October 21, 2024

Massive Regularization for Effective Tax Payment: Evidence from Brazil

Working Paper

Yarygina, Anastasiya[1], Iketani, Eduardo[2], Martinez, Andre[3]; Tiburcio, Ernesto[4]

Abstract

In recent years, tax administration around the globe have leveraged digital transformation to enhance processes and services to improve tax compliance. Massive self-regularization platforms, which identify noncompliant taxpayers, notify them about the detected inconsistencies, and allow to amend the situation with the tax authority, are prominent examples of the digital transformation of tax administrations. This study presents the results of the randomized controlled trial evaluating the effectiveness of such self-regularization platform in the Brazilian State of Para. The results show that the platform increased the value of the tax paid by 12.78 times, and the probability of tax payment by 236%. Overall, the effectiveness of the self-regularization at recovering the evaded tax is 60% higher than that of the traditional audit – based approach. The value of the correction in the declared tax increased by 2.33 times and the probability of correction by 300%. Given the low marginal cost of self-regularization, the results suggest that these platforms are remarkable opportunity for tax administrations to leverage digital transformation in an effective and efficient way improving tax compliance and increasing tax revenue.

Keywords: Taxpayer support, tax compliance, digitalization.

JEL Codes: H26, H30, H32, O38

Acknowledgements:
This study was conducted as a collaborative effort between the Inter-American Development Bank (IDB) and the Secretary of Finance of the state of Para (SEFAZ/PA), Brazil. The authors would especially like to express their gratitude to Jose Guilherme de Souza Moitta Koury, and other SEFAZ/PA employees, for their invaluable contributions to the development of this study. We are also a grateful to Monise Estorani de Faria, Eduardo Fajnzylber and Jairo Neira for excellent analytical inputs for this study. This study benefited from funding of RG-E1807 and RG-E1908 IDB's Economic Sector Work projects. Valuable comments were received from the IDB's Institutions for Development Sector during an internal seminar, and from tax administration experts during June 2023 COGEF meeting. The authors declare that they have no financial or material interests in the results of this study. All the opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the Inter-American Development Bank, its Board of Directors, or the countries they represent. Corresponding author: Anastasiya Yarygina anastasiyay@iadb.org

[1] Inter-American Development Bank
[2] Secretariat of Finance of the State of Para (SEFA/PA)
[3] Inter-American Development Bank
[4] Tufts University

I. Introduction

Strengthening state capacity to collect taxes in an effective and fair way is a key priority of the government around the world (Pomeranz and Vila-Bleda, 2019). Still, in Latin America and the Caribbean (LAC) countries inefficient tax collection and high tax evasion remain as daunting challenges of the fiscal management. While progress has been made in recent years, tax evasion in LAC remains high, amounting to US$325 billion in 2018 and representing 6.1% of GDP (Cristia et al., 2022).

While a priority goal of modern tax administration is to strengthen tax collection by promoting voluntary tax compliance (Bando et al., 2021), an important contributing factor to high levels of tax evasion in LAC are deficiencies in facilitation of compliance and performing tax administration activities efficiently (Jenkins et al., 2023). In terms of cost of compliance, LAC countries remain above the world average, registering 325 hours per year that firms must devote to paying taxes, versus 234 hours per year for the world average (PricewaterhouseCoopers, 2020). In Brazil, there is no official estimate for the aggregated tax gap, but studies show that the country loses up to R$417 billion a year due to tax evasion[5], and the cost of tax compliance remains high, showing the value of 1492 hours per year in 2019[6].

Digitalization and use of information by Tax Administrations (TA) has emerged as a powerful tool that can help improve the institutional competence of TA and encourage taxpayer compliance (Lee, 2016). The transformative force of digitalization is also driving a transition to what is called "Tax Administration 3.0", a more advanced stage of digitalization which includes use of the data and automation to improve processes and services (Calijuri, 2023). Altogether, these transformations prompted TA to implement new techniques and tools that exploit large volumes of the data that TA receive nowadays in almost real time to expand its capacity shifting the focus from punitive to preventive approach and facilitating tax compliance (Lopez-Luzuriaga, unpublished).

Massive self-regularization platforms are prominent examples of the TA's digital transformation and of the change in focus from punitive to preventive approach[7]. In Brazil, the implementation of self-regularization platforms is being consolidated at the federal and subnational level in recent years. The federal government has initiated the CONFIA program, and this experience has already inspired other countries to initiate their programs in Latin America (Calijuri and Oliveira, 2023; Tostes and Calijuri, 2023; Calijuri, 2023). Currently,

[5] The Brazilian Institute of Planning and Taxing has published a study available at: https://ibpt.com.br/estudo-autos-de-infracao-e-sonegacao-fiscal/

[6] According to the World Bank Doing Business Report 2021 - https://subnational.doingbusiness.org/content/dam/doingBusiness/media/Subnational/DB2021_SNDB_Brazil_Full-report_Portuguese.pdf].

[7] Pioneering countries have started implementing preventive approach since the early 2000s, such as Australia, Ireland, the Netherlands, the United Kingdom and the United States. In recent years, this alternative approach to the standard tax system have been implemented in more than 20 countries, with varying degrees of maturity.

22 states in Brazil are implementing some kind of compliance facilitation programs, of which 14 are based on self-regularization platforms (BID, unpublished). Self-regularization platforms use massive data crossings techniques to identify inconsistencies in tax documents (for example, tax declarations), notify taxpayers about these inconsistencies, and allow them to amend the inconsistency and/or pay the tax. The main advantages of the self-regularization platforms, in comparison to a standard audit focused approach, are lower inspection and administrative costs to tax administrations, lower compliance cost to taxpayers, and prevention of litigations.

This paper presents the results form an experimental evaluation of a self-regularization system implemented by a subnational tax administration - Secretariat of Finance of the State of Para (*Secretaria da Fazenda do Estado do Para*, SEFAZ/PA). The study included firms from 1030 business clusters[8] for which SEFAZ/PA identified 8529 declaration with sub-declared Tax on the Circulation of Goods and Services (ICMS)[9]. Half of the clusters was assigned to receive the self-regularization system treatment, which consisted in notifying firms about detected discrepancies in tax returns, and encouraging them to self-regularize through the self-regularization platform. The other half of firms was assigned to the control group and did not receive any self-regularization intervention.

The data was collected three months after the intervention started and was used to analyze whether the self-regularization system affected tax compliance. The main study outcomes are the amount of tax paid and the probability of payment. In addition, the study analyses the effect of the intervention on correction of tax returns, as a mediating pathway for tax payment. For estimation of the intervention effects, we use the Ordinary Least Square (OLS) model for binary and normally distributed variables (e.g., the probability of tax payment), and Poisson model[10] for the outcomes skewed at zero (e.g., the amount of tax paid).

The results of the study show a very large effect of the self-regularization system on tax payment. Specifically, in the main estimations of the Poisson model we find that the amount of tax paid by firms in the treatment group was 12.78 times higher than the amount of tax paid by firms in the control group. The average recover tax represents 8% of the evaded value, which is 60% more than the yield of the alternative audit-based approach. The intervention resulted in an increase in the probability of paying the tax by 236%, from 0.011 percent points (pp.) in the control group to 0.037 pp. in the treatment group. Coupled with the above-mentioned results for the amount of tax paid, these findings show that the intervention was more effective at the intensive margin than the extensive margin[11]. The analysis of the

[8] In this study, a business cluster is an intervention assignment cluster. A business cluster consists of the matrix firm and its subsidiaries. In the study sample business clusters include from 1 to 3 firms.

[9] The ICMS is the most relevant tax collected by states and accounts for almost 85% of revenue at this level of government (Data for 2022, from the State Tax Collection Bulletin - CONFAZ). It is similar to a VAT, and it is imposed on the circulation of goods, electricity, interstate and intermunicipal transportation services and communications.

[10] See, e.g., Wooldridge (2010).

[11] The intensive margin refers to the decision whether to amend or to pay the tax, and the extensive margin refers to the decision on how much to amend or to pay.

effect heterogeneity indicates that the effect of the intervention was stronger for firms with higher detected evasion registered in smaller locations.

The effect on corrections in declarations was also substantial: in the main specification, the value of the tax correction (increase in the declared tax) was 2.33 times higher in the treatment group in comparison to the control group, and the probability of correcting the declared tax increased by 300%, from 0.033pp. in the control group to 0.12 pp. in the treatment group. The compliance with the treatment was not perfect: only 85.9% of the treatment group firms received the message, and only 36.69% of the firms read it. On the other hand, 0.8 % of the control group firms received the notification too. The imperfect compliance with the treatment limited the intervention results, which could have been larger had the treatment compliance and message reading rate been higher.

This study contributes to several strands of literature. First, to the growing body of evidence on the use of third-party information to improve tax compliance. Almunia y López Rodríguez (2016) find that the availability of information works as a complement to the audit of companies and in the same way, Alligham y Sandmo (1972) and Kuchumova (2017) suggest that the amount of information available to the government is essential for the effectiveness of tax enforcement. High-quality and timely third-party information increases the cost of evading taxes by bolstering the effectiveness of monitoring actions like audits. It can also decrease the indirect cost of taxation by enabling the tax authority to provide the electronic prepopulating filling. In general, an electronic billing system can be a powerful tool to promote tax compliance and there is an increase in compliance when higher-quality third-party information is received (Lopez-Luzuriaga, unpublished). Electronic transactions, distinct from cash transactions, involve a third-party intermediary, creating a traceable trail accessible to governments for tax enforcement. This trail, when coupled with an effective tax audit system, deters taxpayers from underreporting taxable transactions (Kleven et al., 2011; Pomeranz, 2015; Naritomi, 2019; Brockmeyer and Sáenz Somarriba, 2022). Notably, taxes subject to third-party reporting, like wages and capital income, exhibit negligible evasion rates, contrasting with higher evasion observed for taxes reliant on self-reported liabilities, where third-party reporting is absent (Kleven et al. 2011; Okunogbe and Santoro, 2023). Self-regularization platforms are a remarkable example of the use of third-party information in massive data crossings to detect tax evasion and noncompliance.

Second, this study contributes to the literature on the cost of tax compliance. The cost of tax filing and complexity of tax codes had been increasing in the recent years, are perceived by taxpayers as encouraging evasion (Benzarti and Wallossek, 2023). Other than that, compliance costs are disproportionately borne by smaller taxpayers (Okunogbe and Santoro, 2023). At the same time, traditional collection methods are sometimes costly, time-consuming, and inefficient for the administration. In Brazil[12], in particular, the volume of tax litigation due to tax complexity is very high - it is estimated, with figures from 2018, that national administrative and judicial tax litigation pending resolution corresponded to

[12] https://blogs.iadb.org/gestion-fiscal/pt-br/estudo-oito-desafios-relacionados-ao-contencioso-tributario-brasileiro/

50.4% of Brazilian GDP. The integration of electronic invoicing and prefilling systems emerges as a pivotal strategy in mitigating tax compliance and administration costs, as highlighted in recent research (Hesami, Jenkins and Jenkins, 2023; Okunogbe and Santoro, 2023). Self-regularization platforms facilitating tax returns amendment aim at reducing tax compliance cost. These technologies not only alleviate burdens associated with tax compliance but also cater to the evolving needs of businesses, fostering economic development.

Third, this study contributes to the large body of literature that analyses the effective communication for tax compliance. Using the data available to the tax authorities, it is possible to experiment with different means of communication, message content, rewards and incentives, among others. Studies by Brockmeyer et al. (2018) and Kettle et al. (2016) underscore the importance of tailored messaging and contextual factors in enhancing tax compliance efforts[13]. Brockmeyer et al. (2018) emphasizes the effectiveness of tax authority messages when they provide specific details about tax inconsistencies, enhancing the credibility of tax authorities and fostering a perception of increased scrutiny. Similarly, Kettle et al. (2016) highlights the significant impact of message design on taxpayer behavior, influenced by contextual factors such as trust in institutions and social norms. Further, Ortega and Scartascini's (2015a, 2015b) show that reinforcing messaging with personal visits and telephone calls can yield better results on compliance. Moreover, experiments conducted in Argentina (Castro and Scartascini, 2013) and Peru (Castro et al., 2020) reveal that deterrence messages significantly influence taxpayer compliance and emphasize the importance of distinguishing between willingness to pay and actual payment behavior. Additionally, Pomeranz's (2015) study with Chilean VAT taxpayers and Bando et al.'s (2021) study of a response of the simplified tax regime taxpayers on messages provide further evidence of the importance of providing detailed information on detected inconsistencies to achieve better compliance outcomes. These findings jointly highlight the critical role of tailored messaging and contextual factors in shaping taxpayer behavior and improving tax compliance outcomes. The design of messages sent by the tax authority in this study too, into account the recommendation from the literature, and included the detailed information on the identified tax discrepancy, clear step-by-step instructions on how to self-regularize, and informed about the fines and penalties for taxpayers if they do not regularize.

More broadly, this study is also related to the literature analyzing the impact of digitalization of tax administration, in particular, to the literature focused on massive auditing and regularization and its relationship with the state capacity (Antinyan and Asatryan, 2019; Bando et al., 2021; Cafe et al. (2024); Coolidge and Yilmaz, 2016; Dom et al., 2022; Pomeranz and Vila-Belda, 2019). By leveraging digitalization interventions, tax administrations can integrate compliance processes seamlessly into daily business, ultimately making compliance more frictionless for taxpayers. Two recent studies and most related to

[13] Mascagni et al (2017) and Mogollon et al (2019) also explore different methods of delivery of the message, such as e-mails, texts, phone calls and letters.

ours are Bando et al. (2021) and Cafe at al. (2024), evaluating self-regularization systems for simplified tax regime taxpayers in Brazilian estates of Piaui and Bahia. Our study adds to this literature by providing the first rigorous evidence on the effectiveness of self-regularization systems for regular regime taxpayers in Brazil, and by looking at the tax payment outcomes.

Finally, this study also relates to the literature on the estimation methodology for the outcomes skewed to zero. This is a methodological challenge for evaluations estimating program effects on the outcome variables with a large proportion of zero values. As recent literature documents (Chen and Roth, 2022; Mullahy and Norton, 2022; Norton, 2022), the traditional strategies to estimate impacts for censored distributions have important shortcomings, showing high sensitivity of estimation results to the units of measurement, both in logarithm transformation and Inverse Hyperbolic Sine Transformations. Unlike OLS, Poisson model employs the logarithm for the difference in expectations, so it is robust to the units of measurement. This study contributes to this literature by implementing the Poisson model for the highly skewed at zero tax compliance outcomes.

The remainder of this paper proceeds as follows. Section II describes the background and institutional context. Section III presents the methodology and the data, Section IV presents the empirical results, and Section V concludes.

II. Study Background

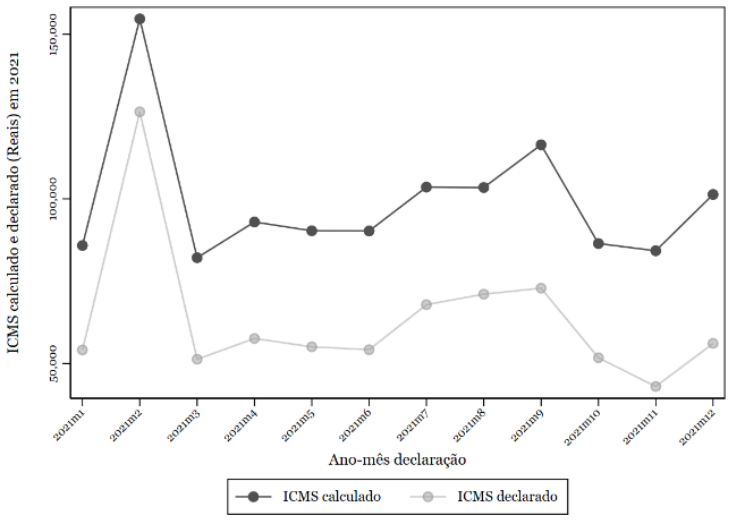
II.1 ICMS tax evasion in Para and self-regularization system as a possible solution

In Brazil, the most important tax collected by State Tax Administrations (SEFAZ) is the ICMS, which accounts for almost 85% of revenue of Brazilian States[14]. In 2023, the state of Para had more than 300,000 active[15] ICMS taxpayers, with this tax accounting for 87% of the state's tax revenue and more than 54% of total state's budget revenue. Of the ICMS revenue, 98% comes from taxpayers registered in the regular tax regime, with the remaining 2% corresponding to the simplified tax regime Simples Nacional[16].

The State of Para has around 20,000 regular tax regime taxpayers. The figure below shows the value of the ICMS tax declared by taxpayers and calculated by tax administration based on third-party information in 2021[17]. It shows that the value of tax declared by taxpayers is, on average, lower than the tax calculated by the tax administration, with the average tax gap between the declared and calculated tax of R$36,446.78 (more than US$ 7,000) per monthly ICMS declaration.

Figure 1. ICMS tax gap in declarations of regular tax regime taxpayers in 2021

[14] Data for 2022, from the State Tax Collection Bulletin – CONFAZ. Brazilian states are also responsible for the Tax on the Ownership of Motor Vehicles (IPVA) and the Tax on *Causa Mortis* Transmissions and Donations (ITCMD).
[15] By way of comparison, the state's population is around 8 million.
[16] Simples Nacional simplified tax regime was established by the Federal Complementary Law No. 123, in December 14, 2006, it is targeted to small and micro firms with the gross revenue of R$ 3.6 million or less per year.
[17] The results are for taxpayers with identified discrepancies between the declared and calculated value of the tax.



Notes: The figure shows the evolution of the tax gap expressed as a divergence between the declared and calculated value of ICMS in the 8,429 declarations of the firms targeted by this study.

Self-regularization platforms have the potential to improve tax compliance and increase tax colection closing the revenue gap. The self-regularization platforms include (i) massive data crossings for detection of inconsistencies in tax returns, (ii) notification of taxpayers about detected inconsistencies, and (iii) the interface to amend the identified inconsistency and pay the tax. The alternative to self-regularization platforms are traditional audits, which are costly and demanding in time and resources. In SEFAZ/PA, an in-depth tax audit can last between 180 and 240 days depending on the taxpayer type and the case difficulty. The effectiveness of the traditional approach is also low, with only 5% of the detected evaded tax being paid. Relying on self-regularization platforms can potentially safe resources to SEFAZ, and reach results improving compliance, which is specifically relevant in the context of the limited institutional capacity which SEFAZ/PA faces: in 2022, the state had approximately 269 auditors who, besides implementing audits, have other tasks and duties. In contrast to the traditional audits, self-regularization does not imply penalties or fines to taxpayers[18], however, failing to self-regularize during the self-regularization period, can trigger penalties and fines.

II.2 The self-regularization intervention

SEFAZ/PA is continually improving its self-regularization systems by including more data crossings and by expanding the scope to more tax obligations and taxpayers. Each expansion of the self-regularization system is usually supported by the pilot stage in which the new data crossings are tested, verified, and the limited number of taxpayers, whose behavior is closely monitored, receive notifications. In collaboration with SEFAZ/PA, we evaluated one such pilot, in which the regular regime taxpayers registered in the State of Para were notified about the

[18] Cafe et al. (2024) discuss the implications of the absence of penalties for self-regularization, in particular, on taxpayer risk perception and future tax compliance behavior.

discrepancies between declared and calculated ICMS tax. The treated group taxpayers received the notification from SEFAZ/PA on their electronic tax mailboxes (Domicílio Tributário Eletrônico, DT-e) and were given 30 days to self-regularize. The notification included the detailed information on the detected inconsistency in the tax return, the instructions to self-regularize, and the information on penalties if they do not self-regularize. The notifications were focused on the tax returns filed in 2018 and 2021,[19] and were sent by regional agencies of SEFAZ/PA between July 29, 2022 and October 19, 2022. SEFAZ/PA closely monitored the responses of taxpayers who participated in the study. It was expected that the intervention would increase tax payments and the declared ICMS value in tax returns.

III. Methodology

III. 1 Study design

The study adopted an experimental evaluation method (randomized controlled trial, RCT), which evaluated the effectiveness of the self-regularization platform implemented by SEFAZ/PA. The firms were assigned to treatment and control groups by business clusters, each cluster corresponding to a *Grupo Empresarial*[20] . This clustered design was implemented to prevent that some tax evading firms of the same cluster would receive the notification and other firms from the same cluster would not.

The information crossings were performed on June 1st of 2022. The study focused on 1392 firms from 1030 business clusters with the value of ICMS in tax returns smaller than the calculated by SEFAZ/PA value of ICMS by R$20.000[21] or more. The following table presents some key information on the universe of the firms subject to information crossings by the self-regularization system, and the firms in the study sample.

Table 1. Comparison of the universe and the study sample

	Universe	**Study sample**
Groups	4378	1030
Firms	5767	1392
Declarations	52073	8529
Discrepancy	R$ 4.039,88	R$ 32.627,24
Declared tax value	R$ 71.896,05	R$ 59.649,78

Notes: This table shows some key statistics for the 2021 declarations for the universe of firms subject to self-regularization system data crossings, and the firms in the study sample. Discrepancy denotes the average value

[19] The evaluation design decisions were taken in early 2022. It was decided to work with the declarations presented in 2021, because of the recency of the tax filing. It was also decided to include the declarations presented in 2018, because SEFAZ/PA was still in the legal period to act upon the inconsistencies detected in these declarations.

[20] A business cluster comprising a matrix firm and its subsidiaries.

[21] About US$ 4000 (Exchange rate September 23, 2023). The threshold was defined based on the capacity restrictions and sample size requirements.

of the difference between the declared tax value and the value computed by SEFAZ/PA based on third-party information.

Because the notifications received by firms were specific on the discrepancies in each tax return, and the amendments were to be done for each specific tax return, the impact estimations were performed at the tax return (or tax declaration) level. In total, the study sample comprised 8529 declarations. The firms were assigned into the treatment and control groups at the business cluster level. The random assignment was performed on June 24, 2022, using Stata 15 software and resulted in 515 clusters assigned to treatment and 515 clusters assigned to the control group.

On July 22, 2022, the regional delegations of the SEFAZ/PA were provided with the list of treatment group firms to notify them about the detected inconsistencies. The first notification was sent on July 29, 2022, and the last one on October 19, 2022. The estimations of the intervention impact were performed with the data collected on October 25, 2022. Figure 1 presents the timeline of the intervention and evaluation.

Figure 1: The study timeline

Source: Prepared by the authors.

III. 2 Empirical model

This study employs two empirical models to estimate the effect of the self-regularization system on tax compliance outcomes, because the variables of interest have different distributions. On the one hand, to study the impact on binary variables, like if the tax was paid, the Ordinary Least Squares (OLS) model – a linear probability model – is employed. The baseline specification does

not include control variables, but the control variables are progressively included to control for whether the year of declaration is 2018, the regional section of SEFAZ/PA where the taxpayer is registered, and the year in which the firm started its economic activity. The equation (1) describes the basic model:

$$Y_{ijg} = \alpha + \beta T_g + X'_{ijg}\delta + \varepsilon_{ijg} \quad (1)$$

Where Y is the outcome of interest of declaration i, taxpayer j, and cluster g denotes a business cluster. T_g denotes whether the business cluster was treated, X_{ijg} is the vector of control variables and ε_{ijt} is the error term. The standard errors are clustered at the business cluster level. The coefficient of interest is β, estimating the change in the conditional expectation of the variables of interest for the treated firms.

Unlike OLS estimators, Poisson model employs the logarithm for the difference in expectations, so it is robust to the units of measurement[22]. Particularly, in Poisson model, for a binary treatment variable, the coefficient estimate denotes the difference in logarithms of the value in two groups: $\ln ln\,(E[Y(1)]) - \ln ln\,(E[Y(0)])$ (Wooldridge, 2010). In this study, the Poisson model is applied for impact estimations for outcome variables with distributions skewed to zero, such as the tax payment, given the large proportion of cases with zero values of tax paid[23]. The interpretation of the Poisson model coefficient estimate is that, in the treatment group, the outcome variable is exp^{β} higher than in the control group, where β is the coefficient estimate. The baseline specification does not include control variables, but the control variables are progressively included to control for whether the year of declaration is 2018, the regional section of SEFAZ/PA where the taxpayer is registered, and the year in which the firm started its economic activity. Equation (2) describes the model:

$$log\left(\mu_{ijg}\right) = \alpha + \beta x_{ijg} \quad (2)$$

Where $\mu = E(Y)$, and x denotes the independent variable(s). It is worth noting that the log of expected value of Y (i.e., $log\,(\mu)$) is a linear function of the independent variables x, and μ is a multiplicative function of x. That said,

$$\mu_{ijg} = \exp exp\left(\alpha + \beta x_{ijg}\right)$$

$$= e^{\alpha} e^{\beta x} \quad (3)$$

III. 3 Data

[22] As the recent literature documents (Chen and Roth, 2022.; Mullahy and Norton, 2022; Norton, 2022), the traditional strategies to estimate impacts for censored distributions have important shortcomings, showing high sensitivity of estimation results for the units of measurement, both in logarithm transformation and Inverse Hyperbolic Sine Transformations.

[23] The proportion of nonzero values in the outcome "value of tax paid" is 1.12% in the control group and 3.74% in the treatment group.

The firms were randomly assigned into treatment and control groups on June 1, 2022, however, the first notification was sent on 29th of July 2022. Because some time elapsed between assignment to treatment and treatment implementation, ten firms corrected some declarations before the start of the intervention. These firms were removed from the study sample,[24] and the final study sample consisted of 512 business clusters with 669 taxpayers in the treatment group and 515 business clusters with 713 taxpayers in the control group, totaling 8429 analyzed declarations (4224 from treated firms and 4255 from control firms).

During the study period, the SEFAZ/PA monitored the treatment and control group firms. The impact estimations were performed using the data collected on October 25, 2022. Some of the variables collected include the value of the tax declared, the divergence between the declared and calculated value of the tax, the tax paid, and firm attributes, such as social capital, the economic activity code, the date of firm constitution, accountant identification. All variables are from the SEFAZ/PA administrative data.

IV. Results

IV.1 Balance

The random assignment to the treatment and control groups guarantees that, on average, there is no significant difference between the treatment and control group, except for the treatment allocation. To check that the groups are balanced, the formal tests of the equality of means of the variables measuring firms' characteristics at baseline were performed. The results are shown in Table 2. Panel 1 shows the results for the characteristics defined at the cluster level, Panel 2 for the characteristics defined at the firm level, and Panel 3 for the characteristics defined at the declaration level.

The results confirm that, on average, the treatment and control groups are balanced with no statistically significant differences in characteristics observed at baseline. Table 2 also shows descriptive statistics of some key variables. For example, it can be seen that the value of evaded tax (the difference between the tax calculated by SEFAZ/PA based on third-party information and the tax declared by a taxpayer) at the baseline was about 30-35 thousand Reais per declaration (US$5.7-6.7 thousand), resulting in a total tax evasion per business cluster of about 270,000 Reais (over US$50,000). Half of the declarations declared zero ICMS. Table B.1 in Annex B shows more descriptive statistics.

IV.2 Program compliance

The results of the assessment of whether the intervention administration followed the random assignment into treatment and control groups are presented in Table 3. As it can be seen, the compliance with the intervention was not perfect: 86% of the treatment group firms were

[24] These firms would not have been eligible for the intervention and would not have been included in the pool of firms subject to the random assignment.

notified, while some of the control group firms (0.8%) also received the notification. This occurred because the decision of whether to notify any given firm was not done by the SEFAZ/PA central office but was delegated to the SEFAZ/PA regional agencies. While all regional agencies received the list of treatment and control firms, it was at their discretion to decide whether to notify any given firm. Before sending notifications to the treatment group firms, regional agencies evaluated whether notifying a firm was appropriate, and could decide not to notify if the firm was already under audit or investigation for the same tax delinquency. On the other hand, regional agencies could notify some control group firms if they considered that the firm's delinquency supported by the information provided by SEFAZ/PA warranted the investigation. These decisions of regional agencies resulted in an imperfect compliance with the intervention assignment, with the coefficient estimate for the binary indicator of the notification sent of 85 percentage points (pp).

Another informative indicator is the rate of message reading. As seen in Table 3, only 37% of the firms in the treatment group read the notification (no one in the control group did). The share of the firms that read the message is not large and limits the impact of the intervention. However, this result is in line with the value found in the recent studies by Bando et al. (2021) and Cafe et al. (2024) in self-regularization interventions focused on simplified tax regime taxpayers.[25]

[25] In this study, the proportion of the firms who read the message is 37%, similar to the findings in Cafe et al. (2024) who find a similar value of this indicator for the firms in the "message" treatment. In this study, the ratio message read/firm notified is 37%/85%=0.44, similar to the results reported in Bando et al. (2021), where all firms were notified and 47% read the message.

Table 2: Balance at baseline

Variable description	Treatment mean	Control mean	Difference treatment – control	Standard deviation	P-value	Treatment sample	Control sample
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel 1: Business cluster level							
Value of ICMS tax declared at baseline	572161.79	407458.64	164703.145	[140562.435]	0.2416	512	515
Value of ICMS tax evaded at baseline[1]	285,143.51	251,118.78	34024.729	[33300.907]	0.3071	512	515
Number of firms per business cluster	1.31	1.38	-0.078	[0.087]	0.3694	512	515
Number of declarations per business cluster	8.25	8.17	0.085	[0.502]	0.8657	512	515
Panel 2: Firm level							
Has accountant (yes =1)	0.77	0.78	-0.010	[0.032]	0.7605	669	713
Years firm operates	12.21	13.19	-0.977	[0.980]	0.3186	669	713
Year in which firm begins operating	2010.11	2009.15	0.960	[0.982]	0.3284	669	713
Firm is a limited company (yes=1)	0.61	0.63	-0.017	[0.041]	0.6808	669	713
Firm is an individual entrepreneur (yes=1)	0.24	0.22	0.013	[0.031]	0.6745	669	713
Firm's Social Capital	50173320.72	44680787.84	5492532.880	[45439401.380]	0.9038	669	713
Firm is registered in the State capital (Regional Section Belem)	0.19	0.20	-0.017	[0.032]	0.6048	669	713
Number of declarations presented by the firm	6.31	5.90	0.416	[0.366]	0.2551	669	713
Panel 3: Declaration level							
Declaration month	6.89	6.85	0.044	[0.106]	0.6825	4224	4205
Value of ICMS tax declared at baseline	69352.94	49902.78	19450.162	[16890.802]	0.2498	4224	4205
Value of ICMS tax evaded at baseline[1]	34,562.85	30,755.33	3807.517	[3840.497]	0.3217	4224	4205
Value of declared ICMS at baseline is zero (yes=1)	0.57	0.54	0.028	[0.037]	0.4592	4224	4205
Declaration is in 2021 (yes=1)	0.64	0.68	-0.038	[0.027]	0.1632	4224	4205
Declaration is in 2018 (yes=1)	0.36	0.32	0.038	[0.027]	0.1632	4224	4205

Notes: 1: The evaded tax is calculated as a difference between the value calculated by the tax administration, based on third-party information, and the value declared by the firm. This table presents the statistics and estimates of the differences between the treated group and the control group, at the level of business cluster (Panel 1), firm (Panel 2) and declaration (Panel 3). Columns (1) and (2) show the means for the treated and control groups, respectively. Column (3) shows the difference between these values estimated by regressing the variable indicated in the "variable description" column of the table on the random assignment indicator. Column (4) shows the standard error. The standard errors are clustered at the business cluster level in the regressions in panels 2 and 3. Column (5) shows the P-values of the test of equality of means in columns (1) and (2). Columns (6) and (7) show the number of observations in the treatment control group samples, respectively. Statistical significance: * 10%, ** 5%, *** 1%.

Table 3. Treatment compliance

Variables description	(1) Difference	(2) Standard deviation	(3) P-value	(4) Sample	(5) Treatment mean	(6) Control mean
Received notification (yes=1)	0.851***	[0.017]	0.0000	1382	0.859	0.008
Read notification (yes = 1)	0.369***	[0.026]	0.0000	1382	0.369	0.000

Notes: This table presents statistics and estimated differences between the treatment and control groups at the firm level. Column (1) presents estimated coefficients, column (2) presents standard errors, clustered at the business cluster level, column (3) shows the P-value of the statistical test of equality of the parameter in column (1) to zero, column (4) shows the number of observations, columns (5) and (6) present the means of the treated and control groups, respectively. The regressions do not include control variables. Statistical significance * 10%, ** 5%, *** 1%.

IV.3 Main results: tax payment

The results for the main outcome of interest of this study – tax payment - are reported in Table 4. The program effect coefficient estimated in the OLS model are reported in Panel 1 of Table 4. They are very large, showing the difference in the tax paid between the treatment and control group means of about 2500 Reais (approximately US$ 500) per declaration. However, the estimates in the basic specification without controls and in specification with additional control variables are very imprecise. Given that the distribution of the right-hand side variable deviates from the normal variable distribution and presents large proportion of zeros, a more appropriate model for the effect estimation is a Poisson model. The results of the Poisson model estimations are presented in Panel 2 of Table 4. In comparison to the OLS, all coefficient estimates in Poisson model are precise and statistically significant at 5% level of statistical significance.

The size of the effect estimated in Poisson model ranges between 2.5 and 1.7 demonstrating a large program effect. The results of the basic specification show that, on average, the tax paid by firms in the treatment group is $exp(2.548) = 12.78$ times greater than the tax paid by the firms in the control group (treatment mean = 2796.1, and control group mean =218.7). The inclusion of the control variables results in a variation in the effect from 6.01 to 13.38. Taking a difference between the control and treatment mean of about 2,500 Reais, this intervention, at almost zero marginal cost to tax administration, has a potential to increase the tax revenue by about US$ 500 (2500 / 5) per tax return, or US$ 3150 (500 x 6.3 declarations) per firm. Given the average value of the evaded tax per declaration of about 32 thousand Brazilian Reais[26], the intervention is able to recover about 8% (2500/32000) of the evaded tax, which is 60% higher than the average yield of the audits in SEFAZ/PA.

In addition to being more effective, the self-regularization system is also more efficient than traditional audits. While the cost of audit is difficult to estimate, it is arguably higher than the cost of self-regularization, given that auditing involves hours of work of tax auditors to investigate, process, reach the firms, and try to recover the tax due. Taking as a reference the per firm cost of self-regularization system reported in Bando et al. (2021) of US$28 per firm, current

[26] The average of the treatment and ctonrol group means reported in the Balance Table 2.

interventions allows to recover one US$ dollar of ICMS payment at a less than one cent of a dollar (US $0.0089 = 28/3,150).

Table 4: Intervention effect on tax payment

	No controls		Controls: declaration year 2018		Controls: declaration year 2018, regional delegation, year of firm establishment	
	Effect estimate	Standard deviation & p-value	Effect estimate	Standard deviation & p-value	Effect estimate	Standard deviation & p-value
	(1)	(2)	(3)	(4)	(5)	(6)
Panel 1: OLS model estimates						
	2577.397	2084.365	2649.582	2145.965	2436.109	1907.859
		0.217		0.217		0.202
Control mean	218.7		218.7		218.7	
Control sample	4205		4205		4205	
Treatment mean	2796.1		2796.1		2796.1	
Treatment sample	4224		4224		4224	
Panel 2: Poisson model estimates						
	2.548**	1.078	2.594**	1.078	1.794**	0.822
		0.018		0.016		0.029
Control mean	218.7		218.7		276.9	
Control sample	4205		4205		3321	
Treatment mean	2796.1		2796.1		3494.3	
Treatment sample	4224		4224		3380	

Notes: This table presents statistics and the treatment effect for the value of the tax paid. Panel 1 shows the results for the OLS model, and Panel 2 shows the results for the Poisson model. The interpretation of the coefficient in Panel 1 a is that the average values of the treatment group, conditional on the controls, are β higher than the control group. The interpretation of the coefficient in Panel 2 is that on average the values of the treatment group, conditional on the controls, are equivalent to exp(β) in relation to the control group. The columns (1), (3), and (5) show the coefficient estimates, standard deviation, control and treatment groups means and sample sizes, the columns (2), (4), and (6) show the p-values of the test of equality to zero of the coefficient estimate. The estimations result in columns (3) and (4) are from models which include the indicator of year of declaration = 2018. The estimation results in columns (5) and (6) are from models which include the year of declaration = 2018 indicator, the regional section and firm establishment year fixed effects. Statistical significance: * 10\%, ** 5\%, *** 1\%.

In addition to the value of the tax paid, we also explore the results for the rate of tax payment (that is, the probability of paying the tax). These results are reported in Table 5. As seen, on average, the probability of paying the tax in treatment group is about 0.025 – 0.027 percentage points larger than the probability of paying the tax in the control group. While the effect represents 240% of the control group mean (0.0112), it is also worth noting that the proportion of

delinquent firms who pay the tax, in both treatment and control groups is very low, and the program effect of more than doubling the proportion of those who pay results in a proportion of 0.0374 of compliant taxpayers in the treatment group. In this sense, the intervention results are mostly driven by the effects at the intensive margin rather than extensive margin.

Table 5: Intervention effect on tax payment rate

	No controls		Controls: declaration year 2018		Controls: declaration year 2018, regional delegation, year of firm establishment	
	Effect estimate	Standard deviation & p-value	Effect estimate	Standard deviation & p-value	Effect estimate	Standard deviation & p-value
	(1)	(2)	(3)	(4)	(5)	(6)
	0.026**	0.010 0.010	0.027***	0.010 0.008	0.025***	0.009 0.008
Control mean	0.0112		0.0112		0.0112	
Control sample	4205		4205		4205	
Treatment mean	0.0374		0.0374		0.0374	
Treatment sample	4224		4224		4224	

Notes: This table presents statistics and the treatment effect on the rate of tax payment (the probability of paying tax). The interpretation of the coefficient estimate is that the average value in the treatment group, conditional on the controls, is β percent points higher than the control group. The columns (1), (3), and (5) show the coefficient estimates, standard deviation, control and treatment groups means and sample sizes, the columns (2), (4), and (6) show the p-values of the test of equality to zero of the coefficient estimate. The estimations result in columns (3) and (4) are from models which include the indicator of year of declaration = 2018. The estimation results in columns (5) and (6) are from models which include the year of declaration = 2018 indicator, the regional delegation and firm establishment year fixed effects. Statistical significance: * 10\%, ** 5\%, *** 1\%.

The coefficient estimates presented in Table 4 and 5 correspond to the "intention-to-treat" (ITT) parameter for participation in the intervention. To estimate a parameter that represents the full effect of the intervention, in is necessary to account for imperfect compliance[27]. For OLS models, we can approximate the Local Average Treatment Effect (LATE) of the intervention on the sample of compliers by dividing the coefficient estimates in Tables 4 and 5 by the compliance rate reported in Table 3, getting the LATE effect for the rate of tax payment of about 1.17 larger for taxpayers who received the message (1/0.85) and 2.7 times larger for taxpayers who read the message (1/0.37). For the Poisson model we present the Instrumental Variables (LATE) estimates in Annex Table B.2. According to these results, in the basic specification, the LATE coefficient

[27] Imperfect compliance arises when not all units assigned to receive the treatment get it, and (or) when some unites assigned not to receive the treatment end up getting it.

for the tax paid among compliers who received the message is 2.83, and 3.52 for compliers who read the message.

Finally, in Table 6 we report the results for the tax paid and the rate of tax payment in different subsamples. According to the obtained results, the firms tend to pay tax more intensively if it corresponds to older declaration (tax return in 2018), however, we cannon conclude the same for the value of tax paid[28]. The value of tax paid, and the rate or tax payment tend to be higher in a subsample of taxpayers with large value of tax evasion at baseline[29]. Finally, there does not seem to be any statistically significant program effect in the subsample of declarations presented in large regional agencies of SEFAZ/PA, namely, Belem and Martuba. On the contrary, the effect is high and statistically significant in all other regional agencies. These results suggest that in the context of capacity restrictions, tax administrations might be interested in targeting specific groups of taxpayers to achieve better results. In case of this intervention, it appears that the program had higher effect in smaller regional sections, among taxpayers who evade more.

Table 6: Heterogeneous effects on the paid tax value and payment rate.

Variable name	Full Sample	Tax return in 2018	Tax return in 2021	Low tax evasion at baseline	High tax evasion at baseline	Belem and Martuba	Not Belem and Martuba
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Value of tax paid (Poisson)	2.548**	NA	2.518**	1.180	2.672**	-0.055	3.759***
Rate of tax payment (OLS)	0.026**	0.035***	0.022	0.021**	0.032**	0.010	0.034***

Notes: This table shows the estimated impact parameters and statistical significance for the Poisson model estimates for the value of tax paid, and OLS model estimates for the rate of tax payment. Each cell in the table corresponds to one regression. Column (1) shows the estimated parameters for the whole sample, while columns (2) - (7) show the estimated parameters in a model without covariates for the different sub-samples, as indicated in the column names. The Poisson model estimation for value of tax paid is not reported because the algorithm did not converge. Statistical significance * 10%, ** 5%, *** 1%.

IV.4 Mediating pathways: tax return correction

As it was previously explained, the message that the taxpayers received specified the discrepancy between the declared and calculated value of the ICMS tax for each tax declaration with this inconsistency, and invited taxpayers to amend the declared value, which would result in a higher value of tax due. In this section, we explore how the intervention affected taxpayers' behavior regarding declared tax corrections. In Table 7 we present the results for the value of declared tax correction, and in Table 8 for the correction rate (that is, the probability to correct the tax in a tax return).

[28] The Poisson estimation for value of tax paid is not reported for a subsample of declarations presented in 2018, because the algorithm did not converge.

[29] The tax returns with high value of tax evasion at baseline are those with the divergence between the calculated and declared tax value higher than the median.

Table 7 shows the results for the amount of corrected tax. The declared tax value is also a truncated variable exhibiting bunching of values at zero, however, the concentration of values at zero is less pronounced than for the value of tax paid variable[30]. Table 7 presents the results for the OLS model (Panel 1), and the Poisson model which takes into account the outcome variable distribution (Panel 2).

As seen, the intervention, on average, resulted in a larger correction of the declared tax in the treatment group firms (mean value of R$ 6,280) than in the control group (mean value R$ 2,700.8), with the effect estimate ranging from 3.5 thousand (about US$700) to 4.2 thousand Brazilian Reais (about US$800). The effect estimates are statistically significant at 10% in the basic specification, and in the specification with the year of declaration control variable. The statistical significance of the estimates improves and reaches the 5% level in the specification with additional controls for the firm's geographical location and years of operation.

The estimation results for the Poisson model show that the value of correction in the treatment group firms, in the basic specification without control variables, is $exp(0.844) = 2.33$ times larger than in the control group (at 10% level of statistical significance). In the specifications with control variables, the effect estimate increases up to $exp(1.23) = 3.42$, and becomes statistically significant at 5% level of statistical significance.

Table 7: Intervention effect on the value of tax correction

	No controls		Controls: declaration year 2018		Controls: declaration year 2018, regional delegation, year of firm establishment	
	Effect estimate	p-value	Effect estimate	p-value	Effect estimate	p-value
	(1)	(2)	(3)	(4)	(5)	(6)
Panel 1: OLS model estimates						
	3579.840*	0.091	3753.597*	0.079	4257.929**	0.039
	(2113.849)		(2135.848)		(2058.253)	
Control mean	2700.8		2700.8		2700.8	
Control sample	4205		4205		4205	
Treatment mean	6280.7		6280.7		6280.7	
Treatment sample	4224		4224		4224	
Panel 2: Poisson model estimates						
	0.844*	0.058	0.882**	0.046	1.230**	0.015
	(0.445)		(0.443)		(0.504)	
Control mean	2700.8		2700.8		2789.7	
Control sample	4205		4205		4071	
Treatment mean	6280.7		6280.7		6451.8	
Treatment sample	4224		4224		4112	

[30] The probability of correcting the declaration after the intervention is 3% in the control group and 12% in the treatment group, in comparison to a probability of paying the tax of about 1% in the control group and 4% in the treatment group.

Notes: This table presents statistics and treatment effect estimates for the value of correction in tax declarations in an OLS model (Panel 1) and Poisson Model (Panel 2). The interpretation of the coefficient in Panel 1 a is that the average values of the treatment group, conditional on the controls, are β higher than the control group. The interpretation of the coefficient in Panel 2 is that on average the values of the treatment group, conditional on the controls, are equivalent to exp(β) in relation to the control group. The columns (1), (3), and (5) show the coefficient estimates, standard deviation, control and treatment groups means and sample sizes, the columns (2), (4), and (6) show the p-values of the test of equality to zero of the coefficient estimate. The estimations result in columns (3) and (4) are from models which include the indicator of year of declaration = 2018. The estimation results in columns (5) and (6) are from models which include the year of declaration = 2018 indicator, the regional section and firm establishment year fixed effects. Statistical significance: * 10\%, ** 5\%, *** 1\%.

Table 8 presents the results for the rate of correction of declared tax value (the probability of correcting the declared tax). The results show that, on average, the rate of correction was higher in the treatment group by 0.086-0.09 percent points.

Table 8: Intervention effect on the rate of tax correction

	No controls		Controls: declaration year 2018		Controls: declaration year 2018, regional delegation, year of firm establishment	
	Effect estimate	p-value	Effect estimate	p-value	Effect estimate	p-value
	(1)	(2)	(3)	(4)	(5)	(6)
	0.086***	0.000	0.089***	0.000	0.090***	0.000
	(0.018)		(0.017)		(0.017)	
Control mean	0.0328		0.0328		0.0328	
Control sample	4205		4205		4205	
Treatment mean	0.119		0.119		0.119	
Treatment sample	4224		4224		4224	

Notes: This table presents the statistics and treatment effects for the rate of tax return correction (probability of the tax declaration to be corrected). The interpretation of the coefficient estimate is that the average value in the treatment group, conditional on the controls, is β percent points higher than the control group. The columns (1), (3), and (5) show the coefficient estimates, standard deviation, control and treatment groups means and sample sizes, the columns (2), (4), and (6) show the p-values of the test of equality to zero of the coefficient estimate. The estimations result in columns (3) and (4) are from models which include the indicator of year of declaration = 2018. The estimation results in columns (5) and (6) are from models which include the year of declaration = 2018 indicator, the regional delegation and firm establishment year fixed effects. Statistical significance: * 10\%, ** 5\%, *** 1\%.

V. Conclusions

State capacity to collect taxes efficiently is critical to ensures sustainability of state finances and provide fiscal space for productive investment and social spending, ensuring growth and development. Promoting tax compliance and strengthening tax collection is therefore a priority goal for modern tax administration. In achieving this goal, tax administration can leverage digitalization and use of administrative information to effectively detect evasion and lower taxpayer compliance cost.

This study evaluates the effectives of a self-regularization system, which leverages the new technological capabilities of tax administration to perform massive data crossings, inform taxpayers about detected inconsistencies, and facilitate self-regularization through a dedicated platform. The implementation of the self-regularization platform for a group of delinquent firms who sub-declared the tax value in their tax returns show that the self-regularization system has very large effects on the payment of taxes and corrections in tax returns, in particular, the effectiveness of the self-regularization system is 60% higher than the effectiveness of the traditional approach based on personalized audits. This study presents the results three months after the intervention, and its long-run impacts depend on the dynamics of the taxpayer behavioral. As discussed in the recent study in Brazil (Café et al. 2024), the expected change in compliance regarding future tax obligations can be dependent on the perceived risk and penalty. Given the very low marginal cost of the self-regularization in comparison to auditing, the self-regularization intervention is arguably more cost-efficient than the traditional audit-based approach. The self-regularization messages were based on the lessons learned from previous literature, and included detailed information on the detected inconsistencies, clear step-by-step instructions on how to amend the tax return, and the information on penalties and fines of not doing so.

The results of this study highlight the potential of self-regularization systems to improve tax compliance and increase revenue at low cost, which is in line with the results reported in recent studies in Brazil for simplified tax regime taxpayers (Bando et al., 2021; Cafe at al., 2024). In addition, this study also provides new evidence for the effects of self-regularization systems in Brazil for payment of taxes for regular tax regime taxpayers. The results of this study are particularly important in Brazil, where in recent years sub-national jurisdictions have been implementing self-regularization systems, and tax administration in other countries and regions interested in substitutions or complementation of the traditional audit-based approach by self-regularization platforms.

References

Alm, J. et al.(1992) Why do people pay taxes? Journal of public Economics, v. 48, n. 1, p. 21-38.

Almunia, M., López-Rodrıguez, D., & Moral-Benito, E. (2016). Using Firm-Level Data to Build a Macro-Representative Micro-Dataset: An Application for the Spanish Economy. Mimeo.

Allingham, M. G., & Sandmo, A. (1972). Income tax evasion: A theoretical analysis. Journal of public economics, 1(3-4), 323-338.

Antinyan, A., & Asatryan, Z. (2019). Nudging for tax compliance: A meta-analysis. ZEW-Centre for European Economic Research Discussion Paper, (19-055).

Bando, R., Canozzi, O., Martínez Carrasco, J., & Dezolt, A. L. (2021). May I Calculate Your Taxes?: The Effect of Bookkeeping on Tax Compliance under a Simplified Regime (No. 11381). Inter-American Development Bank.

Benzarti Y. and L. Wallossek (2023) Rising Income Tax Complexity. No. w31944. National Bureau of Economic Research, 2023.

Brockmeyer, A., & Sáenz Somarriba, M. (2022). Electronic Payment Technology and Tax Compliance: Evidence from Uruguay's Financial Inclusion Reform.

Brockmeyer, A., Hernandez Hernandez, M. A., Kettle, S., & Smith, S. D. (2018). Casting the tax net wider: experimental evidence from Costa Rica. American Economic Journal: Economic Policy, 1–56.

Cafe, R., Yarygina, A., Escalante, L. (2024). Leveraging Data to Improve Tax Compliance for Micro and Small Firms: Evidence from Brazil. IDB Working Paper Series 1609.

CAGALA, Tobias; RINCKE, Johannes; and CUEV, Amanda Tuset (2017). Administrative Efficiency and Tax Compliance Proceedings. Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association. Vol. 110.

Calijuri, M. S. (2023). Revolutionizing Taxation: Tax Administration 3.0 is already here. Review of International and European Economic Law, 2(4), A4-1.

Calijuri, M.S.; Oliveira, P.T.P. (2023) Cooperative compliance in Brazil: what does the future hold? IDB Technical note 2793.
https://publications.iadb.org/en/cooperative-compliance-brazil-what-does-future-hold

CASTRO, Juan Francisco et al. (2020) Spillovers and long-run effects of messages on tax compliance: experimental evidence from Peru. IZA DP, n. 13974. Available at: https://docs.iza.org/dp13974.pdf. Accessed on: 19 nov. 2022.

CASTRO, Lucio; SCARTASCINI, Carlos. (2013) Tax Compliance and enforcement in the pampas. evidence from a field experiment. IDB Working Paper Series, v. 472. Available at: https://publications.iadb.org/en/tax-compliance-and-enforcement-pampas-evidence-field-experiment. Accessed on: 20 nov. 2022.

Chen, J. and Roth, J., 2022. Log-like? Identified ATEs defined with zero-valued outcomes are (arbitrarily) scale-dependent.

Coolidge, J., & Yilmaz, F. (2016). Small business tax regimes. Brief. Viewpoint; No. 349. World Bank, Washington, DC. World Bank

Cristia, J. P., & Vlaicu, R. (2022). Digitalizing Public Services Opportunities for Latin America and the Caribbean.

Dom, R., Custers, A., Davenport, S., & Prichard, W. (2022). Innovations in tax compliance: Building trust, navigating politics, and tailoring reform. World Bank Publications.

Eichfelder, S., and Kegels, C. (2013). Compliance costs caused by agency action? Empirical evidence and implications for tax compliance. Journal of Economic Psychology, 40, 200–219. https://doi.org/10.1016/j.joep.2012.08.012

Hesami, S., Jenkins, H. P., & Jenkins, G. P. (2023). Emerging Digital Technologies to Improve Tax Compliance and Administration Efficiency: A Systematic Literature Review. Available at SSRN 4580004.

Inter-American Development Bank. 2017. "Fiscal Management Modernization Project for the State of Pará." Loan Proposal BR-L1499. Inter-American Development Bank. Washinton, DC. https://www.iadb.org/en/whats-our-impact/BR-L1499.

Jenkins, G.P. et al. (2023). A Cost-Benefit Analysis Methodology for Administrative Prefilling of Value-Added Tax Returns: An application for Chile. IDB Technical Note n. IDB-TN-2765.

KETTLE, Stewart et al.(2016) Behavioral interventions in tax compliance: evidence from Guatemala. World Bank Policy Research Working Paper, n. 7690.

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.

Kuchumova, Y. P. (2017). The optimal deterrence of tax evasion: The trade-off between information reporting and audits. Journal of Public Economics, 145, 162-180.

Lee, H. C. 2016. Can Electronic Tax Invoicing Improve Tax Compliance? A Case Study of The Republic of Korea's Electronic Tax Invoicing for Value-Added Tax. World Bank Policy Research Working Paper (7592). Washington, D.C.: World Bank.

Lopez-Luzuriaga, (unpublished) Tax authority information as a tool. Unpublished manuscript.

Mascagni, G., et al. (2017). One size does not fit all: a field experiment on the drivers of tax compliance and delivery methods in Rwanda.

Mogollón, Mónica, Daniel Ortega and Carlos Scartascini. (2019). Who's Calling? The Effect of Phone Calls and Personal Interaction on Tax Compliance. Inter-American Development Bank Working Paper Series No. IDB-WP-1084.

Mogollon, M., Ortega, D., & Scartascini, C. (2021). Who's calling? The effect of phone calls and personal interaction on tax compliance. International Tax and Public Finance, 1-27.

Mullahy, J. and Norton, E.C., 2022. Why Transform Y? A Critical Assessment of Dependent Variable Transformations in Regression Models for Skewed and Sometimes-Zero Outcomes (No. w30735). National Bureau of Economic Research.

Naritomi, J. (2019). Consumers as tax auditors. American Economic Review, 109(9), 3031-3072.

Newman, W. et al. (2018). Literature review on the impact of tax knowledge on tax compliance among small medium enterprises in a developing country. International Journal of Entrepreneurship, 22(4), 1-15.

Norton, E.C., 2022. The inverse hyperbolic sine transformation and retransformed marginal effects. The Stata Journal, 22(3), pp.702-712.

Okunogbe, O., & Santoro, F. (2023). The Promise and Limitations of Information Technology for Tax Mobilization. The World Bank Research Observer, 38(2), 295-324.

ORTEGA, Daniel; SCARTASCINI, Carlos. (2015a) Don't blame the messenger: a field experiment on delivery methods for increasing tax compliance. Inter-American Development Bank (IDB), Washington, DC. (IDB Working Paper Series, 627). Available at: https://doi.org/10.18235/0000204. Accessed on: 3 nov. 2022.

ORTEGA, Daniel; SCARTASCINI, Carlos. (2015b) Who's calling? The effect of phone calls as a deterrence mechanism. Washington, DC; United States: Inter-American Development Bank; Mimeographed document.

POMERANZ, Dina. (2015) No taxation without information: deterrence and self-enforcement in the value added tax. American Economic Review, v. 105, n. 8, p. 2539-2569. Available at: https://www.aeaweb.org/articles?id=10.1257/aer.20130393. Accessed on: 8 nov. 2022.

Pomeranz, D., & Vila-Belda, J. (2019). Taking state-capacity research to the field: Insights from collaborations with tax authorities. Annual Review of Economics, 11, 755-781.

PricewaterhouseCoopers. 2020. Paying Taxes 2020: The Changing Landscape of Tax Policy and Administration across 190 Economies.

Slemrod, Joel, Brett Collins, Jeffrey Hoopes, Daniel Reck, and Michael Sebastiani (2015). "Does Credit-Card Information Reporting Improve Small-Business Tax Compliance?" NBER Working Paper

Tostes, J.B; Calijuri, M.S. (2023). Transformando la relación contribuyente-administración tributaria en América Latina: el papel de los programas de cumplimiento cooperativo. (IDB Monografia; 1150). https://publications.iadb.org/es/transformando-la-relacion-contribuyente-administracion-tributaria-en-america-latina-el-papel-de-los.

Wooldridge, J.M., 2010. Econometric analysis of cross section and panel data. MIT press.

Annexes

A. Notifications

The notifications sent to taxpayers were prepared based on the knowledge generated by previous studies. In terms of content, SEFAZ/PA decided to choose a dissuasive message, informing taxpayers that the failure to comply will result in actions by the tax administration, fine, and interests for late payment (Castro et al., 2022) Another study that informed the message design is the one carried out by the São Paulo Secretaria de Fazenda (SEFAZ/SP), for collection of the Vehicle Tax. The study divided taxpayers into a control group and 5 other groups that received different types of messages. The best results were obtained for the dissuasive messages (DICAR, 2022) In addition, the message design also considered the experience of the study in Piaui Brazilian State (Bando et al., 2021), which found that giving to taxpayers a detailed information of the delinquency in the notification improves the message effect.

In this study, the delinquent taxpayers received the notification on their fiscal electronic address in pdf format, along with the online spreadsheet detailing the calculations of the tax due. In addition, the taxpayers could download all electronic documents used to compute the due tax by SEFA/PA from the SEFAZ/PA Service Portal. The objective of giving to taxpayers the detailed information was to proof that SEFAZ/PA has a thorough knowledge of the delinquency and convincing taxpayers to make corrections and tax payments.

GOVERNO DO ESTADO DO PARÁ
SECRETARIA DE ESTADO DA FAZENDA
DIRETORIA DE FISCALIZAÇÃO

Identificação do sujeito passivo:

Inscrição Estadual: [redacted]
Nome Empresarial: [redacted]

Mediante consulta ao sistema da Secretaria da Fazenda do Pará, detectamos que há divergência entre os documentos fiscais eletrônicos de saída emitidos (NFe/NFCe) e o valor total declarado em DIEF conforme relatório abaixo. É possível realizar uma consulta que contém as chaves do documentos considerados no Portal de Serviços da Sefa (http://app.sefa.pa.gov.br/e-procedimentos-fiscais).

Referência	Total ICMS NFC-e	Total ICMS NFe	Total ICMS NFe Descontado Lançamento efetuado em decorrência de emissão de Cupom Fiscal (CFOP 5.929)	Total ICMS Notas Fiscais	Valor ICMS Vendas Mercadorias Declarado DIEF (exceto CFOP 5.929)	Diferença ICMS Vendas Mercadorias Não Declaradas DIEF
02/2021	R$ 0,00	R$ 292.186,10	R$ 292.186,10	R$ 292.186,10	R$ 145.715,40	R$ 146.470,70
03/2021	R$ 0,00	R$ 301.877,90	R$ 301.877,90	R$ 301.877,90	R$ 122.369,60	R$ 179.508,30
04/2021	R$ 0,00	R$ 337.767,30	R$ 337.767,30	R$ 337.767,30	R$ 280.007,00	R$ 57.760,24
05/2021	R$ 0,00	R$ 341.848,40	R$ 341.848,40	R$ 341.848,40	R$ 111.848,40	R$ 230.000,00
06/2021	R$ 0,00	R$ 373.134,50	R$ 373.134,50	R$ 373.134,50	R$ 138.239,90	R$ 234.894,50
07/2021	R$ 0,00	R$ 349.191,80	R$ 349.191,80	R$ 349.191,80	R$ 105.352,10	R$ 243.839,70
08/2021	R$ 0,00	R$ 367.589,30	R$ 367.589,30	R$ 367.589,30	R$ 143.544,30	R$ 224.045,00
09/2021	R$ 0,00	R$ 387.749,80	R$ 387.749,80	R$ 387.749,80	R$ 172.680,60	R$ 215.069,20
10/2021	R$ 0,00	R$ 362.155,60	R$ 362.155,60	R$ 362.155,60	R$ 157.685,10	R$ 204.470,50
11/2021	R$ 0,00	R$ 426.523,20	R$ 426.523,20	R$ 426.523,20	R$ 191.569,60	R$ 234.953,60
12/2021	R$ 0,00	R$ 459.361,70	R$ 459.361,70	R$ 459.361,70	R$ 136.614,50	R$ 322.747,30
						R$ 2.293.759,04

Conforme Art.11-A da lei 6.182/98, o contribuinte ficará a salvo das penalidades previstas na legislação de regência do tributo, desde que sane a irregularidade especificada dentro do prazo concedido. Contudo, a manutenção da espontaneidade se restringe às inconsistências descritas neste comunicado.

É possível, através do serviço "Procedimentos Fiscais" (app.sefa.pa.gov.br/e-procedimentos-fiscais), a apresentação de justificativa que comprove com base na legislação tributária a não pertinência da irregularidade informada.

Decorrido o prazo concedido na comunicação sem a devida regularização, o contribuinte estará sujeito ao início de procedimento administrativo e às penalidades previstas na legislação.

Para informações e esclarecimentos entre em contato com o atendimento da Sefa através do chat no site www.sefa.pa.gov.br, pelo call center ou WhatsApp 0800-725-5533, ou envie email para atendimento@sefa.pa.gov.br.

Fundamentação Legal: LEI 5.530/89, ART. 62 E ART. 65

Penalidade em caso de descumprimento: multa equivalente a 40 % sobre Imposto (LEI 5.530/89, ART. 78, I, "L").

Prazo para Autorregularização: 30 dias.

Identificação da Autoridade Administrativa

Nome: [redacted]
AUDITOR FISCAL DE RECEITAS ESTADUAIS-A
Matrícula: [redacted]

B. Additional results

Table B1: Descriptive statistics of the key tax variables

Variable description	Mean	Standard Deviation	Median	Min	Max	Sample size	% zeros
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Value of ICMS tax declared at baseline	59,649.78	354,476.81	0.00	0.00	2,211,9595.88	8429	55.72
Value of ICMS tax evaded at baseline[1]	32,663.38	84,248.17	11157.84	1000.51	3,192,400.00	8429	0.00
Value of ICMS correction at follow-up	4,390.83	48,818.36	0.00	-869,371.66	2,548,606.27	8429	92.40
Tax paid at follow-up	1,510.28	32,242.56	0.00	0.00	1,592,312.88	8429	97.57

Notes 1: The evaded tax is calculated as a difference between the value calculated by the tax administration, based on third party information, and the value declared by the firm. This table presents descriptive statistics of key tax variables at baseline and follow-up. All statistics are measured at the level of tax declaration. Sample size corresponds to the total number of declarations in the study, filed in 2018 and 2021.

Table B.2: LATE Effect estimates for the value of paid tax

	ITT	LATE: received notification	LATE: read message notification
	(1)	(2)	(3)
Coefficient estimate	2.548**	2.832**	3.518**
Standard deviation	(1.078)	(1.163)	(1.133)
p-value	0.018	0.015	0.002

Notes: This table presents statistics and effect estimates for Local Average Treatment Effect (LATE) for the value of tax paid in Poisson model. The first column shows the results of the Intetn-to-treat (ITT) effect estimate in the basic specification without covariates and is equivalent to the results presented in Table 4 Panel 2 column (1). The results in column (2) are for the LATE estimation where the endogenous indicator of receiving the notification is instrumented with the assignment to treatment, while the results in column (3) are for the LATE estimation where the endogenous indicator of having read the message is instrumented with the assignment to treatment. All regressions correspond to a basic specification without covariates. Statistical significance: * 10%, ** 5%, *** 1%.