

## Article

Thomas Lange\*

# Hard versus Soft Intervention: Compliance Effects among Firms<sup>1</sup>

<https://doi.org/10.2478/ntaxj-2025-0001>

Received Mar 19, 2023; accepted Dec 5, 2024

**Abstract:** Audits and letters are two of the enforcement strategies available to a tax administration to ensure compliance. In this paper, we use a unique set of experiments to determine which of the two enforcement strategies is more effective. We use firm-level data from 1,974 randomized audits and 8,000 information letters. We find that audits result in an immediate and significant increase in firms' payroll tax remittance, whereas letters have a weaker effect. Updated, perceived audit probability seems to sustain adjustments in payroll tax remittance two years post-treatment (on-site audit or standardized, electronic information letters). Firms receiving the information letters also adjust their remittance upwards, and more so when the letters are actually read. Our "back of the envelope" cost-benefit calculations suggest that tax administrations could save resources by partially switching to cheaper enforcement strategies, like information letters.

**Keywords:** Public economics, taxation, Business taxes, tax evasion, behavioral economics, treatment effect Models, instrumental variables.

**JEL Classification:** H2; H25; H26; H32; C31; C36

## 1 Introduction

Audits and information letters are two enforcement strategies with distinct effects on a firm's behavior and significantly different administrative costs for tax

authorities. On-site audits are more intrusive than impersonal letters requesting needed information to clarify a firm's report or alerting the firm about possible errors in their report. While a scheduled visit from a tax auditor typically evokes a sense of being subject to the disclosure of potential errors, a letter stating the firm's reporting liabilities is perceived more as a reminder of the applicable rules than as a sense of being inspected. Thus, behavioral changes following an audit likely stem from a sense of enforcement, whereas changes after receiving a letter requesting information may feel more voluntary. From the tax administration's perspective, audits are time-consuming, depending on their scope and nature, while letters, once written, are practically cost-free to distribute. A change in enforcement strategy could therefore result in significant cost savings for the tax administration.

Firms play a central part in the economy and account for a significant fraction of the collected tax revenue. In 2021, the tax revenue ratio of firms to individual taxpayers in Norway was 0.75, not including value added tax (VAT) (Ministry of Finance, 2022). The payroll tax represents a significant contribution to the Norwegian tax revenue. In 2021, Norwegian firms paid NOK 216.4 billion in the total payroll tax, equaling 14.2 percent of the total tax revenue, or 22.9 percent of taxes liable to firms only (Ministry of Finance, 2022). Employers are responsible for the remittance and the reporting of the payroll tax on behalf of their employees, as part of the financing of the National Insurance Scheme. The tax is levied on employees' salaries and other taxable remuneration for work and assignments in and outside of an employment relationship. The tax rates are set by the Norwegian Parliament every year and range from 5.1 percent to 14.1 percent depending on the sector and geographical location.

Firms' tax behavior and exposure to the tax administration differ from those of the individual taxpayers. While the latter is obligated to submit their tax return only once a year, a firm has several reporting and payment obligations throughout the year, such as the bi-monthly reporting of the payroll tax remittance, and the monthly submission of the VAT return. Thus, the

\*Corresponding Author: Thomas Lange: The Norwegian School of Economics, Norway; Email: thomaslmh@gmail.com

<sup>1</sup> Thanks to Pablo Garriga, Evelina Gavrilova-Zoutman, Andreas Kotsadam, Knut Løyland, Anne May Melsom, Hanne Beate Næringsrud, Nina Serdarevic, Floris Zoutman, and Arnstein Øvrum for useful insights and comments.

higher frequency of encounters with the tax authorities may affect the firms' compliance in different ways than in the case of individual taxpayer compliance, requiring different enforcement strategies.

This paper contributes to recent literature on the effects of different tax enforcement strategies. Our experimental setting captures central aspects of the real-world reporting environment, such as the presence of firms' bimonthly remittance of the payroll tax. We compare the effects on firms' remittance of payroll tax of a scalable strategy, namely, comparing the effects of standardized information letters with the effects of on-site audits. Unlike most other contributions to the literature, we study the effects of these two instruments on the same population and in the same institutional setting. The attention is directed to a population of 30,961 firms in different labor-intensive businesses in the Norwegian economy.

Since our two treatments (on-site audits and standardized, electronic information letters) are directed toward the firm's reporting obligations, we seek to uncover whether the firm responds to and corrects their payment of payroll tax after treatment with an audit or a letter stating the required obligations. We measure compliance through remitted amounts of payroll tax. We believe payroll tax is of particular relevance in this context, both because it is a tax only levied on firms, and the message conveyed in the treatments is targeted towards employee reporting obligations, which indirectly influence payroll tax remittance.

The audits were implemented on a stratified random sample of 1,974 firms during the filing and auditing season of 2018. The audited firms were given an up-front notification about the content of the audit so that the firms could prepare relevant documentation. The letters were sent to the same population, but to a different sample of 8,000 recipients simultaneously in June 2018. A 14 percent fraction (1,130) of the letter recipients never opened the letter and were therefore not treated.

The empirical analysis is divided into two main parts. The first part studies the average treatment effects (ATE) of audits and letters on the firms' remittance of payroll tax. As we are interested in a comparison with the article by Bjørneby, Alstadsæter, and Telle (2021) on employee effects, we have kept the number of employees as a dependent variable alongside the analysis of the payroll tax. While on-site audits are more comprehensive in that they involve more scrutiny than standardized letters, we expect considerable variation in remitted payroll tax across firms depending on the treatment assigned.

We find that the audit treatment increased the firm's payroll tax remittance in 2018 by 13.20 percent compared to the pre-treatment level in 2017, and we find a statistically significant increase in remitted payroll tax in all post-treatment years (2018–2020) for the audited firms compared to the reference group receiving no treatment. Furthermore, we find a similar but less strong effect of 1.82 percent for the firms receiving a letter. This confirms the findings of Ortega and Scartascini (2020) who found stronger effects from physical visits than emails. The effect seems to be more stable for the audited firms, two years following the treatment (2020). This is consistent with previous findings such as Kleven et al. (2011), DeBacker et al. (2018) and Advani, Elming, and Shaw (2023).

The second part of the analysis studies the local average treatment Effects (LATE) of the letters.<sup>2</sup> This allows us to isolate the effects of the firms that read the letters, i.e., the treated (or "compliers") in this experiment, from the entire group of firms that were intended to be in the treated group (ITT). The nature of the letter treatment differs from that of audits in this respect: there is often one-sided noncompliance in that a fraction of firms receiving a digital letter will leave it unread. Thus, we study the effects of being treated, using an IV model capturing this effect. We find that the compliers remit significantly higher payroll taxes in all subsequent years following the treatment. This is consistent with the findings of Bjørneby, Alstadsæter, and Telle (2021).

The policy implications of the results in this paper are twofold. First, our analysis reveals that there is a significant compliance effect for alternative enforcement strategies like information letters. Second, since information letters are scalable to a larger population at low cost, such enforcement may be preferable to traditional, more expensive on-site audits. We include in this part a limited cost-benefit analysis, indicating that letters may be preferable in the short run.

The paper is organized as follows; Section 2 reviews adjacent literature, Section 3 describes the institutional background and the random audit program, Section 4 describes the data, Section 5 lays out the experimental design, Section 6 estimates the effect of the two treatments, Section 7 suggests some policy implications, and Section 8 provides the conclusion.

<sup>2</sup> Sometimes referred to as complier average causal effect (CACE) (Gerber and Green, 2012).

## 2 Related Literature

The rational agent-based theory of Allingham and Sandmo (1972) has shaped the modern tax administration's approach to enforcement by increasing taxpayer's perceived costs of evasion, decreasing the cost of compliance, and tailoring enforcement strategies towards different taxpayer segments (Baer and Silvani, 1997). More recently, behavioral and moral aspects of tax compliance have also found their way into tax research (e.g., Dhimi and Al-Nowaihi 2007 and Luttmer and Singhal 2014) and have gained increased attention among tax administrations. Rather than increasing audit frequency, modern tax administrations have sought to gain knowledge of compliance effects from other, prospectively cheaper and more scalable enforcement strategies (Murphy 2019, Alm 2019a, Keen and Slemrod 2017).

The compliance effects of enforcement strategies in general, and audits in particular, have recently been studied in experimental designs, using randomized samples (Kotsadam et al. 2021; Kleven et al. 2011; Advani, Elming, and Shaw 2023; DeBacker et al. 2018; Hebous et al. 2020; Bjørneby, Alstadsæter, and Telle 2021). Most of these studies use samples of individual taxpayers except Almunia and Lopez-Rodriguez (2018), which uses a nonrandom sample of firms from large taxpayers unit (LTU) in Spain; Boning et al. (2020), which uses total employer tax deposits; and Bjørneby Alstadsæter, and Telle (2021), which use on-site audits of Norwegian firms from a stratified sample. Unlike these contributions, which study the effects of enforcement on the individual taxpayer, this article studies the effects of audits and letters as enforcements on the firm.

The novelty of this contribution is the study of how the firm as an entity may respond in a very different way than individual taxpayers. Furthermore, the effects of two enforcements on the payroll tax are studied—a tax to which only the firm is liable—in the same environment. D'Agosto et al. (2018) also studied the compliance effect on small businesses of two different enforcement strategies, namely on-site and desk-based audits. They use risk-based audits, however, not random selection. Our approach also differs from random designs like that of Boning et al. (2020) who studied the effects on overall employment taxes, i.e., payroll taxes and employee income taxes withheld and remitted by IRS-assigned at-risk firms. Since the latter can be manipulated by the employee, even in a third-party remittance regime (Bjørneby et al., 2021), part of

the effects in Boning et al. (2020) may be explained by employee noncompliance or collusive actions involving both employer and employee. We overcome these issues by focusing on the payroll tax alone.

The effects of audits on compliance vary across studies and are sensitive to which dependent variable is chosen, but Kleven et al. (2011); DeBacker et al. (2018); and Advani, Elming, and Shaw (2023) find lasting effects on subsequent tax compliance among audited taxpayers. These effects are confirmed by Løyland et al. (2019) using risk-based audits and Hebous et al. (2020). Because audits are costly, tax administrations also use cheaper and less intrusive enforcement policies such as information campaigns, reminders, enforcement emails (Brockmeyer et al. 2019), letters (Pomeranz 2015, Doerrenberg and Schmitz 2015), and encouragements (Kotsadam et al. 2021). The effects of such soft treatments are also mixed, depending on the message portrayed in the treatment (Alm 2019b, Slemrod 2019, Meiselman 2018, Pomeranz and Vila-Belda 2019). Information that increases perceived detection probability seems to have positive short-term effects on reporting (Slemrod, Blumenthal, and Christian 2001; Kleven et al. (2011); Fellner, Sausgruber, and Traxler 2013; Bott et al. 2020), while general appeals to tax morale and social norms seem to have little or no effect (Hallsworth et al. 2017), or even negative effects (De Neve et al. 2021). Using four different letter treatments, Bergolo et al. (2023) found that information on audits *decreased* the perceived probability of being audited, while, at the same time, inducing a significant deterrent effect on tax evasion.

Our design is similar to that of Boning et al. (2020), except that the 12,172 firms they studied were ex-ante suspected of noncompliance, selected by an algorithm based on payments before and during the fourth quarter of 2014, i.e., firms showing signs of noncompliance before treatment. While it is highly useful for tax administrations to acquire knowledge about treatment effects on potentially noncompliant firms, our contribution differs in that our random selection is not limited to a population of firms suspected of noncompliance, but rather to a fraction of the economy, irrespective of previous risk-based selection. Our findings are representative of a population of 30,961 firms from labor-intensive businesses. Our approach provides a broader view of the compliance effects of the treatments within the entire sector. Boning et al. (2020) is a more targeted approach, allowing for confirmation or rejection of the initial suspicions, and providing evidence of different behavior among high-risk firms. In their setting, a fraction of the firms also received an information letter,

and another fraction had an on-site IRS Revenue Officer visit.

The unit of study in Kotsadam et al. (2021) is a sample of individual taxpayers claiming deductions. Taxpayers claiming deductions are more prone to errors than the rest of the taxpayer population, and so the external validity of the effects observed in Kotsadam et al. (2021) may not extend to a population of firms. What motivates individual taxpayer behavior may also differ from the determinants behind firm compliance.

Bjørneby et al. (2021) study a compliance mechanism evolving from third-party reporting using randomized audits of firms in Norway, and they find compliance effects with the audit treatment. We may expect similar effects from the audits in our study, but we move beyond their setup and compare the effects from the audits with prospective effects from letters. This may give tax administrations information about more cost-efficient enforcement strategies.

## 3 Institutional Background

### 3.1 The Random Audit Program

In 2017, the Norwegian Tax Administration (NTA) introduced the random audit program to build more systematic knowledge about tax compliance. The program is conducted along three thematic strands: labor market regulations, VAT compliance, and quality of third-party data reporting. We focus on data from the first strand.

The audits were directed toward disclosing the scope and magnitude of formal compliance errors in labor-intensive businesses in Norway. This focus originated from risk-based audits, where experience indicates that formal noncompliance is particularly high where foreign labor is involved. Risk-based audits are biased, however, and so the inference from the audit sample to the general population is limited. This random audit program, therefore, sought to measure the compliance gap in labor-intensive industries to make more robust conclusions about compliance risk in the population of these industries. The audits had reporting objects on, for example, firms' accounting and salary systems, recruitment routines, board and lodging for foreign employees, use of foreign workers and subcontractors, tax withholding accounts, salary reporting, and staff registers, if applicable.

The program was thus established to gain knowledge of reporting noncompliance rather than enforcing

tax remittance. However, correct reporting will inevitably entail correct tax remittance, but rather than disclosing the noncompliance as evaded taxes, the program aimed at disclosing errors in the firms' reporting procedures.

The random audit program started with a test pilot in 2017. A total of 60 test audits were performed in this initial phase. These audits are not included in the data used in this paper. Following an evaluation of this pilot, adjustments were made before the start of the program. In its full scope, 187 auditors from the NTA have been involved in executing the audits. Each audit averaged three to five days of work, with an average cost of 12,740 NOK per audit. All 1,974 audits were executed in 2018.

### 3.2 Information Letters

The second treatment in the experiment was information letters. The NTA sent 8,000 randomly selected firms letters containing descriptive information on seven relevant reporting duties for the firms in these sectors but contained no "moral" statements (e.g., Bott et al. 2020). The letter recipients were randomly drawn from the same target population as the audited firms. While the audits were more extensive than the letters, the reporting duties stated in the letters reflected some of the same procedures that the audited firms were asked to provide information about, namely accounting standards, monthly reporting (A-melding), employee tax deductions, documentation of salary expenses, documentation of elapsed time, reporting duties to the International Tax Collection Office, and staff register obligations. A copy of the letter can be found in Appendix 3. The marginal cost of electronically distributing the letters, and the average cost per letter are both negligible.

### 3.3 The Target Population and Sample Selection

The target population for both the audits and the letters is based on the selection criteria in Table 1 and consists of 30,961 firms.

The 8,000 firms receiving letters were randomly drawn from this target population using simple random selection and no stratification or any further selection criteria. With this increased sample size, compared to the audits, the likelihood of obtaining a representative and diverse sample that reflects the overall population

**Table 1:** Population selection criteria

Criteria
1. Total revenue in 2017 > NOK 100,000
2. If not 1, then total revenue in 2016 $\geq$ NOK 200,000.
3. Mean active work relations per month is $\leq$ 20
4. Mean active employees per month is $\geq$ 5
5. If not 4, then registered > NOK 100,000 on subcontracts and/or foreign services in 2016
6. Private limited companies (AS) self-employed (ENK) and Norwegian foreign-registered enterprises (NUF)
7. No termination date registered
8. Sectors in public service, defense and public social security are excluded

*Note:* The table describes the selection criteria for firms in the strata described in Appendix Table A2, which resulted in the 30,961 firms in the population of study. The selection criteria are set to secure representative firm activity (criteria 1–2; 7), employment (3–5) organizational form (6), and exclude the public sector (8).

characteristics increases, and hence stratification becomes less important (Solon, Haider, and Wooldridge 2015). No audited firms received a letter, however. Thus, no firms received two treatments, and there was no interference between treatment groups. The electronic letters were simultaneously sent out through the NTA e-portal in the digital government dialogue (“Altinn”) to all 8,000 firms on 1 June 2018, among which 6,870 firms did open/read the letters. The NTA can read off which firms opened the letters and which firms left them unopened.

Out of 2,000 randomly selected firms, 26 were found “unworthy” of an audit for different, unsystematic reasons. The 1,974 audited firms were selected through a stratified random sample to capture the relative sizes of the different strata in the total population of 30,961 firms. A total number of 22 strata representing different industries were constructed from 65 different NACE codes on 1- and 2-digit levels, cf. Appendix 1. The sample size allocated to each stratum was determined by the method of proportional allocation (based on the number of foreign employees in each industry). The samples were then selected from each industry (stratum) for each of the five tax regions in Norway separately, considering the audit resource capacity in each region. The aim was to perform at least five audits in each stratum in each region, but the number of audits in each stratum could not exceed 15 percent of the total resource capacity in the region (proportional allocation method with lower and upper cut-off). A list of the industry sector and stratum is included in the Appendix.

## 4 Data Description

### 4.1 Descriptive Statistics

The data set covers the period from 2017 to 2020. While the shortcoming of this limited period is of some concern for establishing post-treatment trends, it is evident from other studies that the strongest effects occur immediately or within a couple of years post-treatment (e.g. Advani, Elming, and Shaw 2023; Brockmeyer et al. 2019; Pomeranz 2015), DeBacker et al. (2018) being an exception, finding stronger effects on total income in years 3 and 4 compared to years 1 and 2 post audit. Our data set covers a wider period than, for example, Boning et al. (2020) and Brockmeyer et al. (2019), and an additional post-treatment year compared to Bjørneby, Alstadsæter, and Telle (2021).

To provide an overview of the different sectors’ representation in the sample, the descriptive statistics are broken down by stratum, as displayed in Appendix Table A1, while Table A2 displays which sectors are included within each stratum. Construction of buildings and civil engineering (S26); human health, residential care, and social work (S15); publishing, broadcasting, telecommunications, and information services (S17); food and beverage service activities (S31); and manufacturing (S16) are by far the largest strata in the population concerning the number of firms in each stratum.

Payroll tax represents a significant tax revenue contribution from the firms in our population, where the

total reported contributions vary from NOK 8.5 to 9.8 billion over the years of study. Table 2 shows reported remittance of payroll tax in NOK, and firms' workforce, broken down by population, reference, audit, and letter groups for the years 2017–2020.

The letter group is four times as large as the audit group, and so we expect lower standard errors on the estimated effects of letters than of the audits. The reference group is the same population of firms for both audit and letter arms. The standard deviation is high for all groups, and similar across groups. All groups' payroll tax remittances increased from the pretreatment year 2017 to the treatment year 2018. The decrease in firms' payroll tax remittance in 2020 is an effect of COVID. As expected for a true experimental design, we found no evidence of the Ashenfelter dip when adding a pretreatment year (2016) in the context of any interventions wherein those assigned to either treatment have a temporarily depressed payroll tax remittance that would revert upward toward their longer-term mean, absent treatment (Ashenfelter, 1978), *cf.* Appendix Figure AF1. As expected, due to randomization, the treatment groups are similar before treatment, except that the audit group had a systematically higher number of employees compared to all other groups during the entire period of study. This is due to the stratification of the audit sample, where overrepresented strata had firms with higher numbers of employees, *cf.* t-test statistics with strata fixed effects in Table 3.

## 4.2 Graphical Evidence

A simple way to test the effect of a treatment is to track the means of the outcome variables, pre- and post-treatment, for the treatment and control groups respectively, *cf.* Figure 1.

From 2018 to 2019, all three groups slightly increased their payroll tax remittance compared to the pretreatment year of 2017, and then decrease toward 2020. Thus, on average, we cannot infer that the treatments affect firms' payment of payroll tax. The firms' workforce is also downward sloping from 2017 onwards, with a sharper decline from the treatment year of 2018 for all three groups. On average, we do not see any effect of treatment on the firms' workforce.

As we can infer from both panels, randomization fails on raw data because there are significant differences between the groups, pretreatment. Since the audit sample was stratified, however, we need to control for strata fixed effects when testing randomization.

## 4.3 Proof of Randomization

To test whether the treatment groups and reference group were significantly different in the pre-treatment year of 2017, we ran a simple OLS regression with fixed effects on strata for the audit arm to control for the stratification bias, but no fixed effects for the letter arm because the letter sample selection was not stratified:

$$(1) Y_{ist} = \alpha_i + \beta_t \text{Audit}_{is}^* 2017 + \varepsilon_{ist}$$

$$(2) Y_{it} = \alpha_i + \gamma_t \text{Letter}_i^* 2017 + \varepsilon_{it}$$

$Y_{ist}$  is the log of payroll tax for firm  $i$  at time  $t = 2017$  in stratum  $s$ .  $\beta_t$  and  $\gamma_{ist}$  and  $\gamma_{it}$  are unbiased and consistent estimators for the audit and letter treatments in year 2017,  $\alpha_i$  and  $\eta_t$  are firm constants and  $\varepsilon$  is the error term.

Furthermore, there are reasons for clustering the standard errors, either as the result of the sampling design because we have sampled data from a population using clustered sampling (Liang and Zeger, 1993) or to accommodate the experimental design (Weiss, Lockwood, and McCaffrey 2016) because of the clustered assignment mechanism for the audit treatment. Albeit both reasons are relevant for this study, at this stage of the analysis, we use clustered standard errors adopting the first reason. The test results are displayed in Table 3.

As we can see from the T-statistics in Table 3, randomization holds for both audit and letter arms on payroll tax and Employees.

## 4.4 Pre-treatment Comparison

Because we are also interested in a comparison of treatment effects on the two treatment arms, we have run a two-sample t-test with equal variances on payroll tax and Employees to reveal prospective pre-treatment deviances between the treatment arms on the two variables. The test results are displayed in Table 4.

The audit and letter arms are statistically significantly different on employees, but not on payroll tax.

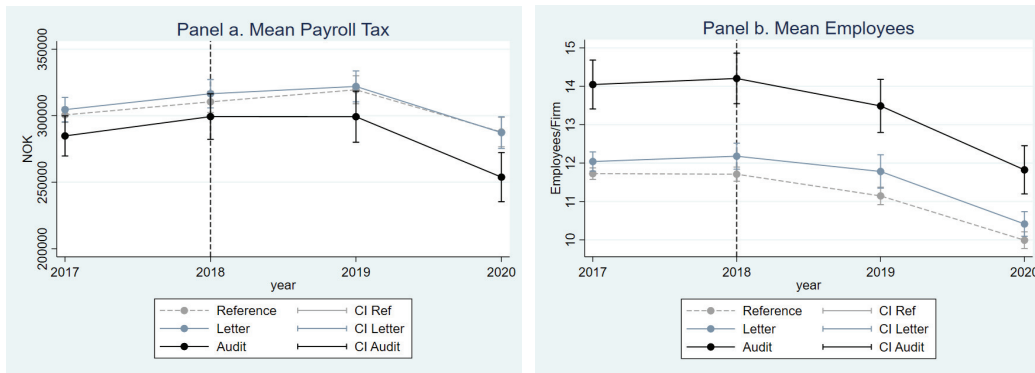
## 5 Experimental Design

In a randomized experiment, where treatments are assigned to test and control groups by random selection, an OLS model with fixed effects would give unbiased estimates. To exploit the panel structure and to adjust for the fact that the two groups have different average levels of compliance in the base year 2017, we use a

**Table 2:** Reported remittance of payroll tax (NOK) and number of employees (2017–2020)

	Year	Payroll Tax			Employees		
		N	Mean	SD	N	Mean	SD
<b>Population</b>	2017	30 869	300 651	416 841	30961	11.95	11.41
	2018	30 813	311 326	550 902	30961	11.99	14.25
	2019	30 807	318 893	702 543	30961	11.46	17.6
	2020	30 667	285 473	733 289	30961	10.22	15.6
	<b>Reference</b>	2017	20 919	300 638	423 045	20987	11.72
	2018	20 893	310 461	586 372	20987	11.71	13.68
	2019	20 884	319 539	776 792	20987	11.15	16.84
	2020	20 786	287 767	815 942	20987	9.99	16.06
<b>Audit Arm</b>	2017	1 970	284 822	342 069	1974	14.05	14.42
	2018	1 964	299 351	388 336	1974	14.21	14.92
	2019	1 966	299 264	434 835	1974	13.49	15.67
	2020	1 964	253 812	417 955	1974	11.82	14.24
	<b>Letter Arm</b>	2017	7 980	304 594	417 213	8000	12.04
2018		7 956	316 552	485 016	8000	12.18	15.47
2019		7 957	322 046	529 603	8000	11.78	19.82
2020		7 917	287 302	539 802	8000	10.41	14.63

Note: Reported remittance of payroll tax in NOK. Column N denotes total number of observations/firms in the respective years given by the first column. Column “Mean” represents sample means, and column SD gives the standard deviation. 735 erroneous, negative values are omitted. Random checks on negative values indicate no evidence that omitted observations are systematic.



**Figure 1:** Means of payroll tax and number of employees. Reference, audit, and letter groups.

Notes: This figure plots point means and 95% confidence intervals from payroll tax remittance (in NOK), and the number of employees for years 2017 to 2020. The treatment, either audits or letters, takes place in year 2018, as visualized by the vertical, dotted line. The black line represents the annual mean of the audit group, the blue line represents the annual mean of the letter group, and the dotted line represents the mean of the reference group. The 95% confidence intervals are represented by the vertical error bars for each group.

difference-in-difference model design (DiD). However, the results are equivalent to an OLS. There are two reasons for this. First, a DiD increases the precision of the estimates (Angrist and Pischke, 2009). Second, we want to estimate LATE for the letter treatment group. As a robustness check, we have run simple OLS fixed effects models. The results are displayed in Appendix Table A3.

The two treatment samples have different properties in three relevant respects. Firstly, the audit sample is stratified, and so we fix the effects on strata for this sample. Secondly, all firms assigned to audit treatment were actually audited. Thus, for the audit treatment group, treatment assignment is identical to the treatment status, and so ITT estimations give us the ATE. Lastly, the letter sample is not stratified but includes

**Table 3:** T-statistics payroll tax and employees in the base year 2017

Arm	Payroll Tax	t	P >  t	Employees	t	P >  t
<b>Audit</b>	3658.342 (6721.756)	0.54	0.592	0.297 (0.321)	0.92	0.366
<b>Letter</b>	5317.589 (3322.063)	1.60	0.124	0.116 (0.190)	0.61	0.548
<b>Observations</b>	30869			30961		
<b>R-squared Audit</b>	0.055			0.124		
<b>R-squared Letter</b>	0.000			0.000		

*Note:* OLS estimation coefficients and test statistics for the audit arm compared to the stratified audit reference arm, and letter arm compared to the unstratified letter reference arm. Estimated coefficients on  $Y = \text{Payroll Tax}$  and  $Y = \text{Employees}$ . Fixed effects on strata for audit arm. Standard errors are adjusted for 22 clusters in strata. Robustness test is run with standard error clusters on firm ID, and the result stands.

**Table 4:** Two sample T-statistics Tax payroll tax and employees in the base year 2017

Arm	Payroll Tax	t	P >  t	Employees	t	P >  t
<b>Audit</b>	284821.6 (342068.8)			14.07107 (14.42238)		
<b>Letter</b>	304594 (417212.9)			12.05564 (11.42717)		
<b>Combined</b>	300679.2 (403508.8)	1.9480	0.0514	12.45467 (12.10517)	-6.6322	0.0

*Note:* Test statistics for Audit (N=1,970) and Letter (N=7,980) arms on  $Y = \text{Payroll Tax}$  and  $Y = \text{Employees}$ . 24 Erroneous, negative values are omitted. Standard deviations in parentheses.

noncompliers, i.e., a known sub-sample of those who received the letter but did not open it. Therefore, they are untreated but nevertheless assigned to the letter treatment sample. We estimate both ATE for this group and LATE for those who were actually treated.

The letter recipients and the audited firms are drawn from the same target population, using simple random sampling for the letters and stratified random sampling for the audits. To ensure that the treatment effects are comparable, we include fixed effects on strata in models including both treatments. This allows for a comparison of treatment effects within each stratum.

## 5.1 ATE of treatments

While evaded tax or changes in reported income tax are commonly used dependent variables (Kotsadam et al. 2021, Bjørneby, Alstadsæter, Telle 2021), we are interested in changes in payroll tax as an expression of

a correction following the treatments. As less reported payroll leads to more profits, more dividends, and then an increase in income, the economic incentive for withholding is present.

Our first model utilizes payroll tax as the dependent variable, with fixed effects on strata, firm ID, and year, to isolate the effects from the treatment:

$$(3) Y_{ist} = \alpha_i + \eta_t + \beta_t \text{Audit}_{is} \times \text{Year} + \gamma_t \text{Letter}_{is} \times \text{Year} + \varepsilon_{ist}$$

$Y_{ist}$  is the log of payroll tax for firm  $i$  at time  $t = 2017 \dots 2020$  in stratum  $s$ .  $\beta_t$  and  $\gamma_t$  are unbiased and consistent estimators for the audit and letter treatments, interacted with years 2017 to 2020,  $\alpha_i$  and  $\eta_t$  are firm and time constants, respectively, and  $\varepsilon$  is the error term. We interact the treatments with year, to estimate the effects of each year, compared with pretreatment year  $t = 2017$ . We also run a model where  $Y_{ist}$  is the log of the workforce for firm  $i$  at time  $t = 2017 \dots 2020$  in stratum  $s$ . Standard errors are clustered on the strata.

The number of clusters is only 22, and so we may underestimate the true standard errors (Angrist and Pischke, 2009). Since a two-sided Wald test would be a poor approximation, we have run the regressions with standard errors clustered on firm ID ( $n = 30,961$ ) and NACE code at a 2-digit level ( $n = 61$ ) as well, without notable deviations, *cf.* Appendix Table A5.

This regression estimates the treatment effect by comparing the changes in outcomes (logged payroll tax remittance and logged number of employees) over time between the treatment group and the reference group. Specifically, it looks at the difference in these changes from the base year (2017) to the subsequent years (2018, 2019, 2020). The DiD approach inherently controls for time-invariant differences between the treatment and reference groups by focusing on the changes over time. The coefficients represent the interaction between the treatment group and the post-treatment periods, capturing the differential effect of the treatment over time. By using 2017 as the base year, the model adjusts for any pre-existing differences between the groups in that year.

There is a difference in effects between the group one intended to treat with the letter and the effects of the group in fact being treated that is, the group opening/reading the letter. The former estimates the effect of being assigned to a treatment group, whereas the latter estimates the effect of being treated. This difference arises from the fact that there may be firms assigned to the treatment group, that end up not getting treated.

To allow for a comparison between ATE and LATE, we run a separate model on the annual effects of the letter treatment compared to the reference group. We include fixed effects on firm ID:

$$(4) Y_{it} = \gamma_t \text{Letter} + \varepsilon_{st}$$

$Y_{it}$  is the log of the payroll tax at time  $t = 2017 \dots 2020$ .  $\beta_t$  and  $\gamma_t$  are unbiased and consistent estimators for the letter treatment in the years 2017–2020, and  $\varepsilon_{st}$  is the error term. Again, we interact treatment year to estimate the effects of each year, compared with pre-treatment year  $t = 2017$ . As for (3), we also run (4) where  $Y_{it}$  is the log of the workforce at time  $t = 2017 \dots 2020$ . Standard errors are clustered on strata.

This approach involves running separate regressions for each year (2017–2020) and comparing the outcomes to the reference group each year. Each regression compares the treatment group to the reference group within the same year, without explicitly accounting for changes over time. Since the regressions are run separately

for each year and group, the coefficients reflect the annual differences between the treatment and reference groups, rather than the changes over time. Unlike the DiD approach, this method does not adjust for pre-existing differences in a baseline year (2017). Each year's coefficient is independent of the others.

## 5.2 LATE of Letter Treatment

Local average treatment effect (LATE) is an estimate that focuses on a specific subgroup within the population, i.e., the compliers. LATE estimates the average effect of the treatment on this subgroup, which is typically smaller than the overall population. It is based on the assumption that there are no unmeasured confounding variables affecting treatment assignment for the compliers. In other words, it assumes that the treatment effect is constant for this subgroup, regardless of the level of treatment received by others. Average treatment effect (ATE), on the other hand, aims to estimate the average effect of the treatment on the entire population, including both compliers and non-compliers. It considers the effect of treatment on all individuals, regardless of whether they fully comply, partially comply, or do not comply with the treatment. For the audits, we have no such differences, since all firms assigned to the audit treatment actually did get audited.

For the letter treatment group, there was one-sided noncompliance, that is, firms who received but didn't open the letters, but no firms who didn't receive the letters and still required/read them. Some of the firms assigned to letter treatment received the letter but never opened them. However, no firms assigned to the reference group received/read the letter. The ATE gives the effects for the whole group intended to get the letters, but there are no reasons to expect effects among firms who never opened the letter, albeit some minor spillovers through informal contacts between firms can occur. To estimate the effect only among firms who were actually treated, that is, the LATE, we use the IV-model described in the following two equations:

$$(5) \text{Letter}_i = \alpha + \beta \text{Treat}_i + u_i$$

$$(6) Y_{it} = \eta + \gamma \text{Letter}_i \text{Year}_{it} + \varepsilon_i$$

$\text{Treat}_i$  is a dummy that takes the value 1 if firm  $i$  was assigned to the letter treatment group, controlled for time,  $t$ , whereas  $\text{Letter}_i$  indicates whether the firm actually opened/read the letter,  $\alpha$  is a constant,  $\beta$  is the letter treatment estimator, and  $u$  is the error term.

Standard errors are clustered in strata. We estimate the parameters  $\beta$  and  $\gamma$  through a two-stage least squares (2SLS) with (5) as the first stage and (6) as the second stage.

Letter assignment is the instrument variable. It affects treatment as only letter recipients may read the letters. It is not associated with the dependent variables (payroll tax and number of employees), however, as firms are randomly assigned to treatment.

Given that the assignment to the letter group was true random on payroll tax, our reduced form ATE estimate (4) can be given causal interpretation of the ATE. The IV estimates (5) and (6) rely on two additional assumptions (Angrist and Pischke (2009); Gerber and Green (2012)).

First, is the noninterference assumption, which consists of two parts. Part A stipulates that whether a firm is treated depends only on the firm’s own treatment group assignment. Because there is no firm receiving a letter outside the letter treatment group, this condition holds. Part B stipulates that potential effects are affected by the firm’s assignment *and* the treatment the firm receives as a consequence of that assignment (Gerber & Green, 2012, p. 138).

Second, the exclusion restriction stipulates that potential effects respond to actual treatments, not treatment assignments. That is, firms respond to the letter only if they open/read it, not by merely receiving it. There are good reasons why we should assume that the exclusion restriction holds because treatment does not affect the outcome variables through any other channels than through the direct reading of the letter. Since communication between the NTA and firms in Altinn is not unusual, it is unlikely that firms might respond by merely receiving a letter from the NTA.

## 6 Results

### 6.1 Main Results

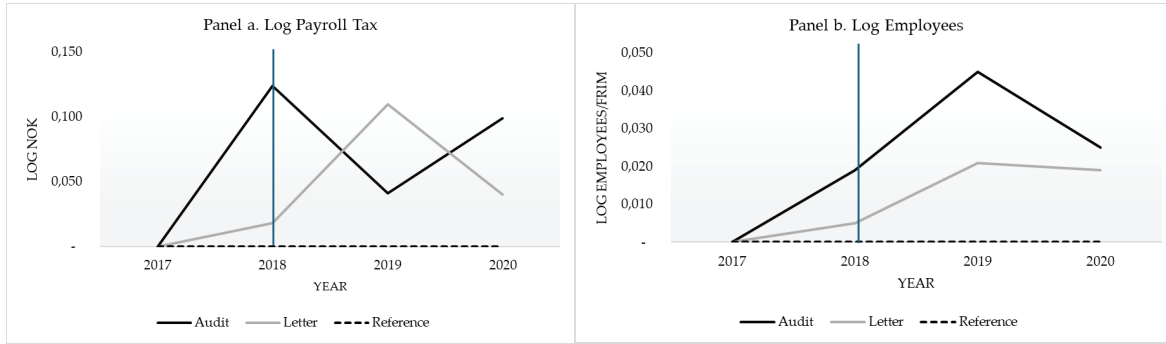
The two treatments may affect a firm’s reporting liabilities other than the payroll tax. Revenue, salary expenses, and number of employees (workforce) may be affected by both the audit and the letter treatments since these are specific items contained in both treatments. We have run all the models in this paper with revenue, salary expenses, and number of employees as dependent variables as well, but we find no major deviations from the main model using payroll tax remittance

as the dependent variable. Audits may affect payroll tax remittance through both better documentation on the payroll tax itself, but also through better documentation of sales and other variables. We are estimating the total effect of all these possible channels, *cf.* correlation matrix in Appendix Table A4. Nevertheless, as we are interested in a comparison with Bjørneby, Alstadsæter, and Telle (2021) on employee effects, we keep the number of employees as a dependent variable alongside the analysis of the payroll tax.

The regression results from our combined model (3) are presented in Figure 2. The results reveal a positive effect of audit and letters compared to the reference group when we use the pretreatment year of 2017 as the base year, but the audit and letter curves have opposite shapes post-treatment.

The audit treatment had an immediate effect in 2018 on firms’ payroll tax remittance compared to the reference group, albeit not statistically significant. In 2018, the audited firms increased payroll tax remittance by .124 log points (equivalent to 13.20 percent), compared to the reference group when we use the pretreatment base year 2017.

Increased remittance following an audit may be considered “mechanical” when firms only adjust their behavior and report accurate information in response to the specific discrepancies or errors uncovered during the audit. The increased remittance may thus not be driven by a genuine compliance commitment or improved understanding of reporting liabilities, but rather a reaction to the fear of penalties or consequences for noncompliance. This identification problem pertains to the distinction between actual behavioral changes in taxpayers (real effects) and mere adjustments made to comply with tax liabilities in response to the audit (reporting effects) (Advani, Eklmng, and Shaw 2023); Kleven et al. 2011). Real effects refer to substantive changes in the behavior of taxpayers resulting from the audit. These changes may include a genuine improvement in tax compliance practices, a better understanding of tax regulations, and a shift towards more accurate and honest reporting of financial information (Kausar, Shroff, and White 2016). Reporting effects, on the other hand, occur solely in response to the audit itself. Firms may correct errors or discrepancies discovered during the audit, but these changes might not reflect any meaningful change in their overall tax compliance behavior beyond the specific issues identified during the audit. Reporting effects are often temporary, and taxpayers may revert to noncompliant behavior once the audit process is over. The immediate effect in



**Figure 2:** Estimated audit and letter effects (base year 2017).

*Notes:* This figure plots point DiD estimates from regressions of measures of logged payroll tax remittance and logged number of employees for the years 2017–2020, compared to the base year 2017, and compared to the reference group. (Difference between treatment and reference in years 2018, 2019, 2020, minus the difference between treatment and reference in 2017). The treatment, either audits or letters, take place in year 2018, visualized by the vertical line. The black line represents the estimated coefficients of the audit group, the grey line represents the estimated coefficients of the letter group, both estimations are compared to the reference group represented by the dotted line. The specification includes fixed effects on strata and firm ID. Standard errors are adjusted for 22 clusters in strata. Table A5 in the Appendix displays the coefficients and their standard errors.

2018 and the slight dip in 2019 may indicate reporting effects, but the following increase in 2020 makes this difficult to establish.

The effect of letters on payroll tax remittance is .018 log points (equivalent to 1.82 percent) in 2018, but neither is this result statistically significant. The audited firms also increased their workforce in 2018 by .019 log points (equivalent to 1.92 percent), compared to the reference group, and the letter group increased their workforce by .005 (equivalent to .50 percent) in the treatment year 2018, but these effects are not statistically significant. We see the strongest effects on payroll tax remittance of audits in 2018, decreasing in 2019, and then increasing in 2020. The letter treatment follows an opposite pattern, where we find the strongest and statistically significant effects in 2019, and then decreasing in 2020. In 2019, the effect of letters on payroll tax remittance is .110 log points (equivalent to 11.63 percent), which is closer to the 2018 estimate for the audited firms. Thus, there appears to be a lag in the letter treatment effect, that is not apparent in the audit treatment.

The time-lagged compliance effect from the information letter, as compared to the immediate effect from the on-site audits, can be attributed to the different nature of these two compliance interventions and how they influence taxpayer behavior over time. The letters are nonbinding and serve as educational tools to inform firms about their compliance obligations. Unlike the audits, the letters do not involve direct enforcement actions or a perceived threat of penalties for noncompliance. As a result, firms may take longer to

internalize the information and voluntarily adjust their behavior. Behavioral change takes time, regardless of the method used to convey information. After receiving an information letter, firms may need time to process the information, assess its implications, and gradually adopt better practices.

An on-site audit typically leads to more immediate compliance adjustments due to the direct scrutiny and enforcement actions involved. The fear of penalties and the immediate presence of auditors can compel firms to address identified compliance issues more promptly.

The treatment effect on the workforce is consistent with the findings of Bjørneby, Alstadsæter, and Telle (2021), who estimate an average increase of 1.12 in the number of employees between treated and untreated firms. The stronger effect from the letter treatment compared to the audit treatment on payroll tax remittance in 2019, can be seen in light of a similar finding in D’Agosto et al. (2018), which reveals a higher effect from the “soft” on-site audit compared to the “deep” desk-based audit treatments in their study. But as our treatments differ in nature, i.e., our audit treatment resembles their on-site audits rather than their desk-based audits, a direct comparison should be considered with caution.

The letter treatment has significant effects on both outcome variables in the first post-treatment year of 2019. Letters affect payroll tax and workforce less than audits two years post-treatment. This confirms the findings of Boning et al. (2020), where letters conveying the same message as on-site visits by the revenue officers in the U.S. have smaller direct effects.

The same pattern is evident in Ortega and Scartascini (2020) who find stronger effects from on-site visits than emails. Although Kotsadam et al. (2021) find a similar response of letter treatment (diminishing with time), the results are not directly comparable; their unit of study is the individual, not the firm, and the letters are more specifically addressing the actual source of error, namely deductions on the tax return. Doerrenberg and Schmitz (2015) also suggest evidence that a letter which reminds small firms of the civic duty to pay taxes and informs about an audit probability following the letter, may increase tax compliance, but their results are not statistically significant.

The over-compliance effects as discussed in, for example, Slemrod Blumenthal, and Christian (2001) and Slemrod and Yitzhaki (2002) cannot be entirely ruled out in our context. When a firm anticipates a potential audit, it may opt to overstate its tax liabilities, and this strategy is often used as a safeguard against the risks of underreporting. However, in a third-party tax reporting regime involving 12 annual report submissions (A-melding), this potential effect seems limited. The regularity of monthly declarations increases the visibility of the firm's financial activities. This frequent oversight encourages more accurate reporting, as discrepancies or patterns of over-reporting are more easily detected over short intervals. Over-reporting taxes consistently across multiple declarations would increase the firm's administrative workload, requiring more effort to track and justify inflated figures. The regular correction process in the next monthly reporting increases transparency and seems to discourage firms from over-compliance due to an awareness of increased audit probability.

## 6.2 Annual Letter Effects

To allow for a comparison between ATE and LATE for the letter treatment, we have run a separate annual regression model (4). The results are displayed in Figure 3.

Receiving a letter would increase the average firm's payroll tax remittance by .127 log points more than the nontreated firms in the treatment year 2018 (equivalent to 13.54 percent). The effect is stronger one year post-treatment (2019) than two years post-treatment (2020). In the post-treatment years 2019–2020, the treatment effect on employees is .039 and .055, respectively (equivalent to 3.98 and 5.65 percent).

The differences in coefficients between the DiD model and the annual regressions are typically due

to temporal dynamics, interaction effects, and time-invariant factors. Whereas the DiD model captures the dynamic changes over time, the annual regressions provide a snapshot comparison for each year. This implies that the DiD coefficients reflect the cumulative effect of the treatment over multiple years, whereas the annual regressions show the effect within each specific year. Furthermore, the DiD model includes interaction terms between the treatment and time, capturing how the treatment effect evolves. Annual regressions do not include these interaction terms. Finally, DiD inherently controls for time-invariant, unobserved heterogeneity between the treatment and reference groups, which may also explain the difference in coefficient estimates compared to annual regressions that do not control for these factors. Nevertheless, the focal point here is to quantify the difference between those who actually read the letters and those who did not.

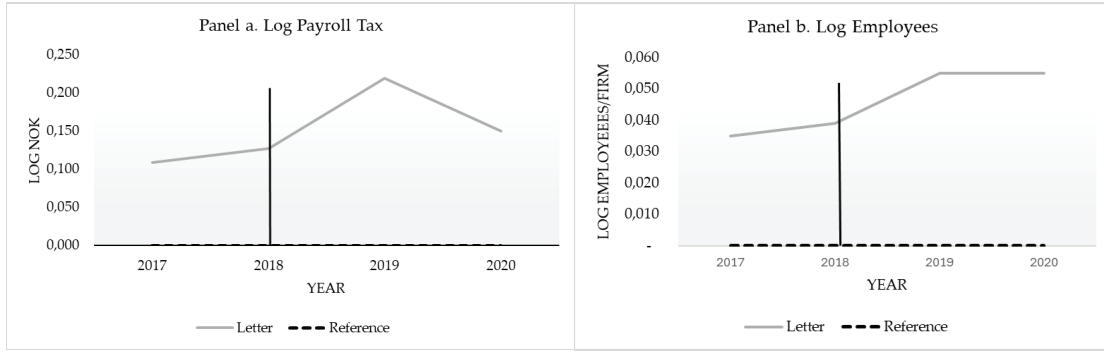
## 6.3 Effects of Reading the Letter (LATE)

To get a clearer picture of the letter treatment effects, we estimate the LATE. Of the 8,000 firms who received a letter, 1,130 never opened/read it. Thus, the ATE estimates presented above may be biased, since just above 14 percent of the letter treatment group was in fact never treated. While the ATE estimates measure the effects of receiving a letter, the LATE estimates measure the effects of actually reading the letter. This is equivalent to the set-up in Børneby, Alstadsæter, and Telle (2021), except that we deal only with one-sided noncompliance, as there were no firms in the reference group receiving a letter.

The regression results, presented in Figure 4, reveal a positive effect of reading the letters (compliers) compared to the reference group.

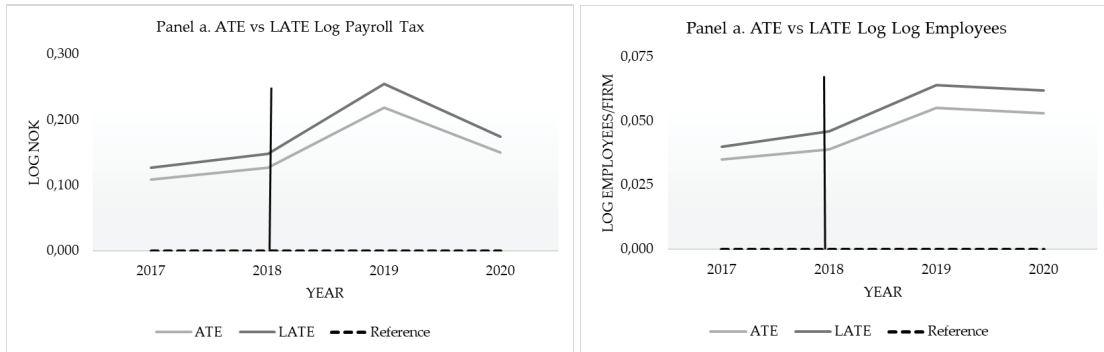
Both in the treatment year 2018, and the post-treatment year 2019, we observe that the compliers report significantly higher payroll tax (.148 and .255 log points, equaling 16.0 and 29.0 percent, respectively) and number of employees (.046 and .064 log points equaling 4.7 and 6.6 percent, respectively) compared to the reference group. The effects seem to remain for two years post-treatment (2020). Furthermore, the statistically significant estimates may indicate that being assigned to the letter treatment is a strong instrument of actually being treated.

The significant difference between the LATE and the ATE estimates tells us that there may be further gains by taking low-cost measures to facilitate letter reading, such as an automated reminder.



**Figure 3:** Estimated annual letter effects (ATE) 2017–2020.

*Notes:* This figure plots point estimates from annual letter regressions of measures of logged payroll tax remittance and logged number of employees for the years 2017–2020, compared to the reference group (x-line). The treatment takes place in the year 2018, visualized by the vertical line. The grey line represents the estimated coefficients of the letter group compared to the reference group, represented by the dotted line. Standard errors are adjusted for 22 clusters in strata. Table A6 in the Appendix displays the coefficients and their standard errors.



**Figure 4:** Estimated LATE of letter treatment (base year 2017).

*Notes:* This figure plots annual point IV estimates from regressions of measures of logged payroll tax remittance and logged number of employees for the years 2017–2020, compared to the reference group (x-line). The treatment takes place in year 2018, visualized by the vertical line. The dark grey line represents the estimated coefficients of the 6,870 letter openers (LATE), and the light grey line represents the estimated coefficients of the 8,000 assigned to the letter group. Standard errors are adjusted for 22 clusters in strata. Table A7 in the Appendix displays the coefficients and their standard errors.

### 6.4 Heterogeneous Effects

In the second chapter of this thesis, we found significant differences in compliance effects between firm types. To study prospective heterogeneous treatment effects, we have run all models on samples restricted to the three firm types represented in the population, namely private limited liability company (private Ltd), self-employed and Norwegian-registered foreign company. By and large, private Ltds appear to drive the main results. The significance levels of estimated coefficients are higher on this subsample, which is over 80 percent of the firms in the population, *cf.* Appendix Table A9. This heterogeneity resembles the results of Almunia and Lopez-Rodriguez (2018) who also find that the impact of monitoring on tax compliance varies across different types of firms. Larger and more profitable firms are more

responsive to increased monitoring, whereas smaller and less profitable firms show a weaker response. Brockmeyer et al. (2019) find an ambiguous heterogeneous effect between corporations and the self-employed in that the filing rate of corporations responds less strongly but their payment rate responds more strongly to the treatment compared to the self-employed. Pomeranz (2015) finds a stronger response on VAT compliance for smaller firms, following a letter providing information about a random audit selection among 400,000 Chilean firms.

We find some small significant negative treatment effects among the self-employed, *cf.* Appendix Table A10. This is contrary to the findings of Mittone, Ramachandran (2017). They suggest that tax audits have a significant short-term impact on increasing compliance. Their treatment group tend to immediately

adjust their behavior and report more accurately to address the specific issues identified during the audit. However, the study also reveals a gradual decline in tax compliance levels after the initial surge following the audit. Over time, the fear of an audit diminishes, and taxpayers may return to their previous, less compliant behavior or adopt tax avoidance strategies to reduce the risk of future audits, known as “the bomb-crater effect” (Mittone, Panebianco, and Santoro 2017). As our significance levels are low, and not reproduced for the letter treatment, this result should be treated with caution. Baer and Silvani (2020) find both pro-deterrent and counter-deterrent effects on future reporting behavior of audits among the self-employed. They suggest that the observed reduction in reported income among self-employed U.S. taxpayers may be associated with dishonesty caused by undetected misreporting during the audit. According to their study, such taxpayers may infer that audits are ineffective and cause the self-employed to understate their income even more aggressively in subsequent years. Since the sub-sample of self-employed persons is limited in our experiment, and the significant negative treatment effect is found on the reported number of employees, not on payroll tax remittance, we cannot infer that the mechanism suggested by Baer and Silvani (2020) is at play in our environment.

We find no significant effects among Norwegian registered foreign companies (*cf.* Appendix Table A11), which is also expected since this company type is less than 1 percent of the total sample.

We have also run all models on samples restricted to firm size in the lower and upper 25th percentiles on firm revenue. We find no significant effects on payroll tax nor employees from either treatment on firms in the lower and upper 25th percentiles measured by firm revenue, *cf.* Appendix Table A8. Using number of employees as a proxy for firm size, Bergolo et al. (2023) finds no substantial or statistically significant difference between firms below and above the median number of employees.

## 7 Policy Implications

These findings may inform resource allocation decisions. We study only the direct effects on payroll tax remittance and the workforce of the firms and assume few or small network effects as studied by, for example, Boning et al. (2020). As only payroll taxes are studied, the benefit estimates should be considered as the lower boundaries.

Any treatment would increase net revenue if the marginal revenue it raises exceeds its marginal administrative costs. In this paper, we limit the revenue component to payroll tax only, even if there may be other benefits, such as network effects, and a general deterrent effect in the overall population of firms. The revenue raised should be compared to the marginal administrative cost of the treatments. The equation which must be true for implementing either treatment is thus:

$$(7) R_A - C_A > 0$$

$$(8) R_L - C_L > 0$$

R is revenue (payroll tax) from either two treatments, audit or letter, denoted by subscripts A and L, respectively, and C is the marginal, administrative cost. The “back of the envelope” calculations are based on the estimated effects from DiD model (3) in 2018 and are displayed in Table 5.

An audit will increase payroll tax remittance by 13.20 percent more than the reference group, whereas the corresponding figure for the letter treatment is 1.92 percent. These changes are due to the treatment. On average, a firm in the audit group increased its payroll tax remittance from 2017 to 2018 by NOK 14,529, whereas the figure for a firm in the letter group was NOK 11,958. By comparison, a firm in the reference group increased its payroll tax remittance by NOK 9,823. Thus, the difference in payroll tax remittance between an audited firm and the reference group is NOK 4,706 on average, and the corresponding figure for a firm receiving a letter is NOK 2,135. If we apply the estimated coefficients to these figures, NOK 1,918 of the payroll tax increment is due to the audit and NOK 230 of the increment is due to the letter. Hence, the administrative cost of the audit exceeds the revenue it generates, and equation (7) is rejected, whereas the opposite is true for the letter treatment, and equation (8) holds.

Some caveats remain, however: First, it is too early to draw any conclusions on the long-term effects of both treatments, and previous literature suggests either declining effects over time (Boning et al., 2020) or mixed long-term effects (Bott et al., 2020). Second, there may also be other effects from both treatments, like general deterrence effects, which may increase the benefits. Third, even without such a deterrence effect, the letters can still help guide the firms that misreport due to honest mistakes, but the additional information will do very little to reduce tax evasion. Fourth, the rejection of (7) might also reflect that randomized

Table 5: Cost-benefit analysis

	Audit (A)	Letter (L)	Reference (R)	A-R	L-R
<b>Marginal cost (MC)</b>	12 740	0	0	12 740	0
$\beta$ (%) 2018	0.1320	0.0192	0	0.1320	0.0192
$\Delta$ (NOK) 2017–2018	14 529	11 958	9 823	4 706	2 135
$\beta^*\Delta =$ Marginal revenue (MR)	1 918	230	0	1 918	230
<b>MR-MC</b>	-10 822	230	0	-10 822	230

Note: Marginal cost of an audit is NOK 12,740.  $\beta$  is the estimated coefficient (%) in year 2018 from models (2) and (3), respectively.  $\Delta$  is the increment in NOK from 2017 to 2018. A-R gives the difference in NOK between audit and reference, and L-R gives the difference in NOK between letter and reference.

audits are wasteful if the objective of those audits is increased compliance measured by revenue on the firm level, which is, however, not the case. Randomized audits are used by tax administrations for objectives *other* than noncompliance disclosure on the firm level, like disclosing new areas of noncompliance or building datasets for predictive modelling (Alm 2019a; Micci-Barreca and Ramachandran 2004). Finally, an electronic letter is easily scalable to a larger population, which, all else being equal, will increase the revenue further.

## 8 Concluding Remarks

There is a need to robustly evaluate different enforcement strategies' effectiveness to choose the most efficient and cost-effective enforcement strategy. Tax authorities may save scarce resources by switching from hard to soft interventions. The main contribution of this paper is a documentation of the firm's response to two interventions in an experimental setting involving the same population of firms. We utilized two randomized experiments, one with stratified on-site audits and one using electronic letters. We demonstrate unbiased, positive effects on firms' remittance of payroll tax from both audits and letters, compared to a reference group with no treatment. While audits have stronger effects than letters, the former is by far the most expensive enforcement strategy.

The results are specific to a population of labor-intensive firms in a relatively advanced tax-reporting environment by international standards. Thus, the external validity of the results is not restricted to the Norwegian setting as such, as one can assume that these effects will be reproduced in several advanced, Western tax jurisdictions with a gold standard system of information reporting. Furthermore, the results suggest

that tax authorities may test and compare these two enforcement strategies in other sectors as well.

A cost-benefit assessment including estimations of deterrence effects and other prospective treatment benefits is necessary to make clear recommendations on which enforcement strategy to use. Such an assessment should still be a core priority of the modern tax administration, in order to increase the effects of limited enforcement resources.

## References

- Advani, A., W. Elming, and J. Shaw. 2023. "The dynamic effects of tax audits." *Review of Economics and Statistics* (2023): 1–45.
- Allingham, Michael G., and Agnar Sandmo. 1972. "Income Tax Evasion: A theoretical analysis." *Journal of Public Economics* 1.3–4 (1972): 323–38.
- Alm, James (2019a). "Tax compliance and administration." In *Handbook on Taxation*, 741–68. Milton Park, Abingdon, Oxfordshire: Routledge.
- Alm, James (2019b). "What motivates tax compliance?" *Journal of Economic Surveys* 33(2): 353–88.
- Almunia, Miguel, and David Lopez-Rodriguez. 2018. "Under the Radar: The Effects of Monitoring Firms on Tax Compliance." *American Economic Journal: Economic Policy* 10(1): 1–38.
- Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, New Jersey: Princeton University Press.
- Ashenfelter, Orley. 1978. "Estimating the Effect of Training Programs on Earnings." *The Review of Economics and Statistics* (1978): 47–57.
- Baer, Katherine, and Carlos Silvani. 1997. *Designing a Tax Administration Reform Strategy: Experiences and Guidelines*. Washington, DC: International Monetary Fund.
- Baer, Katherine, and Carlos Silvani. 2020. "Do Audits Deter or Provoke Future Tax Noncompliance? Evidence on Self-employed Taxpayers." *CESifo Economic Studies* 66(3): 248–64.
- Bergolo, Marcelo, R. Ceni, G. Cruces, M. Giacobasso, and R. Perez-Truglia. 2023. "Tax Audits as Scarecrows: Evidence from a Large-scale Field Experiment." *American Economic Journal: Economic Policy* 15(1): 110–53.

- Bjørneby, Marie, Annette Alstadsæter, and Kjetil Telle. 2021. "Limits to Third-party Reporting: Evidence from a Randomized Field Experiment in Norway." *Journal of Public Economics*, 203: 104512.
- Boning, William C., J. Guyton, R. Hodge, and J. Slemrod. 2020. "Heard It Through the Grapevine: The Direct and Network Effects of a Tax Enforcement Field Experiment on Firms." *Journal of Public Economics* 190: 104261.
- Bott, Kristina M., et al. 2020. "You've Got Mail: A Randomized Field Experiment on Tax Evasion." *Management Science* 66(7): 2801–19.
- Brockmeyer, Anne, et al. 2019. "Casting a Wider Tax Net: Experimental Evidence from Costa Rica." *American Economic Journal: Economic Policy* 11(3): 55–87.
- D'Agosto, Elena, et al. 2018. "The Effect of Audit Activity on Tax Declaration: Evidence on Small Businesses in Italy." *Public Finance Review* 46(1): 29–57.
- De Neve, Jan-Emmanuel, et al. (2021). "How to Improve Tax Compliance? Evidence from Population-wide Experiments in Belgium." *Journal of Political Economy* 129(5), 1425–63.
- DeBacker, Jason, et al. (2018). "Once Bitten, Twice Shy? The Lasting Impact of Enforcement on Tax Compliance." *The Journal of Law and Economics* 61(1): 1–35.
- Dhami, Sanjit, and Ali Al-Nowaihi. 2007. "Why Do People Pay Taxes? Prospect Theory Versus Expected Utility Theory." *Journal of Economic Behaviour & Organization* 64(1): 171–92.
- Doerrenberg, Philipp, and Jan Schmitz. 2015. "Tax Compliance and Information Provision—A Field Experiment with Small Firms." *ZEW-Centre for European Economic Research Discussion Paper* (15-028).
- Fellner, Gerlinde, Rupert Sausgruber, and Christian Traxler. 2013. "Testing Enforcement Strategies in the Field: Threat, Moral Appeal and Social Information." *Journal of the European Economic Association* 11(3): 634–660.
- Gerber, Alan S., and Donald P. Green. 2012. *Field Experiments: Design, Analysis, and Interpretation*. New York: WW Norton.
- Hallsworth, M., List, J. A., Metcalfe, R. D., & Vlaev, I. 2017. "The Behaviouralist as Tax Collector: Using Natural Field Experiments to Enhance Tax Compliance." *Journal of Public Economics* 148: 14–31.
- Hebous, Shafik, et al. 2023. "Do audits improve future tax compliance in the absence of penalties? Evidence from Random Audits in Norway." *Journal of Economic Behavior and Organization* 207: 305–26.
- Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision Under Risk." *Econometrica* 47(2): 263–291.
- Kausar, Asad, Nemit Shroff, and Hal White. (2016). "Real Effects of the Audit Choice." *Journal of Accounting and Economics* 62(1): 157–81.
- Keen, Michael, and Joel Slemrod. 2017. "Optimal Tax Administration." *Journal of Public Economics* 152: 133–142.
- Kleven, Henrik Jacobsen, et al. 2011. "Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark." 79(3): 651–92. <https://doi.org/10.3982/ecta9113>
- Kotsadam, Andreas, et al. 2021. "Does Perceived Risk of Future Audits Explain the Behavioral Effects of Tax Enforcement." 12th Norwegian German Seminar on Public Economics.
- Liang, Kung-Yee, and Scott L. Zeger. 1993. "Regression Analysis for Correlated Data." *Annual Review of Public Health* 14(1): 43–68.
- Løyland, Knut, Oddbjorn Raaum, Gaute Torsvik, and Arnstein Øvrum. 2019. "Compliance effects of risk-based tax audits." Løyland, Knut, Oddbjorn Raaum, Gaute Torsvik, and Arnstein Øvrum. No. 7616. CESifo Working Papers.
- Luttmer, Erzo FP, and Monica Singhal. 2014. "Tax Morale." *Journal of Economic Perspectives* 28(4): 149–168.
- Ministry of Finance. 2022. "Taxes 2022. Prop. 1 LS (2021–2022) Proposition to the Storting (bill and draft resolution) For the Fiscal Year 2022."
- Meiselman, Ben S. 2018. "Ghostbusting in Detroit: Evidence on Nonfilers from a Controlled Field Experiment." *Journal of Public Economics* 158: 180–93.
- Micci-Barreca, Daniele, and Satheesh Ramachandran. 2004. "Improving Tax Administration with Data Mining." *White paper. Elite Analytics LLC*.
- Mittone, Luigi, Fabrizio Panebianco, and Alessandro Santoro. 2017. "The Bomb-crater Effect of Tax Audits: Beyond the Misperception of Chance." *Journal of Economic Psychology* 61: 225–43.
- Murphy, Kristina. 2019. *Moving Towards a More Effective Model of Regulatory Enforcement in the Australian Taxation Office*. Centre for Tax System Integrity (CTSI), Research School of Social Sciences.
- Ortega, Daniel, and Carlos Scartascini. 2020. "Don't Blame the Messenger. The Delivery Method of a Message Matters." *Journal of Economic Behaviour & Organization* 170, 286–300.
- Pomeranz, Dina. 2015. "No Taxation without Information: Deterrence and Self-enforcement in the Value-added Tax." *American Economic Review* 105(8): 2539–69.
- Pomeranz, Dina, and José Vila-Belda. "Taking State-capacity Research to the Field: Insights from Collaborations with Tax Authorities." *Annual Review of Economics* 11: 755–81.
- Slemrod, Joel, Marsha Blumenthal, and Charles Christian. 2001. "Taxpayer Response to an Increased Probability of Audit: Evidence from a Controlled Experiment in Minnesota." *Journal of Public Economics* 79(3): 455–83.
- Slemrod, Joel, and Shlomo Yitzhaki. 2002. "Tax Avoidance, Evasion, and Administration." In *Handbook of Public Economics*, Vol. 3, 1423–70. Amsterdam: Elsevier.
- Slemrod, Joel. 2019. "Tax Compliance and Enforcement." *Journal of Economic Literature* 57(4): 904–54.
- Solon, Gary, Steven J. Haider, and Jeffrey M. Wooldridge. 2015. "What Are We Weighting for?" *Journal of Human Resources* 50(2), 301–316.
- Weiss, Michael J., J. R. Lockwood, and Daniel F. McCaffrey. 2016. "Estimating the Standard Error of the Impact Estimator in Individually Randomized Trials with Clustering." *Journal of Research on Educational Effectiveness* 9(3).

## 9 Appendix

### 9.1 Population

**Table A1:** Descriptive statistics by stratum. Base year 2017

Strata	N	Payroll tax ( $\mu$ )	Employees ( $\mu$ )	SD	Audit (n)	SD	Letter (n)	SD
S1	199	479 174	429 103	11	8	22	51	41
S11	505	110 167	270 168	13	10	34	113	89
S12	1 904	328 739	509 941	14	16	148	472	415
S13	1 097	256 028	173 401	11	6	46	297	261
S15	3 288	255 995	241 945	10	10	67	877	757
S16	2 372	399 841	523 591	12	9	89	637	567
S17	3 065	447 220	637 986	10	12	76	861	746
S19	277	96 385	140 913	17	22	124	47	35
S20	753	147 690	182 643	14	18	53	177	159
S21	1 701	323 216	310 702	11	9	91	425	359
S23	875	342 038	321 279	10	8	91	246	214
S26	5 484	301 787	347 375	9	8	300	1 395	1 211
S27	368	257 976	248 780	25	14	92	79	73
S29	621	243 176	238 360	14	10	132	122	106
S3	1 809	104 121	120 253	10	9	32	505	399
S31	2 815	233 333	197 782	23	13	302	681	564
S33	775	222 703	258 052	8	7	91	203	173
S34	948	493 216	997 972	11	9	48	261	230
S37	146	500 468	669 059	13	10	20	34	34
S38	1 359	342 466	305 586	10	7	77	359	313
S5	151	453 951	607 734	11	10	20	39	36
S6	449	246 548	201 366	11	5	19	119	88

*Notes:* Column “Strata” reflects the strata description in Table A2. Column ‘N’ reflects the number of observations in each stratum. Column “Payroll tax ( $\mu$ )” reflects the sample means of payroll tax (NOK) in each stratum, and column “Employees ( $\mu$ )” reflects the sample mean number of employees in each stratum. Columns “Audit (n)” and “Letter (n)” reflect the number of firms in each stratum receiving audit and letter in treatment year 2018. Column “SD” is standard deviation.

**Table A2:** Stratum and NACE codes

Stratum	NACE 2	Description	N
S1	0	Missing NACE	31
S1	43	Specialized construction activities	144
S1	71	Architectural and engineering activities; technical testing and analysis	21
S1	90	Creative arts and entertainment activities	3
S11	3	Fishing and aquaculture	505
S12	77	Rental and leasing activities	102
S12	78	Employment activities	407
S12	79	Travel agency tour operator and other reservation service and related activities	231

Table A2: Stratum and NACE codes

Stratum	NACE 2	Description	N
S12	80	Security and investigation activities	104
S12	81	Services to buildings and landscape activities	480
S12	82	Office administrative office support and other business support activities	580
S13	96	Other personal service activities	1097
S15	86	Human health activities	2544
S15	87	Residential care activities	28
S15	88	Social work activities without accommodation	716
S16	10	Manufacture of food products	204
S16	13	Manufacture of textiles	59
S16	14	Manufacture of wearing apparel	33
S16	15	Manufacture of leather and related products	6
S16	16	Manufacture of wood and of products of wood and cork except furniture; manufacture of articles of straw and plaiting materials	228
S16	18	Printing and reproduction of recorded media	255
S16	20	Manufacture of chemicals and chemical products	2
S16	21	Manufacture of basic pharmaceutical products and pharmaceutical preparations	5
S16	22	Manufacture of rubber and plastic products	8
S16	23	Manufacture of other nonmetallic mineral products	60
S16	24	Manufacture of basic metals	29
S16	25	Manufacture of fabricated metal products, except machinery and equipment	591
S16	26	Manufacture of computer electronic and optical products	81
S16	27	Manufacture of electrical equipment	55
S16	28	Manufacture of machinery and equipment n.e.c.	152
S16	30	Manufacture of other transport equipment	27
S16	31	Manufacture of furniture	116
S16	32	Other manufacturing	195
S16	33	Repair and installation of machinery and equipment	266
S17	58	Publishing activities	640
S17	59	Motion picture video and television program production sound recording and music publishing activities	421
S17	60	Programming and broadcasting activities	29
S17	61	Telecommunications	87
S17	62	Computer programming consultancy and related activities	1665
S17	63	Information service activities	223
S19	1	Crop and animal production hunting and related service activities	277
S20	90	Creative arts and entertainment activities	480
S20	91	Libraries archives museums and other cultural activities	14
S20	93	Sports activities and amusement and recreation activities	259
S21	49	Land transport and transport via pipelines	1701
S23	43	Specialized construction activities	875
S26	41	Construction of buildings	5182
S26	42	Civil engineering	302
S27	55	Accommodation	368
S29	81	Services to buildings and landscape activities	621
S3	49	Land transport and transport via pipelines	1809
S31	56	Food and beverage service activities	2815

**Table A2:** Stratum and NACE codes

Stratum	NACE 2	Description	N
S33	43	Specialized construction activities	775
S34	50	Water transport	376
S34	51	Air transport	30
S34	52	Warehousing and support activities for transportation	404
S34	53	Postal and courier activities	138
S37	37	Sewerage	45
S37	38	Waste collection treatment and disposal activities; materials recovery	97
S37	39	Remediation activities and other waste management services	4
S38	45	Wholesale and retail trade and repair of motor vehicles and motorcycles	1359
S5	6	Extraction of crude petroleum and natural gas	19
S5	7	Mining of metal ores	2
S5	8	Other mining and quarrying	55
S5	9	Mining support service activities	75
S6	47	Retail trade, except for motor vehicles and motorcycles	449

## 9.2 OLS design fixed effects model results

**Table A3:** OLS design fixed effects model results

Arm	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	2017	2018	2019	2020	2017	2018	2019	2020
Audit	0.198+ (0.114)	0.386* (0.185)	0.337+ (0.168)	0.399* (0.179)	0.000 (0.022)	0.034 (0.031)	0.079* (0.037)	0.082* (0.038)
Letter	0.095* (0.039)	0.114* (0.041)	0.206* (0.075)	0.136+ (0.069)	0.038** (0.011)	0.043** (0.011)	0.058*** (0.013)	0.055*** (0.011)
Observations	30961	30961	30961	30961	30961	30961	30961	30961
R-squared	0.045	0.030	0.026	0.026	0.112	0.089	0.058	0.042

*Note:* Columns (1)–(4) are Log Payroll Tax, Columns (5)–(8) are Log Employees. Fixed effects on strata. Standard errors in parenthesis. Standard Errors are adjusted for 22 clusters in strata. The coefficients are to be interpreted in log points and compared to a reference group with no treatment. +  $p < 0.10$ ; \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ .

## 9.3 Information Letter (English Translation)

### Information on enterprises' reporting obligations

The Tax Administration is working on making it as simple as possible to report correctly. Below, you can read about the most important obligations you have when running a business. You can find more information on each of the obligations at <http://www.skatteetaten.no/rapporteringsplikt>.

## **Obligation to report**

### **Accounting**

Most enterprises must keep accounts according to Norwegian bookkeeping legislation.

### **A-melding**

Enterprises with salaried employees have, among other things, an obligation to report information about the employee's employment. This can be information about when they start, any changes in their employment relationship and when the employment is ended. In addition, they must report salary and other remuneration in the a-melding, along with information about withholding tax and other deductions made upon payment. The a-melding must be submitted every month.

### **Tax deductions**

Enterprises must make tax deductions from salary payments and retrieve their employees' tax deduction cards. Deducted tax must be deposited in a separate bank account by no later than the first working day after the salary payment.

### **Documentation of salary and other benefits subject to a reporting obligation**

The documentation of salary and other benefits subject to a reporting obligation must show remunerations per employee. The following information must follow from the documentation: National identity number, name and position, tax municipality, table number, and/or the deduction percentage as stated on the tax deduction card.

### **Documentation of accrued time**

Enterprises performing services where the payment is based on accrued time must document completed hours. The hours must be specified for each employee per day. The documentation must be finished by the end of the month when the work has been performed.

### **Reporting to the Central Tax Office – Foreign Tax Affairs (SFU)**

All assignments and any subcontracts that are given to foreign contractors must be reported to the Central Tax Office – Foreign Tax Affairs (SFU). The condition is that the assignment has been performed at a place for building assembly work in Norway, at a site that is under the control of the client in Norway or on the Norwegian continental shelf.

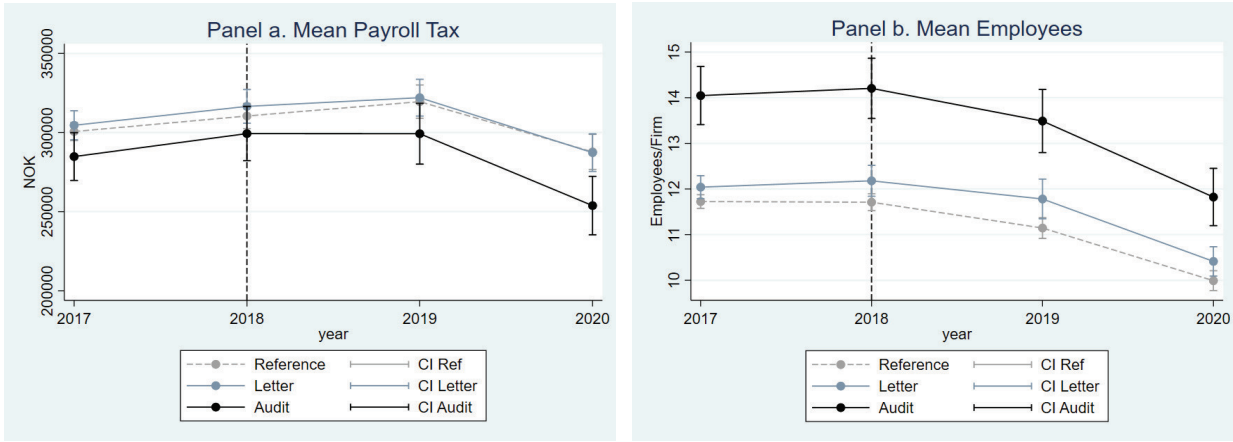
### **Staff register**

Catering establishments, hairdressers, beauticians, car repair workshops and businesses in the car care sector must keep a staff register. A staff register is a list of everyone who works for the enterprise. The register must show when the employees start and finish their workday. The register must be kept at the workplace and be available for inspection during the opening hours of the business.

Yours sincerely,

The Norwegian Tax Administration

### 9.4 Pre-treatment Trend, 2016 Included



**Figure AF1:** Means of payroll tax and number of employees. Reference, Audit and Letter groups.

*Notes:* This figure plots point means and 95% confidence intervals from payroll tax remittance (in NOK), and the number of employees by years 2016 to 2020. The treatment, either audits or letters, take place in year 2018, visualized by the vertical dotted line. The black line represents the annual mean of the audit group, the blue line represents the annual mean of the letter group, and the dotted line represents the mean of the reference group. The 95% confidence intervals are represented by the vertical error bars for each group.

### 9.5 Correlation between Dependent Variables

**Table A4:** Correlation matrix dependent variables

	Log Payroll Tax	Log Revenue	Log Employees	Log Salary
Log Payroll tax	1.000			
Log Revenue	0.613	1.000		
Log Employees	0.495	0.395	1.000	
Log Salary	0.909	0.644	0.530	1.000

## 9.6 Regression Results Tables

**Table A5:** Estimated audit and letter (ATE) effects (base year 2017)

Treatment	(1)	(2)	(3)	(4)	(5)	(6)
Audit*2018	0.124 (0.081)	0.124+ (0.068)	0.124 (0.079)	0.019 (0.023)	0.019+ (0.012)	0.019 (0.022)
Audit*2019	0.041 (0.093)	0.041 (0.098)	0.041 (0.101)	0.045 (0.026)	0.045* (0.018)	0.045 (0.027)
Audit*2020	0.099 (0.097)	0.099 (0.109)	0.099 (0.108)	0.025 (0.022)	0.025 (0.023)	0.025 (0.025)
Letter*2018	0.018 (0.028)	0.018 (0.040)	0.018 (0.034)	0.005 (0.006)	0.005 (0.007)	0.005 (0.006)
Letter*2019	0.110+ (0.056)	0.110* (0.054)	0.110* (0.049)	0.021+ (0.010)	0.021+ (0.011)	0.021+ (0.011)
Letter*2020	0.040 (0.054)	0.040 (0.062)	0.040 (0.043)	0.019 (0.012)	0.019 (0.013)	0.019 (0.012)
Observations	123844	123844	123844	123844	123844	123844
R-squared	0.756	0.756	0.756	0.835	0.835	0.835

*Note:* Regressions with fixed effects on strata and firm ID. The coefficients are to be interpreted in log points. Standard errors in parenthesis. Columns (1) – (3) are log payroll tax and Columns (4) – (6) are log employees. Models (1) and (4) with standard errors adjusted for 22 clusters in strata, models (2) and (5) with standard errors adjusted for 30,961 clusters in firm ID, and models (3) and (6) with standard errors adjusted for 61 clusters in NACE code (2-digit level). +  $p < 0.10$ ; \*  $p < 0.05$ .

**Table A6:** Estimated letter treatment effects vs. no treatment

Year	Dep. Variable	Letter	SE
2017	Log Payroll Tax	0.109*	(0.039)
2018	Log Payroll Tax	0.127**	(0.041)
2019	Log Payroll Tax	0.219**	(0.075)
2020	Log Payroll Tax	0.150*	(0.069)
2017	Log Employees	0.035*	(0.011)
2018	Log Employees	0.039**	(0.011)
2019	Log Employees	0.055***	(0.013)
2020	Log Employees	0.053***	(0.011)

*Note:* The coefficients are to be interpreted in log points. Fixed effects on year. Standard errors are adjusted for 22 clusters in strata.  $N = 28\,987$ . +  $p < 0.10$ ; \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ .

**Table A7:** Estimated ITT vs LATE letter effects

Year	Dep. Variable	ITT	SE	LATE	SE
2017	Log Payroll tax	0.109*	(0.039)	0.127**	(0.045)
2018	Log Payroll tax	0.127**	(0.041)	0.148**	(0.050)
2019	Log Payroll tax	0.219**	(0.075)	0.255**	(0.088)
2020	Log Payroll tax	0.150*	(0.069)	0.174*	(0.082)
2017	Log Employees	0.035*	(0.011)	0.040	(0.015)
2018	Log Employees	0.039**	(0.011)	0.046***	(0.014)
2019	Log Employees	0.055***	(0.013)	0.064***	(0.015)
2020	Log Employees	0.053***	(0.011)	0.062***	(0.013)

*Note:* The coefficients are to be interpreted in log points. Fixed effects on year. Standard errors are adjusted for 22 clusters in strata. N = 28 987. + p<0.10; \* p<0.05; \*\* p<0.01; \*\*\* p<0.001.

## 9.7 Heterogeneous Effects

**Table A8:** Estimated audit and letter (ATE) effects (base year 2017) on firm size 25 percentile

Treatment	(1)	(2)	(3)	(4)	(5)	(6)
Audit*2018	-0.161 (0.265)	-0.161 (0.201)	0.124 (0.079)	-0.011 (0.018)	-0.011 (0.015)	-0.011 (0.021)
Audit*2019	-0.058 (0.364)	-0.058 (0.235)	0.041 (0.101)	-0.002 (0.024)	-0.002 (0.020)	-0.002 (0.030)
Audit*2020	0.178 (0.344)	0.178 (0.242)	0.099 (0.108)	0.009 (0.025)	0.009 (0.024)	0.009 (0.029)
Letter*2018	-0.091 (0.136)	-0.091 (0.114)	0.018 (0.034)	-0.002 (0.007)	-0.002 (0.008)	-0.002 (0.007)
Letter*2019	-0.010 (0.157)	-0.010 (0.135)	0.110* (0.049)	0.001 (0.009)	0.001 (0.012)	0.001 (0.009)
Letter*2020	-0.187 (0.157)	-0.187 (0.143)	0.040 (0.043)	-0.007 (0.012)	-0.007 (0.014)	-0.007 (0.014)
Observations	30711	30711	123844	29439	29439	29439
R-squared	0.826	0.826	0.756	0.922	0.922	0.922

*Note:* Regressions with fixed effects on strata and firm ID. The coefficients are to be interpreted in log points. Standard errors in parenthesis. Columns (1) – (3) are estimated coefficients on log payroll tax for firms with revenue < NOK 735.003 and Columns (4) – (6) are estimated coefficients on log employees for firms with revenue > NOK 9.200.000. Models (1) and (4) with standard errors adjusted for 22 clusters in strata, models (2) and (5) with standard errors adjusted for 30,961 clusters in firm ID, and models (3) and (6) with standard errors adjusted for 61 clusters in NACE code (2-digit level). + p<0.10; \* p<0.05.

**Table A9:** Estimated audit and letter (ATE) effects (base year 2017) on firm type private Ltd

Treatment	(1)	(2)	(3)	(4)	(5)	(6)
Audit*2018	0.191+ (0.093)	0.191** (0.069)	0.191* (0.092)	0.042+ (0.021)	0.042*** (0.011)	0.042* (0.020)
Audit*2019	0.070 (0.094)	0.070 (0.104)	0.070 (0.105)	0.078** (0.024)	0.078*** (0.019)	0.078** (0.025)
Audit*2020	0.077 (0.094)	0.077 (0.119)	0.077 (0.099)	0.045* (0.020)	0.045+ (0.025)	0.045+ (0.025)
Letter*2018	0.020 (0.031)	0.020 (0.043)	0.020 (0.033)	0.005 (0.005)	0.005 (0.007)	0.005 (0.005)
Letter*2019	0.130+ (0.064)	0.130* (0.058)	0.130* (0.055)	0.022+ (0.011)	0.022+ (0.012)	0.022+ (0.012)
Letter*2020	0.088 (0.070)	0.088 (0.066)	0.088 (0.058)	0.025+ (0.013)	0.025+ (0.014)	0.025+ (0.013)
Observations	131460	131460	131460	131460	131460	131460
R-squared	0.654	0.654	0.654	0.771	0.771	0.771

Note: Regressions with fixed effects on strata and firm ID. The coefficients are to be interpreted in log points. Standard errors in parenthesis. Columns (1) – (3) are estimated coefficients on log payroll tax and Columns (4) – (6) are estimated coefficients on log employees. Models (1) and (4) with standard errors adjusted for 22 clusters in strata, models (2) and (5) with standard errors adjusted for 30,961 clusters in firm ID, and models (3) and (6) with standard errors adjusted for 61 clusters in NACE code (2-digit level). +  $p < 0.10$ ; \*  $p < 0.05$ .

**Table A10:** Estimated audit and letter (ATE) effects (base year 2017) on firm type self-employed

Treatment	(1)	(2)	(3)	(4)	(5)	(6)
Audit*2018	-0.183 (0.285)	-0.183 (0.213)	-0.183 (0.286)	-0.072+ (0.036)	-0.072+ (0.040)	-0.072* (0.035)
Audit*2019	0.025 (0.418)	0.025 (0.270)	0.025 (0.411)	-0.103* (0.048)	-0.103+ (0.054)	-0.103* (0.048)
Audit*2020	0.377 (0.344)	0.377 (0.281)	0.377 (0.338)	-0.058 (0.052)	-0.058 (0.060)	-0.058 (0.048)
Letter*2018	0.008 (0.128)	0.008 (0.120)	0.008 (0.114)	0.010 (0.024)	0.010 (0.020)	0.010 (0.019)
Letter*2019	-0.079 (0.182)	-0.079 (0.157)	-0.079 (0.134)	0.007 (0.037)	0.007 (0.029)	0.007 (0.033)
Letter*2020	-0.273 (0.214)	-0.273 (0.172)	-0.273* (0.116)	-0.018 (0.030)	-0.018 (0.033)	-0.018 (0.026)
Observations	22085	22085	22085	22085	22085	22085
R-squared	0.661	0.661	0.661	0.798	0.798	0.798

Note: Regressions with fixed effects on strata and firm ID. The coefficients are to be interpreted in log points. Standard errors in parenthesis. Columns (1) – (3) are estimated coefficients on log payroll tax and Columns (4) – (6) are estimated coefficients on log employees. Models (1) and (4) with standard errors adjusted for 22 clusters in strata, models (2) and (5) with standard errors adjusted for 30,961 clusters in firm ID, and models (3) and (6) with standard errors adjusted for 61 clusters in NACE code (2-digit level). +  $p < 0.10$ ; \*  $p < 0.05$ .

**Table A11:** Estimated audit and letter (ATE) effects (base year 2017) on firm type Norwegian registered foreign company

Treatment	(1)	(2)	(3)	(4)	(5)	(6)
Audit*2018	0.009 (1.179)	0.009 (1.361)	0.009 (1.285)	-0.478 (0.283)	-0.478 (0.323)	-0.478 (0.316)
Audit*2019	-1.474 (1.292)	-1.474 (1.719)	-1.474 (1.454)	-0.339 (0.248)	-0.339 (0.308)	-0.339 (0.289)
Audit*2020	-1.457 (1.462)	-1.457 (1.589)	-1.457 (1.556)	-0.262 (0.266)	-0.262 (0.345)	-0.262 (0.291)
Letter*2018	0.003 (0.588)	0.003 (0.580)	0.003 (0.600)	-0.051 (0.055)	-0.051 (0.062)	-0.051 (0.054)
Letter*2019	1.345 (0.832)	1.345 (0.815)	1.345* (0.573)	0.062 (0.132)	0.062 (0.113)	0.062 (0.136)
Letter*2020	0.589 (1.086)	0.589 (0.932)	0.589 (0.799)	0.011 (0.174)	0.011 (0.153)	0.011 (0.182)
Observations	1260	1260	1260	1260	1260	1260
R-squared	0.685	0.685	0.685	0.768	0.768	0.768

*Note:* Regressions with fixed effects on strata and firm ID. The coefficients are to be interpreted in log points. Standard errors in parenthesis. Columns (1) – (3) are estimated coefficients on log payroll tax and Columns (4) – (6) are estimated coefficients on log employees. Models (1) and (4) with standard errors adjusted for 22 clusters in strata, models (2) and (5) with standard errors adjusted for 30,961 clusters in firm ID, and models (3) and (6) with standard errors adjusted for 61 clusters in NACE code (2-digit level). +  $p < 0.10$ ; \*  $p < 0.05$ .