\text{sign}\left[\frac{\partial \bar{y}}{\partial P_r(E)}\right] = \text{sign}\left[\frac{\partial P_s(E, P_r(E))}{\partial P_r(E)}\right]$  (all the derivations are in the appendix). If the individual assumes that  $P_r(E)$  and  $P_s(E, P_r(E))$  are uncorrelated, then  $\frac{\partial P_s(E, P_r(E))}{\partial P_r(E)} = \frac{\partial \bar{y}}{\partial P_r(E)} = 0$ .

If she assumes that the city government is monitoring her as an individual and not with regards to a particular tax  $P_r = P_s$ ,  $\frac{\partial P_s(E, P_r(E))}{\partial P_r(E)} = 1$  and  $\frac{\partial \bar{y}}{\partial P_r(E)} > 0$ . That is, if taxpayers extrapolate the higher enforcement of one tax to the other taxes they owe, spillovers would be positive.

However, taxpayers could also assume that given limited resources for the tax administration, higher enforcement of one tax might imply lower enforcement of other taxes, which could generate negative spillovers (DeBacker et al., 2015; Maciejovsky et al., 2007; Mittone, 2006). In particular,  $\frac{\partial P_s(E, P_r(E))}{\partial P_r(E)} = -1$  so  $\frac{\partial \bar{y}}{\partial P_r(E)} < 0$ . Adding cash constraints or an overall budget constraint would reinforce these negative effects.

In summary, the model's predictions are as follows: (i) increasing penalties or detection has a positive direct effect (which is consistent with the existing literature), and spillover effects are positive for the penalties (if penalties are correlated across taxes, which is a characteristic common to most countries), but they are ambiguous for detection. Spillovers will be zero if the taxpayer assumes that detection probabilities are uncorrelated across taxes, positive if she assumes a positive correlation, and negative if she assumes a negative correlation. These results are also consistent with the existing literature and help to square off existing results showing positive and negative spillovers.

#### 4. Background and data

Castro and Scartascini (2015) conducted a large field experiment designed to test the determinants of compliance with the property tax in the Municipality of Junín in Argentina. The property tax, formally called the “Public Space Conservation Tax” (Tasa de Conservación de la Vía Pública, or CVP henceforth), is a tax levied on homes, farms, business premises, and most other real estates. The tax is calculated by the city government and is billed every two months to the property owner. The tax is computed according to the front side of the property and the services the city provides, such as public lighting, trash collection, and street cleaning. Because the tax is billed by the city, there is no reporting, and there are no informational asymmetries between the government and the taxpayer. The taxpayers' only choice is whether to pay the billed amount or not, which becomes known to the city government after the due date. Taxpayers have approximately ten days to pay from the moment they receive the bill. By August 2011, which is when the original intervention took place, there were around 26,000 individual taxpayers registered to pay the property tax, equivalent to a third of the population of Junín. The Municipality allows taxpayers to pay on a yearly or monthly basis. However, only around 12% of taxpayers choose either of these options; the rest pay every other month by default. For the experiment, the authors included only individual taxpayers in the sample and dropped firms and corporations. This is exactly the framework of the first stage of our model.

A subgroup of the individual taxpayers who pay the property tax is also liable for a gross-sales tax that is administered by the same municipality. The gross-sales tax is paid by all retail, wholesale, service and industrial businesses in the city. The gross-sales tax is formally called the “Safety and Hygiene Inspection Tax” (“Tasa por Inspección e Higiene”, or SEH henceforth). The tax is calculated based on the gross monthly sales, the number of employees and the size of the establishment where the economic activity is developed (a description of these variables can be found in the appendix). The tax rate depends on the economic activity (see the appendix for the specific rate). Each taxpayer must report their sales once a month, and the number of employees and the size of the establishment once a year. Hence, within a calendar year, the tax has both a fixed and a variable component. Although the Municipality allows taxpayers to pay monthly, only 11% of taxpayers do so; the rest pay every two months. In this tax, there are informational asymmetries: sales are only known to the taxpayer; hence, misreporting is possible. If a business owner fails to fill in the monthly form, it is assumed that the sales were the same as the previous month and taxpayers are fined a penalty of ARS\$250 (equivalent to 7% of the monthly minimum wage and USD\$90.25 PPP<sup>3</sup>) for not filling in the form on time. If a tax form is filled in afterward and the reported sales are higher than those of the previous month, the difference must be paid plus a penalty of 2% compound monthly interest. In contrast, if the reported sales are lower than the sales of the previous month, the taxpayer does not receive a tax credit or a refund for the extra tax that was paid. As such, while there could be incentives for misreporting the actual sales, there is little incentives for not filing the sales declaration form. In this tax, the relevant evasion margin for taxpayers is how much sales to declare, which is not known by the tax authority. In contrast, the Municipality knows whether taxpayers file the form and pay the assessed tax on time, making it easily enforceable.<sup>4</sup> By August 2011, there were around 2500 individual taxpayers registered to pay the gross-sales tax, most taxpayers owning only one business, and just 3% owning more than one business. The median payment was ARS\$98 (equivalent to 2.7% of the monthly minimum wage and USD\$35.38 PPP). The property tax and the gross-sales tax are the two primary sources of tax revenue for the city government. In 2011, the property tax and the gross-sales tax were around 65% of the tax collection for the city of Junín.

<sup>3</sup> We use the purchasing power parity conversion factor, private consumption of 2011 for Argentina of 2.77 local currency units per international dollars. The PPP conversion factor is the number of units of a country's currency required to buy the same amounts of goods and services in the domestic market as the U.S. dollar would buy in the United States. (Source: World Bank, International Comparison Program database).

<sup>4</sup> In this case, only sales matter; therefore, taxpayers cannot offset their liability by increasing costs or claiming any deductions.

The payment scheme is very similar for both taxes. Most taxpayers pay every two months, and there are two due dates for each tax. The first due date is usually in the second week of the month and the second due date takes place the following week. Taxpayers are supposed to pay by the first due date, but if they pay by the second due date, no late fees are charged. The property tax is paid in the first month, and the gross-sales tax is paid in the second month of each two-months pay period. For instance, in the fifth pay period of the year (September and October), the property tax is due in September, and the gross-sales tax is due in October. A cumulative compound monthly interest rate of 2% is applied to any outstanding liabilities with the city government, independently of the tax that generates the debt.

In [Castro and Scartascini \(2015\)](#), approximately 23,000 taxpayers were randomly divided into four groups: three treatment groups and one control group. A message was included on the property tax bill of each treatment group. The messages were designed to test the main determinants of tax compliance: deterrence (beliefs about enforcement and fines), peer effects (beliefs about other taxpayers' behavior), and reciprocity (beliefs about the use of resources by the government). Private companies, social organizations, and taxpayers who paid their dues annually were excluded from the sample. A stratified randomization strategy based on the geographic location was made to select the taxpayers for each treatment. Within each randomization block, one taxpayer was assigned to the control group for each taxpayer randomly assigned to a treatment, so that 60% of taxpayers were randomly assigned to the control group, and the remainder were equally distributed to each of the treatment groups.<sup>5</sup> The results in [Castro and Scartascini \(2015\)](#) show that the deterrence message increased compliance with the property tax by almost five percentage points, which represents an increase in compliance rates of approximately 12%. In this paper, we combine the deterrence message sent to property owners with data from the gross-sales tax. In addition to the information about property tax compliance, we have access to the declared gross-sales tax for each taxpayer for each two-month period in 2011 and their 2010 annual tax return, which includes the total annual sales of 2010, the number of employees in 2010 and the size of the building in meters. We include in our analysis only those taxpayers who pay the gross-sales and property taxes every two months.<sup>6</sup> The subsample of taxpayers who own property and are sole proprietors of a business is small. We have 608 sole proprietors in the control group and 115 in the treatment group.<sup>7</sup> This subgroup of taxpayers was not the focus of the original experiment, yet the randomization was successful in balancing this subgroup of taxpayers between treatment and control ([Table 3](#)). On average, the annual sales of these businesses in 2010 was ARS\$226,380 (USD\$81,725 PPP), and in the billing period before the treatment period (July/August 2011) they paid on average ARS\$ 111 (USD\$40.07 PPP).<sup>8</sup>

The deterrence message sent in the property tax bill had two components. First, one component that tried to increase the salience of the penalty and reduce the computational cost: *“Did you know that if you do not pay the property tax on time for a debt of ARS\$1,0000<sup>9</sup> you will have to pay ARS\$268<sup>10</sup> in arrears at the end of the year?”* The objective of including the example of the cost of noncompliance was to reduce the computational costs derived from the calculation of arrears on unpaid tax liabilities using a compounded interest rate. According to the literature, such a message should increase the salience of the penalty ([Chetty et al., 2009](#); [Congdon et al., 2011](#); [Luttmer and Singhal, 2014](#)). We have anecdotal evidence from focus groups showing that taxpayers' reactions to the information that they have to pay a monthly compound interest of 2% and this alternative way of presenting the same information are quite different. While taxpayers dismiss the 2% interest as being low, they become concerned about the size of the penalty when presented with the example. The second component highlighted the additional consequences that the individual might face for not paying: *“and the Municipality can take administrative and legal action.”* This message was accompanied by an image of a gavel, which intensified the idea of the penalty (see [Table 1](#) for the message included in the tax bill and [Fig. 3](#) for an example of a tax bill).

Following the analytical framework developed in [Section 3](#) we can expect the spillover to be either positive, negative or zero. A positive spillover will occur when the deterrence message sent to the taxpayer on their property tax bill increases their beliefs about the penalty and the probability of enforcement in the gross-sales tax. A negative spillover will occur if the taxpayers increase their belief about the severity of the penalty but they now believe that the increased efforts the government is putting on the property tax will reduce the efforts and resources the government can dedicate to enforcing the gross-sales tax. No spillover will exist if the changes in beliefs compensate for each other or if taxpayers do not update their beliefs in one tax after receiving information in another.

<sup>5</sup> The idea behind this procedure was to reduce contamination of the control group. More details about the randomization can be found in [Castro and Scartascini \(2015\)](#).

<sup>6</sup> We excluded around 12% of taxpayers who pay the property tax on a monthly or yearly basis.

<sup>7</sup> We only look at those who received the deterrence treatment for several reasons. First, while we have an analytical framework we can use as a benchmark for the deterrence message, we have no predictions for the other two messages. Second, we only have a few people in the other two treatment groups. Finally, the samples for those groups are not balanced.


<sup>8</sup> We reproduce the analysis of [Castro and Scartascini \(2015\)](#) in the subgroup of sole proprietors ([Table A.3](#)). Given that the differences in sample size are substantial (23,000 to 700) it is expected that results might differ across exercises. Power calculations for the small sample require large differences between control and treatment. The differences in [Castro and Scartascini \(2015\)](#) were around 5 percentage points. As shown in the [Table A.3](#), the differences are between 2 and 6 percentage points: see Columns 1 and 3.

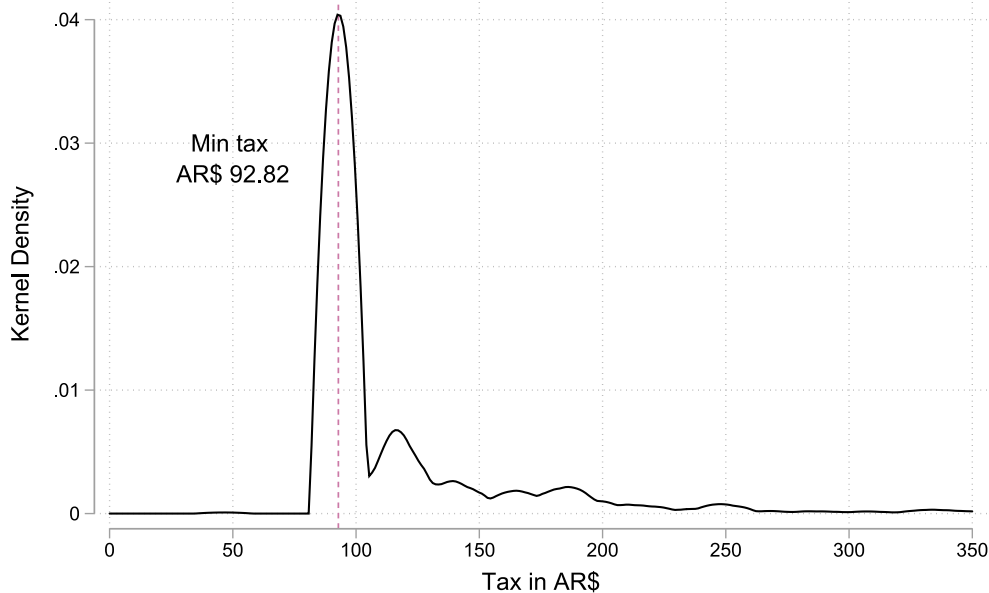
<sup>9</sup> USD\$361.01 PPP.

<sup>10</sup> USD\$96.75 PPP.

**Table 1**

Message included in the property tax bill.

Message/Group	Text	Image
Deterrence	Did you know that if you do not pay the property tax on time for a debt of AR\$ 1000 you will have to pay AR\$ 268 in arrears at the end of the year and the Municipality can take administrative and legal action?	
Control	No message	No image



The top 5% of the observations are excluded from the graph.

Fig. 2. Distribution of the tax before the treatment period (Aug-Jul 2011).

## 5. Empirical strategy

As we described in the previous section, we can exploit the assignment to treatment in [Castro and Scartascini \(2015\)](#) to compare the effect of receiving a deterrence message printed on the bill of the property tax (CVP) on the declaration of the gross-sales tax (SEH). It is important to note that several factors affect the precision of our estimation. First, we have a relatively small treatment group, because the intersection of individuals owning property and having to pay the gross-sales tax is relatively small. Second, we cannot observe reported sales directly, but only the declared tax. The gross-sales tax ( $T_{gs}$ ) is computed by adding a tax rate ( $t_s$ ) times the declared sales ( $\tilde{y}$ ), a tax rate ( $t_e$ ) determined by the number of employees (declared the previous year) times the municipal wage, and a tax rate ( $t_m$ ) determined by the square meters of the establishment (declared the previous year) times a price-per-meter, determined annually by the city government. The first element (declared sales) is the only one that varies within a fiscal year. Consequently, while we cannot observe our variable of interest directly, we can safely assume that a change in the reported tax in any specific period within a calendar year reflects a change in the reported sales. Because declared sales affect only a fraction of the estimated tax, it scales down the overall effect. For the average taxpayer, a 10% change in declared sales implies a 6% change in declared tax. Finally, there is a minimum tax that applies to all taxpayers whose sales are below a certain threshold; that is ( $T_{gs} = \max\{T_{gs}^{\min}, T_{gs}(\tilde{y}, \dots)\}$ ). This minimum tax is binding for a large fraction of taxpayers in our sample because we work with the group of sole proprietors and they tend to smaller businesses than firms. Therefore, the actual distribution of the tax looks truncated compared to what it would have been absent the minimum. As a result, because we cannot observe declared sales directly, we cannot observe the treatment effects on declared sales in the lower part of the distribution (see [Fig. 2](#)). Still, we can observe and measure well the effect of the intervention on actual tax revenues given the tax code.<sup>11</sup> The

<sup>11</sup> Given that we are observing and using declared tax as the dependent variable, it is still appropriate to estimate the model using OLS. A Tobit estimation would overestimate the effect of the intervention on declared taxes. However, if we had declared sales, a Tobit model would be more appropriate.



Fig. 3. Sample tax bills with treatment messages (in Spanish).

minimum tax was updated according to inflation every four months (Table 2). From January to April the minimum tax was ARS\$89.25 (USD\$32.22 PPP), from May to August it was ARS\$92.82 (USD\$33.51 PPP), and from September it was ARS\$96.56 (USD\$34.86 PPP). All of these factors should work against finding any result.

We estimate the minimal detectable effect (MDE) with our sample size and data structure for a significance level of 5% and a power of 0.8. The minimal detectable effect for an OLS estimation with  $\ln(\text{tax})$  as the outcome is 20 percentage points, which is much higher than any result found in the literature. As such, it would be very difficult to find any significant result in such a setting. The MDE becomes more reasonable if we consider instead the first difference of the outcome variable, which becomes our estimation of choice. The power calculations are included in the appendix (Eq. (5) and Table A.2).

To address the challenges generated by the data limitations, including the fact that the original randomization was done in a different and larger sample of taxpayers, our main specification is a difference-in-difference estimator. The difference-



**Table 2**

Descriptive statistics of sole proprietors pre-treatment period (Jul/Aug).

	Retail sector		Other sectors		Total	
Mean annual sales 2010 in ARS\$1,000	274.01	(542.67)	145.59	(290.33)	226.36	(469.16)
Mean number of employees 2010	0.55	(1.10)	0.61	(1.43)	0.57	(1.23)
Mean indoor space in square meters 2010	71.63	(97.99)	131.65	(164.47)	93.90	(129.95)
Mean Gross-sales Tax ARS\$	110.67	(41.58)	113.55	(39.44)	111.74	(40.79)
Percent paid Gross-sales Tax by 1st due date	0.27	(0.45)	0.32	(0.47)	0.29	(0.45)
Percent paid Gross-sales Tax by 2nd due date	0.16	(0.37)	0.13	(0.33)	0.15	(0.36)
Percent Paid Gross-sales Tax in Full	0.70	(0.46)	0.65	(0.48)	0.68	(0.47)
Percent of owners who are men	0.66	(0.47)	0.83	(0.38)	0.72	(0.45)
Mean number of years of the firm	13.12	(10.97)	17.08	(11.17)	14.59	(11.20)
N	417		246		663	

Monetary amounts are in Argentine Pesos (ARS\$). Standard errors are in parentheses.

**Table 3**

Balance test pre-treatment period (Jul/Aug).

	Difference: Deterrence		Control group		N
Ln Tax Gross-sales Tax	0.106	(0.089)	4.817***	(0.026)	723
Ln Tax Gross-sales Tax excluding outliers (1%)	0.036	(0.035)	4.706***	(0.015)	694
1 if retail sector	0.014	(0.051)	0.638***	(0.024)	723
1 if industry	-0.036***	(0.009)	0.044***	(0.007)	723
Annual sales 2010 in ARS\$1,000	36.292	(53.967)	220.454***	(25.439)	669
Num. of employees 2010	0.278*	(0.165)	0.532***	(0.058)	669
Num. of proprietors working 2010	0.036	(0.024)	1.002***	(0.004)	669
Indoor space m2	22.520	(13.920)	91.085***	(6.762)	669
Outdoor space m2	3.010	(3.551)	4.666***	(1.189)	669
Paid Gross-sales Tax by 1st date	0.034	(0.042)	0.288***	(0.027)	723
Paid Gross-sales Tax by 2nd date	-0.012	(0.029)	0.151***	(0.009)	723
Paid Gross-sales Tax in Full	0.028	(0.034)	0.680***	(0.022)	717
Paid Property Tax by 1st date	0.014	(0.042)	0.334***	(0.032)	723
Paid Property Tax by 2nd date	-0.019	(0.031)	0.150***	(0.018)	723
Paid Property Tax in Full	0.055	(0.054)	0.597***	(0.039)	723
Num. lights	0.018	(0.153)	2.955***	(0.101)	723
Manual Sweeping	-0.014	(0.059)	0.414***	(0.077)	723
Mechanical Sweeping	-0.008	(0.066)	0.408***	(0.066)	723
Ln front to street	0.007	(0.067)	2.555***	(0.038)	723
1 if paid Property Tax monthly	-0.005*	(0.003)	0.005*	(0.003)	723

Each row shows a regression of the variable on the treatment. Monetary amounts are in Argentine Pesos (ARS\$). Standard errors are clustered at the randomization block level and in parentheses.

\* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

in-difference design allows us to compare the treatment group over time by controlling the time trend and taking advantage of the panel nature of our data. We estimate the following equation

$$y_{it} = \alpha_0 + \alpha_1 T_i + \gamma t_{Sep/Oct} + \delta D_{it} + X'_{it} \beta + \varepsilon_{it} \quad (3)$$

where the variables are defined as follows.  $y_{it}$  is the variable of interest, the log of the gross-sales reported tax for individual  $i$  in period  $t$ .  $T_i$  is one if the taxpayer received the deterrence letter for the property tax.  $t_{Sep/Oct}$  is the time fix effect equal to one for the fifth pay period (Sep/Oct) and zero from the fourth pay period (Jul/Aug).  $D_{it}$  is the difference-in-difference estimator (interaction of  $T_i$  and  $t_{Sep/Oct}$ ).  $X'_{it}$  is a vector of controls that include characteristics of the business, such as: the annual sales of the previous year, the economic sector, binary variables for the number of employees and the size of the store, which correspond to the categories that are used to calculate the tax, the age of the firm and the gender of the owner. Following [Castro and Scartascini \(2015\)](#), because compliance is highly geographically clustered, we also include the randomization blocks fix effect, and we cluster the standard errors by the same blocks. As discussed in the original paper as well, compliance shows great persistence, so we include a lagged outcome variable.<sup>12</sup>

In addition, we estimate the effect of the treatment on the probability of paying more than the minimum tax using a linear probability model and a probit.<sup>12</sup> Basically, we estimate the following:

<sup>12</sup> In [Castro and Scartascini \(2015\)](#), the probability of paying in period  $t$  given that the taxpayer had paid in  $t-1$  is close to 100%. Similarly, [Dwenger et al. \(2016\)](#) find that those who evaded in 2010 were 87 times more likely to evade in 2011. Adding a lagged outcome in a panel could bias the estimator, but it is not the case in our estimation because the treatment assignment was random, so it is uncorrelated with the outcome of the previous period. Including lagged dependent variables can generate a biased estimator because the residual is correlated with the lagged dependent variable. However, the treatment was randomly assigned, so the beta estimator is consistent since the  $cov(y_{it-1}, D_{it}) = 0$  by construction, because individuals/taxpayers under treatment were selected at random. A discussion of the problem of fixed effect estimators and lagged variables can be found in [Bertrand et al. \(2004\)](#), [Imbens and Wooldridge \(2008\)](#) and [Angrist and Pischke \(2008\)](#). The derivation of the formula for  $\beta$  is in the Appendix.

**Table 4**

Effect of the deterrence letter on the reported tax dependent variable: Ln of the gross-sales tax.

	(1)	(2)	(3)	(4)	(5)	(6)
T: Deterrence	−0.016 (0.010)	−0.014 (0.010)	−0.012 (0.009)	−0.013 (0.010)	−0.014 (0.010)	−0.012 (0.009)
After (t: Sep/Oct)	0.028*** (0.001)	0.031*** (0.001)	0.031*** (0.001)	0.031*** (0.001)	0.031*** (0.001)	0.031*** (0.001)
T: Deterrence x after	0.022* (0.011)	0.021** (0.010)	0.020** (0.010)	0.022* (0.011)	0.021** (0.010)	0.020** (0.010)
Ln Tax Gross-sales $t - 1$	0.990*** (0.002)	0.897*** (0.032)	0.880*** (0.038)	0.931*** (0.027)	0.901*** (0.034)	0.884*** (0.041)
1 if paid the min tax $t - 1$				−0.005 (0.005)	0.009** (0.004)	0.010** (0.004)
Annual sales 2010 100,000 ARSS				0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
1 if owner is male				0.004 (0.004)	0.004 (0.005)	0.003 (0.004)
Age of firm Jan 2012 in years				−0.000 (0.000)	−0.000 (0.000)	−0.000 (0.000)
Constant	0.049*** (0.011)	0.470*** (0.144)	0.573*** (0.184)	0.322** (0.129)	0.441*** (0.153)	0.546** (0.195)
N	1433	1326	1326	1326	1326	1326
Random. Blocks FE	Yes	Yes	Yes	Yes	Yes	Yes
Size Dummies	No	Yes	Yes	No	Yes	Yes
Sector Dummies	No	No	Yes	No	No	Yes

Monetary amounts are in Argentine Pesos (ARSS). Standard errors clustered by randomization block are in parentheses. In specifications from 3 onwards, we include binary variables for the economic sector, and from 4 to 6 we include binary variables for the bins of the tables of the number of employees and the size of the store in square meters.

\* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

In addition, we estimate the effect of the treatment on the probability of paying more than the minimum tax using a linear probability model and a probit.<sup>13</sup> Basically, we estimate the following:

$$Prob(y = 1|X) = \Phi(X\beta + \delta T + \varepsilon) \quad (4)$$

where  $y$  is the binary outcome equal to one if the taxpayer declared a gross-sales tax larger than the minimum,  $X$  is a vector of controls that includes binary variables for the economic sector, business characteristics, the annual sales of 2010, the age of the firm in years, the gender of the proprietor, and fixed effects for randomization blocks.

## 6. Results

The main question of this paper is whether enforcement in one tax creates positive or negative spillovers in other taxes. The evidence coming from the intervention we evaluate seems to reject the hypothesis that taxpayers reduce compliance with other taxes. If anything, the evidence seems to be suggestive of a positive spillover. In particular, we find in a difference-in-difference estimation that the treatment group increases its reported tax on average by 2 percentage points more than the control group (Table 4). The coefficient is stable across specifications with different control variables and statistically significant at the 5% level. In addition to running several different specifications to check the stability of results, we run a placebo regression for the period before the intervention took place, and we find no effect (Table A.4).<sup>14</sup> We also estimate the effect of the treatment in reporting a tax larger than the minimum, and we find that taxpayers in the treatment group are between 7% and 9% more likely to report a tax larger than the minimum than taxpayers in the control group (Table 5). The result is statistically significant at the 10% level for some specifications.<sup>15</sup>

In summary, given the tentative evidence that we have presented, we can conclude that the taxpayers who received the treatment in the property tax declared more and were more likely to pay their gross-sales taxes than those in the control group. This evidence suggests that taxpayers change their beliefs for other taxes when they receive information about deterrence in one tax. It also suggests that taxpayers update those beliefs assuming that higher deterrence in one tax translates equally to other taxes.

Again, it is important to note that our results may well be underestimating the true results for several reasons: (i) the tax is computed according to the declared sales over a two-month period. Most of those in the treatment group could have

<sup>13</sup> We thank an anonymous referee for suggesting we incorporate this analysis into the paper.

<sup>14</sup> If we just compare the treatment and control group only in the post-treatment period, the coefficient is again quite stable and positive across specifications and of similar magnitude—to compare the OLS specification with the difference-in-difference specification, the coefficients of the treatment and the treatment times the period should be added—but in the cross-section estimation, the difference is not statistically significant (Table A.6), probably due as expected to our small sample size (see the power calculation in Table A.2). Again, no results exist in a placebo exercise (Table A.7).

<sup>15</sup> As we add controls and lose observations, statistical significance drops; see previous comments about sample size.

**Table 5**

Effect of the deterrence letter on the probability of reporting more than the minimum tax.

	(1) LPM	(2) Probit-AME	(3) LPM	(4) Probit-AME
T: Deterrence	0.095** (0.045)	0.094** (0.048)	0.071* (0.039)	0.046 (0.029)
N	722	717	669	646
Random. Blocks FE	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes
Period	Sep/Oct	Sep/Oct	Sep/Oct	Sep/Oct

The dependent variable takes the value of one if the declared tax is larger than the minimum tax. In specifications from three onwards, we include binary variables for the economic sector, the bins of the tables of the number of employees and the size of the store in square meters. The business characteristics we include as controls are the annual sales of 2010, the age of the firm in years and the gender of the proprietor. Specifications 1 and 3 show the coefficient of a Linear Probability Model. Specifications 2 and 4 show the average marginal effect from a Probit Model. Monetary amounts are in Argentine Pesos (ARS\$). Standard errors clustered by randomization block are in parentheses.

\* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

received the message after the first month's declaration. Thus, the change in declaration might be one half of what it could otherwise have been. (ii) The declared tax—the variable we observe—is only partially affected by the level of declared sales, which also reduces the size of the estimates. (iii) Many of the taxpayers pay the minimum tax; if there is any effect in this group, we may be unable to observe their response.

## 7. Conclusion

The empirical literature on tax compliance has grown exponentially in the last few years. Greater access to administrative data, a better predisposition of authorities toward impact evaluations, and the relatively low cost of behavioral interventions have made this possible. However, most of this literature has focused almost exclusively on the direct effect of the interventions. However, an intervention could have effects on compliance beyond the tax under study (spillover effects). If spillovers are negative, they can reduce or completely negate the impact of the intervention. Consequently, it is very important to understand the conditions that determine the existence and sign of the spillover. Our simple analytical model predicts that the size and sign of the spillover depends on: (i) the effect of the deterrence message on the salience of the penalty, and (ii) the effect of the deterrence message on how people evaluate the ability of the government to enforce several taxes at the same time. If taxpayers think that enforcement in one tax implies higher enforcement in all taxes, spillovers will most likely be positive. If taxpayers think that higher enforcement in one tax implies lower enforcement in other taxes because resources are limited, then spillovers should be zero or negative. Cash or financial constraints could exacerbate the negative spillover. This simple model can help to explain the seemingly contradictory results found in [Ortega and Scartascini \(2015\)](#), [Carrillo et al. \(2017a\)](#) and [Slemrod et al. \(2017\)](#).

The evidence in this paper, which combines data from a treatment designed to increase compliance with the property tax and gross-sales tax declarations, shows positive spillover effects. This evidence is in line with the model, since the main component of the treatment was highlighting the size of the penalties, which are the same across both taxes.

Given that most taxpayers in most countries are liable for more than one tax, tax authorities should design their control strategies taking into account the possible spillover effect across taxes as well as the fact that penalties and detection may not be interchangeable policies. Moreover, tax authorities should be mindful of the signal that their enforcement strategy sends. If taxpayers evaluate that the resources of the tax authority are limited, then increasing detection in one tax may lead to reductions in compliance with others. In any case, ignoring the interconnectivity of compliance across taxes is inadvisable.

This study is also a cautionary tale for optimal tax policy design. Taxpayers who are liable for several taxes might be different from taxpayers who are not in regards to risk perception and budget constraints. In order to get a full picture of the effects of any intervention, it is important to analyze the taxes that are not the main target of the intervention as well. Researchers should also be well aware of this when designing their interventions in order to make sure that if spillovers are possible, they will not have a negative effect.

Finally, it is important to note that this paper raises several important points regarding the analytical determinants of spillovers and the impact they can have on actual interventions. Still, this is not the last word but only a building stone upon which future studies should build. More sophisticated models that take into account other taxes and strategies as well as empirical papers that have fewer data constraints than this one are encouraged.

## Appendix A

### A1. SEH tax definitions and tables

The SEH tax has three components that correspond based on the gross-sales, on the number of employees and on the building size.

$$T_{SEH} = T_{Bim}^{sale} + T_{y-1}^{employees} + T_{y-1}^{buildingsize}$$

**Table A.1**  
Brackets for the components of the gross-sales tax.

Volume of sales							
Range	0 to \$6,000	\$6,001 to \$10,000	\$10,001 to \$18,000	\$18,001 to \$30,000	\$30,001 to \$80,000	\$80,001 to \$150,000	Higher than to \$150,001
<b>Industry</b>							
Food	0.136%	0.190%	0.285%	0.456%	0.798%	1.556%	3.423%
Goods	0.114%	0.160%	0.240%	0.384%	0.672%	1.310%	2.882%
Other	0.125%	0.175%	0.263%	0.421%	0.737%	1.437%	3.161%
<b>Whole commerce</b>							
Food	0.125%	0.175%	0.263%	0.421%	0.737%	1.437%	3.161%
Goods	0.105%	0.147%	0.221%	0.354%	0.620%	1.209%	2.660%
Other	0.115%	0.161%	0.242%	0.387%	0.677%	1.320%	2.904%
<b>Retail</b>							
Food	0.109%	0.153%	0.230%	0.368%	0.644%	1.256%	2.763%
Goods	0.091%	0.127%	0.191%	0.306%	0.536%	1.045%	2.299%
Other	0.100%	0.140%	0.210%	0.336%	0.588%	1.147%	2.523%
<b>Services</b>							
Personal	0.100%	0.140%	0.210%	0.336%	0.588%	1.147%	2.523%
Others	0.091%	0.127%	0.191%	0.306%	0.536%	1.045%	2.299%
<b>Number of employees</b>							
Range	1	2 to 3	4 to 7	8 to 15	16 to 30	31 to 100	More than 101
<b>Industry</b>							
Food	5.924%	7.109%	9.597%	14.396%	23.034%	39.158%	72.442%
Goods	5.129%	6.155%	8.309%	12.464%	19.942%	33.901%	62.717%
Other	5.386%	6.463%	8.725%	13.088%	20.941%	35.600%	65.860%
<b>Whole commerce</b>							
Food	4.937%	5.924%	7.997%	11.996%	19.194%	32.630%	60.366%
Goods	4.274%	5.129%	6.924%	10.386%	16.618%	28.251%	52.264%
Other	4.488%	5.386%	7.271%	10.907%	17.451%	29.667%	54.884%
<b>Retail</b>							
Food	4.114%	4.937%	6.665%	9.998%	15.997%	27.195%	50.311%
Goods	3.562%	4.274%	5.770%	8.655%	13.848%	23.542%	43.553%
Other	3.740%	4.488%	6.059%	9.089%	14.542%	24.721%	45.734%
<b>Services</b>							
Personal	3.740%	4.488%	6.059%	9.089%	14.542%	24.721%	45.734%
Others	3.400%	4.080%	5.508%	8.262%	13.219%	22.472%	41.573%
<b>Surface in Square Meters</b>							
Range	0 to 40	41 to 60	61 to 90	91 to 120	81 to 120	501 to 1,500	More than 1,501
<b>Industry</b>							
Food	5.032%	6.038%	8.151%	12.227%	19.664%	33.429%	61.844%
Goods	4.375%	5.250%	7.088%	10.632%	17.011%	28.919%	53.500%
Other	4.594%	5.513%	7.443%	11.163%	17.864%	30.369%	56.183%
<b>Whole commerce</b>							
Food	4.193%	5.032%	6.793%	10.190%	16.304%	27.717%	51.276%
Goods	3.646%	4.375%	5.906%	8.859%	14.174%	24.096%	44.578%
Other	3.828%	4.594%	6.202%	9.303%	14.885%	25.305%	46.814%
<b>Retail</b>							
Food	3.494%	4.193%	5.661%	8.492%	13.587%	23.098%	42.731%
Goods	3.038%	3.646%	4.922%	7.383%	11.813%	20.082%	37.152%
Other	3.190%	3.828%	5.168%	7.752%	12.403%	21.085%	39.007%
<b>Services</b>							
Personal	3.190%	3.828%	5.168%	7.752%	12.403%	21.085%	39.007%
Others	2.900%	3.480%	4.698%	7.047%	11.275%	19.168%	65.461%

$T_{Bim}^{sale}$  is calculated by multiplying the total sales of the two-month period by the tax rate.  $T_{y-1}^{employees}$  is the result of the product the tax rate determined by number of employees (paid or unpaid) who worked last year for the businesses, times the city government administrative wage.  $T_{y-1}^{building\ size}$  is the tax rate determined by the indoor space and half of the outdoor space in square meters reported last year, times the cost of a meter of construction. The tax rates according to economic activity and size are described in Table A.1. The city government determines the administrative wage and the cost of a meter of construction in the city by January of each year.

A2. First and second order conditions of the model

$$\max_{\tilde{y}} : (1 - P_s(E, P_r(E)))U(y - t\tilde{y}) + P_s(E, P_r(E))U(y - t\tilde{y} - \theta t(y - \tilde{y}))$$

For notation convenience  $X = y - t\tilde{y}$  and  $\hat{X} = y - t\tilde{y} - \theta t(y - \tilde{y})$ . Let denote  $V$  as the expected utility function. The first order conditions ( $\frac{\partial V}{\partial \tilde{y}} = V' = 0$ ) is:

$$-t(1 - P_s(E, P_r(E)))U'(X) + tP_s(E, P_r(E))U'(\hat{X})(\theta - 1) = 0$$

**Table A.2**  
Power calculation.

	ln tax $y=\alpha+\beta T+u$	$\Delta$ ln tax: $\Delta y=\alpha+\beta T+u$	\$ tax $y=\alpha+\beta T+u$	$\Delta$ \$ tax: $\Delta y=\alpha+\beta T+u$
$\sigma^2(y)$	0.449	0.003	2284.31	93.38
$\rho$	0.006	0.008	0.001	0.008
$\hat{\beta}_{MDE}$	0.191	0.015	13.607	2.751
$a$	1.078	1.106	1.007	1.11
$\hat{\beta}_{MDE} \times a$	0.206	0.017	13.707	3.054
$t_{\frac{\alpha}{2}}$	1.96		$p_T$	0.16
$t_{1-\kappa}$	0.84		$p_C$	0.84
$N$	723		$N_{cluster}$	25

Since the utility function is concave the second order conditions are satisfied:

$$D = t^2(1 - P_s(E, P_r(E)))U''(X) + t^2P_s(E, P_r(E))U''(\hat{X})(\theta - 1)^2 \leq 0$$

### A3. Comparative statics

Differentiating the first order conditions with respect to  $\theta$  and solving for  $\frac{\partial \tilde{y}}{\partial \theta}$ , where  $D = \frac{\partial^2 V}{\partial y^2}$ :

$$\frac{\partial \tilde{y}}{\partial \theta} = \frac{-t[P_s(E, P_r(E))U'(\hat{X})]}{D}$$

$$\text{sign}\left[\frac{\partial \tilde{y}}{\partial \theta}\right] = \text{sign}[tP_s(E, P_r(E))U'(\hat{X})]$$

$$\frac{\partial \tilde{y}}{\partial \theta} > 0$$

Differentiating the first order conditions with respect to  $P_r(E)$  and solving for  $\frac{\partial \tilde{y}}{\partial P_r(E)}$ :

$$\frac{\partial \tilde{y}}{\partial P_r(E)} = \frac{-1}{D} [tU'(X) + tU'(\hat{X})(\theta - 1)] \frac{\partial P_s(E, P_r(E))}{\partial P_r(E)}$$

$$\text{sign}\left[\frac{\partial \tilde{y}}{\partial P_r(E)}\right] = \text{sign}\left[\frac{\partial P_s(E, P_r(E))}{\partial P_r(E)}\right]$$

### A4. Power calculation

$$\hat{\beta}_{MDE} = (t_{\frac{\alpha}{2}} + t_{1-\kappa}) \sqrt{\left(\frac{1}{p_T(1 - p_T)}\right) \frac{\sigma^2(y)}{N}} \quad (5)$$

$$a = \sqrt{1 + \left(\frac{N}{N_c} - 1\right)\rho} \quad (6)$$

Where  $\alpha$  is the significance level,  $\kappa$  is the power,  $p_T$  is the proportion of individuals in the treatment group,  $p_C$  is the proportion of individuals in the control group,  $\sigma^2(y)$  is the variance of the outcome,  $N$  is the number of observations,  $N_c$  is the number of clusters, and  $\rho$  is the intracluster correlation.

### A5. Difference-in-difference estimator

$$y_{it} = \alpha_0 + \alpha_1 T_d + \gamma t_{bim5} + \delta D_{it} + \theta y_{it-1} + X'_{it} \beta + \varepsilon_{it}$$

$$\hat{\beta} = \frac{\text{var}(y_{it-1})\text{cov}(y_{it}, D_{it}) - \text{cov}(y_{it-1}, D_{it})\text{cov}(y_{it}, y_{it-1})}{\text{var}(D_{it})\text{var}(y_{it-1}) - \text{cov}(y_{it-1}, D_{it})^2}$$

Notice that  $cov(y_{it-1}, D_{it}) = 0$  because the treatment was random. So,  $\hat{\beta}$  becomes

$$\hat{\beta} = \frac{var(y_{it-1})cov(y_{it}, D_{it})}{var(D_{it})var(y_{it-1})}$$

$$\hat{\beta} = \frac{cov(y_{it}, D_{it})}{var(D_{it})}$$

$$plim[\hat{\beta}] = plim\left[\frac{cov(y_{it}, D_{it})}{var(D_{it})}\right]$$

$$plim\hat{\beta} = \beta$$

A6. Tables

**Table A.3**

Effect of the deterrence letter on the probability of paying each tax according to the estimation by Castro and Scartascini (2015) tax: in the title of each column.

	(1) Property	(2) Sales	(3) Property	(4) Sales
T: Deterrence	0.060 (0.048)	0.035 (0.035)	0.019 (0.016)	0.021 (0.022)
N	722	718	718	658
Random. Blocks FE	Yes	Yes	Yes	Yes
Lagged output	No	No	Yes	Yes
Controls	No	No	Yes	Yes
Period	Sep/Oct	Sep/Oct	Sep/Oct	Sep/Oct

The dependent variable would take the value one only if the taxpayer paid in full the total tax liabilities for the period of the experiment. The tax is identified in the header. The controls are binary variables for the sector, indicators for having paid the minimum variables in the previous period, variables from the annual declaration of 2010 (annual sales, binary variables for the bins of the tables of the number of employees and the size of the store in square meters), the age of the firm in years and the gender of the proprietor. Monetary amounts are in Argentine Pesos (AR\$). Standard errors clustered by randomization block are in parentheses.  
\*p < 0.10, \*\*p < 0.05, \*\*\*p < 0.01.

**Table A.4**

Effect of the deterrence letter on the reported tax - placebo test dependent variable: Ln of the gross-sales tax.

	(1)	(2)	(3)	(4)	(5)	(6)
T: Deterrence	-0.010 (0.015)	-0.001 (0.002)	0.000 (0.003)	-0.000 (0.002)	-0.000 (0.003)	0.001 (0.003)
After placebo (t: Jul/Aug)	-0.042*** (0.006)	-0.028*** (0.002)	-0.028*** (0.002)	-0.029*** (0.001)	-0.029*** (0.001)	-0.028*** (0.001)
T: Deterrence x after placebo	-0.006 (0.018)	-0.014 (0.011)	-0.014 (0.011)	-0.014 (0.012)	-0.014 (0.011)	-0.014 (0.011)
Ln Tax Gross-sales t - 1	1.007*** (0.009)	0.915*** (0.029)	0.905*** (0.035)	0.949*** (0.026)	0.921*** (0.033)	0.910*** (0.039)
1 if paid the min tax t - 1				0.002 (0.006)	0.012** (0.005)	0.013** (0.005)
Annual sales 2010 100,000 AR\$				0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
1 if owner is male				0.004 (0.004)	0.003 (0.004)	0.003 (0.004)
Age of firm Jan 2012 in years				-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Constant	0.008 (0.039)	0.416*** (0.130)	0.483*** (0.171)	0.266** (0.122)	0.375** (0.147)	0.448** (0.189)
N	1431	1322	1322	1322	1322	1322
Random. Blocks FE	Yes	Yes	Yes	Yes	Yes	Yes
Size Dummies	No	Yes	Yes	No	Yes	Yes
Sector Dummies	No	No	Yes	No	No	Yes

Monetary amounts are in Argentine Pesos (AR\$). Standard errors clustered by randomization block are in parentheses. In specifications from 3 onwards, we include binary variables for the economic sector, and from 4 to 6 we include binary variables for the bins of the tables of the number of employees and the size of the store in square meters.

\*p < 0.10, \*\*p < 0.05, \*\*\*p < 0.01.

**Table A.5**

Effect of the deterrence letter on the probability of reporting more than the minimum tax - placebo test.

	(1) LPM	(2) Probit-AME	(3) LPM	(4) Probit-AME
T: Deterrence	0.071 (0.043)	0.070 (0.049)	0.050 (0.035)	0.029 (0.026)
N	723	718	669	646
Random. Blocks FE	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes
Period	Jul/Aug	Jul/Aug	Jul/Aug	Jul/Aug

The dependent variable takes the value of one if the declared tax is larger than the minimum tax. In specifications from three onwards, we include binary variables for the economic sector, the bins of the tables of the number of employees and the size of the store in square meters. The business characteristics we include as controls are the annual sales of 2010, the age of the firm in years and the gender of the proprietor. Specifications 1 and 3 show the coefficient of a Linear Probability Model. Specifications 2 and 4 show the average marginal effect from a Probit Model. Monetary amounts are in Argentine Pesos (AR\$). Standard errors clustered by randomization block are in parentheses.

\* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

**Table A.6**

Effect of the deterrence letter on the reported tax - OLS estimation dependent variable: Ln of the gross-sales tax.

	(1)	(2)	(3)	(4)	(5)	(6)
T: Deterrence	0.006 (0.006)	0.007 (0.005)	0.008 (0.005)	0.008 (0.006)	0.007 (0.005)	0.009 (0.005)
Ln Tax Gross-sales $t - 1$	0.984*** (0.003)	0.900*** (0.030)	0.883*** (0.038)	0.937*** (0.025)	0.913*** (0.032)	0.896*** (0.039)
1 if paid the min tax $t - 1$				0.002 (0.006)	0.015** (0.007)	0.017** (0.007)
Annual sales 2010 100,000 AR\$				0.000 (0.001)	0.000 (0.001)	0.000 (0.001)
1 if owner is male				0.002 (0.004)	0.001 (0.005)	0.001 (0.005)
Age of firm Jan 2012 in years				0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Constant	0.105*** (0.013)	0.489*** (0.135)	0.596*** (0.181)	0.321** (0.118)	0.413*** (0.145)	0.520** (0.189)
N	718	665	665	665	665	665
Random. Blocks FE	Yes	Yes	Yes	Yes	Yes	Yes
Size Dummies	No	Yes	Yes	No	Yes	Yes
Sector Dummies	No	No	Yes	No	No	Yes
Period	Sep/Oct	Sep/Oct	Sep/Oct	Sep/Oct	Sep/Oct	Sep/Oct

Monetary amounts are in Argentine Pesos (AR\$). Standard errors clustered by randomization block are in parentheses. In specifications from 3 onwards, we include binary variables for the economic sector, and from 4 to 6 we include binary variables for the bins of the tables of the number of employees and the size of the store in square meters.

\* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

**Table A.7**

Effect of the deterrence letter on the Rreported tax - OLS placebo test dependent variable: Ln of the gross-sales tax.

	(1)	(2)	(3)	(4)	(5)	(6)
T: Deterrence	-0.016 (0.010)	-0.014 (0.009)	-0.012 (0.007)	-0.013 (0.009)	-0.013 (0.009)	-0.012 (0.007)
Ln Tax Gross-sales $t - 1$	0.996*** (0.003)	0.895*** (0.058)	0.880*** (0.069)	0.926*** (0.045)	0.890*** (0.061)	0.875*** (0.072)
1 if paid the min tax $t - 1$				-0.013 (0.008)	0.001 (0.006)	0.003 (0.007)
Annual sales 2010 100,000 AR\$				0.002 (0.002)	0.002 (0.002)	0.002 (0.002)

(continued on next page)

Table A.7 (continued)

	(1)	(2)	(3)	(4)	(5)	(6)
1 if owner is male				0.007 (0.008)	0.006 (0.007)	0.006 (0.006)
Age of firm Jan 2012 in years				–0.000 (0.000)	–0.000 (0.000)	–0.000 (0.000)
Constant	0.021 (0.015)	0.475* (0.259)	0.573 (0.338)	0.354 (0.213)	0.495* (0.271)	0.595 (0.349)
N	715	661	661	661	661	661
Random. Blocks FE	Yes	Yes	Yes	Yes	Yes	Yes
Size Dummies	No	Yes	Yes	No	Yes	Yes
Sector Dummies	No	No	Yes	No	No	Yes
Period	Jul/Aug	Jul/Aug	Jul/Aug	Jul/Aug	Jul/Aug	Jul/Aug

Monetary amounts are in Argentine Pesos (ARSS). Standard errors clustered by randomization block are in parentheses. In specifications from 3 onwards, we include binary variables for the economic sector, and from 4 to 6 we include binary variables for the bins of the tables of the number of employees and the size of the store in square meters.

\* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

Table A.8

Effect of the deterrence letter on the probability of paying each tax according to the estimation by Castro and Scartascini (2015) - placebo test tax: in the title of each column.

	(1) Property	(2) Sales	(3) Property	(4) Sales
T: Deterrence	0.052 (0.048)	0.023 (0.031)	–0.002 (0.020)	–0.005 (0.019)
N	723	717	715	654
Random. Blocks FE	Yes	Yes	Yes	Yes
Lagged output	No	No	Yes	Yes
Controls	No	No	Yes	Yes
Period	Jul/Aug	Jul/Aug	Jul/Aug	Jul/Aug

The dependent variable would take the value one only if the taxpayer paid in full the total tax liabilities for the period of the experiment. The tax is identified in the header. The controls are binary variables for the sector, indicators for having paid the minimum tax in the previous period, variables from the annual declaration of 2010 (annual sales, binary variables for the bins of the tables of the number of employees and the size of the store in square meters), the age of the firm in years and the gender of the proprietor. Monetary amounts are in Argentine Pesos (ARSS). Standard errors clustered by randomization block are in parentheses.

\* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

## References

- Abeler, J., Jäger, S., 2015. Complex tax incentives. *Am. Econ. J.* 7 (3), 1–28. doi:10.1257/pol.20130137.
- Advani, A., Elming, W., Shaw, J., 2017. *The Dynamic Effects of Tax Audits*. IFS Working Paper W17/24. Institute for Fiscal Studies.
- Allingham, M.G., Sandmo, A., 1972. Income tax evasion: a theoretical analysis. *J. Public Econ.* 1 (3), 323–338. doi:10.1016/0047-2727(72)90010-2.
- Angrist, J.D., Pischke, J.-S., 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.
- Bérgolo, M., Ceni, R., Cruces, G., Giacobasso, M., Perez-Truglia, R., 2018. Misperceptions about tax audits. *AEA Pap. Proc.* 108, 83–87. doi:10.1257/pandp.20181039.
- Bernheim, B.D., Rangel, A., 2007. Behavioral public economics: welfare and policy analysis with non-standard decision makers. In: Diamond, P., Vartiainen, H. (Eds.), *Economic Institutions and Behavioral Economics*. Princeton University Press, Princeton, pp. 7–77.
- Bernheim, B.D., Rangel, A., 2009. Beyond revealed preference: choice-theoretic foundations for behavioral welfare economics. *Quart. J. Econ.* 124 (1), 51–104. doi:10.1162/qjec.2009.124.1.51.
- Bertrand, M., Duflo, E., Mullainathan, S., 2004. How much should we trust differences-in-differences estimates? *Quart. J. Econ.* 119 (1), 249–275.
- Boning, W.C., Guyton, J., Ronald H. Hodge, I.I., Slemrod, J., Troiano, U., 2018. Heard it Through the Grapevine: Direct and Network Effects of a Tax Enforcement Field Experiment. National Bureau of Economic Research, Inc. Working Paper.
- Brockmeyer, A., Kettle, S., Smith, S.D., 2016. Casting the Tax Net Wider: Experimental Evidence from Costa Rica. SSRN Scholarly Paper. Social Science Research Network, Rochester, NY.
- Carrillo, P., Pomeranz, D., Singhal, M., 2017. Dodging the taxman: firm misreporting and limits to tax enforcement. *Am. Econ. J.* 9 (2), 144–164. doi:10.1257/app.20140495.
- Carrillo, P.E., Castro, E., Scartascini, C., 2017. Do Rewards Work?: Evidence from the Randomization of Public Works. Working Papers. Inter-American Development Bank IDB-WP-794 doi:10.18235/0000673.
- Castro, L., Scartascini, C., 2015. Tax compliance and enforcement in the pampas evidence from a field experiment. *J. Econ. Behav. Organ.* 116, 65–82. doi:10.1016/j.jebo.2015.04.002.
- Chetty, R., Looney, A., Kroft, K., 2009. Salience and taxation: theory and evidence. *Am. Econ. Rev.* 99 (4), 1145–1177. doi:10.1257/aer.99.4.1145.
- Chirico, M., Inman, R.P., Loeffler, C., MacDonald, J., Sieg, H., 2016. An experimental evaluation of notification strategies to increase property tax compliance: free-riding in the city of brotherly love. *Tax Policy Econ.* 30 (1), 129–161. doi:10.1086/685595.
- Congdon, W.J., Kling, J.R., Mullainathan, S., 2011. *Policy and Choice: Public Finance through the Lens of Behavioral Economics*. Brookings Institution Press, Washington, D.C.
- DeBacker, J., Heim, B., Tran, A., Yuskavage, A., 2015. Legal enforcement and corporate behavior: an analysis of tax aggressiveness after an audit. *J. Law Econ.* 58 (2), 291–324. doi:10.1086/684037.
- Doerrenberg, P., Schmitz, J., 2017. Tax compliance and information provision. A field experiment with small firms. *J. Behav. Econ. Policy* 1 (1), 47–54.



- Drago, F., Mengel, F., Traxler, C., 2015. Compliance Behavior in Networks: Evidence from a Field Experiment. Working Paper. Centre for Studies in Economics and Finance (CSEF), University of Naples, Italy. 419.
- Dwenger, N., Kleven, H., Rasul, I., Rincke, J., 2016. Extrinsic and intrinsic motivations for tax compliance: evidence from a field experiment in Germany. *Am. Econ. J.* 8 (3), 203–232. doi:10.1257/pol.20150083.
- Fellner, G., Sausgruber, R., Traxler, C., 2013. Testing enforcement strategies in the field: threat, moral appeal and social information. *J. Eur. Econ. Assoc.* 11 (3), 634–660. doi:10.1111/jeea.12013.
- Hallsworth, M., 2014. The use of field experiments to increase tax compliance. *Oxf. Rev. Econ. Policy* 30 (4), 658–679. doi:10.1093/oxrep/gru034.
- Imbens, G.M., Wooldridge, J.M., 2008. Recent Developments in the Econometrics of Program Evaluation. Working Paper. National Bureau of Economic Research 14251. doi:10.3386/w14251.
- Kleven, H.J., Knudsen, M.B., Kreiner, C.T., Pedersen, S., Saez, E., 2011. Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark. *Econometrica* 79 (3), 651–692.
- LaLumia, S., Sallee, J.M., 2013. The value of honesty: empirical estimates from the case of the missing children. *Int. Tax Public Finance* 20 (2), 192–224.
- Luttmer, E.F.P., Singhal, M., 2014. Tax morale. *J. Econ. Perspect.* 28 (4), 149–168. doi:10.1257/jep.28.4.149.
- Maciejovsky, B., Kirchler, E., Schwarzenberger, H., 2007. Misperception of chance and loss repair: on the dynamics of tax compliance. *J. Econ. Psychol.* 28 (6), 678–691. doi:10.1016/j.joep.2007.02.002.
- Mascagni, G., 2018. From the lab to the field: a review of tax experiments. *J. Econ. Surv.* 32 (2), 273–301. doi:10.1111/joes.12201.
- Meiselman, B.S., 2018. Ghostbusting in Detroit: evidence on nonfilers from a controlled field experiment. *J. Public Econ.* 158, 180–193. doi:10.1016/j.jpubeco.2018.01.005.
- Mittone, L., 2006. Dynamic behaviour in tax evasion: an experimental approach. *J. Socio-Econ.* 35 (5), 813–835. doi:10.1016/j.socsec.2005.11.065.
- Naritomi, J., 2019. Consumers as tax auditors. *Am. Econ. Rev.* (Forthcoming). doi:10.1257/aer.20160658.
- Ortega, D., Scartascini, C., 2015. Don't blame the messenger: a field experiment on delivery methods for increasing tax compliance. IDB Working Paper Series. Inter-American Development Bank IDB-WP-627. doi:10.18235/0000204.
- Pomeranz, D., 2015. No taxation without information: deterrence and self-enforcement in the value added tax. *Am. Econ. Rev.* 105 (8), 2539–2569. doi:10.1257/aer.20130393.
- Rinke, J., Traxler, C., 2011. Enforcement spillovers. *Rev. Econ. Stat.* 93 (4), 1224–1234.
- Slemrod, J., Blumenthal, M., Christian, C., 2001. Taxpayer response to an increased probability of audit: evidence from a controlled experiment in Minnesota. *J. Public Econ.* 79 (3), 455–483.
- Slemrod, J., Collins, B., Hoopes, J.L., Reck, D., Sebastiani, M., 2017. Does credit-card information reporting improve small-business tax compliance? *J. Public Econ.* 149 (Supplement C), 1–19. doi:10.1016/j.jpubeco.2017.02.010.
- Slemrod, J.B., 2016. Tax Compliance and Enforcement: New Research and Its Policy Implications. SSRN Scholarly Paper. Social Science Research Network, Rochester, NY ID 2726077.