# Does deterrence treatment have a dynamic impact on tax compliance? Evidence from experimental study in Indonesia [*]

Agung Satyadini [†]
Crawford School - Australian National University

First draft: Feb 4, 2024
This draft: October 14, 2024

**Abstract**

This paper analyses the dynamic effects of interventions on small businesses' tax compliance in Indonesia. Specifically, it examines whether the deterrence intervention produce sustainable and long-lasting effects rather than just immediate outcomes compared to non-deterrence interventions. We estimate the long-term impact of these interventions and resend the initial treatment letters one year after our initial trials. Our findings show that deterrence letters lead to a significant increase in tax payment rates and have a more enduring impact on the likelihood of timely filing during the observation periods. Additionally, in our repeated experiments, we observe a significant reduction in the average treatment effects across all interventions; however, deterrence letters continue to substantially increase the likelihood of tax payment.

**JEL-Classification:** C93, H25, H26
**Keywords:** Tax Compliance, Natural Field Experiment, Randomised Controlled Trials, Behavioral Insights

[*]This draft is for the Economics PhD Seminar purpose. Please do not cite.

[†]I thank Mathias Sinning, Sarah Dong, Arifin Rosid, Donghyun Park, Tanisa Tawichsri, Yothin Jinjarak, Muhammad Ali Nasir, Arya Gaduh, Dainn Wie, Yon Arsal, and Christopher Hoy for helpful comments. This paper was presented at the 2024 Asian Economic Development Conference, organised by the Asian Development Bank (ADB) in Seoul, July 2024. I gratefully acknowledge financial support from the Indonesia Endowment Fund for Education (LPDP). The opinions presented in this paper belong to the author and may not necessarily represent the stance of the Indonesian Government. All correspondence to Agung Satyadini, Crawford School of Public Policy, College of Asia and the Pacific, Stanner Building 1.07, Lennox Crossing, The Australian National University, Canberra ACT 2601, Tel: +61 416 959 833, E-mail: agung.satyadini@anu.edu.au.

# 1 Introduction

Tax compliance is crucial for the functioning of any economy. It refers to the either voluntary or compulsory behaviour of taxpayers to meet the rules and regulations set by the government – not only in paying taxes, but also in reporting tax obligations timely and accurately. Following World Bank (2022), innovations in pursuing a long-run tax compliance is essential not only to increase sustainability of revenue collection, but also to build state capacity and ultimately create a sustainable and stronger fiscal capability.

In most developing countries, establishing a sustained tax complaint behaviour is a challenge. The difficulties are exacerbated by the economic structure of the developing countries which is dominated by informal and small businesses. Following Schneider et al. (2010) and Pope and Abdul-Jabbar (2008), the administrative and compliance costs of taxation risk the competitiveness of small businesses, limit their tax awareness, increase the propensity towards the shadow economy, and downturn their long-term tax compliance. In this context, some current studies, including World Bank (2022) and OECD (2004) highlighted the importance of innovations in tax compliance by adopting three main pillars – enforcement, literacy (or facilitation), and trust.

It is worth acknowledging that tax compliant behaviour is mostly a product of both quasi-voluntary compliance and enforcement. This is due to the fact that tax compliance exists in a dynamic environment, which in turn can produce various types of compliance outcomes (McKerchar and Evans, 2009), and also the phenomenon that most countries, particularly developing countries, do not have the capacity to control and audit a large number of taxpayers (Mascagni and Nell, 2021). Therefore, mix strategy between enforcement (or deterrence) – related to the probability of audit and penalties, literacy – related to tax information and facilitation and tax morale – related to trust and warm glow effect are relevant in pursuing a sustainable tax compliance.

The topic of tax compliance has been extensively researched and covers a broad range of approaches, from quantitative and qualitative analysis to surveys and experiments. The research mainly builds on the economic deterrence model of tax compliance originally introduced by Allingham and Sandmo (1972). Abundant empirical studies also tested the determinants of tax compliance in practice. However, experimental research offers a step further in exploiting the true taxpayers' behaviour in a real-world setting rather than perceptions; but in most experimental designs, the long-term observation, dynamic treatment effects and non-deterrence model of tax compliance are still relatively unexplored.

This study explores three types of nudges – enforcement, facilitation, and trust – in the form of deterrence, literacy, and public goods provision treatment letters sent by the tax authority, and their dynamic impacts on small businesses' tax compliance. We examine the compliance proxies throughout the tax mechanisms – including tax payment and tax reporting – by considering their perceived behavioural control, transaction costs and risk aversion. We compare the potential drivers of tax compliance (i.e., turnover, business domain, locations, internet access, and geographical factors) and put them on equal footing by differentiating three types of letters. We ran a natural field experiment – a Randomised Controlled Trial (RCT) – in collaboration with Indonesia tax department – the Directorate General of Taxation (DGT) during 2021 and observe the taxpayers' responses until 2022 fiscal year. Further, we also resend our initial trials after 18 months to see whether there are sustained effects rather than immediate effect after treatments. The randomisation procedure allows us to produce an even distribution of taxpayers across group based on their baseline characteristics, and a relatively long observation period enable us to capture the dynamics of the real taxpayers' behavioural responses towards different treatment letters.

The analysis on this paper is designed to exploit the long-term treatment effects over a period of observation to see its dynamic impact on two compliance proxies: tax payment and tax reporting. To achieve this, the paper first constructs a simple theoretical model grounded in the deterrence tax compliance model, extending it with non-deterrence factors, a time-to-event dynamic model, and a reference-dependent model. This framework is used to predict how different treatment effects on taxpayers' compliance evolve over time. Subsequently, the paper estimates the progression of compliance proxies (i.e., payment and filing) over more than one fiscal year as influenced by different treatment letters, revealing the dynamic nature of taxpayers' decision-making processes in response to audits, as evidenced by Advani et al. (2023), Kotsogiannis et al. (2024), DeBacker et al. (2018), and Christiansen (2024). Thirdly, this paper aims to examine the magnitude of changes in the effects of identical trials, as we are interested in investigating the mechanisms through which taxpayers' memories and reference points, formed from their past expectations and experiences with the initial treatments, influence their responses. As the extension, this paper also attempts to investigate the dynamics of heterogeneous treatment effects based on the taxpayers' reference point of tax compliance (i.e., perceived behavioural controls), transaction costs, and risk aversion. Departed from the deterrence model, these analyses are essential to extend the non-deterrence model following reference-dependent behavioural model as suggested by Kőszegi and Rabin (2006).

Our research design, conceptual framework and empirical results thus allows us to push the frontier on the discussion of tax compliance model departed from the deterrence model in accordance with the economic framework streamlined by Allingham and Sandmo (1972) and Yitzhaki (1974a) which focused primarily on the role of deterrence, penalties and expected economic returns to explain compliance. We establish the connection between the economic deterrence model and the social-psychological studies as conducted by Kornhauser (2007), Hasseldine and Bebbington (1991), Blumenthal and Slemrod (1992), Torgler and Schaltegger (2006), which underlined the importance in improving of taxpayers' perception and manipulating of their motivation to be more tax compliant. More broadly, our conceptual framework and research design aim to make a significant contribution to the relatively underdeveloped literature on non-deterrence models. We seek to provide a comprehensive analysis of the dynamic patterns exhibited by these models at the individual level. By examining how deterrence and non-deterrence mechanisms influence taxpayer behavior over time, we aim to deepen the understanding of their long-term effectiveness and their potential to promote sustained tax compliance.

In our analysis, firstly we introduce the crucial factor of the dynamics of tax compliance to elaborate more on deterrence model: the "perceived" probability of detection – the subjective assessment of risk, that can vary substantially across individuals and societies. This framework holds significance because we are not merely viewing the taxpayer's decision from a pure economic cost and benefit rationality, but also incorporates the inclination towards risk-averse behaviour in response to deterrence. Further, we also interested in exploiting the impact of a reference point on taxpayers' decision-making. Following (Kahneman and Tversky, 1979), taxpayers decision are influenced by the perceived utility of gains and losses relative to a reference point, exhibiting a diminishing sensitivity. Consequently, taxpayers' inclination to increase their tax payments is shaped by their risk aversion, featuring a concave utility function for gains (such as tax refunds) and a convex utility function for losses (such as tax penalties).

Secondly, we introduce a linkage between the framework of transaction costs (as part of compliance costs) – which is occurred when taxpayers attempt to pay and file their returns, and the dynamics of compliant behaviour in practice. We incorporate the reduction of transaction costs in our literacy letter and see how the taxpayer's response to it. This design elaborates the question on how the tax information delivery affect the compliance mechanism in practice, as previously discussed by Slemrod et al. (2001), Kopczuk and Pop-Eleches (2007), Dyreng et al. (2016), and Dwenger et al. (2016).

Third, we elaborate the tax morale and the dynamic of payment and filing. We propose

a strong assumption that moral cost of evasion is proportional to the amount of evasion and the prevalence of evasion in a society. In this sense, we focus on intrinsic motivation, peer effects and social influences as highlighted by Luttmer and Singhal (2014) and Torgler (2004). This approach is motivated by *'join the compliant majority'* concept as studied by Blumenthal et al. (2001) where taxpayers tend to be compliant due to the motivation to be part of the majority. Our public goods provision letter attempts to put a shed of light on tax morale studies, as most of the study including Castro and Scartascini (2013) found no significant effect from such letter on equity and public services. Our conceptual framework and empirical evidence on moral cost of evasion is essential to tackle the lack of empirical support as argued by Mascagni (2018), as we also provide the empirical evidence about the dynamic effects of 'warm glow' nudges on tax compliance and its heterogeneous effects based on elasticity of compliance to perception (i.e., based on filing and penalty history).

Our natural experimental design also enables us to tackle the paucity of evidence in the literature body in terms of population-wide experimental studies. According to Mascagni (2018), thus far, there are relatively few experimental studies have attempted to quantify the influence of behavioural elements on tax compliance in developing countries; the vast majority of research has been conducted only in high-income nations. For instance, Hallsworth et al. (2017), performed a large-scale field experiment in the United Kingdom to link the moral factors, social norms and concerns over public goods. Bott et al. (2017) performed a field experiment carried out in Norway among taxpayers who earned income in a foreign country. However, in terms of experimental research that incorporating non-deterrence approach, including Coleman (2007); Slemrod et al. (2001); Wenzel and Taylor (2004); Dwenger et al. (2016) and Hallsworth et al. (2017), thus far, the empirical evidence is equivocal. Further, the lack of experimental reference in developing countries is unfortunate due to the fact that developing countries have a larger risk of tax noncompliance because of limited enforcement ability, low tax awareness, and minimum literacy levels OECD (2014).

The remainder of the paper is structured as follows. Section 2 comprises the model, framework and mechanism. Section 3 depicts data and experimental strategy. Section 4 and 5 discusses the result, and Section 7 concludes.

# 2 Model, Framework and Mechanism

This section's objective is to develop a conceptual framework and model for comprehending the dynamics of tax filing and payment behaviour as long term implication of nudges. Our approach yields observable consequences that impact the empirical model of our interventions, the dynamics of treatment effects, and the intepretation of our trial findings.

## 2.1 Tax Compliance Model

This paper begins with the classical Allingham and Sandmo (1972) model, in which taxpayers rationally select how much income to conceal from tax authorities by evaluating the benefit of not paying the evaded tax against the cost of being detected and being required to pay a penalty. Empirically we denote the evasion amount as $e$, probability of being detected as $\phi$, and income level on two states of situation: (1) $x^a$, when evasion is caught and penalised and (2) $x^b$, when evasion is not caught. Following Traxler (2010) and Luttmer and Singhal (2014), we also incorporate the intrinsic cost of noncompliance given their tax morale $S$, this encompasses the overall burden associated with the process of completing and submitting tax returns, the level of focus required to make the necessary decisions, and the intangible dissatisfaction that comes with fulfilling tax obligations. We model the utility function of tax compliance behaviour as follows:

$$U(e) = \Big(\phi u(x^a(e)) + (1-\phi)u(x^b(e))\Big) + S(e) \tag{1}$$

Simplifying $y$ as taxable income and $y'$ as reported income, we can modify $e$ as $y$ - $y'$. In this sense, $y' \in [0, y]$ captures the behavioural response in terms of payment. In general, denoting $T(y')$ as cost of paying taxes, $N$ as non-compliance element, $C$ as compliance element, and $M$ as tax morale component, the tax payment behaviour solves:

$$\min_{y' \in [0,y]} T(y') + C(y') + N(y-y') + M(y-y') \tag{2}$$

From 1 and 2, we can breakdown the total cost of compliance may include: cost of resources from paying taxes $T(y')$, intrinsic cost of evasion given taxpayer's morale $M(y-y')$, enforcement costs $N(y-y')$ and transaction cost $C(y')$.

## 2.2 Dynamic Modelling Approach

This paper investigates the long-term impact of interventions by analyzing the dynamics of treatment effects on tax payments and tax filing behavior. Additionally, it examines the cumulative risk of tax compliance events – specifically focusing on the number of individuals who: (1) reached or accumulated a certain threshold of tax payments, and (2) experienced specific tax filing events, such as filing late.

We assumed that the long-run average treatment effects depend on the instantaneous rate, which represents the probability of compliance (or non-compliance) events occurring within a very short time interval $dt$, given that these events have not occurred before $t$. Therefore the incremental instantaneous risk that taxpayers will become non-compliant over a small time interval $dt$ is the negative of the ratio between the change in the survival function over $dt$ and the survival probability just before time $t$.

The objectives of data analysis can encompass tasks such as estimating the distribution of failure times and exploring how the failure times relate to various factors, including characteristics, exposures, and other variables commonly referred to as covariates in the context of the study subjects' treatments.

It is worth noting that in time-to-event data analysis, the central analysis could be on single or multiple time variate $t > 0$, characterised by survivor function $F$, which usually assumed to be continuous from the right with left-hand limits. The probability function corresponding to $F$ is defined at time t by $-F(dt)$. Further, denoting the follow up time as $t$ and timeframe of failure event as $t_i$, we can specify the cumulative hazard function, or the failure probability $\Lambda$ as:

$$\Lambda(dt) = -F(dt)/F(t^{-}) \tag{3}$$

The survival function $\hat{F}(t)$ gives the probability that a taxpayer remains compliant up to time $t$, can be defined as:

$$F(t) = \exp\left(-\int_0^t h(s)\,ds\right) \tag{4}$$

This function can be used to estimate the probability of remaining compliant taxpayers for treatment and control groups.

## 2.3 Reference-dependent Model

As outlined earlier, the primary aim of this study is to explore the long-term sustainability of the effects generated by an intervention, rather than merely concentrating on the immediate outcomes. The study seeks to determine how enduring these effects are over time, providing a deeper understanding of their lasting impact. Additionally, a key objective is to evaluate the magnitude of changes in the effectiveness of repeated interventions when compared to the initial interventions. This comparison will help to identify whether the impacts of the interventions diminish, remain consistent, or perhaps even increase with repetition, thereby offering valuable insights into the dynamics of intervention efficacy over time.

In this analysis, we propose that taxpayers' rational decision-making concerning tax compliance is shaped not solely by their inherent preferences for specific outcomes, but also by the reference points they establish. These reference points are likely influenced by the initial interventions, which we expect to embed themselves in taxpayers' memories, experiences, or expectations. As a result, taxpayers' responses to subsequent interventions are expected to be relative to these established reference points. This suggests that the initial intervention plays a critical role in framing how future interventions are perceived and evaluated, thereby affecting the overall decision-making process. Understanding this dynamic is crucial, as it implies that the effectiveness of repeated interventions may be contingent on the benchmarks set by earlier experiences, potentially leading to variations in compliance behavior over time.

In our framework, we view that taxpayers will assume two measures of probability in their decision including probability measure of detection and probability measure of reference point. Further, we depict that reference point is a perceived outcome of compliant decision, which is could be stochastic. Adopting Kőszegi and Rabin (2006), we model taxpayers' utility of their decision with respect to the reference point as follows:

$$U(P|R) = \int \int \; u(e|r) \;\; dR(r) \;\; dP(e) \tag{5}$$

where $P$ is probability measure related to evasion (i.e., detection and penalty probability). $R$ is a probability measure of reference point. $e$ is evasion level and $r$ is reference point level - we allow a stochastic reference point.

# 3 Context, Data and Experimental Strategy

## 3.1 The Indonesian SME Taxation

As aforementioned, small firms have a crucial role in resolving the issues of poverty, inequality, and job creation, particularly in emerging economies. In the case of Indonesia, micro, small, and medium-sized businesses play a pivotal part in the socio-economic structure, dominating around 98% of all company segments and generating at least 62% of total GDP in 2020 (Ministry of Cooperatives and Small and Medium Enterprises, 2020).[1]

Theoretically, a number of studies and policy references struggle with defining the terminology "small business" (Freedman and Crawford, 2010). In the context of Indonesia, Law Number 20/2008 concerning Micro, Small and Medium Enterprises Number 20/2008 is the applicable reference for qualifying small businesses. According to this provision, micro, small, and medium-sized firms are enterprises or sole proprietorships with less than IDR10 billion in assets and IDR50 billion in turnover. This study, however, employs a more practical classification of small businesses for tax purposes based on GR-46/2013 and GR-32/2018.[2]. Effectively, this legislation went into effect on July 1, 2018. Following these laws, the yearly turnover requirement for small firms is set at no more than IDR4.8 billion, inclusive of all branches and business units.[3]

Small businesses – individuals or enterprises, with less than IDR4.8 billion annual turnover are eligible to employ the simplified tax procedures and subject to a single tax rate of 0.5% of gross monthly turnover. This tax is uncreditable in nature, must be self-assessed or withheld and paid electronically or manually by the $15^{th}$ of the following month. Regarding the tax submission, the monthly tax returns substantially due on the $20^{th}$ of the following month. As the simplification procedure, small businesses are eligible to utilise their monthly payment stub as monthly return. However, a regular annual tax return form is still mandatory. Failure to meet the submission deadline will consequence

[1] According to the Ministry of Cooperatives and Small and Medium Enterprises, Indonesia has about 63 million SMEs, with micro-enterprises comprising roughly 98% of this total. Micro-enterprises employ around 97% of the workforce and contribute significantly to GDP, highlighting their crucial role in local economic development and poverty reduction.

[2] GR-23/2018 concerning Income Tax on Operating Income Received or Acquired by Taxpayers with Certain Gross Turnover was issued on 8 June 2018 and revoked the previous GR-46/2014. See: https://jdih.kemenkeu.go.id/en/dokumen/peraturan/04afa793-f519-49e9-b34b-58eed378fc8a and https://jdih.kemenkeu.go.id/in/dokumen/peraturan/ac35bf6e-252a-43d4-8379-c68965b2c634

[3] This qualification seems to be more relevant because the utilization of turnover rather than asset as the measurement criteria. Following Freedman and Crawford (2010), using size of asset might inaccurate because it does not directly relate with actual small business characteristics.

an administrative penalty of IDR100,000.[4]

The tax department provides a set of cutting-edge, comprehensive support in the form of the adoption of internet-based services, including e-registration, e-payment, e-filing, mobile application *m-pajak*, and web-based services. The deployment of these services aims to promote long-term tax compliance and revenue. [5] It is important to note, however, that the government's initiative for the digitalisation of tax services is rolled out to all local tax offices, regardless of their locations and accessibility, through widespread campaign efforts. [6]

## 3.2 Data

To estimate the long-term effects of both initial and repeated interventions, we draw on a comprehensive dataset containing two key types of measurements: outcome measures and covariates. The data was collected monthly over a fourteen-month period, sourced from de-identified records from a range of local tax offices. Specifically, we use data from thirteen offices to assess the impact of initial interventions and data from five offices to evaluate the effects of repeated interventions.

The outcome measures are designed to capture various aspects of taxpayer behaviour and compliance. These include: (1) a dummy variable indicating whether a tax return was filed, (2) a dummy indicating whether a non-zero tax payment was made, (3) a dummy for timely filing of tax returns, (4) a dummy representing both non-zero payment and timely filing, (5) a dummy indicating improvement in tax filing, based on the Compliance Risk Management (CRM) dashboard. The CRM is an automated system used by the DGT to evaluate taxpayer compliance by considering both fiscal importance and risk profile. Fiscal importance reflects the taxpayer's contribution to revenue, while the risk profile accounts for the taxpayer's historical compliance across multiple dimensions such as registration, filing, payment, correct reporting, and penalty history[7], (6) a dummy indicating whether

[4]According to the provisions of Article 3 paragraph (3) b and c of Law No. 6 of 1983 on General Provisions and Tax Procedures that has been amended lastly by Act No. 16 of 2009.

[5]Tax compliance and long-run revenue are two main components of DGT's medium-term strategic plan as depicted in the Strategic Plan document 2018-2020 and 2020-2024. See: https://www.pajak.go.id/sites/default/files/2020-11/Buku%20Renstra%20DJP%202020-2024.pdf.

[6]According to this program, there will be 132 services digitized from 2019 to 2024, which consists of 59 automated services, 50 services with the support of a contact center, and 32 services that use the tax office's support services.

[7]The CRM system is a predictive tool adopted by the DGT to assess taxpayer compliance. It evaluates fiscal importance, which measures revenue contribution, and the risk profile, which tracks the taxpayer's history of compliance, including filing, payment, and any penalties received. By combining these factors, the CRM system allows the DGT to prioritize enforcement efforts, focusing resources on high-risk

taxpayers used electronic payment systems such as e-billing, (7) the number of inbound communications to the tax office, which captures how frequently taxpayers contact the tax office in the month following an intervention, and (8) the amount of additional tax payment[8]. Among these, the indicators of tax filing and tax payment are of primary interest in our analysis.

In addition to these outcome measures, the covariates used in the analysis are critical for understanding the broader context of taxpayer behaviour and how it may change in response to interventions. Table 1 presents the summary statistics of the covariates included in the long-term analysis, while Table 2 provides summary statistics for the repeated interventions. These covariates offer insights into various demographic, economic, and behavioural factors that could influence compliance, helping us control for external variables that might otherwise confound the results.

In Tables 1 and 2, we present the $p$-values from statistical tests comparing the sample means of each treatment group to those of the control group. These tests are conducted to assess whether there are any significant differences in covariates between the groups, providing insight into the effectiveness of the randomisation process. The $p$-values indicate the likelihood that any observed differences are due to chance rather than systematic differences between the groups.

The results demonstrate that the randomisation technique successfully achieved covariate balance between the treatment and control groups – ensuring that the baseline characteristics of taxpayers in both the treatment and control groups are comparable, minimizing potential bias. The majority of the differences in covariate means between the treatment groups and the control group are not statistically significant at the 5% level, which suggests that any observable discrepancies are likely due to random variation rather than inherent differences between the groups.

This non-statistically significant differences indicates that the randomisation was performed correctly and supports the validity of any causal inferences drawn from the analysis. By ensuring that the treatment and control groups are well-matched in terms of key covariates, we can attribute any differences in outcomes to the interventions themselves rather than pre-existing group characteristics. This balance is a critical component in ensuring the reliability of our findings, particularly when examining the long-term effects of the interventions.

taxpayers while encouraging voluntary compliance.

[8]The additional payment data is obtained from the "key performance indicator" evaluation of each tax office, calculated as the increase in payments compared to the previous period. It's important to note that our request for access to payment data from the National Payment Module (MPN) was denied due to strict confidentiality policies.

**Table 1:** Baseline Characteristics - Long Term Effects

| | Control | Deterrence | | Literacy | | Public Goods Provision | |
|---|---|---|---|---|---|---|---|
| | Mean | Mean | $p$-value | Mean | $p$-value | Mean | $p$-value |
| Registration age: | | | | | | | |
| <4 years | 0.204 | 0.207 | 0.775 | 0.209 | 0.647 | 0.206 | 0.852 |
| | (0.403) | (0.405) | | (0.407) | | (0.405) | |
| 4-8 years | 0.219 | 0.225 | 0.604 | 0.217 | 0.806 | 0.221 | 0.879 |
| | (0.414) | (0.418) | | (0.412) | | (0.415) | |
| >8 years | 0.576 | 0.568 | 0.503 | 0.574 | 0.865 | 0.573 | 0.779 |
| | (0.494) | (0.495) | | (0.494) | | (0.495) | |
| Sector: | | | | | | | |
| Service | 0.506 | 0.512 | 0.672 | 0.512 | 0.690 | 0.519 | 0.325 |
| | (0.500) | (0.500) | | (0.500) | | (0.500) | |
| Trading and industry | 0.494 | 0.488 | 0.672 | 0.488 | 0.690 | 0.481 | 0.325 |
| | (0.500) | (0.500) | | (0.500) | | (0.500) | |
| Region: | | | | | | | |
| Java | 0.565 | 0.563 | 0.866 | 0.556 | 0.477 | 0.559 | 0.604 |
| | (0.496) | (0.496) | | (0.497) | | (0.497) | |
| Sumatera | 0.307 | 0.308 | 0.969 | 0.315 | 0.503 | 0.309 | 0.904 |
| | (0.461) | (0.462) | | (0.465) | | (0.462) | |
| Nusa Tenggara | 0.128 | 0.129 | 0.844 | 0.129 | 0.897 | 0.133 | 0.549 |
| | (0.334) | (0.336) | | (0.335) | | (0.339) | |
| Previous Turnover: | | | | | | | |
| <Rp2 Billion | 0.855 | 0.855 | 0.995 | 0.855 | 0.997 | 0.859 | 0.681 |
| | (0.352) | (0.352) | | (0.352) | | (0.348) | |
| Rp2-10 Billion | 0.141 | 0.139 | 0.878 | 0.140 | 0.887 | 0.135 | 0.529 |
| | (0.348) | (0.346) | | (0.346) | | (0.342) | |
| Rp11-20 Billion | 0.004 | 0.005 | 0.457 | 0.005 | 0.702 | 0.005 | 0.483 |
| | (0.063) | (0.073) | | (0.068) | | (0.072) | |
| >Rp20 Billion | 0.000 | 0.000 | 0.997 | 0.001 | 0.319 | 0.001 | 0.325 |
| | (0.018) | (0.018) | | (0.032) | | (0.031) | |
| $N$ | 41,664 | 41,874 | | 41,846 | | 42,560 | |

*Note:* $p$-values refer to the comparison of means between treatment and control groups. The random assignment resulted in the following allocation of taxpayers: (1) Deterrence: 2,991 observations; (2) Literacy: 2,989 observations; (3) Public goods provision: 3,040 observations; and (4) Control: 2,976 observations, $N$ comprises total for fourteen months period.

## 3.3 Experimental Strategy

In general, we compare the responses of small businesses on each treatment group and control group under separate regression models. Each model includes $n_t$ taxpayers

**Table 2:** Baseline Characteristics - Repeated Interventions

| | Control | Deterrence | | Literacy | | Public Goods Provision | |
|---|---|---|---|---|---|---|---|
| | Mean | Mean | $p$-value | Mean | $p$-value | Mean | $p$-value |
| Reg. Age | 13.0 | 12.9 | 0.873 | 12.7 | 0.320 | 12.9 | 0.738 |
| | (8.7) | (8.6) | | (8.5) | | (8.6) | |
| Sector is Service | 0.540 | 0.536 | 0.814 | 0.534 | 0.731 | 0.552 | 0.537 |
| | (0.499) | (0.499) | | (0.499) | | (0.497) | |
| Sector is Non-Service | 0.460 | 0.464 | 0.814 | 0.466 | 0.731 | 0.448 | 0.537 |
| | (0.499) | (0.499) | | (0.499) | | (0.497) | |
| Region is Java | 0.821 | 0.808 | 0.391 | 0.814 | 0.657 | 0.810 | 0.465 |
| | (0.384) | (0.394) | | (0.389) | | (0.392) | |
| Region is Non-Java | 0.179 | 0.192 | 0.391 | 0.186 | 0.657 | 0.190 | 0.465 |
| | (0.384) | (0.394) | | (0.389) | | (0.392) | |
| Turnover (log) | 20.098 | 20.102 | 0.954 | 20.072 | 0.699 | 20.078 | 0.768 |
| | (1.719) | (1.809) | | (1.818) | | (1.814) | |
| Distance (km) | 5.791 | 6.132 | 0.392 | 5.946 | 0.697 | 6.056 | 0.506 |
| | (10.405) | (10.682) | | (10.530) | | (10.599) | |
| BTS density/50km2 | 62.842 | 64.014 | 0.325 | 65.142 | 0.054 | 63.232 | 0.749 |
| | (32.822) | (30.085) | | (29.794) | | (31.601) | |
| $N$ | 1,379 | 1,424 | | 1,379 | | 1,417 | |

*Note:* $p$-values refer to the comparison of means between treatment and control groups. Distance measures the distance from taxpayer's address to the nearest tax office. BTS density refers to the number of Base Transceiver Station from taxpayer's location. The random assignment resulted in the following allocation of taxpayers: (1) Deterrence: 1,424 observations; (2) Literacy: 1,379 observations; (3) Public goods provision: 1,417 observations; and (4) Control: 1,379 observations, $N$ comprises total for one months period for repeated trials. BTS density is slightly imbalance for literacy treatment.

assigned on treatment group $t$, $t \in [1,\ 2,\ 3]$ and $n_0$ taxpayers assigned to the control group. We control for a set of baseline characteristics including region, sector, age, and turnover. We also anticipate the longitudinal analysis of covariance in estimating the treatment effect as suggested by Twisk et al. (2018). The empirical model is formalised as follows:

$$Y_i^t = \beta_0^t + \beta_1^t T_i^t + \beta_2^t X_i^t + \beta_3^t P_i^t + \beta_4^t \mathrm{T}_i^t P_i^t + \varepsilon_i^t \tag{6}$$

where $Y_i^t$ is a given outcome of taxpayer $i$. $T_1^t$ is the treatment indicator for the comparison of treatment group $n_t$ to the control group $n_0$, $P_i^t$ is period, $X_i^t$ is a vector of baseline characteristics, and $\varepsilon_i^t$ is the model error term.

Moreover, this study leverages taxpayers' baseline characteristics to investigate the underlying mechanisms driving the observed outcomes by evaluating the treatment effects across various subgroups that are pertinent to taxpayers' compliance decisions. Specifically, we consider factors such as the cost of resources associated with paying taxes, the intrinsic costs of tax evasion, transaction costs, and elements related to tax morale. By doing so, we are able to conduct a nuanced analysis of the heterogeneity in treatment effects among different taxpayers.

This analysis takes into account key variables such as the taxpayer's distance to the nearest tax office, their level of internet connectivity, their history of penalties, and their filing history. By examining these factors, we aim to identify how the impact of the intervention may vary across different taxpayer profiles, offering insights into the differential effects of compliance strategies. This approach allows for a more comprehensive understanding of how specific characteristics influence the effectiveness of interventions, thereby enabling more targeted and effective policy design.

### 3.3.1 Long-term Effects Analysis

This study aimed to enhance monthly tax payments and improve timely tax filings. We initiated our treatment in December 2021, sending three types of letters and collecting outcome data for fourteen months, covering monthly returns from January 2022 to February 2023. Our intervention comprised three distinct types of letters: (1) a deterrence letter, which emphasized the submission deadline, the imposition of an administrative late-filing penalty of IDR100,000, the possibility of computerised audit utilisation, and the potential for further enforcement measures against non-compliant entities; (2) a literacy letter, characterized by simplified language and a less formal style, with a primary focus on providing clear instructions for tax payment and reporting, accompanied by a QR code facilitating access to official mobile apps and the DGT webpage; and (3) a public goods provision letter, which underscored the taxpayer's contribution to the national budget, specifically targeting COVID-19 expenditures and investments in future generation education. Additionally, informative flyers were included as attachments to the literacy and public goods provision letters.[9]

The field experiment was conducted across thirteen local tax offices, which are managed by three distinct Regional Tax Offices (RTOs). These RTOs include the North Jakarta RTO, which oversees tax administration in the northern part of Jakarta; the South Sumatera and Bangka Belitung RTO, which covers a broad area encompassing

[9]See Appendix A.1, A.2, and A.3 regarding the treatment letter formats.

South Sumatera and the Bangka Belitung Islands; and the Nusa Tenggara RTO, responsible for the tax affairs in the Nusa Tenggara region.[10] Each RTO manages several local tax offices within its jurisdiction, allowing the experiment to address a wide array of geographical and economic contexts.[11] The target population for this study comprised approximately 12,000 small businesses, specifically those with an annual turnover of less than IDR4.8 billion within the period spanning from 2017 to 2019.

The target population of our study encompasses individuals who have been classified as zero-filers or late-filers for a minimum of three months within the preceding fiscal year. To comply with the stipulations outlined in Article 34 of Indonesia's General Provision and Tax Procedure Law, we took meticulous measures to de-identify all data, ensuring the strict preservation of confidentiality and anonymity. Furthermore, we obtained ethical clearance from the Australian National University, and the project was duly registered in the RCT Registry of the American Economic Association.[12]

In order to achieve a robust stratified randomisation, taxpayers exhibiting similar baseline characteristics were organized into distinct strata. Within each stratum, every case was randomly allocated to either one of the treatment groups or the control group. This approach ensured a balanced distribution of taxpayers across groups, accounting for their baseline characteristics. Randomization was performed using a random variable generator in STATA® and a random selection of the seed. For stratified randomisation, the user-written command *randtreat* version 1.4 was utilised.

The random assignment resulted in the following allocation of taxpayers: (1) Treatment Group 1 (deterrence letters): 2,991 observations; (2) Treatment Group 2 (literacy letters): 2,989 observations; (3) Treatment Group 3 (public goods provision letters): 3,040 observations; and (4) Control Group: 2,976 observations. Table 1 depicts the $p$-values comparation between sample means of each treatment group to those of the control group.

In exploring the dynamics induced by the treatment letters, our approach involves a straightforward comparison of the average treatment effects across each period throughout

[10]Due to administrative procedures and other technical constraints, we were unable to send the treatment letters on the same day. However, they were dispatched within a relatively short timeframe. Appendix A.5 provides the details of the dispatch dates for both the initial and repeated trials. Time frame is selected to include the application of previous regulation i.e., GR-46/2013 and to minimise the selection bias due to special arrangement related to COVID-19 in 2020.

[11]In terms of total population, Statistics Indonesia reports that North Jakarta comprises approximately 0.4% of Indonesia's population, South Sumatra accounts for about 3.2%, and Nusa Tenggara represents around 2.0%.

[12]ANU human ethics protocol number 2021/508, titled Small Businesses Tax Compliance Behaviour: Deterrence, Literacy, or Social Norm? Evidence from RCT. AEA registry number AEARCTR-0009460 and AEARCTR-0010721

the observation window. By doing so, we are able to measure the progression and evolution of these effects over time. Our analysis specifically focuses on two key treatment effects: the frequency and amount of monthly tax payments and the timeliness of tax submissions.

We also present a time-to-event analysis framework focused on two tax compliance measures. This analysis offers a more comprehensive understanding of the likelihood that a taxpayer will remain compliant throughout the observation period. First, we utilised payment proxy. In this approach, we adopt the cumulative hazard function using non-parametric estimator from Nelson (1972) and Aalen (1978). To be more precise, we compute distinct Nelson-Aalen cumulative hazard functions separately for monthly tax payments within the treatment and control groups. The cumulative hazard function, denoted as $H(t)$, quantifies the number of "failures" per subject at each observed time point $t$, with "failure" being interpreted as a positive payment. These values are then summed to form $\hat{H}(t)$.

Subsequently, we introduced the filing proxy into our analysis. In this approach, we employ the survival function derived from the Kaplan-Meier model, as outlined in Kaplan's seminal work (Kaplan and Meier, 1958). We consider non timely tax filing, completed after the 15th of the following month, as an "event of failure". Accordingly, in the Kaplan-Meier model, we calculate the probability of each subject's survival beyond a given time point $t$, based on observed instances of "failure".

### 3.3.2 Repeated Interventions

The repeated interventions in this study aims to examine the magnitude of significant changes in the impact of identical trials, taking into consideration the assumption that taxpayers have already established a reference point based on their recollection, past experiences, or expectations from the initial trials, aligning with the reference-dependent behavioural model as suggested by Kőszegi and Rabin (2006).

In doing so, during week 2 of June 2023, we replicated and resent the identical treatment letters that were originally sent in December 2021. However, this time, we limit their distribution to just five local tax offices.[13] The target observation for these interventions consists of approximately 5,599 taxpayers who had previously received the initial trial in December 2021.

We followed an identical procedure for handling and processing the data, ensuring that all data were de-identified in accordance with the stringent requirements for data

[13]This is due to experimental budget constraints and the fact that some tax offices are no longer issuing permission letters related to the initial permission letter number S-2543/RISET/PJ.09/2021, dated October 19, 2021.

confidentiality and anonymity mandated by Indonesia's General Provision and Tax Procedure Law. In addition, the ethics clearance was obtained from the Australian National University and the project was registered in the RCT Registry of the American Economic Association.[14]

We also incorporated the similar randomisation technique as the first trial. For the repeated trial, the random assignment resulted in the following allocation of taxpayers: (1) Treatment Group 1 (deterrence letters): 1,379 observations; (2) Treatment Group 2 (literacy letters): 1,424 observations; (3) Treatment Group 3 (public goods provision letters): 1,379 observations; and (4) Control Group: 1,417 observations. Table 2 presents the $p$-values for comparing the sample means of each treatment group to those of the control group.

# 4 Experimental Result

This section explores the findings resulting from implementing the experimental approach detailed in the preceding section. Initially, we center our attention on the comprehensive average treatment effects over a fourteen-month duration, taking into account various behavioural nudges such as deterrence, literacy, and public goods provision letters. Table 3 and 4 display the estimation of overall treatment effects. This procedure is essential as the identifying stage to see the total treatment effects within a fiscal period.

In order to capture the intricate dynamics of long-term effects, we employ a multivariate analysis of tax payment and reporting, examining the progression and lapse of compliance over time. Figure 3 depict the progression of payments accumulation, whereas Figure 4 displays the progression rate of timely filing over fourteen months period. This finding is essential to view the sustainable treatment effects rather than immediate effects following an intervention.

To provide a more holistic perspective, we also assess the impacts of repeated and identical trials, operating under the assumption that taxpayers have already established a connection with a reference point as a consequence of our previous experimental interventions. Table 5 demonstrates the treatment effect of replicated letters. This procedure is of significance in ascertaining whether the decisions made by taxpayers are not solely determined by their intrinsic preferences but are also influenced by their reference point, encompassing elements related to their memory, past experiences, or anticipations.

[14] AEA registry number AEARCTR-0011485.

**Table 3:** Average Treatment Effects for Long-Term Analysis

| | Treatment 1 VS Control | Treatment 2 VS Control | Treatment 3 VS Control |
|---|---|---|---|
| File Tax Returns | 0.238*** | 0.229*** | 0.084*** |
| | (0.003) | (0.003) | (0.003) |
| File Tax Returns Timely | 0.452*** | 0.120*** | 0.099*** |
| | (0.003) | (0.002) | (0.002) |
| Payment Amount (Rp) | 38,644.17*** | 5,232.13*** | 3,230.94*** |
| | (254.88) | (181.74) | (162.87) |
| Pay Non-zero | 0.395*** | 0.078*** | 0.120*** |
| | (0.003) | (0.003) | (0.003) |
| Use e-Payment | 0.290*** | 0.266*** | 0.106*** |
| | (0.003) | (0.003) | (0.003) |
| Full sample $N$ | 167,944 | 167,944 | 167,944 |
| Non-zero filers (Rp) | 20,067.45*** | -1,115.35*** | -14,799.78*** |
| | (226.65) | (218.62) | (198.76) |
| Non-zero filers (log) | 0.224*** | -0.055*** | -0.278*** |
| | (0.005) | (0.007) | (0.006) |
| Non-zero filers $N$ | 52,618 | 52,618 | 52,618 |

*Note:* Treatment 1: Deterrence Letters, Treatment 2: Literacy Letters, Treatment 3: Public Goods Provision Letters. Robust standard errors in parentheses.
* $p$ <0.1, ** $p$ <0.05, *** $p$ <0.01.

## 4.1 Average Treatment Effect

The treatment effects accumulation over 14-month period on the outcome variable indicating tax filing decision, payment amount, filing improvement, filing timely, utilisation of electronic payment, and whether taxpayers contacted tax office prior to the treatment are presented in Table 3.

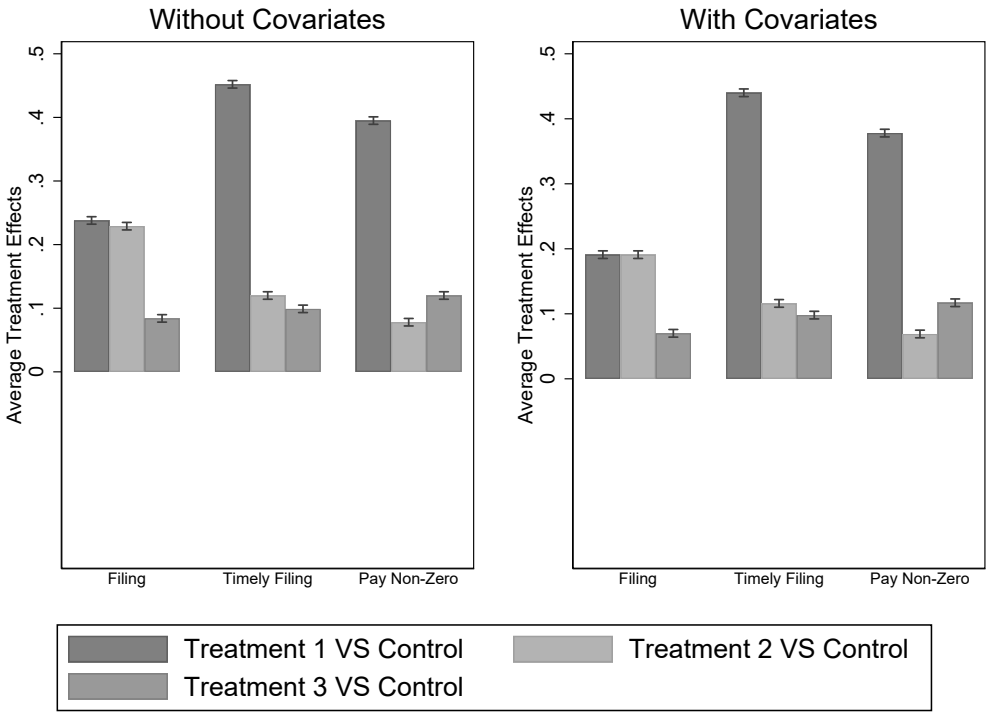
Regarding tax payments, our findings show that within the 14 months following the dispatch of the letters, all groups displayed a substantial increase in payments compared to a control group that did not receive treatment letters. As illustrated in Table 3, the deterrence letter had the most significant impact, leading to a roughly 22-percentage point increase in the likelihood of payment. In this context, the deterrence letter resulted in an additional tax revenue of IDR 38,644 (approximately 2.43 USD) per taxpayer on average, potentially signifying a substantial boost in tax revenue, given the scale of the small

**Table 4:** Average Treatment Effects for Long-Term Analysis (with Covariates)

| | Treatment 1 VS Control | Treatment 2 VS Control | Treatment 3 VS Control |
|---|---|---|---|
| File Tax Returns | 0.191*** | 0.191*** | 0.070*** |
| | (0.003) | (0.003) | (0.002) |
| File Tax Returns Timely | 0.440*** | 0.116*** | 0.098*** |
| | (0.003) | (0.002) | (0.002) |
| Payment Amount (Rp) | 37,120.03*** | 4,689.06*** | 3,107.29*** |
| | (252.02) | (178.07) | (154.99) |
| Pay Non-zero | 0.378*** | 0.069*** | 0.117*** |
| | (0.003) | (0.003) | (0.003) |
| Use e-Payment | 0.240*** | 0.225*** | 0.090*** |
| | (0.003) | (0.003) | (0.002) |
| Full sample $N$ | 167,944 | 167,944 | 167,944 |
| Non-zero filers (Rp) | 19,928.09*** | -571.04** | -14,413.55*** |
| | (262.36) | (247.09) | (210.43) |
| Non-zero filers (log) | 0.228*** | -0.034*** | -0.261*** |
| | (0.006) | (0.008) | (0.006) |
| Non-zero filers $N$ | 52,618 | 52,618 | 52,618 |

*Note:* Treatment 1: Deterrence Letters, Treatment 2: Literacy Letters, Treatment 3: Public Goods Provision Letters. The set of covariates includes registration age, turnover, tax return type, distance to tax office, and internet density.
Robust standard errors in parentheses.
* $p$ <0.1, ** $p$ <0.05, *** $p$ <0.01.

**Figure 1:** Average Treatment Effects




*Note:* Treatment 1: Deterrence Letters, Treatment 2: Literacy Letters, Treatment 3: Public Goods Provision Letters. Each bar represents an estimated treatment effect, with 95% confidence intervals shown as lines around the point estimates. The set of covariates includes year since registration, turnover, tax return type, distance to tax office, and internet density.

business sector in Indonesia. Meanwhile, the literacy letter and public goods provision letter contributed approximately IDR 5,232 (about 0.34 USD) and IDR 3,230 (around 0.2 USD) respectively.

With respect to tax filing, all treatment letters clearly demonstrated a significant effect on increasing filing rates and concurrently enhancing the timeliness of submissions. Both deterrence letters and literacy letters yield nearly identical substantial effects in encouraging taxpayers to submit their tax returns. Nonetheless, in the case of timely tax filing, in stark contrast to the realm of tax payments, it is the literacy letter that now exerts the largest effect, rather than the deterrence letter. These results are consistent with the economic deterrence model streamlined by Allingham and Sandmo (1972) and Yitzhaki (1974a), and to some extent, in line with Torgler and Schaltegger (2006) and Alasfour et al. (2016), which argued that both formal and informal education significantly impacted the tax compliant behaviour particulary in terms of filing rather than payment.

## 4.2 Dynamics

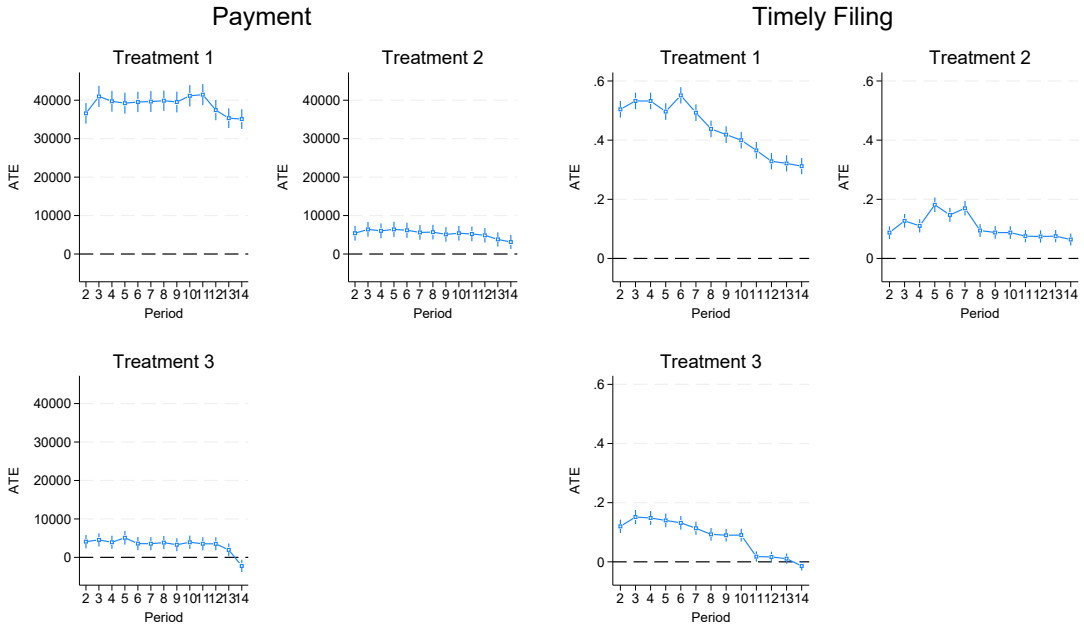
Given that small businesses must meet their tax payment and filing obligations monthly, based on their income for each period, we are interested in examining how the treatment effects evolve each month throughout the observation period. This interest is further driven by the dynamic nature of taxpayers' decision-making processes, particularly in relation to risk and psychological factors. For example, lab experiments by Mittone et al. (2017) highlight how psychological deterrence is influenced by moral constraints and prior experiences. Additionally, Advani et al. (2023) and Kotsogiannis et al. (2024) demonstrates how audits have a dynamic impact on reported income, with variations depending on the income source and audit scope. Furthermore, DeBacker et al. (2018) and Christiansen (2024) examine the long-term effects of audits on under-reporting and filing.

Our approach allows us to capture any trends or changes in the treatment effects as they unfold over time, providing a more nuanced understanding of the intervention's impact on tax compliance behavior. Figure 2 presents the average treatment effect estimates for each group over the entire observation period. This includes treatment group 1, which received the deterrence letter; treatment group 2, which received the literacy letter; and treatment group 3, which received the public goods provision letter. The control group, which did not receive any intervention, is also included for comparison.

The figure shows that focusing on payment and timely filing outcomes, deterrence letter initially has the strongest impact on both payment and timely filing, but its effectiveness declines over time. The literacy letter shows a moderate and gradually decreasing effect, while the public goods provision letter has the least impact, with its effects nearly fading to zero by the end of the observation period. Overall, all treatments lose the treatment effect magnitude over time, with treatment 1 having the most substantial but diminishing impact, especially on tax filing.

The negative effect indicated by the average treatment effect values dropping below zero suggests that the public goods provision letter might have led to worse outcomes in the long-run compared to the control group. Several potential reasons could explain this negative impact. One possibility is that the letter inadvertently triggered negative psychological reactions, such as resentment or skepticism (Brehm and Brehm, 2013). Taxpayers might have found the message irrelevant, unconvincing, or manipulative, leading to reduced compliance in both payment and timely filing. Additionally, if taxpayers perceived the public goods message as unrelated to their tax obligations or of no direct benefit to them, their motivation to comply might have decreased, diminishing the intended positive effect.

**Figure 2:** Average Treatment Effects - All Observation Periods



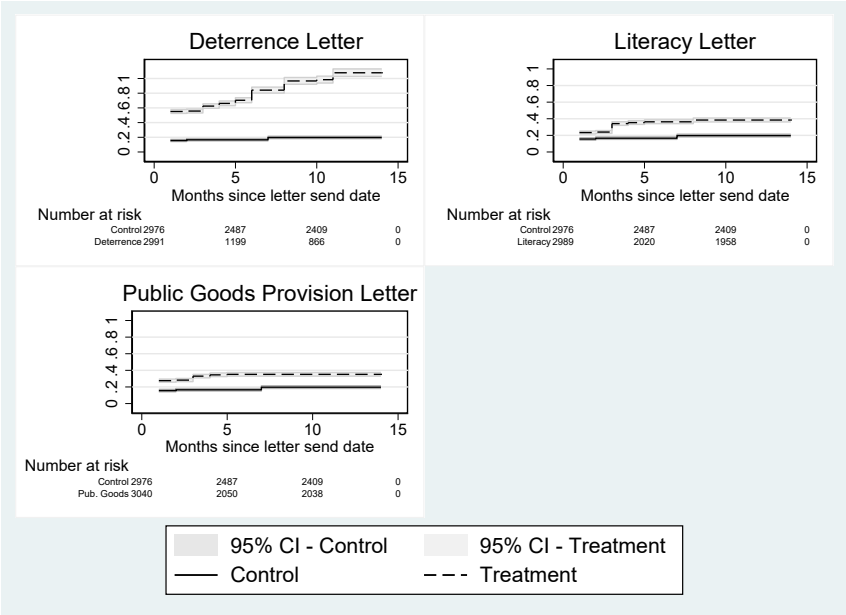

*Note:* Treatment 1: Deterrence Letters, Treatment 2: Literacy Letters, Treatment 3: Public Goods Provision Letters. Each point represents an estimated treatment effect, with 99% confidence intervals shown as lines around the point estimates.

Another potential explanation is the *crowding out* effect, where the external appeal to support public goods undermined the intrinsic motivation to comply with tax obligations as highlighted by Frey and Jegen (2001). Rather than reinforcing compliance, the message could have been perceived as unnecessary or coercive, reducing taxpayers' willingness to meet their obligations. Misinterpretation of the letter's content might also have played a role, causing taxpayers to believe their contribution was less critical or that there was no immediate need to comply (Alm et al., 1992).

## 4.3 Time-to-Event Analysis

Employing time-to-event analysis, we can illustrate the progression of treatment effects on tax payment and tax filing, quantified in terms of months subsequent to the initiation of treatment. Figure 3 illustrates the Nelson-Aalen cumulative hazard estimates for treatment group 1 – the deterrence letter, treatment group 2 – the literacy letter, treatment group 3 – the public goods provision letter and the control group. The payment rate, defined as the proportion of partial or full payments within each group.

**Figure 3:** Nelson– Aalen Cumulative Hazard Estimates (Payment)

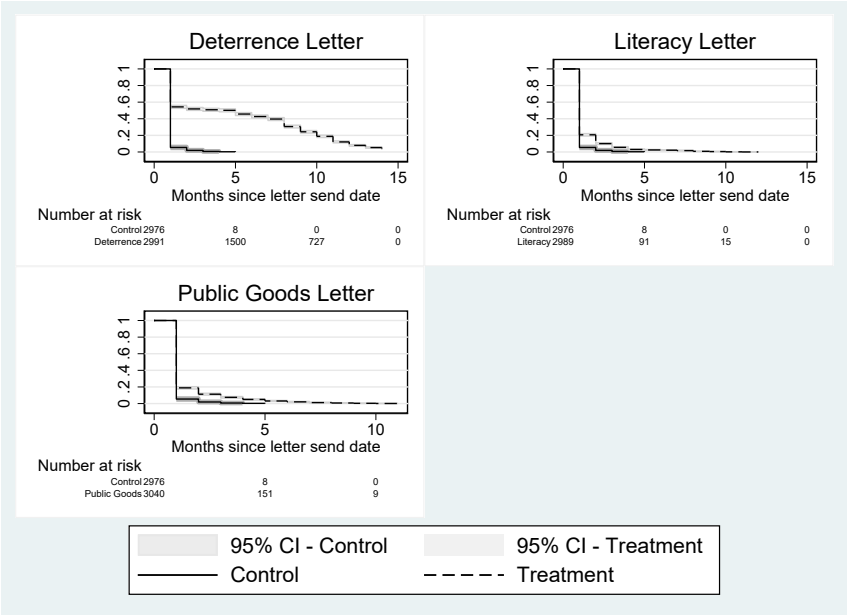



*Note:* Each line represents a rate of respective payment amount of all taxpayers. The progression rate shows that compared to control group, the deterrence letter led to a particularly strong increase in payment rates during 14 months observation period. The corresponding increase in payment rates resulting from the public good provision letter is weaker.

The cumulative hazard estimates reveal a notable disparity in payment rates accumulation approximately after six months of the mailing of the deterrence letter, showing that payment rate of treatment group 1 is significantly higher than those in the control group. This observation suggests that the deterrence letter yielded a substantial and sustained upsurge in payment rates from the time of mailing, and significantly increase after six months, until the conclusion of the trial period, spanning approximately fourteen months. In contrast, the control group's payment rate throughout the trial period remains below 20 percent and remain unchanged until the fourteen months after the observation period.

The payment rates of treatment group 2 , the literacy letter exhibit a significant increase over the control group's rates, manifesting around four months after the letter's dispatch, and relatively constant until the end of the observation period. Although the literacy letter elicits a noteworthy increase in payment rates in comparison to the control group, this effect is somewhat substantially lower than the impact of the deterrence letter. The discrepancy in payments between treatment group 2 and the control group amounts to approximately 20 percentage points within the fourteen months following the letter's send date.

**Figure 4:** Kaplan-Meier Survival Estimates (Timely Filing)




*Note:* Each line represents a rate of respective timely tax filing of all taxpayers . The failure event criteria is the dummy whether monthly tax return submission is not timely i.e., after the 15th of the following month. The progression rate analysis reveals that, in comparison to the control group, the deterrence letter engendered a more enduring impact on the propensity for timely filing over the course of a 14-month observational period. We also managed to perform the test equality of survivor functions using log-rank test according to Mantel and Haenszel (1959) and Savage (1956) and we obtained $p$-value $<0.01$.

Whereas, the payment rates in treatment group 3, which received the public goods provision letter, exhibit a pronounced increase relative to those of the control group, becoming evident approximately four months after the mailing of the letter and maintaining a relatively consistent pattern throughout the duration of the observation period. While the public goods provision letter prompts a substantial enhancement in payment rates when compared to the control group, it is notably less impactful than the deterrence letter, and its effect is relatively similar to that observed in the literacy letter group. The differential in payment rates between treatment group 3 and the control group accumulates to approximately 20 percentage points within the fourteen months subsequent to the letter's dispatch.

It is noteworthy that the Nelson-Aalen cumulative hazard estimation profiles for the literacy and public goods provision letters exhibited minimal fluctuation after approximately four months, indicative of the fact that the majority of payments engendered by these letters transpired within a relatively concise timeframe following their dispatch. In

contrast, the profile for the deterrence letter displayed a more dynamic trajectory during the fourth to tenth month after the mailing of the letters, signifying that the vigorous and sustained effects on payments triggered by this letter persisted for nearly an entire fiscal year subsequent to its dispatch.

With respect to the dynamic treatment effects on tax filing, Figure 4 displays the Kaplan-Meier survival estimates for timely tax filing during 14-months observation period across four groups: treatment group 1 (represented by the dashed line, receiving deterrence letters), treatment group 2 (receiving literacy letters), treatment group 3 (receiving public goods provision letters), and the control group (represented by the solid line). In this diagrammatic representation, the failure event criteria is the non-timely filing (i.e., after the 15th of the following month). Overall, the survival estimates show a significant discrepancy in terms of tax filing starting from the first month after sending treatment letters. Specifically, the timely filing rate in the deterrence letter group is considerably higher compared to the control group, the literacy letter group, and the public goods provision letter group.

Treatment group 1 (deterrence letter) exemplifies more sustainable effects with regards to the timely tax filing. This effect experiences a pronounced reduction of nearly 50 percent subsequent to the tenth month, resulting in the preservation of approximately 21 percent of the total number of entities after the end of the observation period. In contrast, treatment group 2 (literacy letters) and treatment group 3 (public goods provision letter) manifest nearly identical progression patterns. It becomes evident that the effect remains sustained only for five months since sending the treatment letters, gradually disappear thereafter until the end of the observation period.

## 4.4 Repeated Interventions Effect

The treatment effects for the replicated trials, as evidenced by the linear regression analysis of tax payment amounts and the punctuality of tax filings, are delineated in Table 5. Our findings reveal that one month following the dispatch of the replicated letters (approximately 18 months subsequent to the completion of the initial trials), only the deterrence letter has demonstrated a statistically significant impact on tax payments in comparison to the control group. Intriguingly, all treatment interventions remain effective in eliciting statistically significant effects on the timely submission of tax filings.

Table 5 illustrates that the deterrence letter continues to exert a substantial influence, increasing the likelihood of payment by approximately 24 percentage points among non-zero filers and enhancing the likelihood of on-time filing by approximately 6 percentage

**Table 5:** Average Treatment Effects for Repeated Trials

| | Treatment 1 VS Control | Treatment 2 VS Control | Treatment 3 VS Control |
|---|---|---|---|
| File Tax Returns | 0.217*** | -0.017 | 0.058*** |
| | (0.018) | (0.016) | (0.017) |
| File Tax Returns Timely | 0.348*** | 0.089*** | 0.102*** |
| | (0.014) | (0.011) | (0.012) |
| Payment (Rp) | 24,607.77*** | 925.67 | 1,732.84 |
| | (1,489.72) | (1,177.36) | (1,193.03) |
| File Non-Zero Returns Timely | 0.173*** | -0.038** | -0.010 |
| | (0.018) | (0.016) | (0.016) |
| Use e-Payment | -0.004 | 0.051*** | -0.015 |
| | (0.017) | (0.016) | (0.017) |
| Full sample $N$ | 5,599 | 5,599 | 5,599 |
| Non-zero filers (Rp) | 23,921.50*** | 1,432.07 | 1,319.91 |
| | (1,072.38) | (944.09) | (1,108.05) |
| Non-zero filers (log) | 0.266*** | -0.001 | -0.015 |
| | (0.019) | (0.017) | (0.017) |
| Non-zero filers $N$ | 2,107 | 2,107 | 2,107 |

*Note:* The trials is conducted in June 2023, we limit the distribution to approximately 5,599 taxpayers who had previously received the identical letters in December 2021. Treatment 1: Deterrence Letters, Treatment 2: Literacy Letters, Treatment 3: Public Goods Provision Letters. Robust standard errors in parentheses.
* $p$ <0.1, ** $p$ <0.05, *** $p$ <0.01.

**Table 6:** Average Treatment Effects for Repeated Trials (with Covariates)

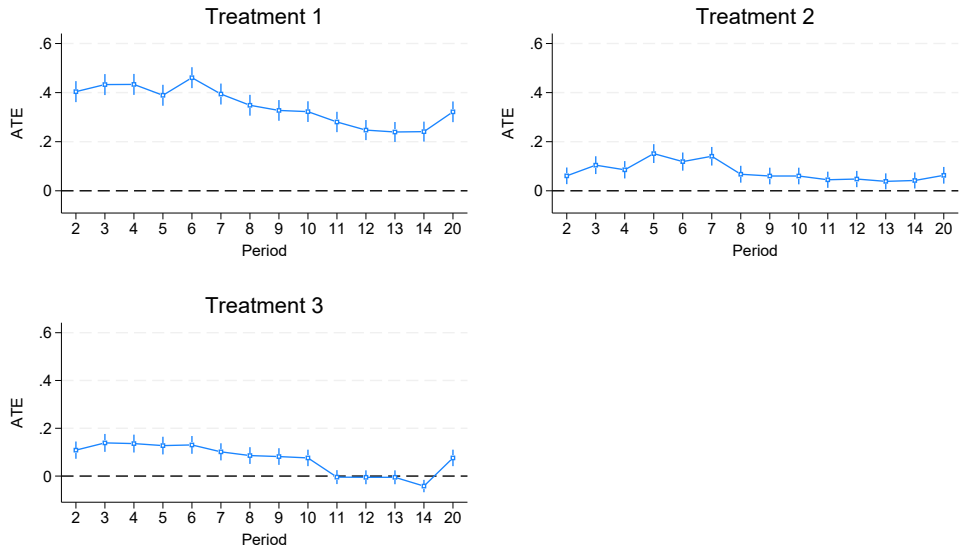
| | Treatment 1 VS Control | Treatment 2 VS Control | Treatment 3 VS Control |
|---|---|---|---|
| File Tax Returns | 0.219*** | -0.017 | 0.060*** |
| | (0.017) | (0.016) | (0.016) |
| File Tax Returns Timely | 0.347*** | 0.089*** | 0.103*** |
| | (0.014) | (0.011) | (0.011) |
| Payment (Rp) | 24,614.38*** | 976.85 | 1,872.45 |
| | (1,450.72) | (1,156.70) | (1,155.88) |
| File Non-Zero Returns Timely | 0.176*** | -0.038** | -0.008 |
| | (0.017) | (0.016) | (0.016) |
| Use e-Payment | -0.006 | 0.043*** | -0.012 |
| | (0.013) | (0.012) | (0.012) |
| Full sample $N$ | 5,599 | 5,599 | 5,599 |
| Non-zero filers (Rp) | 24,296.17*** | 1,729.80 | 1,475.06 |
| | (1,107.81) | (967.46) | (1,125.98) |
| Non-zero filers (log) | 0.279*** | 0.006 | -0.013 |
| | (0.020) | (0.018) | (0.018) |
| Non-zero filers $N$ | 2,107 | 2,107 | 2,107 |

*Note:* The trials is conducted in June 2023, we limit the distribution to approximately 5,599 taxpayers who had previously received the identical letters in December 2021.Treatment 1: Deterrence Letters, Treatment 2: Literacy Letters, Treatment 3: Public Goods Provision Letters. Robust standard errors in parentheses. Set of covariates are island, sector, registration age, turnover, and distance.
* $p$ <0.1, ** $p$ <0.05, *** $p$ <0.01.

points. Consequently, the recurrent deterrence letter still yields an additional tax revenue of IDR 24,607 (equivalent to 1.60 USD) per taxpayer on a monthly basis, on average.

In contrast to the initial trials conducted in December 2021, we acknowledge a notable decrease in the average treatment effect concerning tax payments. More specifically, the effects of literacy letters and public good provision letters in the subsequent trials appear to be statistically insignificant in improving tax payments. This trend is also reflected in our time-to-effect analysis of the impact of literacy and public goods provision letters, which demonstrates a more dynamic profile and less sustained impact on tax payments. This signifies a potentially noteworthy upswing in tax revenue only for the replicated deterrence letter, especially when considering the scale of the small business sector in Indonesia. Concerning tax filing, it is noteworthy that the literacy letters continue to yield the most pronounced effects on the likelihood of timely tax filing.

**Figure 5:** Average Treatment Effects on Timely Filing (Restricted Sample)




*Note:* Treatment 1: Deterrence Letters, Treatment 2: Literacy Letters, Treatment 3: Public Goods Provision Letters. Each point represents an estimated treatment effect, with 99% confidence intervals shown as lines around the point estimates. We limit the distribution to approximately 5,599 taxpayers who had previously received the identical letters in December 2021.

As previously noted, our study aims to investigate the pattern of repeated treatment effects as the initial treatment effect progresses over the observation period. However, the sample size for the repeated trials is smaller than that of the initial trial, primarily due to attrition and constraints related to permissions. To accurately capture the dynamics of

these effects and ensure consistency in our analysis, we have decided to restrict the overall sample into 5,599 observation to align with the randomisation assignment of the repeated trials. This approach allows us to maintain the integrity of the comparison and better understand how the treatment effects evolve over time. Since taxpayers may spread their payments across several months, making it unclear if payments correspond to specific months, we have chosen to focus on the dynamics of timely filing rather than payment amounts.

Figure 5 illustrates the long-term average effect before and after the repeated trial in period 20. The results indicate that the impact of the treatments may change with repetition. The deterrence letter shows a modest recovery in the average treatment effect, increasing by about 5-10% after earlier declines. In contrast, the literacy letter experiences a slight increase of roughly 3-5%, suggesting that repetition may modestly enhance its effectiveness. However, the public goods provision letter remains largely ineffective, with only a minor increase of about 2-3% in treatment effect, even with repetition. These findings imply that while some treatments, particularly those affecting payment behavior, might benefit from repeated application to improve their effectiveness, others show minimal responsiveness to repeated trials. This underscores the potential value of repeated interventions in achieving sustained or enhanced outcomes.

## 4.5 Effects Heterogeneity

### 4.5.1 SMEs Eligibility

It is important to highlight that under the tax simplification programs outlined in Government Regulations GR-46/2013 and GR-23/2018, only taxpayers with an annual turnover of less than IDR 4.8 billion are eligible for tax simplification and reduced rates. However, as noted by Satyadini and Rosid (2024), while this tax simplification program has successfully improved taxpayer filing compliance, it has not been as effective in increasing actual tax payments.[15]

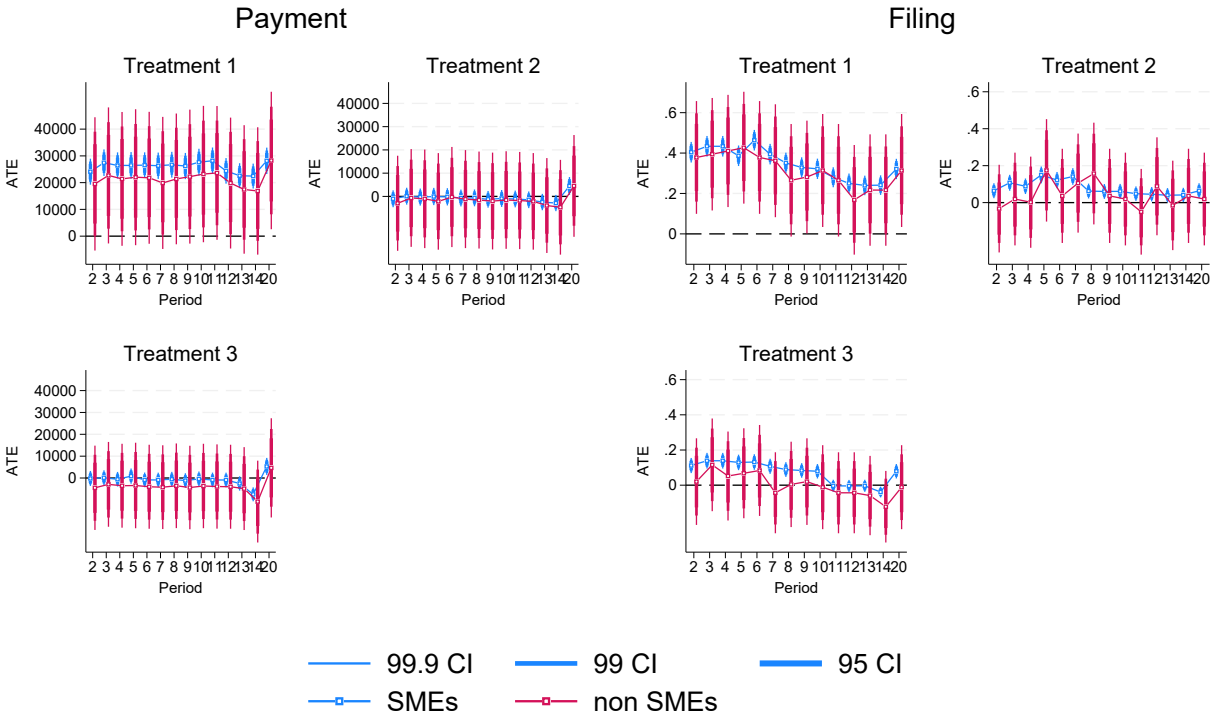
In our dataset, although we initially categorised businesses as SMEs based on their status in the baseline year (2020), we observed that some taxpayers have since grown their turnover beyond IDR 4.8 billion.[16] These taxpayers are still included in our study,

[15]GR23/2018 provides a 0.5% income tax on gross turnover for small businesses in Indonesia with annual revenue up to IDR 4.8 billion, replacing GR46/2013, which imposed a 1% tax. Both regulations aim to simplify tax compliance and support small enterprises.

[16]It is possible that the annual turnover reported by taxpayers may not reflect their actual turnover and is not perfectly aligned with their tax liability. This is due to the nature of the self-assessment system and the "bunching" phenomenon, when taxpayers intentionally setup their turnover slightly below the

whether in the control or treatment groups. Given this scenario, it is particularly interesting to analyze the treatment effects for those who remain eligible for the tax simplification program versus those who have outgrown this eligibility. This comparison will provide insights into how the tax policy's effectiveness varies across different business sizes and growth trajectories, particulary for taxpayers that should not be relevant with our treatment letters.

**Figure 6:** Treatment effects by eligibility SMEs




*Note:* The subset of SMEs in this study comprises those with an annual turnover of less than Rp4.8 billion.These businesses fall inside the eligibility criteria for tax simplification features as defined by GR46 and GR23.

We divided our sample based on whether the taxpayers' annual turnover is below IDR 4.8 billion or exceeds this threshold. Figure 6 presents the estimated average treatment effects for those eligible for the SMEs simplification program and those who are not. The findings indicate that all three treatments generally have a positive impact on tax filing for SMEs, particularly in the earlier periods (2-8), where the treatment effect remains consistently above zero. Among the treatments, treatment 1 demonstrates the strongest positive effect on timely filing and payment, especially for SMEs. In contrast, non-SMEs

threshold, as demonstrated in Appendix A.14 and A.15.

exhibit more variability and less consistent responses, with treatment effects frequently closer to zero, especially in the later periods.

The results also show a declining trend in the treatment effect over time across all treatments. Initially, there is a positive effect on filing compliance for SMEs, but this diminishes in later periods. Non-SMEs exhibit a similar pattern, although their treatment effects tend to be closer to zero, indicating a smaller impact. The confidence intervals around the estimates show more uncertainty for non-SMEs, suggesting greater variability in their responses. Overall, the treatments appear to have a more significant and sustained impact on tax payments than on filing compliance, with SMEs responding more than non-SMEs.

In repeated trial period 20, the treatment effects on tax payment remain positive for SMEs, particularly under treatment 1, where the treatment effect is highest, indicating a stronger impact of the treatment on increasing payments for eligible SMEs. In contrast, non-SMEs show more volatility, with treatment effects fluctuating and generally closer to zero, especially in the later periods, including period 20. For tax filing, the treatment effects for both SMEs and non-SMEs decline over time, with the treatment effect nearing zero by period 20, suggesting that the initial improvements in filing compliance are not sustained.

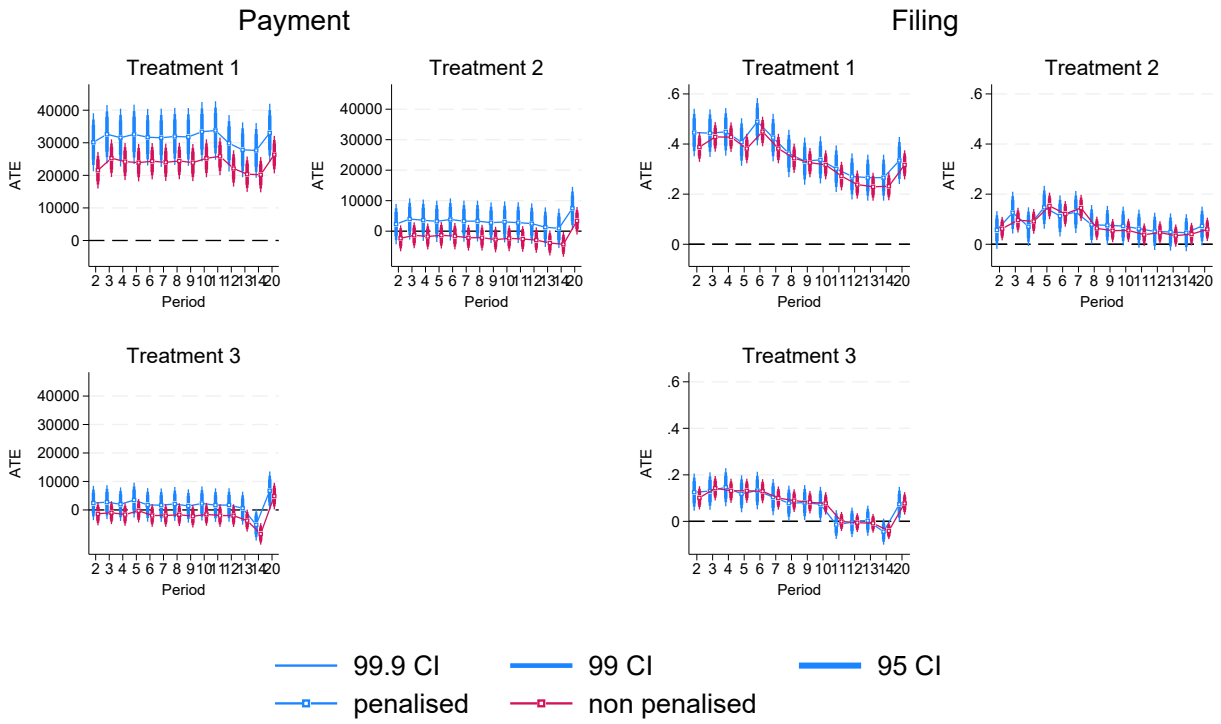
### 4.5.2 Reference Point

As previously discussed, this study's repeated interventions aim to assess the magnitude of significant changes in the impact of identical trials, considering that taxpayers may have established a reference point based on their recollection, past experiences, or expectations from initial trials. This approach aligns with the reference-dependent behavioral model proposed by Kőszegi and Rabin (2006). In our case, we are interested in examining how the experience of being penalised serves as a reference point for taxpayers.

To explore this, we divided the sample into two groups: those who had been penalised before and those who had not. We assumed that taxpayers with prior penalties would be more cognizant of the additional costs associated with penalties and would likely adopt strategies to avoid future penalties.

Overall, the results show that, in terms of payment, previously penalised taxpayers respond more compared to those who have never been penalised. More specifically, examining the payment outcome, treatment 1 and treatment 2 exhibit similar patterns, where the penalised group shows higher values compared to the non-penalised group. For instance, in treatment 2, the penalised group's payment values are approximately 20-25%

**Figure 7:** Treatment effects by penalties




*Note:*The subset of taxpayers in this study includes those who have either been penalised before or have never been penalised. Penalties refer to administrative sanctions only – not related to the result of audit.

higher than the non-penalised group across the observed periods. The effect is even more pronounced in treatment 1, with the penalised group demonstrating payment values that are 30-35% greater than the non-penalised taxpayers.

In terms of filing, all three treatments display a consistent pattern, whereby the penalised group consistently exhibits slightly higher filing values compared to the non-penalised group. The magnitude of this effect varies across the treatments, with treatment 1 and treatment 2 showcasing a more pronounced difference, with the penalised group's filing values being approximately 5-10% higher than the non-penalised group. This indicates that the penalty exerts a positive and less considerable influence on the filing, with a stronger impact observed in treatment 1 and treatment 2.

### 4.5.3 Other characteristics

We conduct several heterogeneity analysis to provide a fuller picture of taxpayers' compliant behaviour with respect to the taxpayers' baseline characteristics and non-economic

rationality proxies. Initially, we employ Kaplan-Meier survival analysis to examine timely filing across different baseline characteristics, including business size, sector, location, and registration age. In this analysis, the failure event is represented by a dummy variable indicating whether the monthly tax return submission is timely or not. Appendix A.16, A.17, A.18, A.19, A.20, and A.21, illustrate the progression rate of timely tax returns for deterrence letters, literacy letters, public goods provision letters, and the control group over a 14-month observation period, categorised by business size, sector, location, and registration age, respectively.

In general, we observe a predominantly identical survival pattern of subsample compared to all samples. Notably, the survival estimates reveal a significant disparity between deterrence letters and other treatments, as well as control groups, concerning the timely filing rate. However, concerning business size, we identify a relatively more sustained likelihood of timely filing among taxpayers with larger businesses. Regarding the business sector, the endurance of timely filing is notable in service-related sectors. Geographically, taxpayers located in Sumatera exhibit relatively greater persistence in maintaining timely filing. Additionally, taxpayers with a registration age exceeding eight years also demonstrate heightened endurance in terms of timely filing.

## 4.6 Robustness Checks

We undertake several robustness checks to confirm the validity of our findings. In the randomization process, we assume that the covariates are evenly distributed between the treatment and control groups. To confirm this assumption, we conduct a simple covariate balance test, the results of which are presented in Table 1 and 2, showing the baseline characteristics of control and treatment groups. This procedure aims to ensure that the individuals in the treatment and control groups have similar attributes. The $p$-values in Table 1 indicate that our sample is indeed well-balanced in terms of baseline characteristics such as age, sector, location, and size. This table provides evidence that the covariates are effectively balanced between the groups, demonstrating the similarity of taxpayers in the treatment and control groups. Our results suggest that no covariates have significantly influenced the estimation of the treatment effect.

As a part of our analysis, we use the variables presented in Table 1 as control variables to account for any potential discrepancies in baseline characteristics. Table 1 and Table 3 present treatment effect estimates under both unconditional and conditional – controlling for baseline characteristics. Our analysis indicates that incorporating baseline characteristics into the model has negligible changes on the estimated treatment effects,

thereby affirming the robustness of our results.

Furthermore, we follow a similar procedure to ensure the reliability of our results under the repeated trials. Table 2 presents the baseline characteristics of the control and treatment groups in these repeated trials. The $p$-values in Table 2 are generally well-balanced, with the exception of the BTS density variable.[17] This imbalance occurs because the repeated interventions reduced the sample size. However, since this variable is not used to estimate the treatment effects, it does not significantly influence the overall results. Additionally, Table 5 and Table 6 reveal that both the conditional and unconditional models yield nearly identical treatment effects.

To examine the factors influencing treatment effects in relation to tax payments, we divided the taxpayer sample based on their payment behavior–specifically, whether they make any payment and report zero or not.[18] Approximately 31 percent of all taxpayers in our sample fall into the category of non-zero filers. Our analysis reveals that the treatment effects of the deterrence letter are statistically significant for non-zero filers but slightly lower than the effects observed in the overall sample. In contrast, both the literacy and public goods provision letters yield negative treatment effects. This suggests, tentatively, that the deterrence letters may have a relatively high impact on the intensive margin for non-zero filers.

# 5 Mechanisms

We are interested in examining the mechanisms contributing to these outcomes by measuring the treatment effects on various subgroups relevant to taxpayers' compliance decisions. These include the cost of resources from paying taxes $T(y')$, the intrinsic cost of evasion and enforcement $N(y')$ , the transaction cost $C(y')$, and tax morale element $M(y')$ according to Equation 1 and 2 .

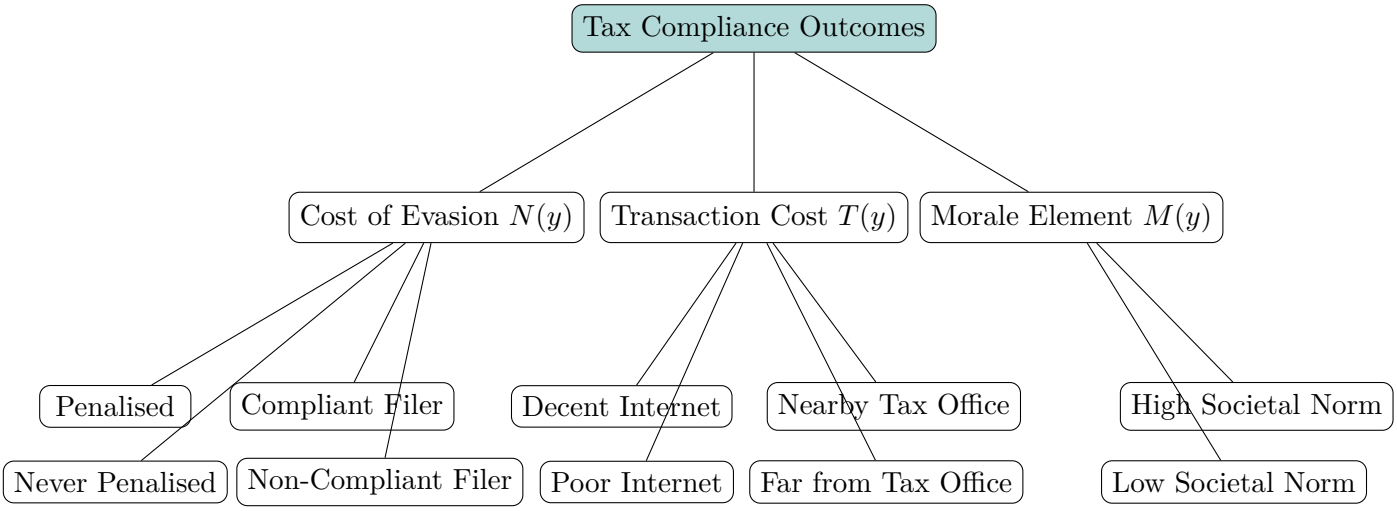
We investigate whether compliance costs influence taxpayers' decisions to pay and report taxes and whether the intervention interacts with the cost of resources from paying taxes, cost of enforcement and tax morale element during our trial period as drawn in Figure 8. In Figure 9 we show the diagrammatic representation of average treatment

[17]BTS density refers to the concentration of Base Transceiver Stations within a 50km radius of the taxpayer's location. This measurement reflects the availability and accessibility of mobile network infrastructure, which can impact communication services in the area.

[18]Zero-filers reported no taxable income, which could suggest either underreporting, tax evasion, or legitimate financial hardship. Studying this group can provide valuable insights into compliance issues and highlight areas where targeted enforcement or support may be necessary.

effects by accessibility (i.e., internet connection and distance to tax office), compliance history, and societal compliance prevalence.

**Figure 8:** Assessment of Treatment Effects




## 5.1 Accessibility and connectivity

We partition the sample into two subcategories based on taxpayers' proximity to the tax office, determined by the distance from the taxpayer's address to the nearest tax office using longitude-latitude coordinates. The "near" group comprises observations within a 5km radius, while the "far" group includes those beyond this distance, with the 5km threshold deemed as an acceptable urban design transporting distance.

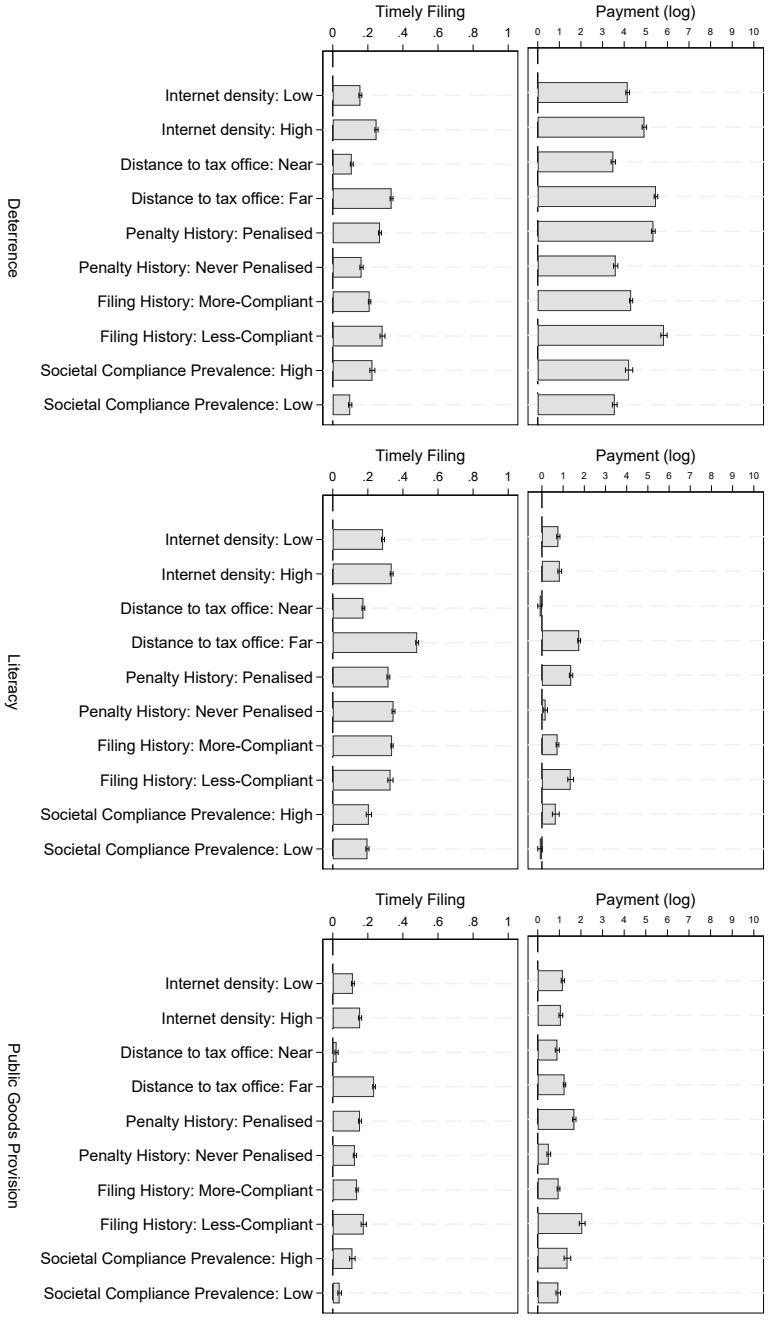
In terms of internet connectivity, the study further categorises the sample into two subgroups based on accessibility. This is measured by a weighted average density of Base Transceiver Stations per 50km square, a fixed broadband internet index across subdistricts, and an internet bandwidth index across areas relative to the taxpayer's location, each scored on a scale from 0 to 10.[19] The "low" group encompasses observations with scores below 5, while the "high" group includes those with scores equal to or exceeding 5. [20]

We find suggestive evidence that taxpayers who received a deterrence letter, along

[19]The number of BTS includes seven major cellular operators, covering both GSM and CDMA technologies: Telkomsel, Indosat, XL Axiata, Hutchison Tri, Telkom, Smartfren, and Esia. Data is obtained from the Ministry of Communication and Information Technology and Indonesia Statistics.

[20]A higher BTS density indicates better mobile network coverage and potentially better communication services in that area, which could influence taxpayer behavior and accessibility.

**Figure 9:** Partial Assessment of Treatment Effects on Relevant Subgroups




*Note:* Each bar represents the average treatment effects for respective outcome variables with 95% confidence intervals shown as lines around the point estimates.

with those who have better access to resources, improved connectivity, and are part of communities with a higher prevalence of societal compliance, are more likely to report and pay higher taxes. This indicates that the deterrence letter may be more effective in environments where compliance is already socially reinforced and supported by infrastructure.

## 5.2 Compliance history

We further categorise taxpayers into four distinct subgroups to account for variations in past compliance behaviours, particularly focusing on their penalty history and total monthly filing frequency within a fiscal year. The "penalised" subgroup consists of taxpayers who have previously been subjected to administrative penalties due to non-compliance, while the "never-penalised" subgroup includes those who have consistently avoided such penalties. To capture filing behaviour, we distinguish between the "more-compliant" and "less-compliant" groups. The "more-compliant" group comprises taxpayers who have submitted six or more monthly tax returns within a fiscal year, while the "less-compliant" group includes those who filed fewer than six returns.

Interestingly, in line with our earlier findings, the results demonstrate that taxpayers with a history of non-compliance—specifically those who have previously been penalised for failing to meet their tax obligations and have exhibited patterns of underreporting or evasion—are more likely to respond positively to deterrence letters. This suggests that even taxpayers who have shown resistance to compliance in the past can be influenced by targeted enforcement actions. The effectiveness of these deterrence letters highlights their potential to improve tax compliance across a wide range of taxpayer profiles, including those who have historically been less compliant. This reinforces the value of enforcement measures in encouraging adherence to tax regulations, particularly among individuals with a track record of non-compliance.

## 5.3 Societal compliance prevalence

The analysis of treatment effects based on societal compliance prevalence, or social norms, is essential to support our proposition that taxpayers incur a moral cost with tax evasion. To investigate this, we grouped taxpayers living within 1,5 kilometers of a tax office into two clusters based on their penalty records, encompassing 6,719 taxpayers across 13 tax offices. Our assumption is that if more taxpayers within a community engage in tax evasion, the social norm (tax morale) among their peers will be negatively

impacted, leading to diminished responses to the treatment.

Our findings indicate that taxpayers residing in communities with higher compliance rates (fewer penalty records) respond significantly better to treatment letters. This suggests that the moral costs incurred by individuals are proportional to the externalities of societal compliance within their group. These externalities are defined by the group's compliance profile, proximity (distance), and the number of taxpayers. Therefore, the moral costs experienced by an individual influence their behaviour and compliance patterns, which reflect the prevalence of tax morale in society. We also anticipate that this approach will be applicable to analysing community-wide spillover effects.

This triple effect—where typically compliant taxpayers, those with histories of non-compliance, and those exposed to higher compliance norms all respond positively to deterrence measures—underscores the mechanisms, broad applicability and significant impact of such interventions in enhancing overall tax compliance rates. The effectiveness of deterrence letters in motivating various taxpayer segments suggests that these measures can play a crucial role in improving compliance behaviour across diverse groups.

However, beyond this promising outcome, our findings indicate no strong evidence that similar effects occur with literacy or public goods provision interventions. These interventions did not show a strong impact on taxpayers' compliance behaviour on subgroups relevant with compliance costs, evasion history and societal prevalence. This lack of effectiveness highlights the unique influence of deterrence letters compared to other methods aimed at improving tax compliance.

Given the absence of effects from treatments other than the deterrence letter, it is unsurprising that we find no significant interaction effects between the non-deterrence interventions and compliance costs. The specific impact of deterrence letters, as opposed to broader educational or infrastructural improvements, suggests that direct enforcement measures may be more immediately effective in altering compliance behaviour. This distinction emphasises the need for targeted strategies when aiming to enhance tax compliance, as not all interventions yield the same results.

# 6 Discussion

**Deterrence and perceived probability of detection.** In general, the findings indicate that the deterrence letter had the most significant effect on increasing the likelihood of tax payment and filing. This paper suggests two primary reasons for this impact. First, the deterrence letter highlights both the enforcement of administrative penalties and the

adoption of computerised tax audits, which are crucial for improving compliance. By emphasizing these enforcement mechanisms, the letter enhances the perceived probability of detection, $\phi$, under circumstances where tax evasion is identified and penalised, $x^a$, as outlined in Equation 1. When taxpayers perceive a higher likelihood of being audited and facing penalties, they are more likely to comply with tax regulations. This increased perceived risk of detection leads to a reduction in the amount of tax evasion, as taxpayers are less willing to underreport income or avoid filing. Consequently, the deterrence letter encourages higher declared income, more timely tax payments, and improved filing rates, as taxpayers adjust their behavior to avoid potential penalties. This demonstrates the effectiveness of deterrence strategies in promoting tax compliance through both psychological and practical channels.

Secondly, we anticipated that the perceived probability of detection in Indonesia, particularly among small businesses, tends to be quite low. In this context, even a modest increase in the perceived likelihood of detection can lead to a significant rise in declared amounts, especially when penalty rates are low and individuals display risk-averse behavior. Our heterogeneity analysis supports this, revealing that taxpayers in regions with an especially low perceived probability of detection, such as Nusa Tenggara, experienced a significant shift in their perceptions. This shift is largely due to the fact that, prior to receiving the deterrence letters, many taxpayers held a skewed view of audit likelihood, often overestimating how low the actual audit probability was. When confronted with the deterrence letter, which highlighted the possibility of penalties and computerised audits, these taxpayers were more responsive to the information provided. This responsiveness stems from the sudden adjustment in their perceived risk of detection, leading to greater compliance.

These findings are consistent with the economic deterrence models introduced by Allingham and Sandmo (1972) and Yitzhaki (1974a), which argue that taxpayers' compliance behavior is influenced by both the perceived probability of detection and the potential penalties involved. In regions where the perceived audit probability was especially low, the deterrence letter had a disproportionately large effect on compliance, as it corrected taxpayers' previous underestimation of risk. This reinforces the importance of perception in tax enforcement strategies, particularly in contexts where actual detection rates may be lower than taxpayers initially assume.

**Transaction costs and tax morale.** The findings also suggest that the literacy letter has a more substantial impact on improving tax filing rates than on actual tax payments. This highlights the importance of various factors that influence compliant behavior, such

as the inherent costs of compliance and the administrative burden of tax reporting. Additionally, the public goods provision letter also has a significant effect, though its impact is smaller compared to the other treatments. This indicates that the intangible sense of satisfaction or duty associated with fulfilling tax obligations also plays a critical role in shaping taxpayers' willingness to comply.

We can interpret these findings through the lens of the compliance model outlined in Equation 2. Specifically, the literacy letter is associated with lowering both the cost of tax payment, $T(y')$, and the transaction costs, $C(y')$. By simplifying the processes and reducing the perceived complexity of paying taxes, the literacy letter reduces barriers to compliance. On the other hand, the public goods provision letter taps into the psychological and social aspects of tax compliance, particularly the tax morale or the social guilt factor, $M$. This letter reminds taxpayers of the broader benefits their contributions provide to society, such as funding public goods, thereby fostering a sense of civic duty and increasing their willingness to comply.

In this scenario, a reduction in transaction costs (through the literacy letter) or an increase in social guilt and tax morale (through the public goods provision letter) makes the act of filing taxes more appealing, resulting in higher filing rates. While these interventions may not immediately translate into increased tax payments, they play a crucial role in improving filing behavior by addressing both practical and psychological barriers to compliance. This highlights the potential for policies that focus on reducing the complexity of tax procedures or enhancing the moral appeal of tax compliance to significantly improve taxpayer engagement.

**Long-term effects.** Using long-term and time-to-event analysis, this study finds that deterrence letters significantly increased the tax payment rate over a fourteen-month observation period. Moreover, these letters had a more enduring effect on the likelihood of timely filing when compared to literacy and public goods provision letters. This can be attributed to the fact that the perceived probability of detection and the threat of penalties have a stronger influence on taxpayers' utility functions than the costs of compliance or moral considerations.

As previously discussed, small businesses in Indonesia generally operate under a very low perceived probability of detection. Therefore, even a slight increase in this perceived probability, along with the associated penalties emphasized in the deterrence letters, results in a higher marginal effect on both payment and filing behaviors. The impact of this marginal effect is not only significant but also more sustainable over time compared to the effects of the literacy and public goods provision letters. The deterrence letters, by

heightening the perceived risks of non-compliance, provide a stronger and longer-lasting incentive for taxpayers to adhere to their obligations, whereas the effects of reducing compliance costs or enhancing tax morale tend to fade more quickly.

**Repeated trial and reference points.** Additionally, our repeated experiments demonstrate that only the deterrence letter continues to show a statistically significant impact on tax payments. However, it is important to note that the magnitude of this effect in the repeated trials is substantially lower than what was observed in the initial experiment. We propose that this reduction can be understood through the lens of rational decision-making, wherein taxpayers not only consider their inherent preferences but also account for their reference points when deciding on tax payment and filing.

In the initial interventions, taxpayers are likely to form reference points based on their experiences, memories, or expectations surrounding tax compliance. These reference points, once established, serve as benchmarks against which they measure subsequent interventions. In other words, the initial deterrence letters likely created a significant impression, altering their perceived probability of detection and their willingness to comply. However, as these interventions are repeated, taxpayers begin to adapt their behavior, responding not only to the continued emphasis on penalties and detection probabilities but also to how these new interventions compare with the reference points created during the first round.

This concept aligns with the idea of reference-dependent decision-making, where individuals assess new stimuli based on prior experiences. For example, if the initial deterrence letter heightened taxpayers' perceived risk of being audited or penalised, this creates a baseline expectation. In repeated trials, even though the deterrence letters continue to reinforce these risks, taxpayers may no longer perceive the same level of threat because they have already adjusted their behavior to a certain extent based on their previous experience. As a result, the marginal impact of these letters diminishes over time.

Moreover, this decline in the impact of repeated interventions suggests that taxpayers are not responding solely to the probability of detection but also weighing the intervention against their past compliance decisions. Their behavior is influenced not just by the immediate threat of penalties but also by a more complex calculation that takes into account their previous interactions with the tax system and their evolving expectations. This helps explain why the deterrence letter, while still effective, produces a lower marginal effect in repeated experiments compared to its initial implementation.

By incorporating reference points into their decision-making process, taxpayers' responses to repeated interventions reflect a more nuanced understanding of compliance,

where both past experiences and future expectations play a significant role. This highlights the importance of considering behavioral dynamics when designing long-term compliance strategies, as the effectiveness of interventions can change over time as taxpayers adapt to new norms and expectations.

Prospect theory, as introduced by Kahneman and Tversky (1979), offers valuable insights into our findings, particularly regarding why only the deterrence letter had a significant impact on tax payments in the repeated experiment. According to prospect theory, individuals tend to evaluate outcomes relative to a reference point, with their utility function being concave for gains and convex for losses. This means that people are more sensitive to potential losses than to equivalent gains, displaying greater aversion to risk when facing losses.

In the context of tax compliance, we can assume that taxpayers, when exposed to repeated interventions, develop a common reference point, possibly using their audited income as a general benchmark, as suggested by Bernasconi (2004). The deterrence letter, which emphasizes the probability of penalties and detection, likely triggers this reference point because it directly relates to potential financial losses—whether through fines, penalties, or the cost of being audited. Given that losses loom larger than gains in taxpayers' decision-making, they are more likely to assign greater weight to deterrence letters than to other types of interventions.

The repeated exposure to deterrence letters may also reflect a learning curve in taxpayers' risk aversion. Initially, the perceived risk of detection and penalties may not be fully internalized. However, as taxpayers encounter deterrence letters multiple times, they adjust their expectations and behaviors to minimise potential losses. This aligns with the core principles of prospect theory, where people become more risk-averse when faced with repeated opportunities to lose, reinforcing their compliance behavior over time.

Moreover, this focus on loss aversion explains why the deterrence letters have a more sustained impact compared to other interventions, such as literacy or public goods provision letters, which are more likely to be perceived as neutral or even positive. These types of letters do not invoke the same level of urgency or concern because they do not directly threaten financial losses, thereby having less of an impact on taxpayer behavior in the long term. In contrast, the deterrence letters constantly remind taxpayers of the potential for penalties, keeping the possibility of loss at the forefront of their decision-making process.

**Proximity and conformity effects.** Our mechanisms and heterogeneity analysis reveal two key effects relevant to small businesses which we term the *proximity effect* and the *conformity effect*. The proximity effect highlights how factors such as internet connec-

tivity, physical distance from the nearest tax office, and access to information or support influence small businesses' tax compliance and amplify (or diminish) the deterrence or non-deterrence effects as outlined by Slemrod et al. (2001), Yitzhaki (1974b), Dyreng et al. (2016) and Dwenger et al. (2016), specifically for new and inexperienced taxpayers, such as small businesses. When small businesses have reliable internet access, they are better able to obtain tax-related information, file returns, and stay informed about regulations, increasing their likelihood of compliance. Interestingly, we observe a counter-intuitive pattern with physical distance. Taxpayers in more remote areas, who face greater challenges accessing in-person services, tend to rely more on digital resources and external support to meet their obligations. This reliance motivates them to actively engage with online platforms and seek information to ensure compliance. Conversely, small businesses located closer to tax offices may feel less urgency to adopt these measures, given their easier access to physical support. Therefore, the proximity effect demonstrates that distance from tax offices can enhance compliance when businesses leverage technology and available resources.

In addition, the conformity effect shows how tax compliance is influenced by the behaviour of peers and the societal norms within a community. When small businesses see others in their area consistently adhering to tax regulations, a social expectation is created to follow suit. This collective behavior establishes compliance as the standard, encouraging small businesses to avoid social disapproval or being seen as outliers, as shown in studies by Torgler (2004), Luttmer and Singhal (2014), and Blumenthal et al. (2005). As a result, tax compliance becomes more than an individual choice – it reflects broader community practices, reinforced by social influence and peer pressure.

# 7 Conclusion

Developing countries often face significant challenges in fostering sustained tax compliance among their populations. Despite this, there is limited understanding of how nudging strategies and its mechanisms can create long-term effects on tax reporting and payment behaviours. This paper seeks to fill that gap by examining the dynamic impacts of deterrence, literacy, and public goods provision letters on taxpayer compliance in Indonesia, with a particular focus on small businesses, through natural field experiments.

Our findings show that, over the observation period following the distribution of the treatment letters, all three treatment groups – those receiving deterrence, literacy, or public goods provision letters – experienced a significant increase in tax payments compared

to the control group, which did not receive any intervention. This suggests that these nudging strategies can effectively encourage compliance. In particular, the deterrence letter stands out as having the most pronounced impact on tax payments, indicating its potential to drive substantial improvements in compliance, especially within the small business sector. Given the importance of small businesses in Indonesia's economy, these results suggest that deterrence-focused strategies could play a pivotal role in increasing tax revenues and promoting long-term compliance within this crucial sector.

Utilising long-term and time-to-event analysis, we able to illustrate the progression of treatment effects on tax payment and tax reporting, offering a broad picture of the enduring effects of different treatments following an intervention. Concerning tax payment, our findings reveal that the deterrence letter resulted in a significant and lasting increase in payment rates. Similarly, with respect to timely tax filing, the deterrence letter demonstrates notably higher sustained effects compared to the other two treatments and the control group.

In our repeated experiment, we find that deterrence letters continue to exert a statistically significant influence on tax payments, although the average treatment effect is notably reduced compared to the initial intervention. Despite this reduction in magnitude, deterrence letters remain effective in prompting taxpayers to make payments, underscoring their sustained impact over time.

Similarly, when examining tax filing behaviour, the repeated trial reveals that deterrence letters consistently yield the most substantial effects on both the likelihood of filing and filing on time. These findings suggest that deterrence letters, even when issued multiple times, continue to serve as a powerful tool in encouraging compliance, particularly in areas related to meeting filing deadlines. While the impact may diminish slightly with repeated exposure, the consistent positive effects highlight the enduring value of direct enforcement measures in promoting taxpayer adherence to their obligations.

It is crucial to acknowledge the potential limitations in our experimental design. Firstly, our dataset confines our analysis to examining the treatment effects within a fourteen-month period after distributing treatment letters. Consequently, this research does not capture the treatment effects during the full fiscal year preceding the submission of annual tax returns – a period that could be of great interest. Additionally, since our experiment focused on only a few regional offices, we cannot claim external validity for all of Indonesia.

Furthermore, potential limitations in our empirical analysis arise from the absence of data on key socioeconomic variables such as education level, access to information,

gender, and ethnicity, as well as other variables related to non-deterrence factors. These limitations may have restricted our ability to comprehensively capture the heterogeneous treatment effects across the diverse taxpayer population in Indonesia.

# A Appendix

## A.1 Treatment Letter Format: Deterrence




**Translation:**

According to the provisions of Article 2 of the Minister of Finance Regulation Number PMK-242/PMK.03/2014 dated December 24, 2014, regarding the Procedures for Tax Payment and Deposit, the MSME Income Tax (PPh UMKM) as regulated by Article 4 paragraph (2) which must be paid by the Taxpayer, must be deposited no later than the 15th of the following month after the Tax Period ends.

In relation to this, we inform you that you are required to deposit the PPh UMKM before the 15th of the following month, which can also be considered as the submission of a Periodic Tax Return (SPT Masa).

**Red box:** We need to inform you that if the Tax Return (SPT) is not submitted within the specified timeframe, an administrative penalty in the form of a fine of IDR 100,000.00 (one hundred thousand rupiah) may be imposed. Furthermore, other active collection measures may be taken, including warning letters, forced collection orders, asset seizures, auctions, and travel restrictions.

As part of our commitment to creating a fairer tax environment, we will be adopting a computerized system for tax audits and the determination of business turnover. We are also currently collaborating with third parties for tax data exchange.

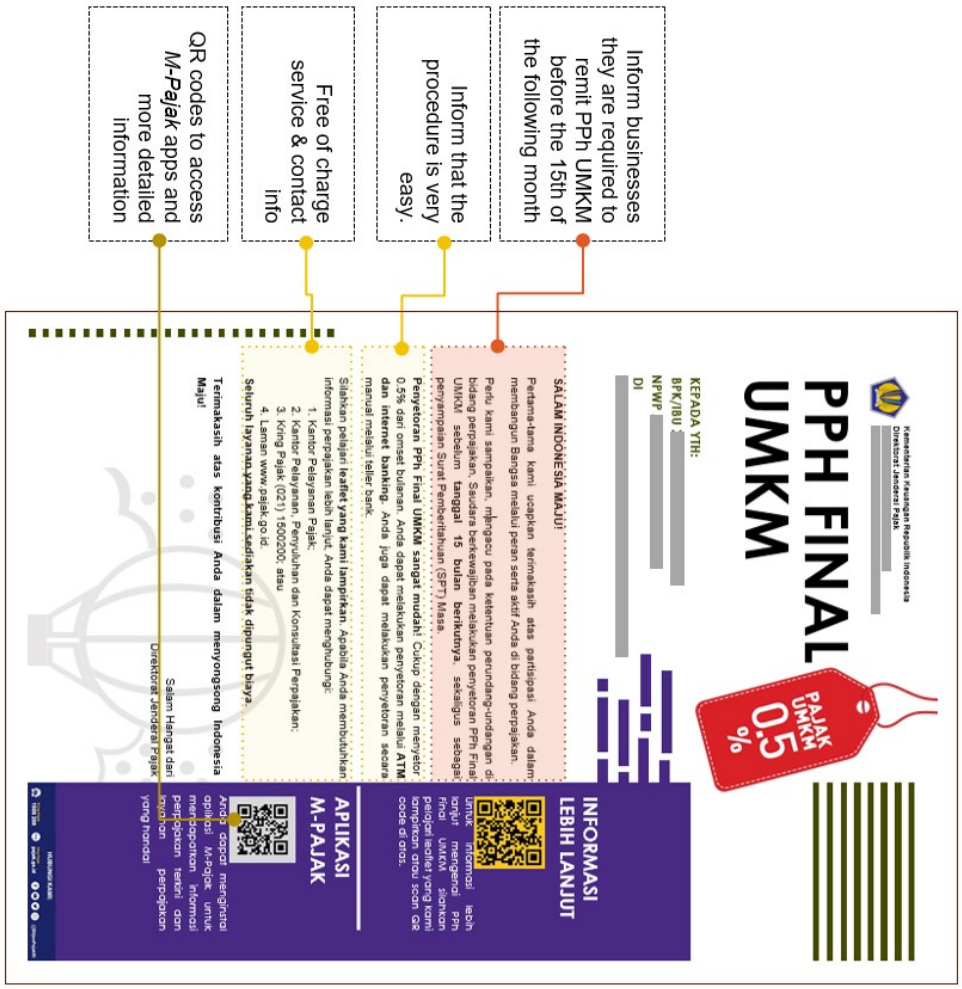
If you require further tax information, you may contact:

1. The Tax Office;
2. The Tax Service, Counseling, and Consultation Office;
3. Tax Hotline (021) 1500200; or
4. The website www.pajak.go.id.

All services we provide are free of charge.

We thank you for your participation in tax payment and reporting.

## A.2 Treatment Letter Format: Literacy




**Translation:**

First of all, we would like to thank you for your participation in building the nation through your active role in taxation.

We would like to inform you that, according to the tax regulations, you are required to make the payment for Final MSME Income Tax (PPh Final UMKM) before the 15th of the following month, which also serves as the submission of your Periodic Tax Return (SPT Masa).

Paying the PPh Final UMKM is very easy! You simply need to deposit 0.5% of your monthly turnover. You can make the payment via ATMs and internet banking. You can also make manual payments through bank tellers.

Please refer to the attached leaflet for more information. If you require further tax information, you can contact:

1. The Tax Office;
2. The Tax Service, Counseling, and Consultation Office;
3. Tax Hotline (021) 1500200; or
4. The website www.pajak.go.id.

All services we provide are free of charge.

Thank you for your contribution in advancing Indonesia towards progress!

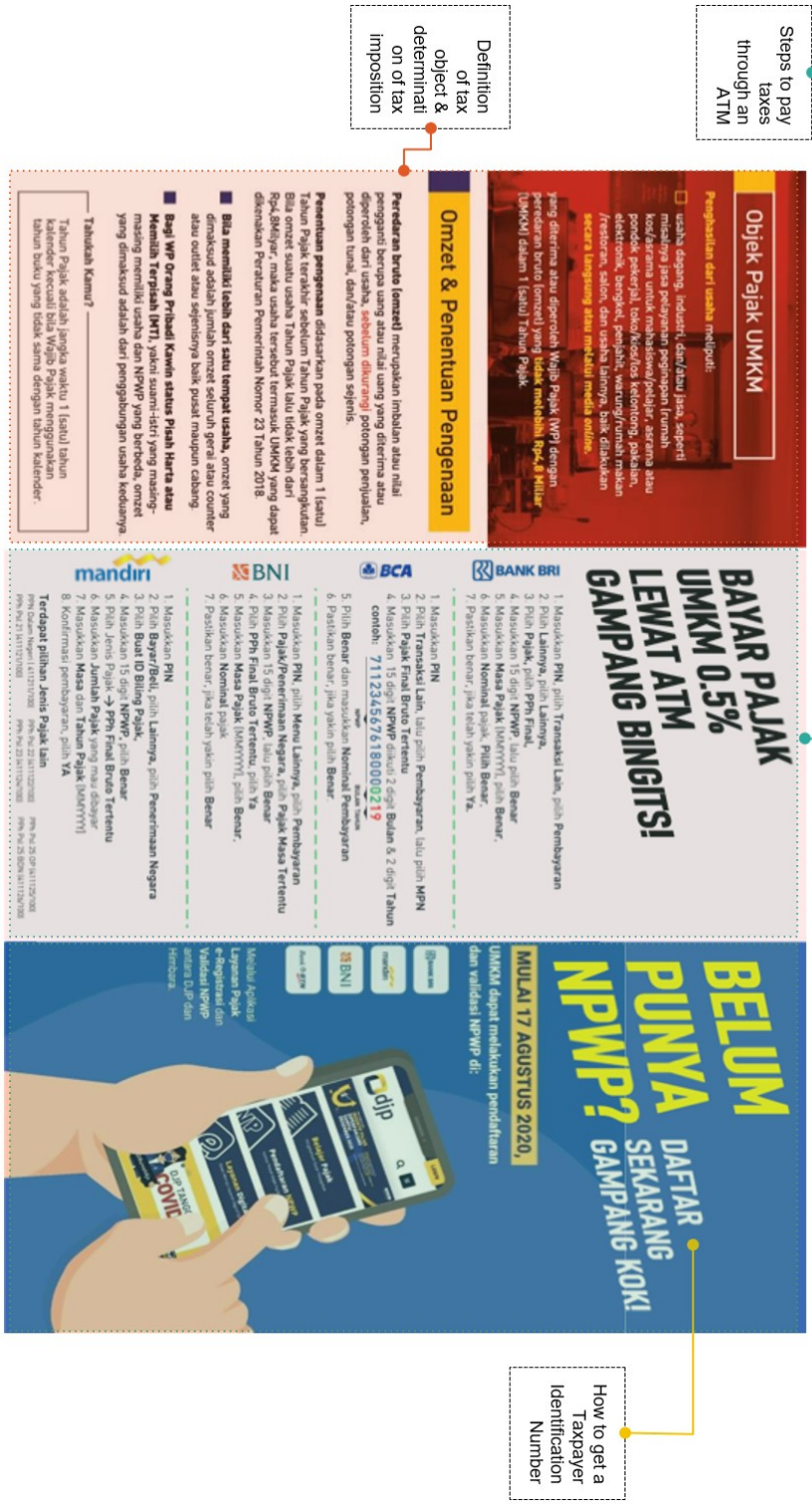
**Blue box:** For more information regarding the Final MSME Income Tax (PPh Final UMKM), please refer to the attached leaflet or scan the QR code above.

You can install the M-Pajak app to access the latest tax information and reliable tax services.

### A.2.1 Treatment Letter Format: Flyer for Literacy Page 1



### A.2.2 Treatment Letter Format: Flyer for Literacy Page 2



## A.3 Treatment Letter Format: Pubic Goods Provision




**Translation:**

According to the provisions of Article 2 of the Minister of Finance Regulation Number PMK-242/PMK.03/2014 dated December 24, 2014, regarding the Procedures for Tax Payment and Deposit, the MSME Income Tax (PPh UMKM) as stipulated in Article 4 paragraph (2), which must be paid by the Taxpayer, must be deposited no later than the 15th (fifteenth) day of the following month after the tax period ends.

In relation to this, we inform you that you are required to deposit the PPh UMKM before the 15th of the following month, which will also be considered as the submission of the Periodic Tax Return (SPT Masa).

**Blue box:** We would like to inform you that your participation and contribution are essential in supporting the government's efforts to handle the COVID-19 pandemic. Your contribution will be allocated to the National Economic Recovery Program post-pandemic to help build a STRONG INDONESIA.

If you require further tax information, you can contact:

1. The Tax Office;
2. The Tax Service, Counseling, and Consultation Office;
3. Tax Hotline (021) 1500200; or
4. The website www.pajak.go.id.

All services we provide are free of charge.

We thank you for your participation in tax payment and reporting.

### A.3.1 Treatment Letter Format: Flyer for Pubic Goods Provision



## A.4 Cost-Benefit Analysis

| | **Deterrence Letter** | **Literacy Letter** | **Public Goods Provision Letter** | **Total** |
|---|---|---|---|---|
| **Initial Trial** | | | | |
| *Estimated Revenues* | | | | |
| Tax paid (USD) | 2.51 | 0.34 | 0.21 | 3.06 |
| *Estimated Costs* | | | | |
| Printing - Letters (USD) | 0.05 | 0.05 | 0.05 | 0.15 |
| Printing - Flyers (USD) | 0 | 0.2 | 0.2 | 0.4 |
| Postage (USD) | 0.58 | 0.58 | 0.58 | 1.74 |
| Total Costs (USD) | 0.63 | 0.83 | 0.83 | 2.29 |
| Net Revenue (USD) | 1.88 | -0.49 | -0.62 | 0.77 |
| *N* | 2,991 | 2,989 | 3,040 | |
| Total Net Revenue (USD) | 5,623 | -1,465 | -1,885 | 2,274 |
| **Repeated Trial** | | | | |
| Estimated Revenue (USD) | 1.6 | 0 | 0 | 1.6 |
| Total Costs (USD) | 0.63 | 0.83 | 0.83 | 2.29 |
| Net Revenue (USD) | 0.97 | -0.83 | -0.83 | -0.69 |
| *N* | 1,379 | 1,424 | 1,379 | |
| Total Net Revenue (USD) | 1,338 | -1,182 | -1,145 | -989 |
| | | **Total Benefits (USD)** | | **1,285** |

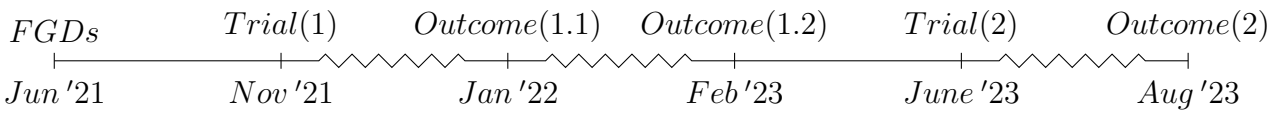
*Note:* The actual postage costs are lower because tax offices have adopted standard postage rates. We did not account for the potential decrease in revenue due to the compliance trade-off, where lower penalties might also lead to a reduction in total revenue after the treatments..

## A.5 Treatment Letters Dispatch Date

| Trials | Tax Office | Location | Dispatch Date | | |
|---|---|---|---|---|---|
| | | | Treatment 1 | Treatment 2 | Treatment 3 |
| **Initial** | T045 | Java | 15-18 Nov 21 | 15-18 Nov 21 | 15-18 Nov 21 |
| | T041 | Java | 15-26 Nov 21 | 15-26 Nov 21 | 15-26 Nov 21 |
| | T047 | Java | 15-18 Nov 21 | 15-18 Nov 21 | 15-18 Nov 21 |
| | T043 | Java | 15-18 Nov 21 | 15-18 Nov 21 | 15-18 Nov 21 |
| | T042 | Java | 19 Nov 21 | 19 Nov 21 | 19 Nov 21 |
| | T044 | Java | 19 Nov 21 | 19 Nov 21 | 19 Nov 21 |
| | T914 | Nusa Tenggara | 11-14 Nov 21 | 11-14 Nov 21 | 11-14 Nov 21 |
| | T911 | Nusa Tenggara | 09-14 Nov 21 | 09-14 Nov 21 | 09-14 Nov 21 |
| | T915 | Nusa Tenggara | 12 Nov 21 | 12 Nov 21 | 12 Nov 21 |
| | T301 | Sumatera | 15 Nov 21 | 15 Nov 21 | 15 Nov 21 |
| | T306 | Sumatera | 15 Nov 21 | 15 Nov 21 | 15 Nov 21 |
| | T307 | Sumatera | 09-15 Nov 21 | 09-15 Nov 21 | 09-15 Nov 21 |
| | T313 | Sumatera | 09-15 Nov 21 | 09-15 Nov 21 | 09-15 Nov 21 |
| **Repeated** | T047 | Java | 7-9 Jun 23 | 7-9 Jun 23 | 7-9 Jun 23 |
| | T043 | Java | 7-9 Jun 23 | 7-9 Jun 23 | 7-9 Jun 23 |
| | T041 | Java | 7-9 Jun 23 | 7-9 Jun 23 | 7-9 Jun 23 |
| | T914 | Nusa Tenggara | 12-16 Jun 23 | 12-16 Jun 23 | 12-16 Jun 23 |
| | T915 | Nusa Tenggara | 12-20 Jun 23 | 12-20 Jun 23 | 12-20 Jun 23 |

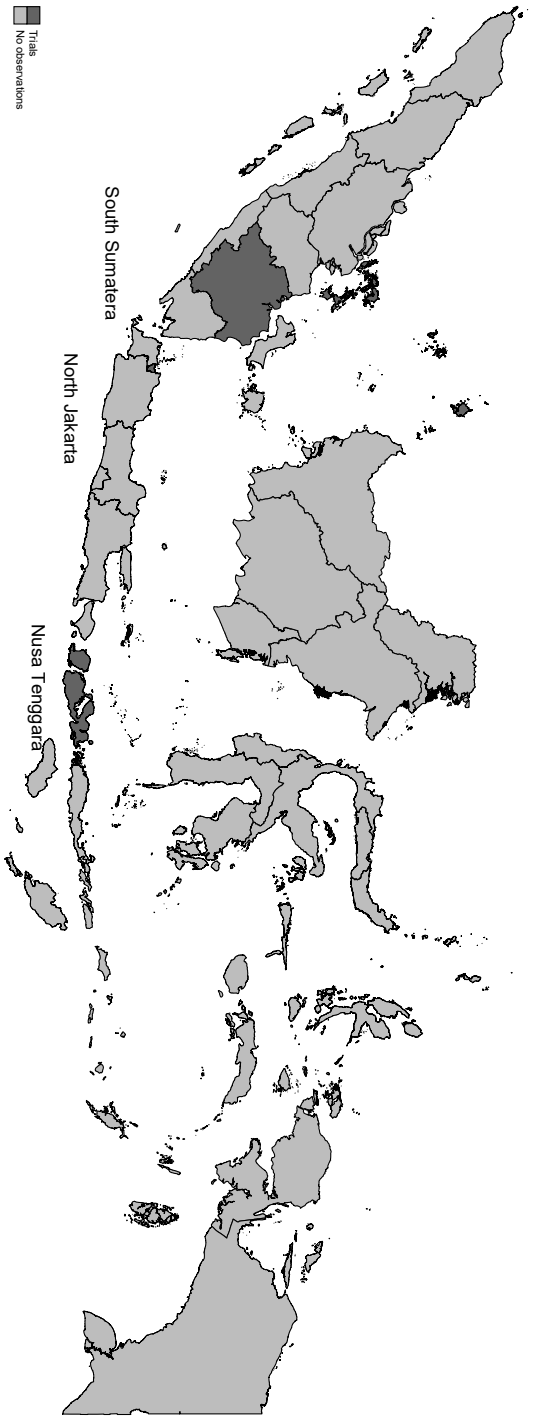
*Note:* Treatment 1: Deterrence Letters, Treatment 2: Literacy Letters, Treatment 3: Public Goods Provision Letters.

## A.6 Experiment Timeframe



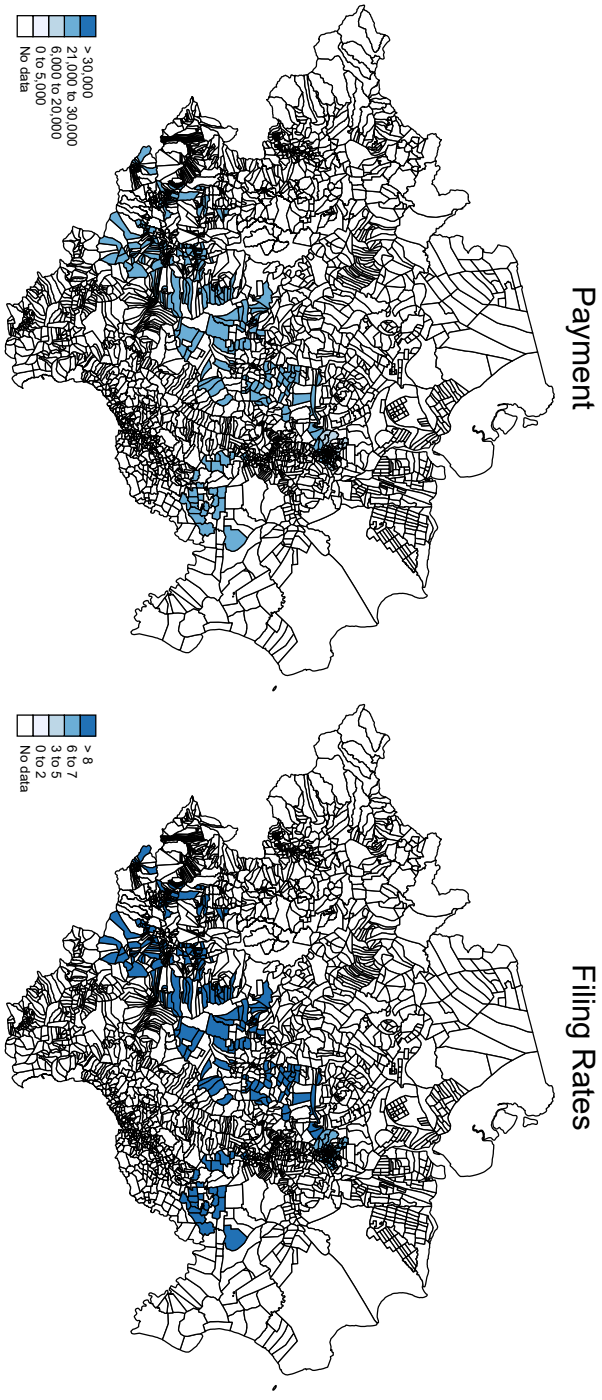

- Focus Group Discussions (FGDs) involved conducting multiple interviews with small business owners, tax officials, and policy actors.
- Trial(1) is the initial trial.
- Outcome(1.1) represents the average treatment effects immediately after the trials are conducted.
- Outcome(1.2) comprises the long-term effects.
- Trial(2) is the repeated trial.
- Outcome(2) represents the average treatment effects of the repeated trials.

## A.7 Field Experiment Location



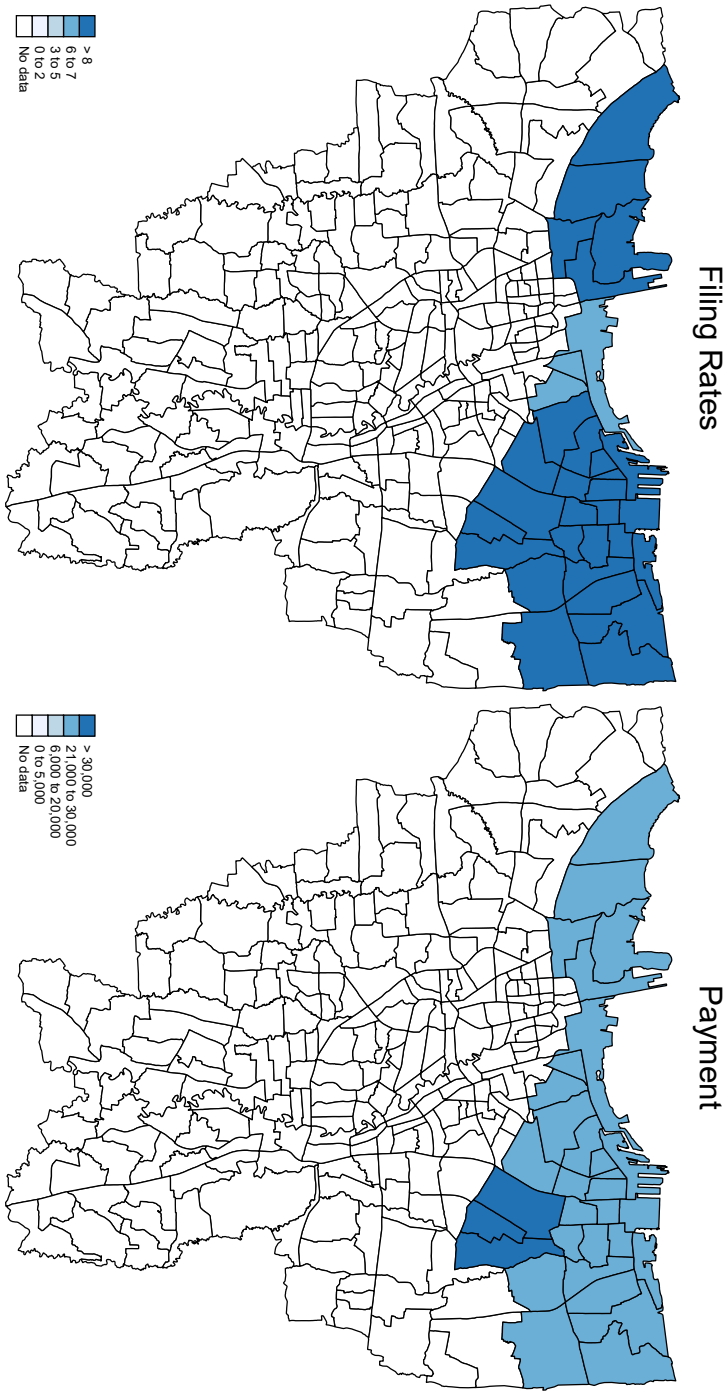

*Note:* Map data is provided by geospatial information from the Indonesia Statistics. Data is merged with subdistrict data from the experimental design.

## A.8 Average Treatment Effect - South Sumatera



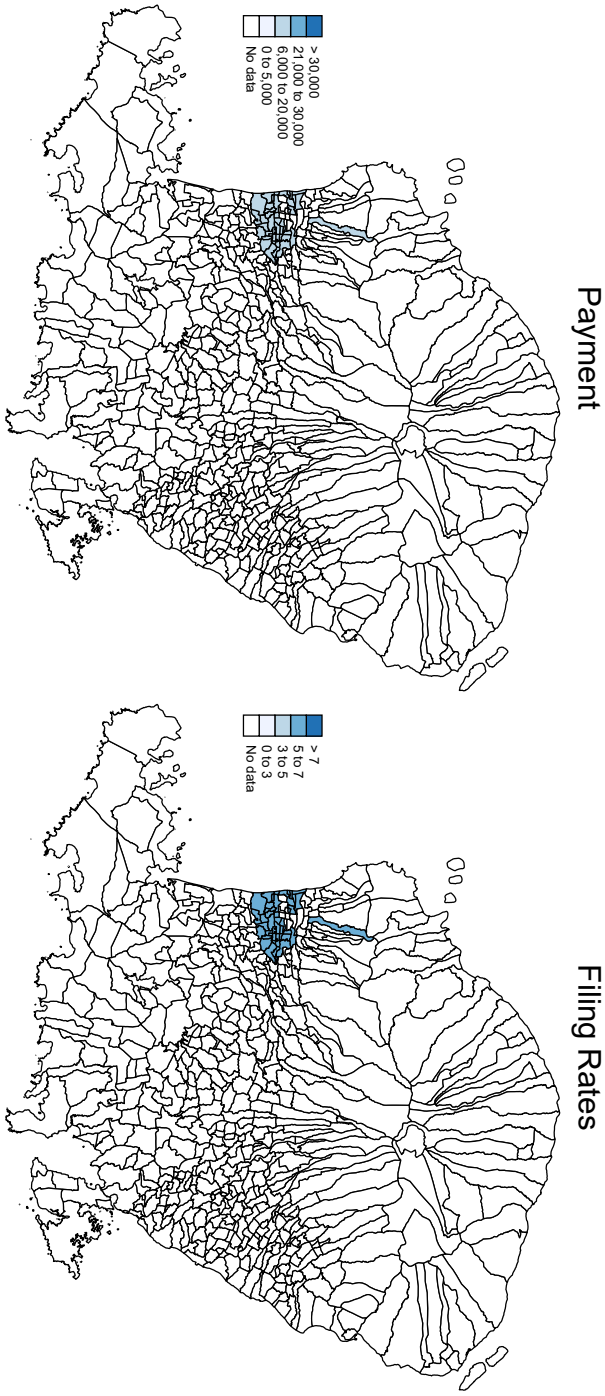

*Note:* Map data is provided by geospatial information from the Indonesia Statistics. Data is merged with subdistrict data from the experimental design.

## A.9 Average Treatment Effect - Jakarta




*Note:* Map data is provided by geospatial information from the Indonesia Statistics. Data is merged with subdistrict data from the experimental design.

## A.10 Average Treatment Effect - Nusa Tenggara




*Note:* Map data is provided by geospatial information from the Indonesia Statistics. Data is merged with subdistrict data from the experimental design.

## A.11 Nelson Aalen Cumulative Hazard Function Estimates

| Timeframe | At Risk | Fail | Survivor function | s.e | [95% conf. int.] |
|---|---|---|---|---|---|
| **Treatment=0** | | | | | |
| 1 | 2976 | 464 | 0.8441 | 0.0066 | 0.8305 0.8566 |
| 2 | 2512 | 25 | 0.8357 | 0.0068 | 0.8219 0.8485 |
| 7 | 2487 | 78 | 0.8095 | 0.0072 | 0.7949 0.8231 |
| 14 | 2409 | 0 | 0.8095 | 0.0072 | 0.7949 0.8231 |
| **Treatment=1** | | | | | |
| 1 | 1991 | 1648 | 0.449 | 0.0091 | 0.4311 0.4667 |
| 2 | 1343 | 8 | 0.4463 | 0.0091 | 0.4284 0.4641 |
| 3 | 1335 | 90 | 0.4162 | 0.009 | 0.3985 0.4339 |
| 4 | 1245 | 46 | 0.4009 | 0.009 | 0.3833 0.4184 |
| 5 | 1199 | 51 | 0.3838 | 0.0089 | 0.3664 0.4012 |
| 6 | 1148 | 158 | 0.331 | 0.0086 | 0.3142 0.3479 |
| 8 | 990 | 124 | 0.2895 | 0.0083 | 0.2734 0.3059 |
| 10 | 866 | 15 | 0.2845 | 0.0082 | 0.2685 0.3008 |
| 11 | 851 | 80 | 0.2578 | 0.008 | 0.2422 0.2736 |
| 14 | 771 | 0 | 0.2578 | 0.008 | 0.2422 0.2736 |
| **Treatment=2** | | | | | |
| 1 | 2989 | 699 | 0.7661 | 0.0077 | 0.7506 0.7809 |
| 2 | 2990 | 14 | 0.7615 | 0.0078 | 0.7458 0.7763 |
| 3 | 2276 | 231 | 0.6842 | 0.0085 | 0.6672 0.7005 |
| 4 | 2045 | 25 | 0.6758 | 0.0086 | 0.6587 0.6923 |
| 5 | 2020 | 19 | 0.6695 | 0.0086 | 0.6523 0.6860 |
| 8 | 2001 | 43 | 0.6551 | 0.0087 | 0.6377 0.6718 |
| 14 | 1958 | 0 | 0.6551 | 0.0087 | 0.6377 0.6718 |
| **Treatment=3** | | | | | |
| 1 | 3040 | 838 | 0.7243 | 0.0081 | 0.7081 0.7399 |
| 2 | 2202 | 15 | 0.7194 | 0.0081 | 0.7031 0.7350 |
| 3 | 2187 | 104 | 0.6852 | 0.0084 | 0.6684 0.7014 |
| 4 | 2083 | 33 | 0.6743 | 0.0085 | 0.6574 0.6907 |
| 5 | 2050 | 12 | 0.6704 | 0.0085 | 0.6534 0.6868 |
| 14 | 2038 | 0 | 0.6704 | 0.0085 | 0.6534 0.6868 |

*Note:* Treatment 1: Deterrence Letters, Treatment 2: Literacy Letters, Treatment 3: Public Goods Provision Letters. The cumulative hazard function is the number of "failures" per subject – interpreted as an accumulative positive tax payment.

## A.12 Stratified Log-Rank Test

| **Strata** | **Treatment** | **Obs. events** | **Expected events** | $\chi^2$ |
|---|---|---|---|---|
| Location: Java | 0 | 1682 | 1248.26 | 2245.51*** |
| | 1 | 1623 | 2619.29 | |
| | 2 | 1662 | 1369.11 | |
| | 3 | 1698 | 1428.35 | |
| Location: Sumatera | 0 | 914 | 701.32 | 1716.13*** |
| | 1 | 894 | 1448.68 | |
| | 2 | 942 | 777.8 | |
| | 3 | 938 | 760.21 | |
| Location: Nusa Tenggara | 0 | 380 | 340.34 | 275.66*** |
| | 1 | 380 | 483.35 | |
| | 2 | 385 | 357.9 | |
| | 3 | 404 | 367.41 | |
| Sector: Service | 0 | 1507 | 1147.1 | 2266.86*** |
| | 1 | 1480 | 2359.56 | |
| | 2 | 1529 | 1269.62 | |
| | 3 | 1578 | 1317.72 | |
| Sector: Trading and Manuf. | 0 | 1469 | 1140.57 | 1872.47*** |
| | 1 | 1417 | 2207.54 | |
| | 2 | 1460 | 1233.02 | |
| | 3 | 1462 | 1226.87 | |
| Business size: Small | 0 | 973 | 757.65 | 1349.91*** |
| | 1 | 980 | 1513.49 | |
| | 2 | 1000 | 839.21 | |
| | 3 | 1018 | 860.65 | |
| Business size: Medium | 0 | 1011 | 782.32 | 1354.89*** |
| | 1 | 942 | 1490.09 | |
| | 2 | 984 | 831.1 | |
| | 3 | 1026 | 859.49 | |
| Business size: Large | 0 | 992 | 747.1 | 1430.55*** |
| | 1 | 975 | 1563.33 | |
| | 2 | 1005 | 832.5 | |
| | 3 | 996 | 825.08 | |

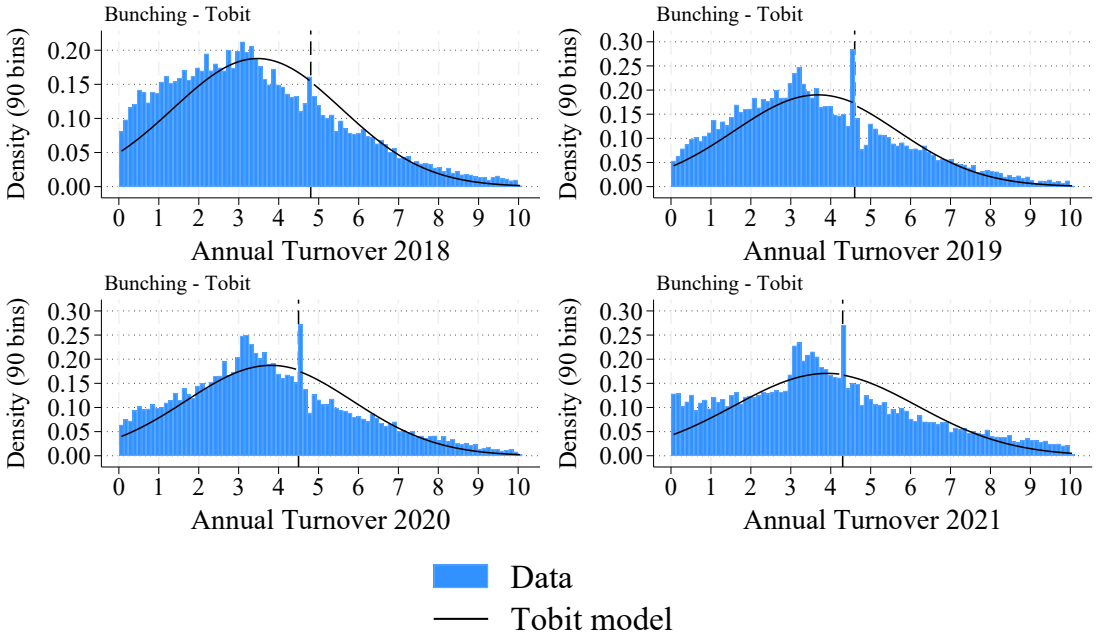
*Note:* Treatment 0: Control, Treatment 1: Deterrence Letters, Treatment 2: Literacy Letters, Treatment 3: Public Goods Provision Letters. * $p$ <0.1, ** $p$ <0.05, *** $p$ <0.01.

## A.13 Estimation of Geographic-Clustered Effects

| **Geospatial cluster** | **Diff. Means** | | **Linear-in-Means** | |
|---|---|---|---|---|
| | **coef** | **s.e.** | **coef** | **s.e.** |
| Kamal Muara ⇔ Kapuk Muara | -0.035 | 0.037 | 0.071 | 0.021 |
| Warakas ⇔ Sunter Jaya | -0.018* | 0.019 | 0.023 | 0.001 |
| Selaparang ⇔ Ampenan | 0.086* | 0.052 | 0.032* | 0.021 |
| Sandubaya ⇔ Mataram Kec. | 0.125 | 0.012 | 0.121 | 0.014 |
| Mataram ⇔ Sekarbela | -0.045* | 0.023 | 0.034 | 0.020 |
| **Average within cluster (sample)** | | | | |
| CL-A (ARH) | 0.121 | 0.025 | 0.021 | 0.025 |
| CL-B (CW) | 0.035* | 0.030 | -0.052 | 0.023 |
| CL-C (PS) | -0.012 | 0.013 | 0.125 | 0.045 |
| CL-D (MK) | 0.235** | 0.072 | 0.085 | 0.032 |

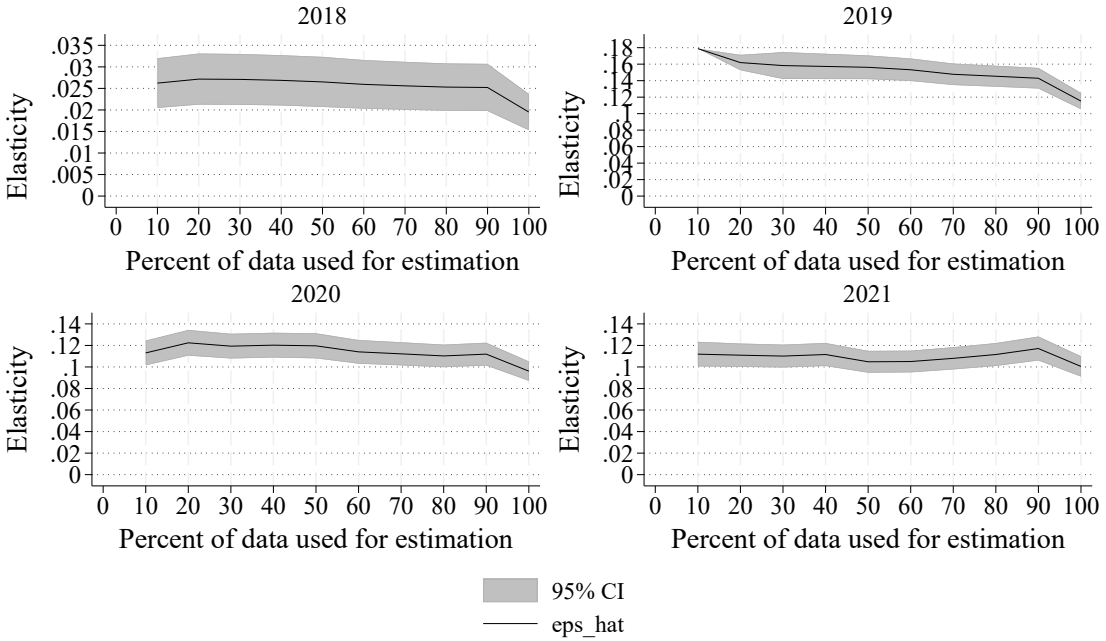
*Note:* Cluster-robust s.e. Regressions include distance FE. ARH is Jalan Arif Rahman Hakim, CW is Catur Warga, PS is Plumpang Semper, MK is Muara Karang. Cluster is approximated based on address and geospatial information from Indonesia Statistics. *$p$ <0.1, ** $p$ <0.05, *** $p$ <0.01.

## A.14 Distibution of Taxpayers - Non Experimental Data



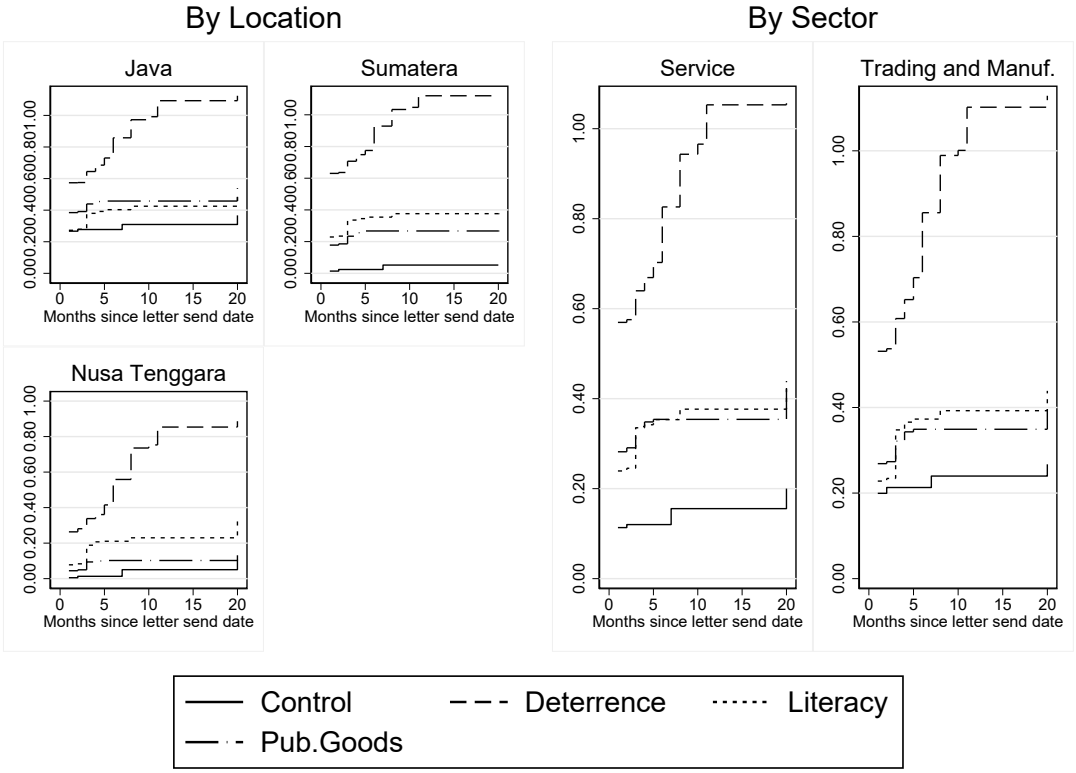

*Note:* The estimation uses a larger, non-experimental dataset and employs data-driven algorithms from Bertanha et al. (2022). The spike represents the mass point, referred to as "bunching," while the drop indicates the "missing mass". Turnover is measured in billions of Rupiahs, with the cut-off point set at Rp4.8 billion, the threshold for SME registration.

## A.15 Elasticity of Bunching Taxpayers - Non Experimental Data



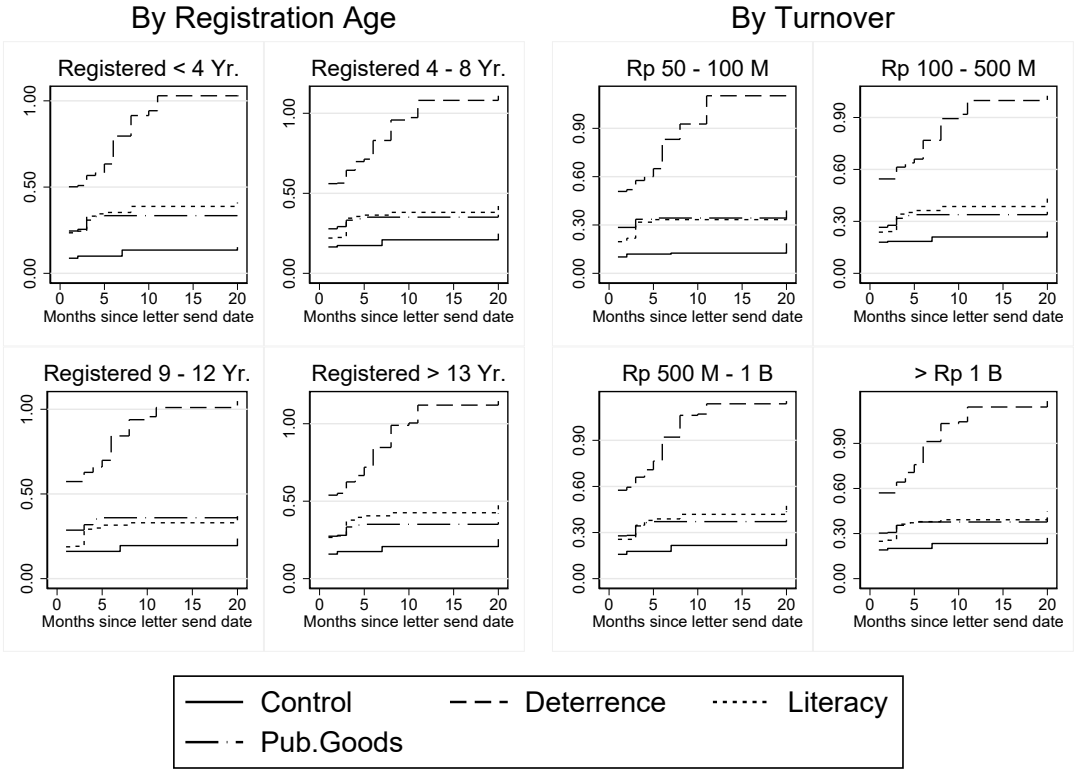

*Note:* The elasticity bunching estimation is based on a larger, non-experimental dataset and is produced using data-driven algorithms from Bertanha et al. (2022).

## A.16 Cumulative Hazard Function of Payment - By Location and Sector



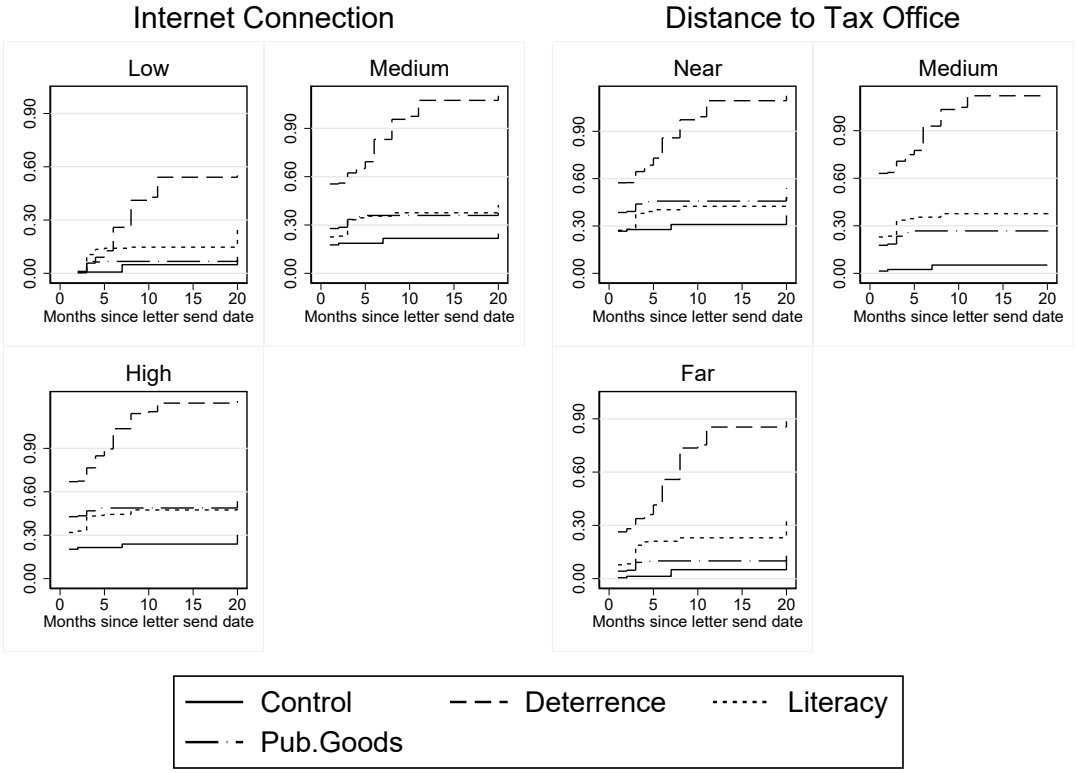

*Note:* Each line represents a rate of respective payment amount of all taxpayers. The progression rate shows that compared to control group, the deterrence letter led to a particularly strong increase in payment rates during 14 months observation period. The corresponding increase in payment rates resulting from the public good provision letter is weaker.

## A.17 Cumulative Hazard Function of Payment - By Registration Age and Turnover



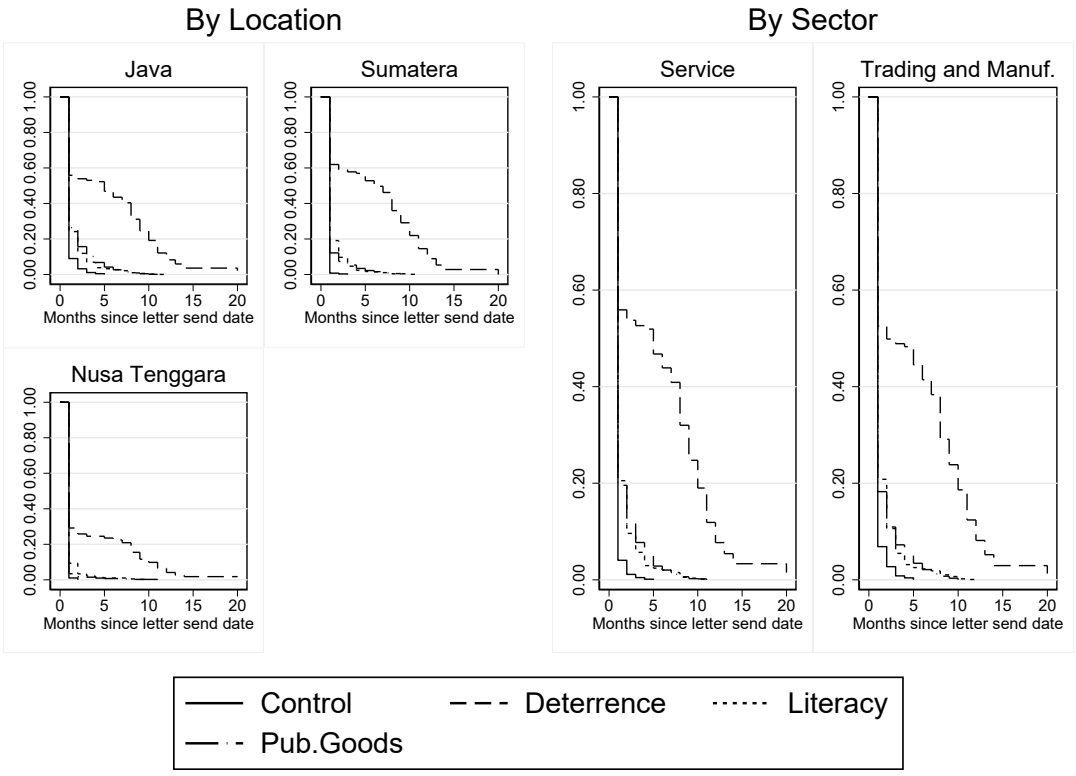

*Note:* Each line represents a rate of respective payment amount of all taxpayers. The progression rate shows that compared to control group, the deterrence letter led to a particularly strong increase in payment rates during 14 months observation period. The corresponding increase in payment rates resulting from the public good provision letter is weaker.

## A.18 Cumulative Hazard Function of Payment - By Internet and Distance



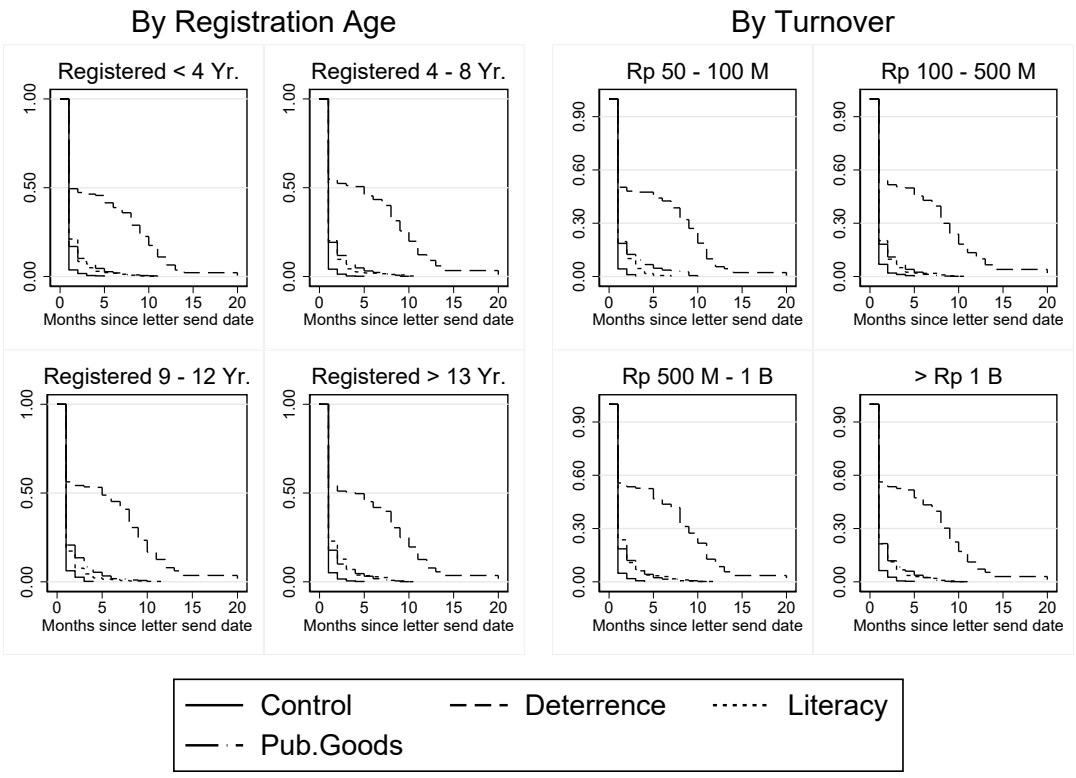

*Note:* Each line represents a rate of respective payment amount of all taxpayers. The progression rate shows that compared to control group, the deterrence letter led to a particularly strong increase in payment rates during 14 months observation period. The corresponding increase in payment rates resulting from the public good provision letter is weaker.

## A.19 Survival Estimates of Filing by Location and Sector



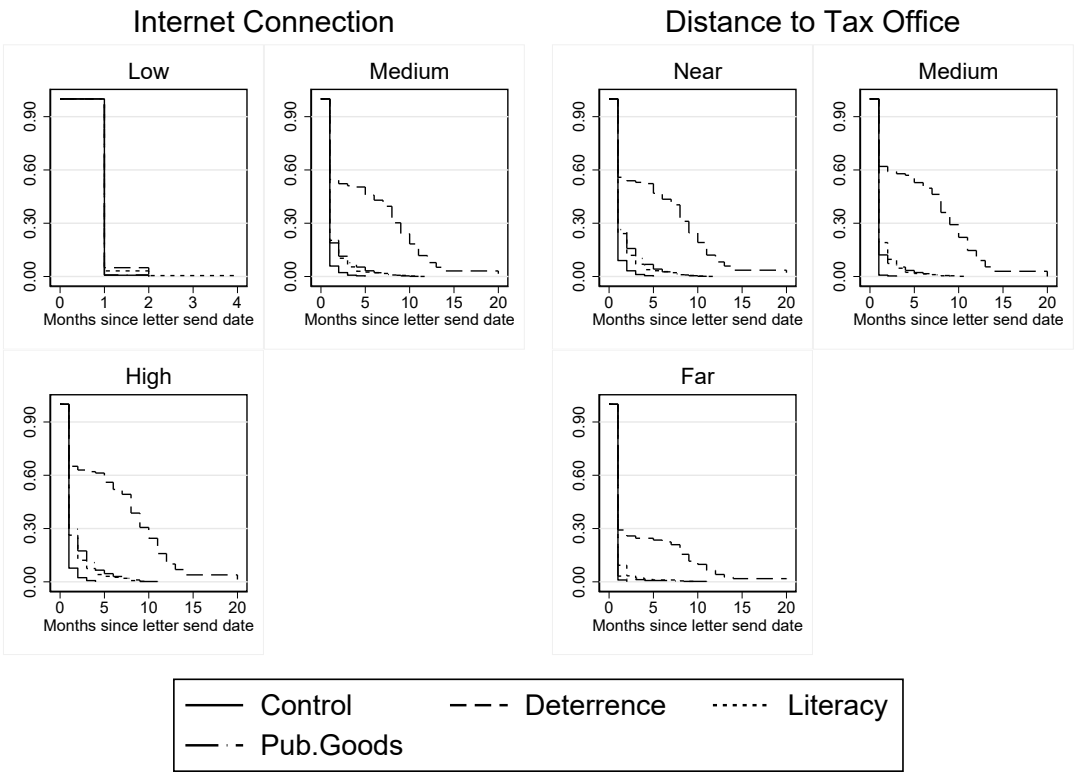

*Note:* Each line represents a rate of respective timely tax filing of all taxpayers . The failure event criteria is the dummy whether monthly tax return submission is not timely i.e., after the 15th of the following month. The progression rate analysis reveals that, in comparison to the control group, the deterrence letter engendered a more enduring impact on the propensity for timely filing over the course of a 14-month observational period.

## A.20 Survival Estimates of Filing by Registration Age and Turnover



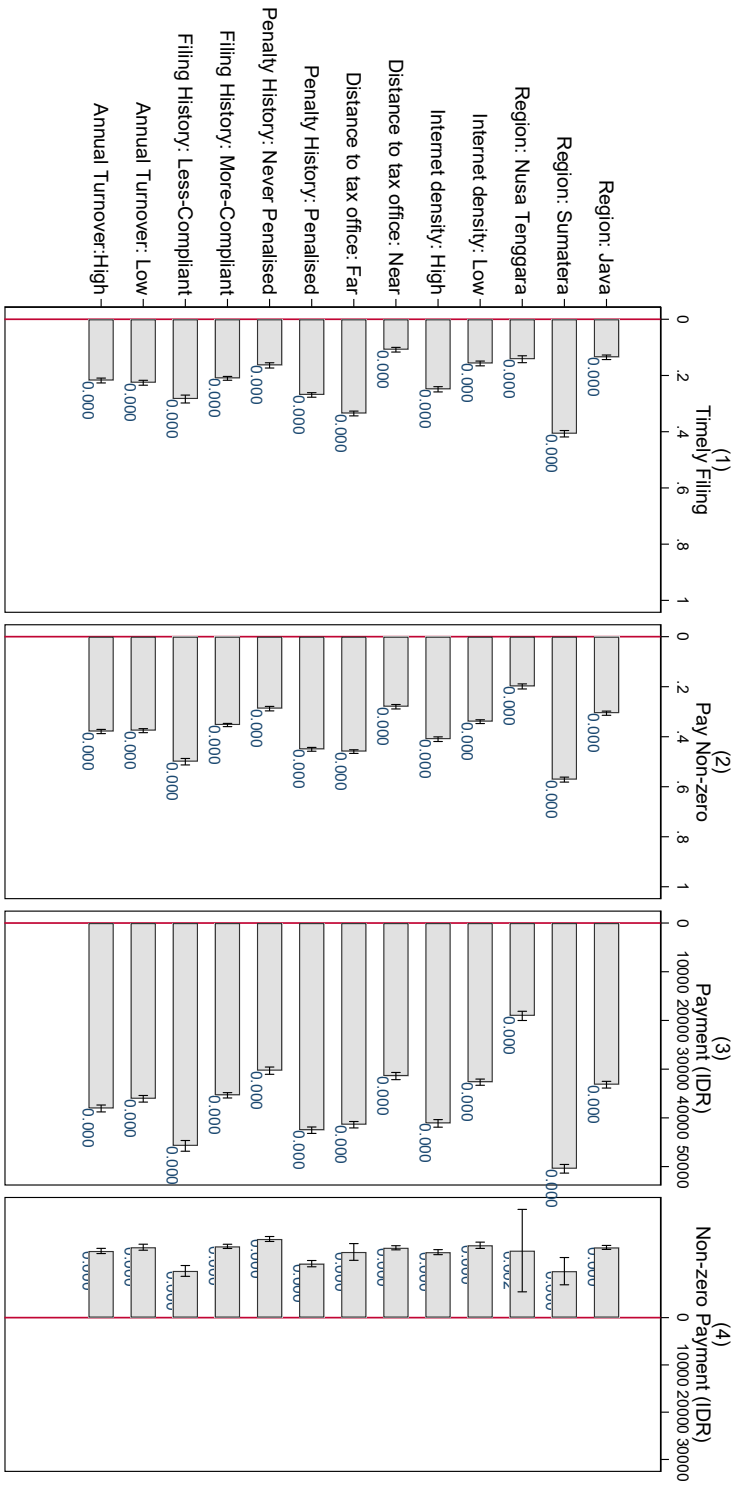

*Note:* Each line represents a rate of respective timely tax filing of all taxpayers . The failure event criteria is the dummy whether monthly tax return submission is not timely i.e., after the 15th of the following month. The progression rate analysis reveals that, in comparison to the control group, the deterrence letter engendered a more enduring impact on the propensity for timely filing over the course of a 14-month observational period.

## A.21 Survival Estimates of Filing by Internet and Distance



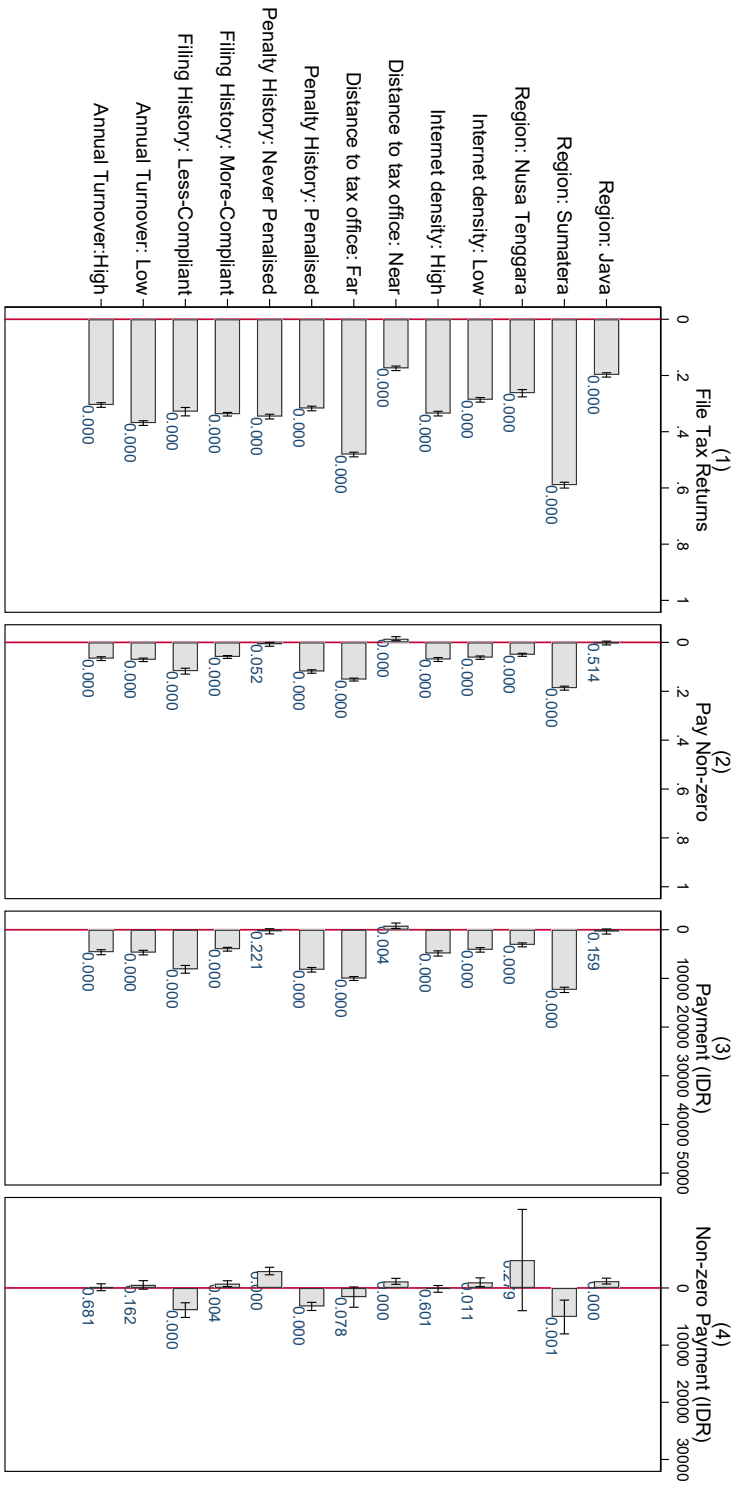

*Note:* Each line represents a rate of respective timely tax filing of all taxpayers . The failure event criteria is the dummy whether monthly tax return submission is not timely i.e., after the 15th of the following month. The progression rate analysis reveals that, in comparison to the control group, the deterrence letter engendered a more enduring impact on the propensity for timely filing over the course of a 14-month observational period.

## A.22 Heterogeneous Treatment Effects: Deterrence



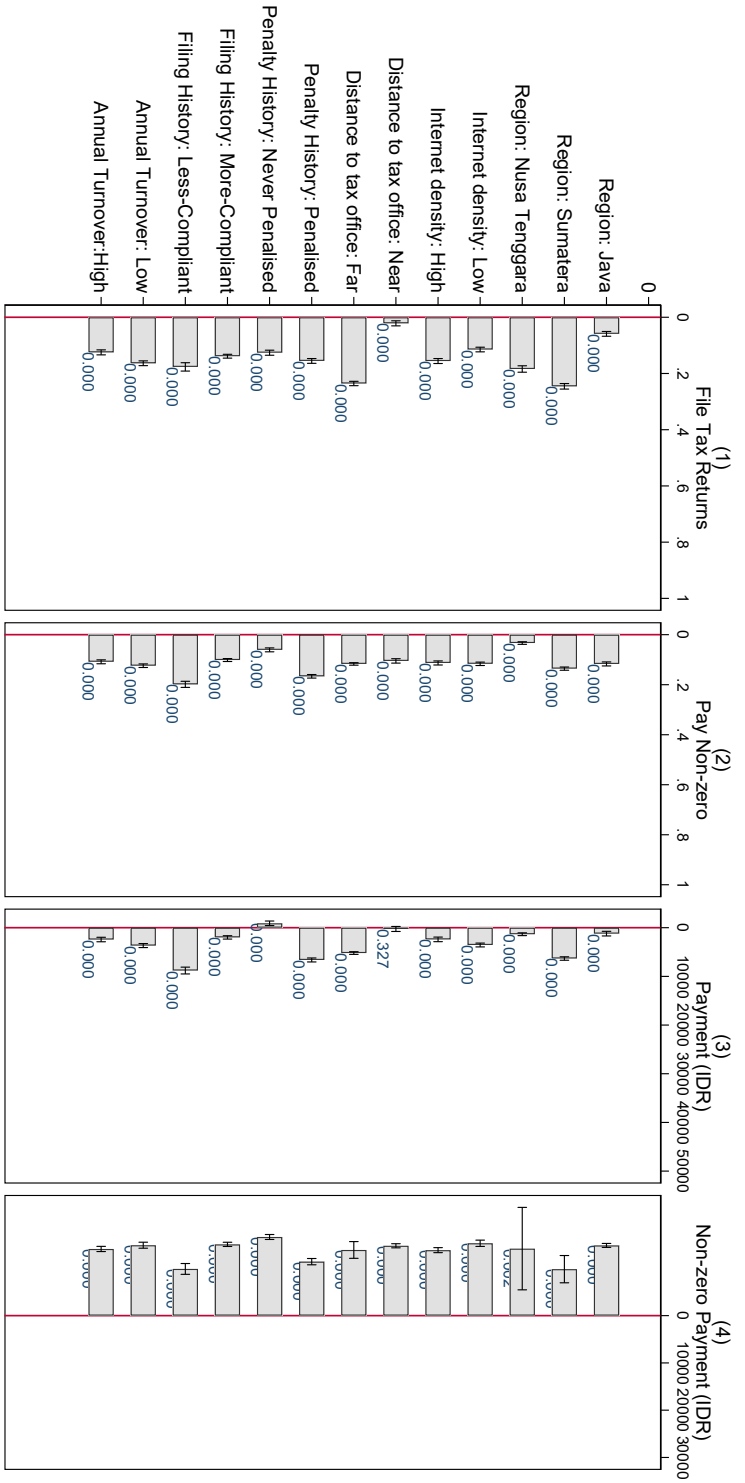

*Note:* Each bar represents the average treatment effects for respective outcome variables with 95% confidence intervals shown as lines around the point estimates.

## A.23 Heterogeneous Treatment Effects: Literacy




*Note:*Each bar represents the average treatment effects for respective outcome variables with 95% confidence intervals shown as lines around the point estimates.

## A.24 Heterogeneous Treatment Effects: Public Goods Provision




*Note:* Each bar represents the average treatment effects for respective outcome variables with 95% confidence intervals shown as lines around the point estimates.

# References


**Aalen, Odd**, "Nonparametric Inference for a Family of Counting Processes," *The Annals of Statistics*, 1978, *6* (4), 701–726. Publisher: Institute of Mathematical Statistics.

**Advani, Arun, William Elming, and Jonathan Shaw**, "The Dynamic Effects of Tax Audits," *The Review of Economics and Statistics*, May 2023, *105* (3), 545–561.

**Alasfour, F., M. Samy, and Roberta Bampton**, "The Determinants of Tax Morale and Tax Compliance: Evidence from Jordan," in "in" 2016.

**Allingham, Michael G. and Agnar Sandmo**, "Income Tax Evasion: A Theoretical Analysis," *Journal of Public Economics*, 1972, *1* (3-4), 323–338.

**Alm, James, Betty R. Jackson, and Michael Mckee**, "Estimating The Determinants of Taxpayers Compliance with Experimental Data," *National Tax Journal*, March 1992, *45* (1), 107–114. Number: 1.

**Bernasconi, Michele**, "Tax Evasion, Tax Rates, and Reference Dependence," *FinanzArchiv (FA)*, 2004, *60* (3), 422–445. Place: Tübingen Publisher: Mohr Siebeck.

**Bertanha, Marinho, Andrew H. McCallum, Alexis Payne, and Nathan Seegert**, "Bunching estimation of elasticities using Stata," *The Stata Journal*, September 2022, *22* (3), 597–624. Publisher: SAGE Publications.

**Blumenthal, Marsha and Joel Slemrod**, "The Compliance Cost of the U.S. Individual Income Tax System: A Second Look After Tax Reform," *National Tax Journal*, 1992, *45* (2), 185–202. publisher: National Tax Association.

**_ , Brian Erard, and Chih chin Ho**, "Participation and Compliance with the Earned Income Tax Credit," *National Tax Journal*, June 2005, *58.*

**_ , Charles Christian, and Joel Slemrod**, "Do Normative Appeals Affect Tax Compliance? Evidence from a Controlled Experiment in Minnesota," *National Tax Journal*, 2001, *54* (1), 125–138.

**Bott, Kristina M., Alexander Cappelen, Erik Ø. Sørensen, and Bertil Tungodden**, "You've Got Mail: A Randomised Field Experiment on Tax Evasion," Working Papers 2017-051, Human Capital and Economic Opportunity Working Group 2017.

**Brehm, Sharon S. and Jack W. Brehm**, *Psychological Reactance: A Theory of Freedom and Control*, Academic Press, October 2013. Google-Books-ID: 6ZxGBQAAQBAJ.

**Castro, Lucio and Carlos Scartascini**, "Tax Compliance and Enforcement in the Pampas: Evidence from a Field Experiment," Research Department Publications IDB-WP-472, Inter-American Development Bank, Research Department December 2013.

**Christiansen, Tobias Gabel**, "Dynamic effects of tax audits and the role of intentions," *Journal of Public Economics*, June 2024, *234*, 105121.

**Coleman, Stephen**, "The Minnesota Income Tax Compliance Experiment: Replication of the Social Norms Experiment," Technical Report 11 2007. [Online; accessed 2021-04-30].

**DeBacker, Jason, Bradley T. Heim, Anh Tran, and Alexander Yuskavage**, "Once Bitten, Twice Shy? The Lasting Impact of Enforcement on Tax Compliance," *The Journal of Law and Economics*, February 2018, *61* (1), 1–35. Publisher: The University of Chicago Press.

**Dwenger, Nadja, Henrik Kleven, Imran Rasul, and Johannes Rincke**, "Extrinsic and Intrinsic Motivations for Tax Compliance: Evidence from a Field Experiment in Germany," *American Economic Journal: Economic Policy*, 8 2016, *8* (3), 203–232.

**Dyreng, Scott D., Jeffrey L. Hoopes, and Jaron H. Wilde**, "Public Pressure and Corporate Tax Behavior," *Journal of Accounting Research*, 2016, *54* (1), 147–186. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1475-679X.12101.

**Freedman, Judith and Claire Crawford**, "Small Business Taxation," Technical Report, Rochester, NY 6 2010. [Online; accessed 2021-07-14].

**Frey, Bruno S. and Reto Jegen**, "Motivation Crowding Theory," *Journal of Economic Surveys*, 2001, *15* (5), 589–611. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1467-6419.00150.

**Hallsworth, Micheal, Jone A. List, Robert D. Metcalfe, and Ivo Vlaev**, "The Behavioralist as Tax Collector: Using Natural Field Experiments to Enhance Tax Compliance," *Journal of Public Economics*, 2017, *148*, 14–31.

**Hasseldine, D.John and K.Jan Bebbington**, "Blending economic deterrence and fiscal psychology models in the design of responses to tax evasion: The New Zealand experience," *Journal of Economic Psychology*, 6 1991, *12* (2), 299–324.

**Kahneman, Daniel and Amos Tversky**, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 1979, *47* (2), 263–291. Publisher: [Wiley, Econometric Society].

**Kaplan, Edward L and Paul Meier**, "Nonparametric Estimation from Incomplete Observations," *Journal of the American Statistical Association*, 1958, *53* (282), 457–481.

**Kopczuk, Wojciech and Cristian Pop-Eleches**, "Electronic Filing, Tax Preparers and Participation in the Earned Income Tax Credit," *Journal of Public Economics*, 2007, *91*, 1351–1367.

**Kornhauser, M.**, "Normative and Cognitive Aspects of Tax Compliance: Literature Review and Recommendations for the Irs Regarding Individual Taxpayers," *undefined*, 2007. [Online; accessed 2021-04-26].

**Kotsogiannis, Christos, Luca Salvadori, John Karangwa, and Theonille Mukamana**, "Do tax audits have a dynamic impact? Evidence from corporate income tax administrative data," *Journal of Development Economics*, September 2024, *170*, 103292.

**Kőszegi, Botond and Matthew Rabin**, "A Model of Reference-Dependent Preferences," *The Quarterly Journal of Economics*, 2006, *121* (4), 1133–1165. Publisher: Oxford University Press.

**Luttmer, Erzo F. P. and Monica Singhal**, "Tax Morale," *Journal of Economic Perspectives*, 2014, *28* (4), 149–168.

**Mantel, N. and W. Haenszel**, "Statistical aspects of the analysis of data from retrospective studies of disease," *Journal of the National Cancer Institute*, April 1959, *22* (4), 719–748.

**Mascagni, Giulia**, "From the Lab to the Field: A Review of Tax Experiments," *Journal of Economic Surveys*, 2018, *32* (2), 273–301.

_ **and Christopher Nell**, "Tax Compliance in Rwanda: Evidence from a Message Field Experiment," *Economic Development and Cultural Change*, 2 2021, p. 713929.

**McKerchar, Margaret and Chris Evans**, "Sustaining Growth in Developing Economies through Improved Taxpayer Compliance: Challenges for Policy Makers and Revenue Authorities," *eJournal of Tax Research*, December 2009, *7* (2), 171–201. Num Pages: 31 Place: Sydney, Australia Publisher: University of New South Wales.

**Mittone, Luigi, Fabrizio Panebianco, and Alessandro Santoro**, "The bomb-crater effect of tax audits: Beyond the misperception of chance," *Journal of Economic Psychology*, August 2017, *61*, 225–243.

**Nelson, Wayne**, "Theory and Applications of Hazard Plotting for Censored Failure Data," *Technometrics*, 1972, *14* (4), 945–966. Publisher: [Taylor & Francis, Ltd., American Statistical Association, American Society for Quality].

**OECD**, "Compliance Risk Management: Managing and Improving Tax Compliance," 2004.

**OECD**, *Tax Compliance by Design: Achieving Improved SME Tax Compliance by Adopting a System Perspective*, OECD, October 2014.

**Pope, Jeff and Hijattulah Abdul-Jabbar**, "Small and Medium Sized Enterprises and Tax Compliance Burden in Malaysia: Issues and Challenges for Tax Administration," *Small Enterprise Research The Journal of SEAANZ*, 1 2008, *16*, 47–60.

**Satyadini, Agung and Arifin Rosid**, "Does Tax Simplification Motivate Small Businesses to be More Compliant? Evidence from a Regression Discontinuity in Indonesia," *Bulletin of Indonesian Economic Studies*, May 2024, *60* (2), 217–238. Publisher: Routledge _eprint: https://doi.org/10.1080/00074918.2023.2252746.

**Savage, I. Richard**, "Contributions to the Theory of Rank Order Statistics -the Two-Sample Case," *The Annals of Mathematical Statistics*, September 1956, *27* (3), 590–615.

**Schneider, Friedrich, Andreas Buehn, and Claudio E. Montenegro**, "New Estimates for the Shadow Economies all over the World," *International Economic Journal*, 12 2010, *24* (4), 443–461.

**Slemrod, Joel, Marsha Blumenthal, and Charles Christian**, "Taxpayer response to an increased probability of audit: evidence from a controlled experiment in Minnesota," *Journal of Public Economics*, 3 2001, *79* (3), 455–483.

**Torgler, Benno**, "Moral Suasion: An Alternative Tax Policy Strategy? Evidence from a Controlled Field Experiment in Switzerland," *Economics of Governance*, 2004, *5* (3), 235–253.

_ **and Christoph Schaltegger**, "Tax Morale: A Survey with a Special Focus on Switzerland," *Swiss Journal of Economics and Statistics (SJES)*, 2006, *142* (III), 395–425. publisher: Swiss Society of Economics and Statistics (SSES).

**Traxler, Christian**, "Social norms and conditional cooperative taxpayers," *European Journal of Political Economy*, 2010, *26* (1), 89–103.

**Twisk, Jos, Lisa Bosman, Trynke Hoekstra, Judith J M Rijnhart, Welten M, and Heijmans M**, "Different ways to estimate treatment effects in randomised controlled trials," *Contemporary Clinical Trials Communications*, March 2018, *10.*

**Wenzel, M. and N. Taylor**, "An Experimental Evaluation of Tax-Reporting Schedules: A Case of Evidence-Based Tax Administration," *Journal of Public Economics*, 2004, *88* (12), 2785–2799.

**World Bank**, "Innovations in Tax Compliance: Building Trust, Navigating Politics, and Tailoring Reform," *World Bank*, 2022.

**Yitzhaki, S.**, "Income Tax Evasion: A Theoretical Analysis," *Journal of Public Economics*, 1974, *3* (2), 201–202.

**Yitzhaki, Shlomo**, "Income tax evasion: A theoretical analysis," *Journal of Public Economics*, May 1974, *3* (2), 201–202. Number: 2.