

WORKING PAPER N° IDB-WP-1621

Massive Regularization for Effective Tax Payment

Evidence from Brazil

Anastasiya Yarygina
Eduardo Iketani
Andre Martinez
Ernesto Tiburcio

Inter-American Development Bank
Institutions for Development Sector
Fiscal Management Division

January 2025



Massive Regularization for Effective Tax Payment

Evidence from Brazil

Anastasiya Yarygina, Inter-American Development Bank

Eduardo Iketani, Secretariat of Finance of the State of Para (SEFA/PA)

Andre Martinez, Inter-American Development Bank

Ernesto Tiburcio, Goldman School of Public Policy, University of California, Berkeley

Inter-American Development Bank
Institutions for Development Sector
Fiscal Management Division

January 2025



Massive regularization for effective tax payment: evidence from Brazil / Anastasiya Yarygina, Eduardo Iketani, Andre Martinez, Ernesto Tiburcio.

p. cm. — (IDB Working Paper Series ; 1621)

Includes bibliographical references.

1. Taxpayer compliance-Brazil. 2. Taxation-Effect of technological innovations on-Brazil. 3. Tax administration and procedure-Brazil. I. Yarygina, Anastasiya. II. Iketani, Eduardo. III. Martínez, André. IV. Tiburcio, Ernesto. V. Inter-American Development Bank. Fiscal Management Division. VI. Series.

IDB-WP-1621

Keywords: digitalization, tax compliance, taxpayer support

JEL Codes: H26, H30, H32, O38

<http://www.iadb.org>

Copyright © 2025 Inter-American Development Bank ("IDB"). This work is subject to a Creative Commons license CC BY 3.0 IGO (<https://creativecommons.org/licenses/by/3.0/igo/legalcode>). The terms and conditions indicated in the URL link must be met and the respective recognition must be granted to the IDB.

Further to section 8 of the above license, any mediation relating to disputes arising under such license shall be conducted in accordance with the WIPO Mediation Rules. Any dispute related to the use of the works of the IDB that cannot be settled amicably shall be submitted to arbitration pursuant to the United Nations Commission on International Trade Law (UNCITRAL) rules. The use of the IDB's name for any purpose other than for attribution and the use of the IDB's logo shall be subject to a separate written license agreement between the IDB and the user and is not authorized as part of this license.

Note that the URL link includes terms and conditions that are an integral part of this license.

The opinions expressed in this work 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.



Abstract

In recent years, tax administrations 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 them 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 a self-regularization platform in the Brazilian State of Para. The results show that the platform increased the amount of the taxes paid by 12.78 times and the probability of tax compliance by 236 percent. Overall, the effectiveness of self-regularization in recovering the evaded tax is 60 percent higher than that of the traditional audit-based approach. The amount of the correction in the declared tax increased by 2.33 times, and the probability of correction by 300 percent. Given the low marginal cost of self-regularization, the results suggest that these platforms are a remarkable opportunity for tax administrations to leverage digital transformation effectively and efficiently, improving tax compliance and increasing tax revenue.

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 thank Jose Guilherme de Souza Moitta Koury, and other SEFAZ/PA employees, for their invaluable contributions to the development of this study. We are also grateful to Monise Estorani De Faria, Eduardo Fajnzylber, and Jairo Neira for their excellent analytical inputs. We also thank Philip Keefer and the anonymous peer reviewer for reviewing the manuscript and providing valuable suggestions. This study received financial support from RG-E1807 and RG-E1908, IDB's Economic Sector Work projects, and RG-T4120, IDB's Technical Cooperation. Valuable comments were received from the IDB's Institutions for Development Sector during an internal seminar, and from tax administration experts during the June 2023 COGEF meeting, and at the XIII Ibero-American Conference on Local Financing (XIII-JIFL). The authors declare that they have no financial or material interests in the results of this study.

1. Introduction

Strengthening state capacity to collect taxes effectively and fairly is a key priority of governments around the world (Pomeranz and Vila-Bleda, 2019). Still, inefficient tax collection and high tax evasion remain daunting challenges of fiscal management in Latin America and Caribbean (LAC) countries. 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 percent of GDP (Cristia and Vlaicu, 2022).

While strengthening tax collection by promoting voluntary tax compliance is a high priority (Bando et al., 2021) of modern tax administrations, some important contributing factors to the high levels of tax evasion in LAC are deficiencies in the facilitation of compliance and inefficient administration of taxation (Jenkins et al., 2023). Regarding the cost of compliance, LAC countries are above the world average. Firms devote 325 hours per year 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. Still, studies show that the country loses up to R\$417 billion a year due to tax evasion,¹ and the cost of tax compliance remains high, reaching 1,492 hours per year in 2019.²

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

Massive self-regularization platforms are prominent examples of TAs’ digital transformation and of the change in focus from punitive to preventive. In Brazil, the implementation of self-regularization platforms is being consolidated at the federal and subnational levels in recent years. The federal government has initiated the CONFIA program³, which has inspired other countries in LAC to initiate similar programs (Calijuri, 2023; Calijuri and Oliveira, 2023; Tostes Neto and Calijuri, 2023). Currently, 22 states in Brazil are implementing compliance facilitation programs, 14 of which are based on self-regularization platforms (Biderman, 2023). Self-regularization platforms use massive data-crossing techniques to identify inconsistencies in tax documents (for example, tax returns), 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

¹ The Brazilian Institute of Planning and Taxation.

² According to the World Bank [Doing Business Report 2021](#).

³ <https://www.gov.br/receitaefederal/pt-br/acesso-a-informacao/acoes-e-programas/confia>

and administrative costs to tax administrations, lower compliance costs to taxpayers, and prevention of litigation.

This paper presents the results from an experimental evaluation of a self-regularization system implemented by a subnational tax administration—the Secretariat of Finance of the State of Para (Secretaria da Fazenda do Estado do Para, SEFAZ/PA). The study included firms from 1,030 business clusters⁴ for which SEFAZ/PA identified 8,529 tax returns that under-reported their Tax on the Circulation of Goods and Services (ICMS).⁵ Half of the clusters with firms that under-reported their ICMS were assigned to receive the self-regularization system treatment, which consisted of notifying firms about detected discrepancies in tax returns and encouraging them to self-regularize through the self-regularization platform. The other half were 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 analyzes the effect of the intervention on the 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 the Poisson model⁶ for the outcomes skewed at zero (e.g., the amount of tax paid).

The study results 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 paid by firms in the control group. The average recovered tax represents 8 percent of the evaded amount, which is 60 percent more than the yield of the alternative audit-based approach. The intervention increased the probability of paying the tax by 236 percent, from 0.011 percentage points (pp) in the control group to 0.037 pp in the treatment group. Coupled with the abovementioned results for the amount of tax paid, these findings show that the intervention was more effective at the intensive margin than the extensive margin.⁷ The analysis of the heterogeneous effects indicates that the effect of the intervention was stronger for firms with higher detected evasion registered in smaller regional delegations of SEFAZ/PA.

⁴ 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 one to three firms.

⁵ The ICMS is the most important tax collected by states and accounts for almost 85 percent of revenue at this level of government (data for 2022, from the State Tax Collection Bulletin - CONFAZ). It is similar to a value-added tax (VAT), and it is imposed on the circulation of goods, electricity, interstate and intermunicipal transportation services, and communications.

⁶ See, for example, Wooldridge (2010). Poisson distribution characterizes rare occurrences, in which a large share of values are zero, like in our setting. When the underlying data is distributed in this way, it is more accurate to estimate expected marginal change using the Poisson regression.

⁷ The intensive margin refers to the decision of whether to amend or to pay the tax, and the extensive margin refers to the decision of how much to amend or to pay.

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

This study contributes to several strands of literature. First, it contributes to the growing evidence on using third-party information to improve tax compliance. Almunia and López Rodríguez (2016) find that the availability of information works as a complement to the audit of companies. Similarly, Allingham and 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 tax evasion by bolstering the effectiveness of monitoring actions like audits. It can also reduce the indirect cost of taxation by enabling the tax authority to provide prepopulated tax returns. 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, as 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 (Brockmeyer and Saenz Somarriba, 2022; Kleven et al., 2011; Naritomi, 2019; Pomeranz, 2015). Notably, evasion rates on taxable income subject to third-party reporting, such as wages and capital income, are negligible. This contrasts with higher evasion observed for taxes that rely on self-reported liabilities, where third-party reporting is absent (Kleven et al., 2011; Okunogbe and Santoro, 2023). Self-regularization platforms are a notable example of using third-party information in massive data crossings to detect tax evasion and noncompliance.

Second, the study contributes to the literature on the cost of tax compliance. The cost of tax filing and the complexity of tax codes have been increasing in recent years, and taxpayers perceive them as encouraging evasion (Benzarti and Wallossek, 2024). Other than that, smaller taxpayers bear compliance costs disproportionately (Okunogbe and Santoro, 2023). At the same time, traditional collection methods are sometimes costly, time-consuming, and inefficient to administer. In Brazil,⁸ 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 amounted to 50.4 percent of Brazilian GDP. Integrating electronic invoicing and prefilling systems is emerging as a pivotal strategy in mitigating the costs of tax compliance and administration, as highlighted in recent research (Hesami, Jenkins, and Jenkins, 2023;

⁸ IDB Blog: [Oito desafios relacionados ao contencioso tributário brasileiro e possíveis soluções](#).

Okunogbe and Santoro, 2023). Self-regularization platforms that facilitate the amendment of tax returns aim to reduce tax compliance costs. These technologies not only alleviate burdens associated with tax compliance but also cater to the evolving needs of businesses, fostering economic development.

Third, the study contributes to the large body of literature that analyzes 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. Studies by Brockmeyer, Kettle, and Smith (2018) and Kettle et al. (2016) underscore the importance of tailored messaging and contextual factors in enhancing tax compliance efforts.⁹ Brockmeyer, Kettle, and Smith (2018) emphasize the effectiveness of tax authority messages that provide specific details about tax inconsistencies, enhancing the credibility of the tax authorities and fostering a perception of increased scrutiny. Similarly, Kettle et al. (2016) highlight 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 (2015a, 2015b) show that reinforcing messaging with personal visits and telephone calls can yield higher compliance. 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 to messages provide further evidence of the importance of providing detailed information on inconsistencies detected 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 considered the recommendation from the literature and included detailed information on the identified tax discrepancy with 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 tax administrations' digitalization, 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, Yarygina, and Escalante, 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 closely related to ours are Bando et al. (2021) and Cafe, Yarygina, and Escalante (2024), evaluating self-regularization systems for simplified tax regime taxpayers in the Brazilian States of Piauí and Bahia. Our study adds to this literature by providing the first rigorous evidence on the effectiveness of self-

⁹ Mascagni et al. (2017) and Mogollon, Ortega, and Scartascini (2019) also explore different methods of message delivery, such as e-mails, text messages, phone calls, and letters.

regularization systems for regular regime taxpayers in Brazil and by looking at tax payment outcomes.

Finally, this study also relates to the literature on the estimation methodology for the outcomes skewed to zero. Estimating the impact on the outcome variables with a large proportion of zero values is a methodological challenge. As recent literature documents (Chen and Roth, 2022; Mullahy and Norton, 2024; 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, the 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.

2. Study Background

In Brazil, the most important tax collected by State Tax Administrations (SEFAZ) is the ICMS, which accounts for almost 85 percent of the revenue of Brazilian States.¹⁰ In 2023, the state of Para had more than 300,000 active¹¹ ICMS taxpayers, with this tax accounting for 87 percent of the state's tax revenue and more than 54 percent of the total state's budget revenue. Of the ICMS revenue, 98 percent comes from taxpayers registered in the regular tax regime, with the remaining 2 percent corresponding to the simplified tax regime Simples Nacional.¹²

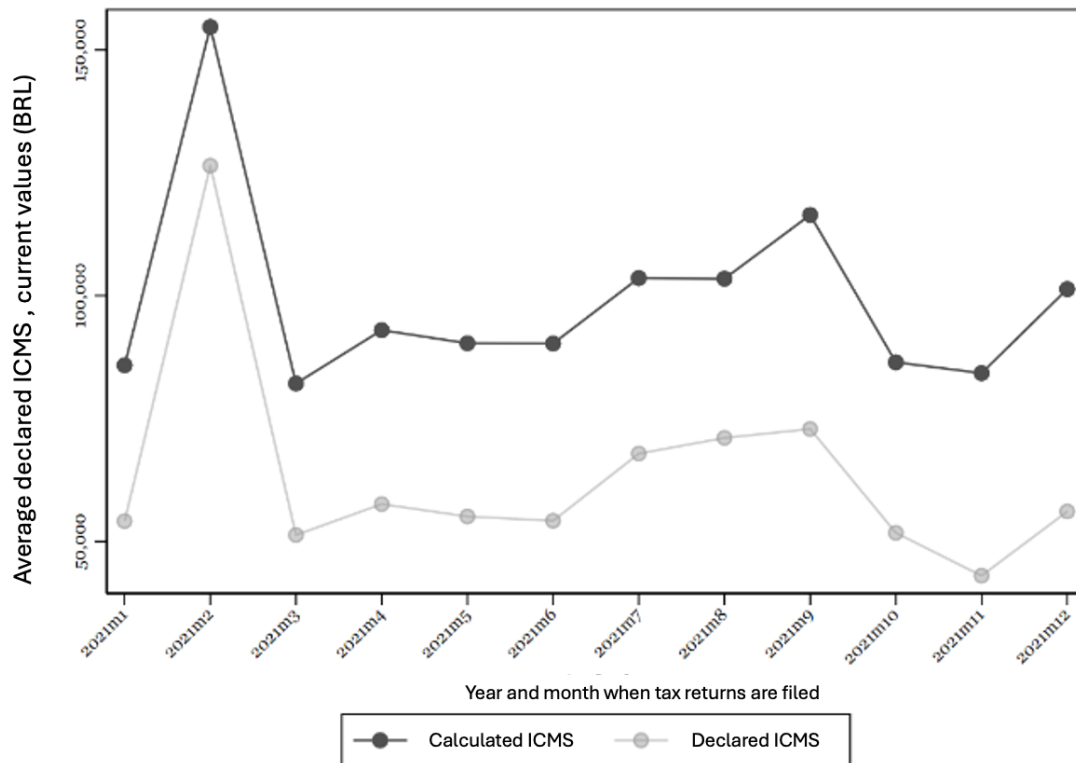
The State of Para has around 20,000 regular tax regime taxpayers. Figure 1 shows the amount of the ICMS tax declared by regular regime taxpayers and calculated by the tax administration based on third-party information. It shows that the amount of the tax declared by taxpayers is, on average, lower than 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 one monthly ICMS tax return.

¹⁰ 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).

¹¹ By way of comparison, the state's population is around 8 million.

¹² The Simples Nacional simplified tax regime was established by Federal Complementary Law No. 123 on December 14, 2006. It is targeted to small and micro firms with gross revenue of R\$3.6 million or less per year.

Figure 1. ICMS Tax Gap in Returns of Regular Tax Regime Taxpayers in 2021



Note: The figure shows the evolution of the tax gap expressed as a divergence between the declared and calculated amount of ICMS in the study analysis sample. Calculated ICMS is the value of tax obtained by the tax administration based on third-party information (credit card transactions, accounting records, digital tax accounting, and electronic documents). The declared ICMS corresponds to the values in taxpayers' tax returns.

Self-regularization platforms have the potential to improve tax compliance and increase tax collection by 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 time-consuming. 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 percent of the detected tax evaded being paid. Relying on self-regularization platforms can potentially save resources to SEFAZ and reach results improving compliance, which is important in the context of the limited capacity which SEFAZ/PA faces: in 2022, the state had approximately 269 auditors who perform other tasks and duties besides implementing audits. In contrast to the traditional audits, self-regularization does not

impose penalties or fines to taxpayers.¹³ However, failing to self-regularize during the self-regularization period can trigger penalties and fines.

SEFAZ/PA continually improves its self-regularization systems by including more data crossings and expanding the scope to include 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 and 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 discrepancies between the declared and calculated ICMS tax. The taxpayers in the treated group 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 detailed information on the detected inconsistency in the tax return, the instructions to self-regularize, and the information on penalties if they did not self-regularize. The notifications were focused on the tax returns filed in 2018 and 2021¹⁴ and were sent by regional agencies of SEFAZ/PA between July 29 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 amount on tax returns.

3. Methodology

3.1. Study Design

The study adopted an experimental evaluation method (randomized controlled trial, or RCT¹⁵), 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.¹⁶ This clustered design was implemented to prevent some tax evading firms of the same cluster from receiving the notification while other firms from the same cluster did not.

The information crossings were performed on June 1, 2022. The study focused on 1,392 firms from 1,030 business clusters with the amount of ICMS in tax returns smaller than the amount calculated by SEFAZ/PA, with the discrepancy between declared and calculated amount of R\$20,000¹⁷ or more. Table 1 presents some key information on the

¹³ Cafe, Yarygina, and Escalante (2024) discuss the implications of the absence of penalties for self-regularization, in particular, on taxpayer risk perception and future tax compliance behavior.

¹⁴ The evaluation design decisions were taken in early 2022. It was decided to work with the tax returns presented in 2021 because of the recency of the tax filing. It was also decided to include the tax returns presented in 2018 because SEFAZ/PA was still in the legal period to act upon the inconsistencies detected in these returns. An example notification is presented in the Annex A.

¹⁵ The key feature of an RCT is that the units that receive the treatment are otherwise identical to those that do not. Therefore, the difference in their average performance is credited to the treatment.

¹⁶ A business cluster comprises a matrix firm and its subsidiaries.

¹⁷ About US\$ 4,000 (exchange rate September 23, 2023). The threshold was defined based on the SEFAZ/PA capacity restrictions and study sample size requirements.

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

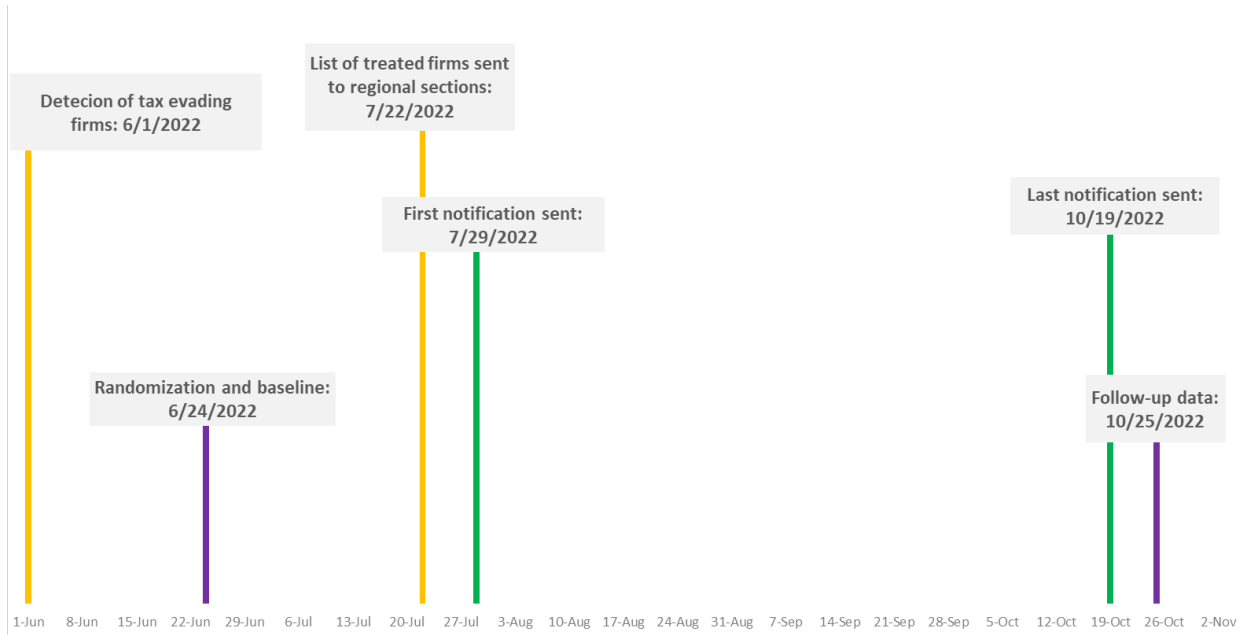
	Universe	Study Sample
Groups	4,378	1,030
Firms	5,767	1,392
Number of Tax Returns	52,073	8,529
Discrepancy	R\$ 4,039.88	R\$ 32,627.24
Declared tax	R\$ 71,896.05	R\$ 59,649.78

Notes: This table shows some key statistics for the 2021 ICMS tax returns for the universe of firms subject to self-regularization system data crossings, and the firms in the study sample. Discrepancy denotes the average difference between the declared tax and the amount 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 level. In total, the study sample comprised 8,529 tax returns. The firms were assigned to the treatment and control groups at the business cluster level. The random assignment, which was performed on June 24, 2022 using Stata 15 software, resulted in 515 clusters assigned to the treatment and 515 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 inconsistencies detected. The first notification was sent on July 29, 2022, and the last on October 19, 2022. The estimations of the intervention impact were performed with the data collected on October 25, 2022. Figure 2 presents the timeline of the intervention and evaluation.

Figure 2: The Study Timeline



Note: The figure shows the study timeline.

3.2 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. To study the impact on binary variables, such as 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. The control variables are progressively included to control for whether the year of tax return is 2018, the regional section of SEFAZ/PA where the taxpayer is registered, and the year in which the firm started its economic activity.¹⁸ 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 return 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.

¹⁸ Throughout the main analysis we include three control variables. The first one is whether the return year is 2018. We do this to account for any time-variant effects, affecting the treated and untreated units similarly. The second one is the regional section of SEFAZ/PA. We do this to account for any regional dynamics affecting treated and untreated units within each region similarly. The third one is the year in which the firm was established. We do this to account for any cohort dynamics affecting treated and untreated units similarly.

Unlike OLS estimators, the Poisson model employs the logarithm for the difference in expectations, so it is robust to the units of measurement.¹⁹ Particularly in the Poisson model, for a binary treatment variable, the coefficient estimate denotes the difference in logarithms of the amount in two groups: $\ln(E[Y(1)]) - \ln(E[Y(0)])$ (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.²⁰ The interpretation of the Poisson model coefficient estimate is that, in the treatment group, the outcome variable is $\exp^\beta$ 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 tax return 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:

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

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,

$$\begin{aligned} \mu_{ijg} &= \exp(\alpha + \beta x_{ijg}) \\ &= e^\alpha e^{\beta x} \quad (3) \end{aligned}$$

3.3 Data

The firms were randomly assigned into treatment and control groups on June 1, 2022, and 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 returns before the start of the intervention. These firms were removed from the study sample.²¹ 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 8,429 analyzed returns (4,224 from treated firms and 4,255 from control firms).

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

¹⁹ As the recent literature documents (Chen and Roth, 2022.; Mullahy and Norton, 2024; 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.

²⁰ The proportion of nonzero values in the outcome variable “value of tax paid” is 1.12 percent in the control group and 3.74 percent in the treatment group.

²¹ These firms would not have been eligible for the intervention or included in the pool of firms subject to the random assignment.

constitution, and accountant identification. All variables are from the SEFAZ/PA administrative data.

4. Results

4.1. Balance

The random assignment to the treatment and control groups guarantees that, on average, there is no significant difference between the treatment and control groups, 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 tax return 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 amount 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 30,000-35,000 reais per return (US\$5,700-6,700), resulting in a total tax evasion per business cluster of about 270,000 reais (more than US\$50,000). Half of the returns declared zero ICMS. Table B.1 in Annex B shows more descriptive statistics.

4.2. 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 can be seen, compliance with the intervention was not perfect: 86 percent of the treatment group firms were notified, while some of the control group firms (0.8 percent) also received the notification. This occurred because the SEFAZ/PA central office did not make the decision to notify any given firm – this task 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 choose 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 imperfect compliance with the intervention, with the coefficient estimate for the binary indicator of the notification sent at 85 pp.

Another informative indicator is the rate of message reading. As seen in Table 3, only 37 percent 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, which limited the impact of the intervention. However, this result is in line with the amount found in the recent studies by Bando et al. (2021) and Cafe, Yarygina, and Escalante (2024) on self-regularization interventions, which focused on simplified tax regime taxpayers.²²

²² In this study, the proportion of the firms who read the message is 37 percent, similar to the findings in Cafe, Yarygina, and Escalante (2024), who find a similar value of this indicator for the firms in the “message” treatment. In this study, the ratio of 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 percent read the message.

Table 2: Balance at Baseline

Variable description	Treatment mean (1)	Control mean (2)	Difference treatment – control (3)	Standard error (4)	P-value (5)	Treatment sample (6)	Control sample (7)
Panel 1: Business cluster level							
Amount of ICMS tax declared at baseline	572,161.79	407,458.64	164,703.145	(140,562.435)	0.241	512	515
Amount of ICMS tax evaded at baseline ¹	285,143.51	251,118.78	34,024.729	(33,300.907)	0.307	512	515
Number of firms per business cluster	1.31	1.38	-0.078	(0.087)	0.369	512	515
Number of returns per business cluster	8.25	8.17	0.085	(0.502)	0.865	512	515
Panel 2: Firm level							
Has accountant (yes =1)	0.77	0.78	-0.010	(0.032)	0.760	669	713
Years firm operates	12.21	13.19	-0.977	(0.980)	0.318	669	713
Year in which firm begins operating	2010.11	2009.15	0.960	(0.982)	0.328	669	713
Firm is a limited company (yes=1)	0.61	0.63	-0.017	(0.041)	0.680	669	713
Firm is an individual entrepreneur (yes=1)	0.24	0.22	0.013	(0.031)	0.674	669	713
Firm's Social Capital	50,173,320	44,680,787	5,492,532	(45,439,401)	0.903	669	713
Firm is registered in the State capital (Regional Section Belem)	0.19	0.20	-0.017	(0.032)	0.604	669	713
Number of returns presented by the firm	6.31	5.90	0.416	(0.366)	0.255	669	713
Panel 3: Return level							
Return month	6.89	6.85	0.044	(0.106)	0.682	4,224	4,205
Amount of ICMS tax declared at baseline	69,352.94	49,902.78	19,450.162	(16,890.802)	0.249	4,224	4,205
Amount of ICMS tax evaded at baseline ¹	34,562.85	30,755.33	3,807.517	(38,40.497)	0.321	4,224	4,205
Amount of declared ICMS at baseline is zero (yes=1)	0.57	0.54	0.028	(0.037)	0.459	4,224	4,205
Return is in 2021 (yes=1)	0.64	0.68	-0.038	(0.027)	0.163	4,224	4,205
Return is in 2018 (yes=1)	0.36	0.32	0.038	(0.027)	0.163	4,224	4,205

Notes: 1: The evaded tax is calculated as a difference between the amount calculated by the tax administration, based on third-party information, and the amount 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 tax return (Panel 3). Columns (1) and (2) show the means for the treated and control groups, respectively. Column (3) shows the difference between these amounts 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

	(1)	(2)	(3)	(4)	(5)	(6)
Variables description	Difference	Standard error	P-value	Sample	Treatment mean	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%.

4.3. Main Results: Tax Payment

Table 4 presents the results for the main outcome of interest of this study—tax payment. The program effect coefficients, 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 return. However, the estimates in the basic specification without controls and in the specification with additional control variables are imprecise. Since the distribution of the right-hand side variable deviates from the normal variable distribution and presents a large proportion of zeros, a more appropriate model for the effect estimation is a Poisson model. Panel 2 of Table 4 presents the results of the Poisson model estimations. Compared to the OLS, all coefficient estimates in the Poisson model are precise and statistically significant at 5 percent statistical significance.

The size of the effect estimated in the 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 = 2,796.1, 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 the potential to increase tax revenue by about US\$500 (2,500/5) per tax return, or US\$3,150 (500 x 6.3 returns) per firm. Given the average amount of the tax evaded per return of about 32,000 Brazilian reais,²³ the intervention is able to recover about 8 percent (2,500/32,000) of the evaded tax, which is 60 percent 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 auditing is challenging to estimate, it is arguably higher than the cost of self-regularization, given that auditing involves hours of work of tax auditors who investigate, process, reach the firms, and attempt to recover the tax due. Taking as a reference the per firm cost of the self-regularization system

²³ The average values of the treatment and control group means are reported in Table 2.

reported in Bando et al. (2021), of US\$28 per firm, the self-regularization evaluated in this study costs less than one U.S. cent (US\$0.0089 = 28/3,150) for every U.S. dollar of ICMS payment recovered.

Table 4: Intervention Effect on Tax Payment

	No controls		Controls: return year 2018		Controls: return year 2018, regional delegation, year of firm establishment	
	Effect estimate	Standard error & p-value	Effect estimate	Standard error & p-value	Effect estimate	Standard error & p-value
	(1)	(2)	(3)	(4)	(5)	(6)
Panel 1: OLS model estimates						
	2577.397	(2084.365) 0.217	2649.582	(2145.965) 0.217	2436.109	(1907.859) 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) 0.018	2.594**	(1.078) 0.016	1.794**	(0.822) 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 treatment effect estimates for the amount 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 is that the average amounts 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 amounts of the treatment group, conditional on the controls, are equivalent to $\exp(\beta)$ in relation to the control group. Columns (1), (3), and (5) show the coefficient estimates, control and treatment groups means and sample sizes, columns (2), (4), and (6) show the standard errors and the p-values of the test of equality to zero of the coefficient estimate. The estimation results in columns (3) and (4) are from the models that include the indicator of the year of return = 2018. The estimation results in columns (5) and (6) are from the models that include the year of return = 2018 indicator, the regional section and the firm establishment year fixed effects. Statistical significance: * 10%, ** 5%, *** 1%.

In addition to the amount of the tax paid, we also explore the results for the tax payment rate (that is, the probability of paying the tax). These results are reported in Table 5. On average, the probability of paying the tax in the treatment group is about 0.025–0.027 pp larger than in the control group. While the effect represents 240 percent of the control group mean (0.0112), it is also worth noting that the proportion of delinquent firms that pay the tax in both the treatment and the control groups is very small, 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 regard, the intervention results are driven mainly by the effects at the intensive margin rather than the extensive margin.

Table 5: Intervention Effect on Tax Payment Rate

	No controls		Controls: return year 2018		Controls: return year 2018, regional delegation, year of firm establishment	
	Effect estimate	Standard error & p-value	Effect estimate	Standard error & p-value	Effect estimate	Standard error & 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 treatment effect estimates on the rate of tax payment (the probability of paying tax). The interpretation of the coefficient estimate is that the average amount in the treatment group, conditional on the controls, is β percentage points higher than the control group. Columns (1), (3), and (5) show the coefficient estimates, control and treatment groups means and sample sizes, columns (2), (4), and (6) show the standard errors and the p-values of the test of equality to zero of the coefficient estimate. The estimations result in columns (3) and (4) are from the models that include the indicator of year of return = 2018. The estimation results in columns (5) and (6) are from the models that include the year of return = 2018 indicator, the regional delegation and the firm establishment year fixed effects. Statistical significance: * 10%, ** 5%, *** 1%.

The coefficient estimates presented in Tables 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, it is necessary to account for imperfect compliance.²⁴ 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. The LATE effect for the rate of tax payment is thus 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 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 results obtained, firms tend to respond more intensively on messages for older returns (tax return filed in 2018); however, the same cannot be said for the amount of the tax paid.²⁵ The amount of the tax paid and the tax payment rate tend to be higher in a subsample of taxpayers with a large amount of tax evasion at baseline.²⁶ Finally, there does not seem to be any statistically significant program effect in the subsample of tax returns presented in large regional agencies of

²⁴ Imperfect compliance arises when not all units assigned to receive the treatment get it and (or) when some units assigned not to receive the treatment end up getting it.

²⁵ The Poisson estimation for the amount of tax paid is not reported for a subsample of returns presented in 2018 because the algorithm did not converge.

²⁶ The tax returns with a high amount of tax evasion at baseline are those with a divergence between the calculated and declared tax amount higher than the median.

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 may be interested in targeting specific groups of taxpayers to achieve better results. In the case of this intervention, it appears that the program had a higher effect in smaller regional sections among taxpayers who evaded more.

Table 6: Heterogeneous Effects on the Amount of Taxes Paid 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)
Amount 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 amount of tax paid and the OLS model estimates for the tax payment rate. 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 the amount of tax paid (line one in column two) is not reported because the algorithm did not converge. Statistical significance * 10%, ** 5%, *** 1%.

4.4. Mediating Pathways: Tax Return Correction

As previously explained, taxpayers received a message specifying the discrepancy between the declared and calculated amount of the ICMS tax for each tax return with this inconsistency and invited taxpayers to amend the declared amount, which would result in a higher amount of tax due. This section explores how the intervention affected taxpayers' behavior regarding tax corrections. In Table 7, we present the results for the amount of declared tax correction, and in Table 8 for the correction rate (that is, the probability of correcting the tax on a tax return). Table 7 shows the results for the amount of corrected tax. The declared tax amount is also a truncated variable exhibiting bunching of values at zero, however, the concentration of values at zero is less pronounced than for the amount of tax paid variable.²⁷ Table 7 presents the results for the OLS model (Panel 1) and the Poisson model (Panel 2).

The intervention, on average, resulted in a larger correction of the declared tax in the treatment group firms (mean amount of 6,280 reais) than in the control group (mean amount 2,700.8 reais), with the effect estimate ranging from 3,500 reais (about US\$700) to 4,200 reais (about US\$800). The effect estimates are statistically significant at 10 percent in the basic specification and in the specification with the year of the tax return control variable. The statistical significance of the estimates improves and reaches the 5

²⁷ The probability of correcting the return after the intervention is 3 percent in the control group and 12 percent in the treatment group, compared to the probability of paying the tax of about 1 percent in the control group and 4 percent in the treatment group.

percent 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 amount of the 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 percent 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 the 5 percent level of statistical significance.

Table 7: Intervention Effect on the Amount of Tax Correction

	No controls		Controls: return year 2018		Controls: return 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						
	3,579.840*	0.091	3,753.597*	0.079	4,257.929**	0.039
	(2,113.849)		(2,135.848)		(2,058.253)	
Control mean	2,700.8		2,700.8		2,700.8	
Control sample	4205		4,205		4,205	
Treatment mean	6,280.7		6,280.7		6,280.7	
Treatment sample	4,224		4,224		4,224	
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	2,700.8		2,700.8		2,789.7	
Control sample	4205		4205		4071	
Treatment mean	6,280.7		6,280.7		6,451.8	
Treatment sample	4224		4224		4112	

Notes: This table presents statistics and treatment effect estimates for the amount of the correction on tax returns in an OLS model (Panel 1) and Poisson Model (Panel 2). The interpretation of the coefficient in Panel 1 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 $\exp(\beta)$ in relation to the control group. Columns (1), (3), and (5) show the coefficient estimates, standard errors, control and treatment groups means, and sample sizes, 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 the models that include the indicator of year of return = 2018. The estimation results in columns (5) and (6) are from the models that include the year of return = 2018 indicator, the regional section and firm establishment year fixed effects. Statistical significance: * 10%, ** 5%, *** 1%.

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

Table 8: Intervention Effect on the Tax Correction Rate

	No controls		Controls: return year 2018		Controls: return 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*** (0.018)	0.000	0.089*** (0.017)	0.000	0.090*** (0.017)	0.000
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 effect estimates for the tax return correction rate (probability that the tax return will be corrected). The interpretation of the coefficient estimate is that the average amount in the treatment group, conditional on the controls, is β percentage points higher than the control group. Columns (1), (3), and (5) show the coefficient estimates, standard errors, control and treatment groups means and sample sizes; columns (2), (4), and (6) show the p-values of the test of equality to zero of the coefficient estimate. The estimation results in columns (3) and (4) are from the models that include the indicator of year of the return = 2018. The estimation results in columns (5) and (6) are from the models that include the year of the return = 2018 indicator, the regional delegation, and firm establishment year fixed effects. Statistical significance: * 10%, ** 5%, *** 1%.

5. Conclusions

The government's ability to collect taxes efficiently is critical to ensuring the sustainability of state finances and provide fiscal space for productive investment and social spending, leading to economic growth and development. Promoting tax compliance and strengthening tax collection are, therefore, priority goals of modern tax administration. To achieve these goals, tax administrations can leverage digitalization and use administrative information to detect evasion and effectively lower taxpayer compliance costs. Moreover, a more efficient tax administration increases state capacity and the availability of resources for financing government and finalistic policies, as well as delivering more productive investment and services to citizens. If supported by a focus on taxpayers and the search for an improved relationship between public administration and citizens, this approach can bolster government legitimacy and trust in institutions.

This study evaluates the effectiveness of a self-regularization system, which leverages the new technological capabilities of the tax administration to perform massive data crossings, inform taxpayers about inconsistencies detected, and facilitate self-regularization through a dedicated platform. The implementation of the self-regularization platform for a group of delinquent firms who under-declared the tax amount owed on their tax returns shows that the self-regularization system has substantial effects on the payment of taxes and corrections of tax returns. Specifically, the effectiveness of the self-regularization system is 60 percent higher than the that of the traditional approach based on personalized audits. This study presents the results three months after the intervention began and finds that that the amount of tax paid by firms in the treatment group was 12.78 times higher than that paid by firms in the control group. The intervention increased the probability of paying the tax by 236

percent (from 0.011 pp in the control group to 0.037 pp in the treatment group). Its long-run impacts depend on the dynamics of the taxpayer's behavior.

The results presented in this study complement the discussion in recent research in Brazil (Cafe, Yarygina, and Escalante, 2024) on whether the expected change in compliance regarding future tax obligations can depend on the perceived risk and penalty. Given the low marginal cost of self-regularization compared to auditing, the self-regularization intervention is arguably more cost-efficient than the traditional audit-based approach. It is also considerably less demanding to taxpayers, decreasing the cost of compliance. The design and content of self-regularization messages sent to delinquent taxpayers were based on the lessons learned from the literature and included detailed information on the detected inconsistencies, clear step-by-step instructions on how to amend the tax return, and information on penalties and fines for 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, Yarygina, and Escalante, 2024). In addition, this study provides new evidence for the effects of self-regularization systems in Brazil on the payment of taxes for regular tax regime taxpayers. These results are particularly important in Brazil, where, in recent years, sub-national jurisdictions have been implementing self-regularization systems. Tax administration in other countries and regions interested in substituting or complementing the traditional audit-based approach with self-regularization platforms may also benefit from the results of this study, given that the intervention is scalable and customizable to different contexts.

References

- Almunia, M., D. López-Rodríguez, and 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., and A. Sandmo. 1972. Income Tax Evasion: A Theoretical Analysis. *Journal of Public Economics* 1(3-4): 323–38.
- Antinyan, A., and Z. Asatryan. 2019. Nudging for Tax Compliance: A Meta-analysis. ZEW-Centre for European Economic Research. Discussion Paper 19-055. Mannheim: Leibniz Centre for European Economic Research.
- Bando, R., O. Canozzi, J. Martínez, and A. L. Dezolt. May I calculate your taxes? The Effect of Bookkeeping on Tax Compliance Under A Simplified Regime. IDB Working Papers No. 11381. Washington, D.C.: Inter-American Development Bank.
- Benzarti, Y., and L. Wallossek. 2024. Rising Income Tax Complexity. *National Tax Journal* 77(1): 143–73.
- Biderman, C. 2023. Portais de Autorregularização dos Estados Brasileiros. Unpublished consultancy product for the Inter-American Development Bank.
- Brockmeyer, A., and M. Sáenz Somarriba. 2022. Electronic Payment Technology and Tax Compliance: Evidence from Uruguay's Financial Inclusion Reform. World Bank Policy Research Working Paper No. 9947. Washington, D.C.: World Bank.
- Brockmeyer, A., S. Kettle, and S. Smith. 2016. Casting the Tax Net Wider: Experimental Evidence from Costa Rica. World Bank Policy Research Working Paper No. 7850. Washington, D.C.: World Bank.
- Cafe, R., A. Yarygina, and L. Escalante. 2024. Leveraging Data to Improve Tax Compliance for Micro and Small Firms: Evidence from Brazil. IDB Working Papers No. 1609. Washington, D.C.: Inter-American Development Bank.
- 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., and P. T. P. de Oliveira. 2023. Cooperative Compliance in Brazil: What Does the Future Hold?. IDB Technical Note, 22793. Washington, D.C.: Inter-American Development Bank.
- Castro, J. F., D. Velásquez, A. Beltrán, and G. Yamada. 2020. Spillovers and Long-Run Effects of Messages on Tax Compliance: Experimental Evidence from Peru. *Institute of Labor Economics* 1(1): 1 – 50. IZA DP, n. 13974.
- Castro, L., and C. Scartascini. 2013. Tax Compliance and Enforcement in the Pampas. Evidence from a Field Experiment. IDB Working Papers, 472. Washington, D.C.: Inter-American Development Bank.

- Chen, J., and J. Roth. 2022. Log-like? ATEs Defined with Zero Outcomes are (arbitrarily) Scale-dependent. arXiv preprint arXiv:2212.06080, 21, 30.
- Coolidge, J., and F. Yilmaz. 2016. Small Business Tax Regimes. Brief. *Viewpoint* No. 349. Washington, D.C.: World Bank.
- Cristia, J. P., and Vlaicu, R. 2022. Digitalizing Public Services Opportunities for Latin America and the Caribbean. IDB Report. Washington, D.C.: Inter-American Development Bank.
- Dom, R., A. Custers, S. Davenport, and W. Prichard. 2022. Innovations in Tax Compliance: Building Trust, Navigating Politics, and Tailoring Reform. World Bank Publications. Washington, D.C.: World Bank.
- Hesami, S., H. P. Jenkins, and G. P. Jenkins. 2023. Emerging Digital Technologies to Improve Tax Compliance and Administration Efficiency: A Systematic Literature Review. Available at SSRN 4580004.
- 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. Washington, D.C.: Inter-American Development Bank.
- Kettle, S., M. Hernandez, S. Ruda, S., and M. A. Sanderson. 2016. Behavioral Interventions in Tax Compliance: Evidence from Guatemala. World Bank Policy Research Working Paper 7690. Washington, D.C.: World Bank.
- Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez. 2011. Unwilling or unable to cheat? Evidence from a Tax Audit Experiment in Denmark. *Econometrica* 79(3): 651–92.
- Kuchumova, Y. P. 2017. The Optimal Deterrence of Tax Evasion: The Trade-off between Information Reporting and Audits. *Journal of Public Economics* 145: 162–80.
- 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 No. 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. ICTD Working Paper N. 58. Brighton, United Kingdom.
- Mogollón, M., D. Ortega, and C. Scartascini. 2019. Who's Calling? The Effect of Phone Calls and Personal Interaction on Tax Compliance. IDB Working Papers No. IDB-WP-1084. Washington, D.C.: Inter-American Development Bank.
- Mullahy, J. and E. C. Norton. 2024. Why Transform Y? The Pitfalls of Transformed Regressions with a Mass at Zero. *Oxford Bulletin of Economics and Statistics* 86(2): 417–47.

- Naritomi, J. 2019. Consumers as Tax Auditors. *American Economic Review* 109(9): 3031–3072.
- Norton, E. C. 2022. The Inverse Hyperbolic Sine Transformation and Retransformed Marginal Effects. *The Stata Journal* 22(3): 702–12.
- Okunogbe, O., and F. Santoro. 2023. The Promise and Limitations of Information Technology for Tax Mobilization. *The World Bank Research Observer* 38(2): 295–324.
- Ortega, D., and C. Scartascini. 2015a Don't Blame the Messenger: A Field Experiment on Delivery Methods for Increasing Tax Compliance. IDB Working Paper Series, 627. Washington, D.C.: Inter-American Development Bank.
- Ortega, D., and C. Scartascini. 2015b Who's calling? The Effect of Phone Calls as a Deterrence Mechanism. Washington, D.C.: Inter-American Development Bank. Mimeographed document.
- Pomeranz, D. 2015. No Taxation Without Information: Deterrence and Self-Enforcement in the Value Added Tax. *American Economic Review* 105(8): 2539–69.
- Pomeranz, D., and J. Vila-Belda. 2019. Taking State-Capacity Research to the Field: Insights from Collaborations with Tax Authorities. *Annual Review of Economics* 11: 755–81.
- PricewaterhouseCoopers. 2020. Paying Taxes 2020: The Changing Landscape of Tax
- Tostes Neto, J., and M. Calijuri. 2023. Transformando la relación contribuyente-administración tributaria en América Latina: el papel de los programas de cumplimiento cooperativo. IDB Monograph 1150. Washington, D.C.: Inter-American Development Bank.
- Wooldridge, J.M., 2010. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: 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 five 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 Piauí Brazilian State (Bando et al., 2021), which found that giving detailed information to taxpayers on the delinquency in the notification improves the message's effect.

In this study, 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 tax due by SEFA/PA from the SEFAZ/PA Service Portal. The objective of giving taxpayers the information was to prove that SEFAZ/PA is aware of the delinquency and to convince taxpayers to correct their returns and pay the taxes owed.



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

Identificação do sujeito passivo:

Inscrição Estadual: [REDAZIDA]

Nome Empresarial: [REDAZIDA]

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 dos 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: [REDAZIDA]

AUDITOR FISCAL DE RECEITAS ESTADUAIS-A

Matrícula: [REDAZIDA]

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)
Amount of ICMS tax declared at baseline	59,649.78	354,476.81	0.00	0.00	2,211,9595.88	8429	55.72
Amount of ICMS tax evaded at baseline ¹	32,663.38	84,248.17	11157.84	1000.51	3,192,400.00	8429	0.00
Amount 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: The evaded tax is calculated as the difference between the amount calculated by the tax administration, based on third party information, and the amount declared by the firm. All statistics are measured at the level of the tax return. Sample size corresponds to the total number of returns in the study, filed in 2018 and 2021.

Table B2: LATE Effect Estimates for the Amount of Tax Paid

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

Notes: This table presents statistics and effect estimates for the local average treatment effect (LATE) for the amount of tax paid in the Poisson model. The first column shows the results of the Intent-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%.