CESifo WORKING PAPERS

8480
2020

August 2020

Do Audits Improve Future Tax Compliance in the Absence of Penalties? Evidence from Random Audits in Norway

Shafik Hebous, Zhiyang Jia, Knut Løyland, Thor O. Thoresen, Arnstein Øvrum

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 8480*

Do Audits Improve Future Tax Compliance in the Absence of Penalties? Evidence from Random Audits in Norway

Abstract

The Norwegian Tax Administration operated multi-year random audits of personal income tax returns. We exploit this exceptional randomized setup to estimate the effects of tax audits on future compliance explicitly distinguishing between dynamic responses of compliant and noncompliant audited taxpayers. A priori, the literature has suggested two competing effects: A post-audit deterrence effect—whereby audits prompt taxpayers to comply in subsequent years—or an "approval effect"—whereby audits lower taxpayers' subjective probability of detecting future evasion and hence weaken compliance. Our results suggest improved future compliance for five post-audit years by those that were found noncompliant in the audits. Those that were found compliant, however, show no signs of behavioral adjustments. Although the findings are consistent with the deterrence effect, we argue that there is also a "learning" effect with the important implication that better information for taxpayers critically complements tax audits.

JEL-Codes: H260, C230.

Keywords: tax administration, tax evasion, tax compliance, tax audits, administrative data.

Shafik Hebous
International Monetary Fund
Washington DC / USA
Shebous@imf.org

Zhiyang Jia
Statistics Norway / Oslo
zhiyang.jia@ssb.no

Knut Løyland
Norwegian Tax Administration / Oslo
knut.loyland@skatteetaten.no

Thor O. Thoresen
Statistics Norway / Oslo
thor.olav.thoresen@ssb.no

Arnstein Øvrum
Norwegian Tax Administration / Oslo
arnstein.ovrum@skatteetaten.no

July 2020
This research has been carried out as a collaboration between the Norwegian Tax Administration and Oslo Fiscal Studies (Department of Economics, University of Oslo). Support from the Norwegian Research Council to Oslo Fiscal Studies is gratefully acknowledged. We thank Edwin Leuwen, Monique de Haan, participants in the "Modelling tax policy and compliance workshop" (TARC, University of Exeter), Skatteforum 2019, Annual Congress of the International Institute of Public Finance (Glasgow) for helpful comments. The views expressed here are those of the authors and do not necessarily represent the views of the IMF, its Executive Board, or IMF management.

1. Introduction

A central premise of tax audits is enhancing tax compliance. This insight—dating to the seminal work by Allingham and Sandmo (1972)—has been predominantly viewed in a static manner: a higher probability of detecting tax evasion, and thus paying a penalty, lowers the expected utility from concealing income, ceteris paribus.[1] The static nature of this prediction and the direct effects of audits on revenues in the year of the audit (through adjusting income and penalties) have been extensively studied and empirically confirmed.[2] However, with a few exceptions, there is little evidence, thus far, on the dynamic effects of tax audits—i.e., to what extent do tax audits encourage or discourage compliance in post-audit years? This question is crucial for an optimal tax administration as it is one element in the evaluation of the cost and effectiveness of administrative interventions (Slemrod and Keen, 2017)

The theoretical prediction regarding post-audit compliance behavior is ambiguous as it ultimately depends on taxpayers' perception of the detection probability, which can go in either direction. Taxpayers may perceive that an audit in this year is likely to be followed by other audits (and thus higher detection probability) in subsequent periods—a deterrence effect (e.g., Hashimzade, Myles and Tran-Nam, 2013)—, but it is also equally foreseeable that an audit today prompts agents to perceive a significant decline in the probability of being audited again in the aftermath of this audit—an approval effect (e.g., Maciejovsky, Kirchler and Schwarzenberger (2007) and Mittone, Panebianco and Santoro, 2017). Thus, empirical evidence based on administrative data is necessary to shed light on the evolution of the effects of audits on taxpayers' behavior over time.

This study estimates the effects of tax audits on future tax compliance using novel Norwegian administrative personal tax records data for about 30,000 individuals and random audits in the period 2009–2011. Our database contains rich information on all taxpayers eligible for the audit—not only those that were selected for the audits—and all are observed for up to six years after the audit and up to three years before the audit. Studying "real-life" random tax audits complements experimental studies on enforcement that are subject to well-known concerns[3] and provides relevant estimates for tax authorities.

Our empirical strategy exploits the random nature of tax audits by the Norwegian Tax Administration, which, critically, randomizes both the assignment into a treated group (audited) and a control (non-audited). The challenge that typically faces this type of analysis—and most studies on audits—is the nonrandom selection for tax audits in practice leading to a severe selection bias. Usually, audits are based on risk scores. In this study, given the random assignment, our main identification strategy for obtaining consistent and unbiased estimates is a difference-in-difference (DiD) research design.

One further advantage of our study stems from features of the data and the institutional setup per se. Most Norwegian taxpayers receive fully prepopulated income tax returns that they digitally file by a mere mouse click by simply approving the tax return. This largely automatic and digitalized process relies on an extensive and highly developed third-party reporting system. However, taxpayers have the option of entering missing information and making some amendments. One particular item of interest in the tax form is "Other deductions", used by both wage earners and the self-employed, for claiming deductions that are not already recorded. This

[1] There are a range of further determinants of tax compliance beyond the Allingham-Sandmo model, which fall under the broad umbrella of tax moral (Luttmer and Singhal, 2014).
[2] For surveys, see, e.g., Andreoni, Erard, and Feinstein (1998), Alm (2019), and Slemrod (2019).
[3] See, e.g., Czibor, Jimenez-Gomez, and List (2019).

particular item has been subject to random audits by the tax administration, which generated the dataset of this study. Misreporting in this item does not exclusively imply deliberate evasion but can also reflect inadvertent errors resulting from misunderstanding or uncertainty about the tax treatment. We provide one of the first empirical study on unintentional errors and how audits can impact compliance, essentially, in the absence of penalties. Despite a general recognition in the literature that confusion and unintentional errors are important factors that can explain noncompliance (e.g., Slemrod, 2007), there have been no empirical studies on this issue.

Our DiD benchmark results suggest that audited taxpayers, overall, reduce their claimed income deductions in the post-audit years. However, this is an average effect for two groups as audit outcomes split audited taxpayers into compliant and noncompliant. Estimates for the compliant and the noncompliant taxpayers separately indicate that the compliance effect is driven mainly by the noncompliant group, lasting for five years but it decreases over time. The decrease in reported deductions is 12 percent in the immediate post-audit year and 5 percent in the fifth post-audit year, measured against the average claimed deduction in the year of the audit. Note that adjusted incomes (after the audits) were subjected to the tax but in principle without imposing penalties. In contrast, the compliant group shows no statistically significant change in reporting during post-audit years.

The validity of our results is based on the common trend assumption, which we can reasonably motivate by simple plots of self-reported deductions by the three groups: unaudited, audited-compliant, and audited -noncompliant taxpayers. One issue is that the unaudited group potentially contains both compliant and noncompliant taxpayers. We openly discuss this aspect and show that the common trend assumption remains the key aspect for valid group-specific estimates. Addititonaly, using different methodologies, we obtain similar results to those from the DiD design. In particular, we provide estimates from: i) a less restrictive approach that derives bounds' estimates of the group-specific audit effect following the methodology in Manski and Pepper (2018); and ii) matching techniques.

Yet, our data analysis reveals a further intriguing pattern of interest in its own right. Claimed deductions by taxpayers in the noncompliant group are, on average, below those by compliant taxpayers. Moreover, there are more self-employed in the compliant group than in the noncompliant group, contrary to common intuition and existing literature on tax evasion by the self-employed. Combining these observations with the de facto absence of penalty, we motivate that the dynamic compliance effect is, at least in part, prompted by taxpayers' learning as a result of audits—i.e., tax audits raise awareness and attention of taxpayers. To some extent, having ruled out an approval effect of audits, from a collection point of view, the estimated ("reduced-form") dynamic effect is what matters irrespective of the motivation to comply in post-audit years—whether it is "deterrence" or "learning"—, but by exploring this compliance motivation further we argue that the learning mechanism is consistent with the Norwegian data. This implies that improving taxpayers' information and strengthening the communication of tax policy are important complements to deterrence strategies.

Our study contributes to various strands of the literature. First, a few studies empirically examine the reaction of taxpayers to audits. Slemrod et al. (2001) find that taxpayers report higher income after being warned about future audits of their income returns. DeBacker et al. (2015) study corporate behavior in the US and find that firms reduce tax payments immediately after audits but increase payments gradually afterwards. DeBacker et al. (2019) find that the effect of audits on future tax payments in the US is short-lived without third-party information. On the other hand, Advani et al. (2019) find that third party information does not predict whether a taxpayer is compliant in the UK. Kleven et al. (2011) find in a field experiment in Denmark that tax evasion is close to zero for income subject to third-party reporting, in stark contrast to self-reported income. Gemmell and Ratto (2012) distinguish between noncompliant and compliant taxpayers in the UK and report evidence suggesting that the "past experience" rather than the threat of audits enhance future tax compliance by the noncompliant. Beer et al. (2020) study the dynamic compliance effect for self-employed taxpayers in the US, but the selection of audits is based on a risk score. They find that compliance depends on the audit outcome and that non-audited taxpayers reduce their reported income in post-audits years. We contribute to this literature by i) distinguishing between the responses of compliant and noncompliant taxpayers, which with the exception of Beer at al. (2020) and Gemmell and Ratto (2012) has not been explicitly studied; ii) using high-quality tax return data with actual random audits; and iii) looking at the role of taxpayer learning in enhancing compliance beyond the threat of penalty.

Second, several studies have examined compliance behavior in a laboratory environment. For example, Alm, Jackson, and McKee (2009) find that pre-announced audit rates improve compliance. This finding is broadly in line with theorical predictions in Snow and Warren (2007), suggesting that as unaudited taxpayers update expectations about the probability of future audits (i.e., Bayesian learning) tax evasion increases. Our empirical paper complements this literature with evidence from high quality administrative data.

Third, a strand of the literature extends the Allingham-Sandmo model in several directions accounting for compliance factors beyond enforcement, among other things, moral sentiments of guilt and shame (Erard and Feinstein, 1994) and social conformity effects (Myles and Naylor, 1996; Fortin, Lacoix, and Villeval, 2007).[4] Somewhat relatedly, Beck and Jung (1989) theoretically model uncertainty about tax lability and predict that, as a result, the impact on compliance is ambiguous. Alm, Jackson, and McKee (1992) in a laboratory experiment find that uncertainty tends to increase tax compliance, but the effect in principle can also go in either direction. In this context, we contribute to the literature by providing suggestive evidence that learning by taxpayers (and thus lower uncertainty) improves tax compliance, feeding into the broad intrinsic motivation of compliance.[5]

The remainder of this paper is organised as follows. Section 2 describes the Norwegian institutional setting that produced the audit data. Section 3 presents descriptive statistics distinguishing between compliant and noncompliant taxpayers. Section 4 presents estimates for the overall behavioral effect of tax audits, for compliant and noncompliant taxpayers, and includes results of several robustness checks. Section 5 concludes.

2. Institutional Setting

Norway adopts a tax system that is close to a dual income tax, thereby broadly business and capital incomes (such as from interest) are subject to a flat tax rate (which was 28 percent in 2011 and is 22 percent as of 2020) whereas gross employment income is subject to a progressive tax schedule with a top personal income tax (PIT) rate of 47.8 percent in 2011 (46.4 percent in 2020). Roughly speaking, the top PIT rate kicks in at a level of income equals to a multiple of 1.6 of the average wage.[6]

[4] Luttmer and Singhal (2014) and Hashimzade et al. (2013) provide surveys.

[5] See, e.g., Dwenger et al. (2016).

[6] As the progressive tax scale applies to gross income and the flat tax rate applies to net income, deductions (the difference between gross and net income), if denied would be subjected to a flat rate which was 18 percent during the sample.

The Norwegian third-party information system—based on data from employers, the financial sector, and others—is advanced, almost fully eliminating the need for contact between taxpayers and the tax administration. Taxpayers simply receive a digitally fully prepopulated tax return for approval. Prefilled tax returns can be amended by taxpayers. For example, charitable donations are deductible up to a threshold in the taxation of ordinary income. The recipients of donations (i.e., recognized charitable organizations) report the individual donations directly to the tax authorities. Of course, in case errors occur or incomes or deductions are not reported, taxpayers can make amendments to the income tax return directly through the internet without the involvement of administrative staff from the tax authorities.

The analysis of this paper focuses on a particular item in the income tax return called "Other deductions", which is frequently used to report additional deductions not already recorded through the third-party reporting system. Common categories of deductions under this item are fees related to capital income—including management fee, and stock exchange account—, expenses for home office, and charitable donations to Civil Society Organizations or scientific entities.[7] This item is filled in by wage earners and the self-employed.

The nature of this item does not rule out the possibility that taxpayers are genuinely uncertain about the qualification of their costs for deduction. Correspondingly, from a tax administrative standpoint, audits are required as taxpayers can make substantiated or unsubstantiated (illegal) claims. All taxpayers with a claim under "Other deductions" above 50,000 (8,300 USD)[8] were audited. Among taxpayers who have claimed "other deductions" in the range from 5,000 to 50,000 NOK (approximately 570–5,700 USD), a subsample of approximately 10 percent is randomly assigned for further audits. In this study, we use information from auditing of income in the years 2009, 2010 and 2011.[9]

Noncompliant taxpayers were informed since the tax authorities adjust their income tax return. For example, if a deduction of 50,000 NOK was denied by the tax authorities, a 28-percent flat rat applies, impling an extra tax burden of a tax of 14,000 NOK (about 1,650 USD). In principle, there is a penalty on top of the regular tax (given by the falt rate) in the case of a deliberate tax evasion. However, among the audited taxpayers in sample of this study, less than one percent received a penalty. From an empirical identification perspective, it is important to note that not only the noncompliant are informed about the attention from the tax administration, but many of the compliant taxpayers too. Some of the compliant taxpayers may go through the process without notification, if they already have provided all the necessary documentation needed. However, many become aware of the audit because they have been asked to provide additional information.

[7] Other, less frequently used categories include judicial costs, newspaper subscriptions, moving costs, and dental costs. Some of these deductions will be limited to one year, while others (annual fees) will be repeated for several years.
[8] Exchange rates for 2010.
[9] After 2011, the tax administration changed the audit procedures.

3. Audit Results

3.1. Data Description

We base our analysis on three waves of audits on "Other deductions", in 2009, 2010, and 2011. For the taxpayers who are eligible for the audits, we have information on the particular deduction claimed and a set of other individual characteristics including gender, age, total gross income, third part reported gross income, total deductions (including the item we study), third party reported deductions, and self-employment status.[10] The observation period starts in 2008 and lasts until 2015. That is, we observe taxpayers in 1–3 years before the audit and in 3–6 years after the audit.

The auditing process generates two distinctively different groups among the treated: those who have been caught not reporting correctly, the noncompliant, and those who can substantiate that their claims are correct and therefore get cleared (compliant). Around 36 percent of the taxpayers are found to have misreported their deductions.

Table1. Descriptive Statistics for Compliant, Noncompliant, and Non-audited (Averages), 2008–2015

	Compliant	Noncompliant	Audited in total	Non-audited
Claimed "Other deductions"	22,275	23,153	23,091	23,104
(in the year of the audit)	(8,914)	(10,463)	(11,520)	(11,290)
Direct correction in "Other deductions" due to audit	–	25,948	-	–
		(22,182)		
Self-employed†	25	14	21	21
	(29.3)	(24.3)	(41.0)	(40.6)
Female†	25	34	28	29
	(45.8)	(049.0)	(45.0)	(45.4)
Age (years)	52	43	49	48
	(12.5)	(12.9)	(13.0)	(13.0)
Temporary work migrant†	3.4	7.3	4.8	6.00
	(19.9)	(24.3)	(21.0)	(23.7)
Total deductions	232,364	191,417	217,780	217,410
	(79,232)	(70,882)	(214,060)	(205,798)
Third-party rep. deductions	153,390	128,443	144,505	141,596
	(52,507)	(46,071)	(116,134)	(108,897)
Total gross income	859,141	617,731	773,161	770,830
	(355,130)	(243,154)	(733,130)	(736,564)
Total third-party rep. gross income	655,057 (296,625)	526,447 (235,434)	609,252 (480,575)	607,691 (484,428)
Observations	2,238	1,238	3476	26,775

Notes: Standard deviations are in parentheses. † Figures are in percent. The rest of the values are in NOK except for age (years).

Table 1 shows descriptive statistics for the non-audited and the audited distinguishing for the latter group between compliant and noncompliant taxpayers. The average claimed 'other

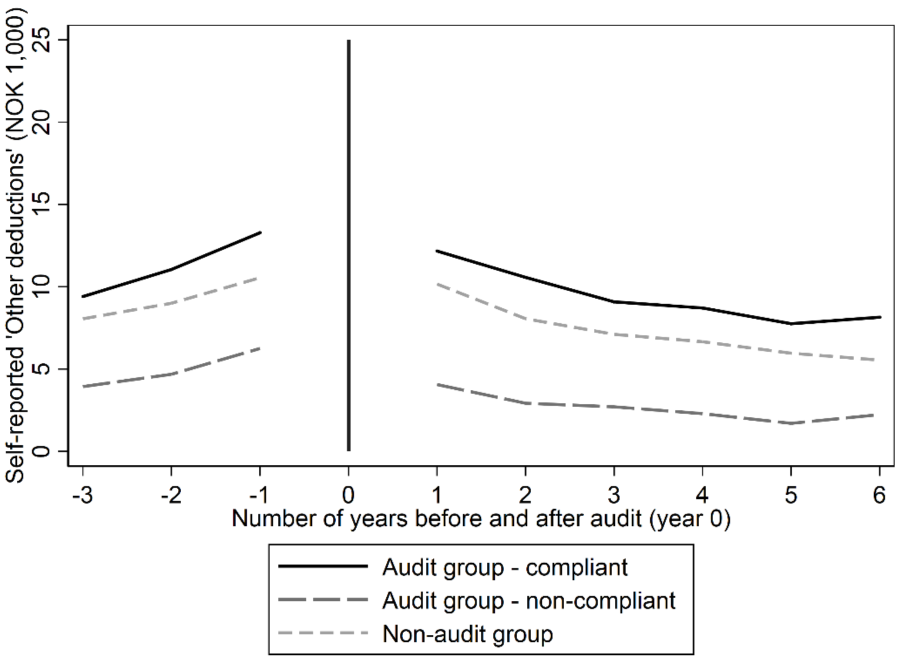
[10] An individual is classified as a self-employed if the income from the business is at least 10 000 NOK, and even if the individual is a wage earner as well.

deductions' is close to 23,000 NOK (3,800 USD) for the three groups. Appendix A presents the distribution of deductions by groups of taxpayers. The average correction of deductions for the noncompliant group is about 26,000 NOK (4,300 USD). Additionally, the summary statistics reveal some notable differences between the two subgroups of the audited. For example, the compliant taxpayers are older and richer than individuals in the noncompliant group. Moreover, somewhat surprisingly, in the noncompliant group there are more self-employed (25 percent vs. 14 percent) and more females (34 percent vs. 25 percent) than in the compliant group. This is in contrast to some studies that find that males are more likely to evade taxes than females (Torgler and Valev, 2010). The self-employed are typically the focus of tax evasion studies because they have more scope to misreport income than wage earners that are subject to wage withholding taxes. However, this item of deductions is self-reported for both the self-employed and the non-self-employed, and thus the self-employed do not have more flexibility to misreport this item.

3.2. Deduction Behavior by Group

Figure 1 shows the trend in the deduction behavior for all subgroups: non-audited, audited-compliant, and audited-noncompliant. Recall that the three waves of audits were in 2009–2011 (coded as year 0). We observe taxpayers for up to six years after the audit and three years pre-audit. Note that Figure 1 shows the average values for the pre- and post-t_0, whereas Table 1 displays the average values of "Other deductions" for year t_0.

Figure 1. Average Deductions before and after the Audits: Compliants, Noncompliants and Non-audited



Notes: Left-hand diagram includes deductions in the year of the audit, year 0. Right-hand diagram is representative for the observations used in the regressions, i.e., year 0 is excluded.

Figure 1 shows that individuals in the noncompliant group (on average) have much lower claimed deduction level than individuals in the compliant group, despite that the levels are almost the same at the year of audit (as shown in Table 1). In Appendix C, we sketch a simple model that rationalizes different levels of deductions as observed in Figure 1, and show that it is consistet with unintentional errors in the noncompliant group.

One important question is: Which characteristics can explain the assignment into compliant or noncompliant group? We estimate a simple Probit model of the form:

(1) $$Pr(\text{ non-compliance}) = \Phi(\beta_i x_i),$$

where x_i represents the observed characteristics and Φ is the standard Normal Cumulative distribution function (CDF). While the results from such a model cannot be interpreted as causal, they do, are informative, at least regarding two aspects. First, this kind of information can be used to help to design a more efficient audit program by targeting those who are more likely to be noncompliant. Second, these results may shed lights on possible mechanism underlying the compliance behavior. Table 2 presents the average marginal effects of a range of characteristics on the probability of being noncompliant. The main findings broadly confirm the summary statistics in Table 1. For example, self-employment and male have negative significant effects on the likelihood of being a noncompliant. Also, the estimation results suggest that immigrant status does not have a significant effect on the probability of belonging to the noncompliant group (after controlling for other variables).

Table 2. The Probability of Noncompliant: Average Marginal Effects (Probit Model)

	AME	Std
Self-employed	-0.07	0.021
Female	0.04	0.017
Age	-0.01	0.001
Temporary work migrant	-0.02	0.349
Total deductions (1 million NOK)	0.04	0.052
Third-party rep. deductions (1 million NOK)	-0.26	0.101
Total gross income (1 million NOK)	-0.04	0.016
Total third-party rep. gross income (1 million NOK)	-0.04	0.023
Observations		3,476
McFadden R-square		0.107

4. Estimates of the Dynamic Effects

4.1. Overall Audit Effect

The overall, average, post-audit effect can be estimated either by comparing the outcome of the audited and unaudited—given that the assignment of audit is random—or using a simple DiD design with unaudited taxpayers as the control group. These two methods are in principal equivalent, but the latter has the advantage that it can be applied when we want to separately estimate effects for the compliant and noncompliant groups (as discussed below).

We estimate the overall post-audit effect of the audit in year zero on the deduction behavior of individual i (y_i) using the following DiD specification:

$$(2) \qquad y_{is} = \alpha_i + \lambda_t + \theta D_i + \delta_s D_i \times 1(s > 0) + \varepsilon_{ist} \quad s \in \{-3, -2, -1, 1, \ldots, 6\},$$

where s measures the distance in years to the year of the audit ($s = 0$). The binary regressor, D_i, takes the value 1 if the taxpayer was audited in year $s = 0$; and it is zero otherwise. α_i denotes the individual fixed effect and λ_t represents the calendar year effect. The year of the audit is excluded from the dataset used in the estimation. Thus, δ_s measures the treatment effect of audit at a specific year s after the audit.

Table 3 reports the estimated audit effects on deductions in six post-audit years. There is statistically significant negative effect in the first year after the audit suggesting that "Other deductions" decline by approximately 1,300 NOK, on average. For the rest of the post-audit years (i.e., for $s > 1$), point estimates suggest that taxpayers reduce their reporting of deductions because of the audit and depict a diminishing effect overtime, but they are insignificant. As the results indicate that there is a negative shift in the mean deduction after the audit, in Appendix B we examine whether the shape of the deduction distribution has been changed.

Table 3. Effects of Tax Audit on Post-Audit Deduction Behavior

Year after audit	Coefficient	Estimate	*t*-value
First	δ_1	-1,272*** (460)	-2.76
Second	δ_2	-572 (454)	-1.26
Third	δ_3	-626 (460)	-1.36
Fourth	δ_4	-557 (465)	-1.20
Fifth	δ_5	-479 (482)	-0.99
Sixth	δ_6	-189 (592)	-0.32
Observations		177,161	

Notes: Fixed effect estimation based on panel data 2008–2015. Robust standard errors in parentheses.
$*p < 0.1, **p < 0.05, *** p < 0.01$

4.2. Separate Estimates for the Compliant and Noncompliant

The above estimated average effect is informative, but it masks heterogenous responses of those who were found compliant and noncompliant as a result of the audit, calling for separate estimates for both groups. The identification of a group-specific effect is, however, challenging because the behavior of taxpayers in the non-audited group is latent —i.e., it is not necessarily exclusively representing the compliant group or the noncompliant group. It is rather a "mixed" control group, consisting of both compliant and noncompliant taxpayers.

We discuss here the question as to what extent it is possible to use the whole group of non-audited as the control group for obtaining treatment estimates for the compliant and the noncompliant. Advani, Elming, and Shaw (2019) use a before-after setup comparing the compliance behavior prior to and and after the audit. This method overcomes a possible endogeneity problem, but it only allows identifying the effect for the compliant group (and not for the noncompliant group). Moreover, it does not have a control group for the post-audit period even for the compliant group.[11]

We show, below, that the common trend (for the non-audited and for the group-specific trends) suffices to obtain consistent reliable estimates. Let $Q_i = 1$ denote that individual i is of type noncompliant and $Q_i = 0$ denote the noncompliant type. As above, we have $D_i = 1$ if the individual is audited and $D_i = 0$ if not. Let δ_1 and δ_0 be the DiD estimates for the noncompliant and compliant taxpayers, respectively, using all the non-audited as the control group. $\Delta Y_i(Q_i, D_i)$ denotes the difference between the post-audit and the pre-audit deduction of individual i. Then the DiD estimator for the noncompliant group can be written as:

$$\begin{aligned}\delta_1 &= E[\Delta Y_i(Q_i = 1, D_i = 1) - \Delta Y_i(D_i = 0)] \\ &= E[\Delta Y_i(Q_i = 1, D_i = 1)] - E[\Delta Y_i(D_i = 0)] \\ &= E[\Delta Y_i(Q_i = 1, D_i = 1)] - qE[\Delta Y_i(Q_i = 1, D_i = 0)] - (1-q)E[\Delta Y_i(Q_i = 0, D_i = 0)],\end{aligned} \tag{3}$$

where q is the probability for individual i being a noncompliant taxpayer. Given the random assignment of audit, it can be consistently estimated. δ_1 will be a consistent estimator of the type specific audit effect, $\gamma_1 = E[\Delta Y_i(Q_i = 1, D_i = 1) - \Delta Y_i(Q_i = 1, D_i = 0)]$, if and only if the following condition holds: $E[\Delta Y_i(Q_i = 1, D_i = 0)] = E[\Delta Y_i(Q_i = 0, D_i = 0)]$. That is, the change in outcome variable in absence of the treatment does not depend on the unobserved types, or, in other words, the common trend assumption holds. In our data, the common trend assumption seems to hold for both groups (Figure 1), supporting the use the non-audited group as a control group for both types of the audited taxpayers (compliant or noncompliant).

Following Autor (2003), we formally test the common trend assumption by regressing deductions in the two groups prior to the audit, $s \in \{-1, -2, -3\}$, against time dummies and dummies for type of taxpayer, compliant or noncompliant taxpayer, denoted by Q_i (as established after the audit),

$$y_{i,s<0} = \alpha_i + \lambda_t + \beta Q_i + \kappa Q_i \times 1(s = -1) + \pi Q_i \times 1(s = -2) + \xi Q_i \times 1(s = -3) + \varepsilon_{ist}, \tag{4}$$

where α_i is a set of individual fixed effects and λ_t represents year fixed effects. We estimate Equation (4) omitting years in an alternate manner, $s \in \{-1, -2, -3\}$. We obtain statistically insignificant estimates of κ, π and ξ for all pre-audit years (results are not reported). Hence, we conclude that there is no statistical support for the hypothesis of rejecting the common trend.

Thus, overall, given the above discussions and results, we argue that the non-audited group is a reasonable control group. Nonetheless, we return to to this methodological challenge below, in terms of results from a less restrictive partial identification method.

[11] Gemmell and Ratto (2012) simply used the non-audited as the control group for both the compliant and noncompliant groups.

4.3. Separate Estimates for the Two Compliance Groups

Following the above discussion, we extend specification (2) to estimate group-specific effects. In particular, we introduce a further distinction in the post-treatment years distinguishing between two types of taxpayers: those who were able to substantiate the claimed deductions versus those who were not. Let subscript $j \in \{0,1\}$ denote that the treated taxpayers are in subgroups 0 or 1, the compliant and noncompliant taxpayers, respectively. We estimate the following equation separately for both groups:

(5) $$y_{is} = \alpha_i + \lambda_t + \theta_j D_{ij} + \delta_{js}\left[D_{ij} \times 1(s > 0)\right] + \varepsilon_{ijst} \quad s \in \{-3,-2,\ldots,6\} \wedge j \in \{0,1\}.$$

The estimates of effects of audits on "Other deductions", δ_{sj}, in Equation 5 measure average group-specific effects, separately for the compliant and for noncompliant groups.

Table 4 displays the estimation results for Equation 5. The findings suggest that there are large differences between cleared taxpayers and those that were requested to adjust their claims due to the lack of sufficient substantiation. First, compliant taxpayers do not alter their deduction behavior after the audit. Estimated coefficients are statistically insignificant, and they change signs. Second, noncompliant taxpayers reduce their deductions by 2,876 NOK (480 USD) in the first year and 8,089 NOK (1,340 USD) over the five post-audit years. The compliance effect diminishes over time, and turns insignificant in the sixth post-audit year. Thus, overall, the results are not consistent with an approval effect, but they rather lend support to an improved compliance effect for the noncompliant.

Table 4. Effects of audit on post-audit deduction behavior. Compliant and noncompliant taxpayers

	Year after audit	Coefficient	Estimate	t-value
Compliant	First	δ_{01}	-400(611)	-0.65
	Second	δ_{02}	123(602)	0.21
	Third	δ_{03}	-384(603)	-0.64
	Fourth	δ_{04}	-302(617)	-0.49
	Fifth	δ_{05}	-85(638)	-0.13
	Sixth	δ_{06}	-90(794)	-0.11
Noncompliant	First	δ_{11}	-4,313***(599)	-4.88
	Second	δ_{12}	-3,161***(553)	-3.15
	Third	δ_{13}	-2,878*** (606)	-1.75
	Fourth	δ_{14}	-2,877*** (573)	-1.74
	Fifth	δ_{15}	-3,451*** (606)	-1.99
	Sixth	δ_{16}	-2,537*** (763)	-0.55

Notes: Fixed effect estimation based on panel data 2008–2015. Robust standard errors in parentheses.

$*p < 0.1, **p < 0.05, *** \; p < 0.01$

The results are striking as only 1.5 percent of the noncompliant taxpayers were fined because of their unverified claims. This suggests that audits can have positive dynamic effects on compliance even in the absence of penalties. Uncertainty about the tax treatment of some deductions lead some, but not all, taxpayers, to claim the deductions. This behavior is mainly driven by inadvertent errors or "wishful thinking". A crucial aspect of this, as discussed in detail in the Appendix C, is that in the absence of audits compliance behavior is not correlated over time. One important implication, consistent with our data, is that being noncompliant in a given audit per se would not provide additional information on compliance behavior in the pre-audit years. The finding that audits lowers future unintentional tax evasion suggests that not only deterrence but also learning of taxpayers plays a key role in enhancing tax compliance.

4.4. Robustness Tests

Establishing Bounds Based on a Partial Identification Method

Our results so far are derived from a DiD framework critically dependent on the common trend assumption. In the following we discuss results from an empirical approach that seek to obtain results under less restrictive conditions, a version of a partial identification method, where the ambition is to derive bounds to the group-specific audit effects. The same underlying idea is used by Manski and Pepper (2018) under the name the Bounded Variation Assumptions approach.

When the common trend fails, we can use Equation (3) and some additional assumptions to derive bounds for the true audit effects for the compliant group (γ_0) and for the noncompliant group (γ_1). The true effects can be defined as

$$\gamma_0 = E[\Delta Y_i(Q_i = 0, D_i = 1)] - E[\Delta Y_i(Q_i = 0, D_i = 0)],$$

$$\gamma_1 = E[\Delta Y_i(Q_i = 1, D_i = 1)] - E[\Delta Y_i(Q_i = 1, D_i = 0)].$$

Further, we introduce two assumptions, with respect to average behavior, that our empirical approach relies on. Firstly, in absence of audit the noncompliant taxpayers will not reduce their deduction claims more than the compliant taxpayers, and, secondly, there are more reductions for those who get caught than for the compliant taxpayers not being audited.

In the following we formalize how bounds can be derived based on these relatively mild assumptions. The exercise is primarily helpful in order to clarify in which direction one would expect results to move when not relying on a subgroup common trend. Then, one should be aware that the first assumption, that the noncompliant taxpayers will not reduce their deduction claims more than the compliant taxpayers, basically states that the γ_1 is not above the average treatment effect of the noncompliant, δ_{1s} in Table 4.

Nevertheless, let us see how the bounds can be derived. The two assumptions imply that we have

$$E[\Delta Y_i(Q_i = 1, D_i = 0)] \geq E[\Delta Y_i(Q_i = 0, D_i = 0)] \text{ and}$$

$$E[\Delta Y_i(Q_i = 0, D_i = 0)] \geq E[\Delta Y_i(Q_i = 1, D_i = 1)].$$

This means that we have the following conditions for audit effect of the noncompliant group:

(7) $$\gamma_1 = \delta_1 - (1-q)(E[\Delta Y_i(Q_i = 1, D_i = 0)] - E[\Delta Y_i(Q_i = 0, D_i = 0)]) \le \delta_1 ,$$

which gives

(8) $$\gamma_1 = \frac{1}{p}\left\{\delta_1 + (1-q)(E[\Delta Y_i(Q_i = 1, D_i = 1)] - E[\Delta Y_i(Q_i = 0, D_i = 0)])\right\} \ge \frac{\delta_1}{q} .$$

Thus, we bound the true treatment effect for the noncompliant group as

$$\frac{\delta_1}{q} \le \gamma_1 \le \delta_1 .$$

For the compliant group we correspondingly have,

$$\delta_0 = \gamma_0 + q(E[\Delta Y_i(Q_i = 0, D_i = 0)] - E[\Delta Y_i(Q_i = 1, D_i = 0)]) .$$

When we use the same assumption as employed to restrict γ_1 in Eq. (8), we get $\gamma_0 \ge \delta_0$. To obtain the upper bound, we can then use the identity

$$q\gamma_1 + (1-q)\gamma_0 = \mathrm{ATT} ,$$

where ATT is the average effect of audit on the audited group, of which estimation results already have been obtained. Thus,

$$\gamma_0 = \frac{\mathrm{ATT} - p\gamma_1}{1-q} \le \frac{\mathrm{ATT} - \delta_1}{1-q} ,$$

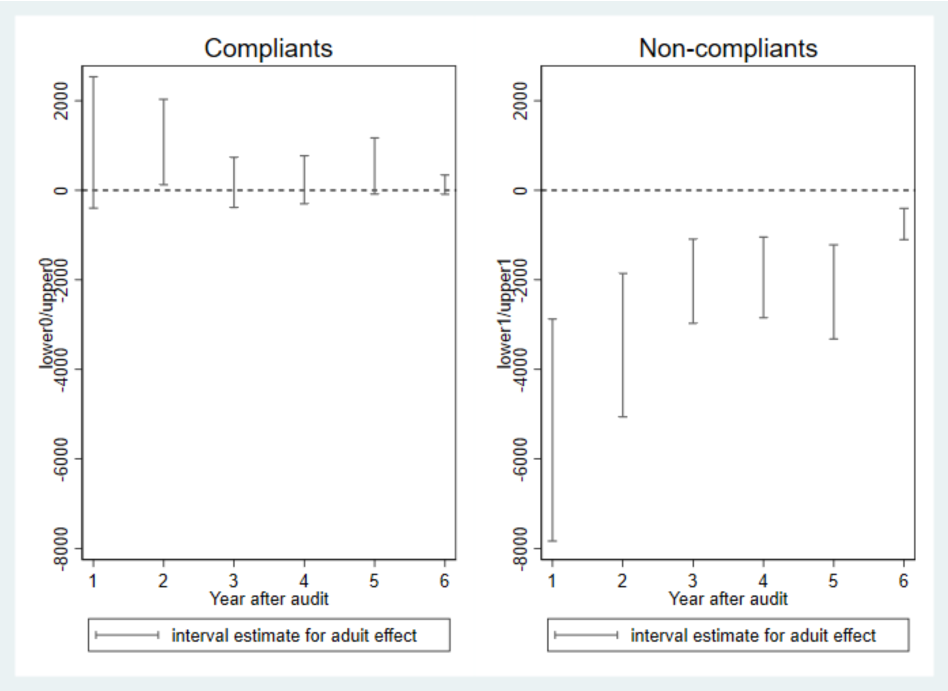
which follows from $\gamma_1 \ge \frac{\delta_1}{q}$.

Under these assumptions the bounds the type-specific audit effects can be seen as

(9) $$\gamma_0 \in [\delta_0, \frac{\mathrm{ATT} - \delta_1}{1-q}] \text{ and } \gamma_1 \in [\frac{\delta_1}{q},\ \delta_1].$$

Hence, based on estimates reported in Section 4.1 and Section 4.3 we obtain empirical estimates of the bounds for the group specific audit effects. It follows from our two conditions that the point estimates, δ_{0s} and δ_{1s} in Table 4, represent the lower and upper bound for the compliant and noncompliant , respectively. Intuitively, the tightness of the bounds for γ_1 is an increasing function of the share of individuals belonging to the noncompliant group (q). When there is are no noncompliant individuals in the population, that is when $q = 0$, there is no information in the data to identify γ_1, while the exact identification is obtained when $q = 1$. In this case the interval is reduced to a single point.

Figure 2. Bounds for the Effects of Auditing



The bounds are reported in Figure 2 (without standard errors), showing that bounds are relatively wide for the noncompliant taxpayers. However, as one would expect, given the two assumptions that found the basis for obtaining them, the results point to possible directions if one leaves the common trend assumption. If anything, noncompliant taxpayers may reduce their deduction claims more after being audited, whereas the upper bounds of the compliant signify a possibility for approval.

Matching Method Results

To further check of the robustness of the results, we estimate a new model with a new control group obtained form a matching procedure. In particular, we apply the Coarsened Exact Matching algorithm (CEM) and use pre-audit control variables to obtain better balance between the treated and the control groups.[12] Approximately, 10 percent of the audited individuals were not matched to anyone in the control group, so they were excluded from the matched regression analysis.

Table 5 presents the results using only the matched sample. Compared with the non-matched sample, the estimated effects audits for the noncompliant groups are more clearly identified and the effects are larger. The point estimates for the compliant groups now are all positive but none of them are significant, except for the last year of period. Moreover, we also obtained results for propensity score matching, which are very close to the results reported in Table 5.

[12] See Iacus, King and Porro (2011).

Table 5. Effects of audit on post-audit deduction behavior. Compliant and noncompliant taxpayers, matched sample.

	Year after audit	Coefficient	Estimate	t-value
Compliant	First	δ_{01}	513(671)	0.76
	Second	δ_{02}	1072(668)	1.61
	Third	δ_{03}	377 (664)	0.57
	Fourth	δ_{04}	891(673)	1.32
	Fifth	δ_{05}	1131(700)	1.61
	Sixth	δ_{06}	1662*(850)	1.95
Noncompliant	First	δ_{11}	-4,313***(599)	-7.20
	Second	δ_{12}	-3,161***(553)	-5.71
	Third	δ_{13}	-2,878*(606)	-4.75
	Fourth	δ_{14}	-2,877*(573)	-5.01
	Fifth	δ_{15}	-3,451**(606)	-5.70
	Sixth	δ_{16}	-2,537(763)	-3.32

Notes: Fixed effect estimation based on panel data 2008–2015. Robust standard errors in parentheses. Matching of sample carried out by Coarsened Exact Matching (CEM) algorithm. $*p < 0.1, **p < 0.05, *** p < 0.01$

4.5. Spillover Effects on Other Items

An account of the costs and benefits of an audit should control for audits influencing the reporting on other items. In our case we may ask if the attention received in terms of the check on the item "Other deductions" may cause the agents to adjust their subsequent filing behavior in general. In order to explore this issue further, we estimate Equation (5) after replacing the dependent variable by gross income. Thus, if the dynamic audit effect spreads to the reporting of income to, we expect to see similar patterns as for "Other deductions" for gross income too. However, we find no indications of spread to the gross income reporting. Results are not reported here but are available upon request.

5. Conclusion

The effectiveness of audits is one crucial element of an efficient tax administration. In terms of assessing the revenue implication of audits, to draw the big picture, the calculation should not only account for tax adjustments made in the year of the audit, but also future tax adjustments triggered by behavioral responses to the initial audit. Based on data from random audits by the Norwegian Tax Administration, the findings of this study suggest that audited taxpayers reduced their claimed income deductions in the post-audit years, thereby raising their reported income and hence compliance.

Moreover, the analysis suggests that the increased future compliance effect is driven by the those that were audited and prompted to correct their tax returns. The decrease in their reported deductions is 12 percent in the first post-audit year, then it gradually decreases reaching 5 percent in the fifth post-audit year. However, no dynamic reaction was found for those that were audited and their tax returns were approved by the tax authorities without adjustments.

While this outcome, in general, can be explained by increasing taxpayers' subjective probabilities of future audits and detection (i.e., a deterrence effect), the analysis suggests that the dynamic improvement in compliance can be triggered even in the absence of penalties broadly in line with a learning effect. This implies that improving the information set of taxpayers is one of the key aspects of an efficient tax administration.

References

Advani, A., W. Elming, and J. Shaw (2019). The dynamic effects of tax audits. Working Paper No. 414, University of Warwick.

Allingham, M. and A. Sandmo (1972). Income tax evasion: A theoretical analysis. *Journal of Public Economics*, 1(3-4), 323–338.

Alm, J. (2019). What motivates tax compliance. *Journal of Economic Surveys*, 33(2), 353–388.

Alm, J., B. R. Jackson, and M. McKee (1992). Institutional uncertainty and taxpayer compliance. *The American Economic Review*, 82(4), 1018–1026.

Alm, J., B. R. Jackson, and M. McKee (2009). Getting the word out: Enforcement information dissemination and compliance behavior. *Journal of Public Economics*, 93, 392–402.

Andreoni, J., B. Erard, and J. Feinstein (1998). Tax compliance. *Journal of Economic Literature*, 36(2), 818–860.

Autor, D. H (2003). Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. *Journal of Labor Economics*, 21, 1–42.

Beck, P.J. and W.-O. Jung (1989). An economic model of taxpayer compliance under uncertainty. *Journal of Accounting and Public Policy*, 8(1), 1–27.

Beer, S., M. Kasper, E. Kirchler, and B. Erard (2020). Do audits deter or provoke future tax noncompliance? Evidence on self-employed taxpayers. *CESifo Economic Studies*, 1–17.

Brinch, C., E. Hernæs, and Z. Jia (2017) Salience and Social Security Benefits. *Journal of Labor Economics*, 35(1), 265–297.

Czibor, E., D. Jimenez-Gomez, and J. A. List (2019). The Dozen Things Experimental Economists Should Do (More of). *Southern Economic Journal*, 86(2), 371–432

DeBacker, J., B. T. Heim, A. Tran, and A. Yuskavage (2015). The impact of legal enforcement: An analysis of corporate tax aggressiveness after an audit. *Journal of Law and Economics*, 58(2), 291–324.

DeBacker, J., B. T. Heim, A. Tran, and A. Yuskavage (2018). Once bitten, twice shy? The impact of IRS Audits on Filer Behavior. *Journal of Law and Economics,* 61(1), 1–35.

Dwenger, N., H. Kleven, I. Rasul, and J. Rincke (2016). Extrinsic and intrinsic motivations for tax compliance: Evidence from a field experiment in Germany. *American Economic Journal: Economic Policy*, 8(3), 203–232.

Erard, B. and J. Feinstein (1994). The role of moral sentiments and audits perceptions in tax compliance. *Public Finance/Finances Publiques*, 49(Supplement), 70–89.

Fortin B., G. Lacroix, and M.-C. Villeval (2007). Tax evasion and social interactions. *Journal of Public Economics*, 91, 2089–2112.

Gemmell, N. and M. Ratto (2012). Behavioral responses to taxpayer audits: Evidence from random taxpayer enquiries, *National Tax Journal*, 65(1), 33–57.

Hashimzade, N., G. D. Myles, and B. Tran-Nam (2013). Applications of behavioural economics to tax evasion. *Journal of Economic Surveys*, 27(5), 941–977.

Hernæs, E. and Z. Jia (2013). Earnings distribution and labour supply after a retirement earnings test reform, *Oxford Bulletin of Economics and Statistics*, 75(3), 410–434.

Iacus, S. M., G. King, and G. Porro (2011). Multivariate matching methods that are monotonic imbalance bounding. *Journal of the American Statistical Association*, 106(493), 345–361.

Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez (2011). Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark. *Econometrica*, 79(3), 651–92.

Luttmer, E. F. P. and M. Singhal (2014). Tax morale. *Journal of Economic Perspectives*, 28(4), 149–168.

Maciejovsky, B., E. Kirchler, and H. Schwarzenberger (2007). Misperception of chance and loss repair: On the dynamics of tax compliance. *Journal of Economic Psychology*, 28(6), 678–91.

Manski, C. and J. V. Pepper (2018). How do right-to-carry laws affect crime rates? Coping with ambiguity using bounded-variation assumptions. *The Review of Economics and Statistics*, 100(2), 232–244.

Mayraz, G. (2011): Wishful thining. CEP Discussion Paper No 1092, Centre for Economic Performance, London School of Economics, London.

Mendoza, J. P., J. L. Wielhouwer, and E. Kirchler (2017). The backfiring effect of auditing on tax compliance. *Journal of Economic Psychology*, 62, 284–294.

Mittone, L., F. Panebianco, and A. Santoro (2017). The bomb-crater effect of tax audits: Beyond the misperception of chance. *Journal of Economic Psychology*, 61(C), 225–243.

Myles, G. D. and R. A. Naylor (1996). A model of tax evasion with group conformity and social customs. *European Journal of Political Economy*, 12(1), 49–66.

Slemrod, J., M. Blumenthal, and C. Christian (2001). Taxpayer response to an increased probability of an audit: Evidence from a controlled experiment in Minnesota. *Journal of Public Economics*, 79(3), 455–483.

Slemrod, J. and M. Keen (2017). Optimal tax administration. *Journal of Public Economics*, 152, 133–142.

Slemrod, J. (2007). Cheating ourselves: The economics of tax evasion. *Journal of Economic Perspectives*, 21(1), 25–48.

Slemrod, J. (2019). Tax compliance and enforcement. *Journal of Economic Literature*, 57(4), 904–954.

Snow, A. and R. S. Warren, Jr. (2007). Audits uncertainty, Bayesian updating, and tax evasion. *Public Finance Review*, 35(5), 555–571.

Torgler, B. and N. T. Valev (2010). Gender and public attitudes toward corruption and tax evasion, *Contemporary Economic Policy* 28(4), 554–568.

Appendix A. Distribution of "Other Deductions"

Figure A1. Distribution of "Other deductions" among audited and non-audited. The year of the audit

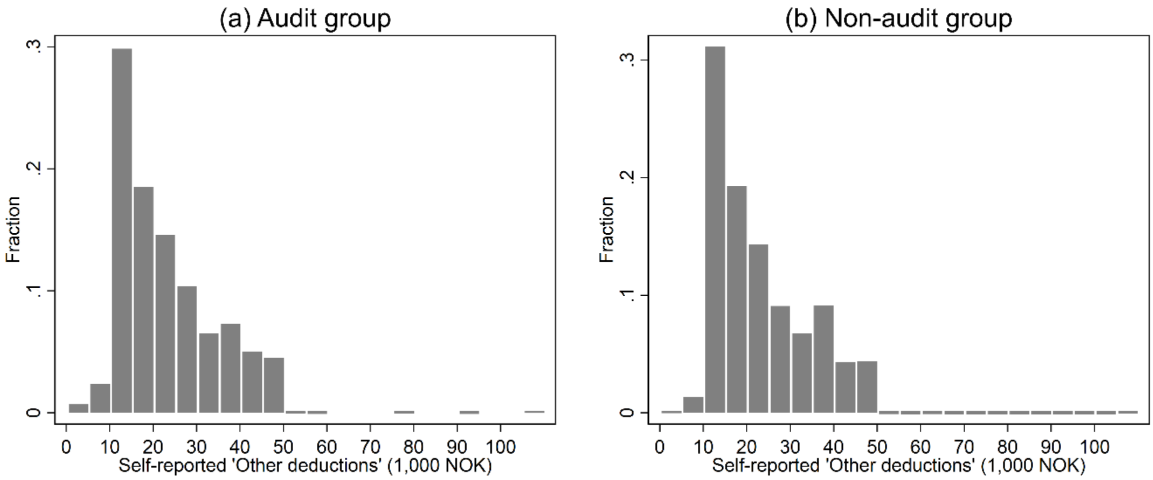
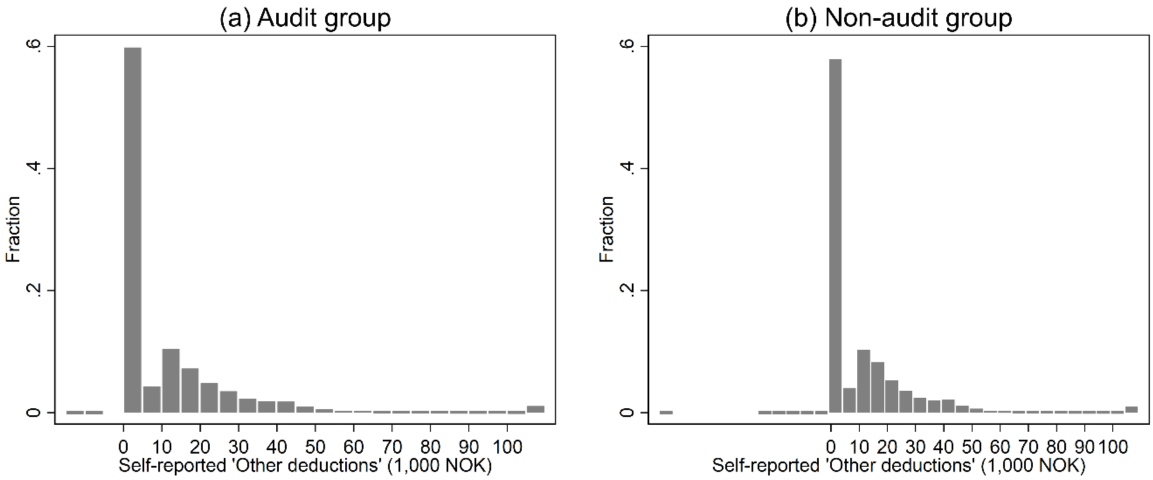


Figure A2. Distribution of "Other deductions" among audited and non-audited. Averages over all years used in the empirical analysis



Appendix B. Distribution of Treatment Effect

As the results indicate that there is a negative shift in the mean deduction after the audit, it is informative to examine to what extent the shape of the deduction distribution has been changed. Following Hernæs and Jia (2013) and Brinch, Hernæs and Jia (2017), we look at the Complementary Cumulative Distribution Functions (CCDF), $\bar{F}(y \mid X) = \Pr(Y > y \mid X)$, before and after audit. In particular, we use a series of logit specifications to model the conditional complementary CDF for a number of values of y. This allows a simple application of the difference in difference technique to identify the treatment effect of the audit.

For any given value of $y \geq 0$, we assume that for individual i at:

(B1) $$\Pr(y_{it} > y_k) = F\left(\alpha_k + X_{it}\beta_k + \lambda_{tk} + \gamma_k D_i + \delta_{tk} D_i \times 1(t>0)\right),$$

where X_{it} denotes individual characteristics and F represents the logit function. We estimate a logit model for y_k varying from NOK 0 to NOK 100,000 by increments of NOK 5,000. The graphical illustration in Figure 2 is based on five separate estimations, one for each of the deduction levels described above. For each estimation, we find the marginal effect of audit evaluated at the covariate value equal to the average of the treatment group. These marginal effects are
equal to the difference in the post-audit and pre-audit probability of a deduction larger than a given level of y: $\Pr(y_{it} > y \mid D_{it} = 1, X_{it}) - \Pr(y_{it} > y \mid D_{it} = 0, X_{it})$.

Figure B1 shows the estimated marginal effects with 95 percent confidence envelops over different deduction levels for the first year after the audit ($t=1$). The effects from the other years are similar but much weaker. The figure shows that the audit affects deduction claims on both the intensive and extensive margin. There are fewer individuals who claim deduction after the audit, and effects on the intensive margin are uneven across deduction levels, with the largest effect observed in the interval $[5000, 25000]$. The corresponding shifts in the probability of being in different intervals of the claimed deduction distribution are reported in Table B1.

Figure B1. Audit effects on the distribution of deductions, the year after audit

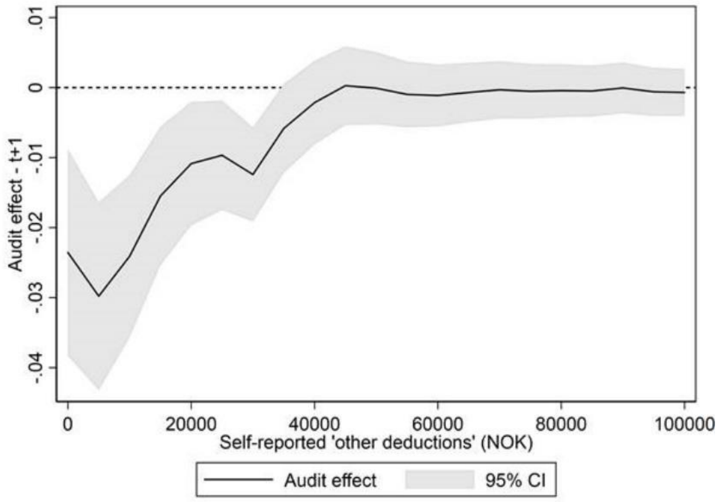


Table B1. Estimate of shift into deduction interval after audit

	Audit effect	
Interval	Estimate	Standard error
No claiming (NOK 0)	0.024	0.008
NOK 0 – NOK 5,000	0.006	0.010
NOK 5,000 – NOK 25,000	-0.020	0.008
NOK 25,000 – NOK 40,000	-0.008	0.005
NOK 40,000 – NOK 50,000	-0.002	0.004
> NOK 50,000	-0.000	0.003

Appendix C: Unintentional Tax Evasion

Model Setup

A main message of the present study is that our results do not seem to comply with intentional fraud behavior. In the following we set up a simple model where we explain the observed differences in deduction claiming behavior before and after the audit observed in the compliant and noncompliant groups. In the model, individuals do not cheat intentionally. The behavior is instead a result of individuals mistakenly claim illegitimate deductions which they are unsure about themselves – for instance due to a lack of understanding of the tax rules.

For a given taxpayer, there is a set of potential tax deductions that they may claim. Among these deductions, there are "risk free" ones that the taxpayer knows correctly to be legitimate. There are also "risky" ones that taxpayer is unsure whether they are legitimate or not. The tax payer chooses which deductions they will claim.

Next, the tax authority conducts a random audit among taxpayers whose claimed deductions are above a given level, denoted as C. While taxpayers are aware that their claimed deductions may be audited. They do not know the rule of the audit selection. Some deductions are limited to one year while others are repeated several years. So we assume that the "risk-free" claim consists of a time invariant part λ_i and time varying part ε_{it} where we assume to be i.i.d. over time and individual with $\mathrm{E}[\varepsilon_{it}] = 0$.

The "risky" claim $u_{it} > 0$ is independent over time and uncorrelated with ε_{it}. The independence assumption is restrictive but not essential for our main results. We assume that there is a subject belief probability $0 < p_i < 1$ that the risky claim is legitimate, which can be seen as a proxy of a self-evaluation of knowledge on the tax system.

The taxpayer will always claim "risk free" deductions. There is, however, a positive cost if the tax payer's claimed deduction is audited and found to be illegitimate. Thus, she will only claim the risky deduction if its amount is above a certain threshold. This threshold should depend on, among others, on two subjective probabilities: the probability of it being legitimate, p_i, and the probability of being audited, q_i. In our model, tax payers may claim when they are not certain about the legitimacy of the claim, and it is argued that the decision is guided by the size of the loss (increased tax burden). Of course, such behaviour can be a result of several misconceptions, such as "wishful thinking" bias; on wishful thinking bias see Mayraz (2011).

Model implications

For simplicity, we have assumed that all tax payers are observably identical. In other words, the implication is valid when we control for observed characteristics. Let us consider a three-period model, $t \in [-1,0,1]$, where $t=0$ is the year of audit and $t=-1$ and $t=1$ are the year before and after the audit, respectively. At any given year t, taxpayer i's claimed deduction is denoted as y_{it}. Among taxpayers, there are two types of individuals. Type I are those who claim only the risk-free deductions ($G_{it}=0$), while Type II individuals claim both types ($G_{it} = 1$). Total claim then can be written as

$$y_{it} = \lambda_i + \varepsilon_{it} + G_{it} u_{it}$$

Implication 1. Noncompliance is incidental, not intentional

We understand immediately that the noncompliant group consists of only type II individuals who claimed "risky" deductions, whereas the compliant group consists of both types. Since u_{it} is uncorrelated over time, then group membership dummy G_{it} is uncorrelated over time. This implies that the compliance behavior is not correlated over time. Namely, being a noncompliant at a given audit gives no additional information on her compliant behavior in years prior to the audit. This result is quite strong, since it rules out intentional fraud and implies that the deduction behavior will be similar for these two groups prior to the audit. This is consistent with what we observe, i.e., that time trends for the deduction claimed prior to the audit are parallel. Note that this is not true for the behavior after audit, since the audit will change the key parameters governing the model, as we will return to below.

Implication 2. Noncompliant group has lower deductions pre-audit

We claimed that the deduction behavior prior to the audit would not be different for the compliant and noncompliant groups. However, we do observe that there are level differences even after we control for observed characteristics. In the following, we will show that this is due to the special eligibility criteria used in the audit we study.

Since we assume that the level of risk-free amount is uncorrelated with the risky amount, we see immediately that

$$E\left[y_{it} \middle| G_{i0}=0\right] < E\left[y_{it} \middle| G_{i0}=1\right].$$

For the audit we study, the taxpayer is eligible to audit only when the total deduction level is above a certain level. This implies that

$$\begin{aligned} &E\left[\lambda_i \middle| G_{i0}=0, \text{ qualified for audit}\right] = E\left[\lambda_i \middle| G_{i0}=0, y_{it} > C\right] \\ &> E\left[\lambda_i \middle| G_{i0}=1, y_{it} > C\right] = E\left[\lambda_i \middle| G_{i0}=1, \text{ qualified for audit}\right]. \end{aligned}$$

Thus, we have

$$E\left[\lambda_i \middle| \text{compliance at } 0\right] > E\left[\lambda_i \middle| \text{non-compliance at } 0\right].$$

Together with the assumption that u_{it} is uncorrelated over time, we have

$$E\left[y_{i,-1} \middle| \text{compliance at } 0\right] > E\left[y_{i,-1} \middle| \text{non-compliance at } 0\right].$$

The intuition is rather straightforward: suppose the there are two individuals who claim the same amount deduction. One is complaint and the other is noncompliant . Since individuals only claim a "risky" deduction when the amount is high, the noncompliant will have lower time invariant risk-free claim than the compliant. Thus, this explains the pattern seen in Figure 1 prior to the audit.

Implication 3. Noncompliant s adjust deductions downward post-audit

As mentioned above, there are mainly two key parameters which define the deduction claim behavior: the subjective belief on his own knowledge of the tax rules, proxied by belief probability, p, and the probability of getting audited, q.

After being found that their "risky" deductions are illegitimate, the tax payers would likely to adjust downward their subject belief probability, p_i. On the other hand, they may adjust upwards/no change on the probability of getting audited, q_i, – it will be irrational to consider the case that the assumption of "bomb crater" would apply here and the tax payer would actually adjust downward q_i. This results in a decrease in the threshold and reduce the claimed deduction after the audit.

Implication 4: Compliants may adjust their deductions in either direction post-audit

For the compliant, the direction is less clear. Assume for now that they are aware of the fact they have been audited and found all deductions are legitimate. It is possible that they will adjust upward the subject belief, p_i. On the other hand, they may also adjust upward the audit probability, q_i. Thus, the overall effect could go either way.

In section 4, the results of the Probit model suggest a systematic difference in the observed individual characteristics (including richer, older, and male who own businesses) probabilities of being found noncompliant . According to the sketch in this appendix, differences across observed characteristics can be explained by differences in the threshold of reporting a "risky" deduction—e.g., richer individuals may tolerate a higher financial loss than poorer or younger individuals who are wage earners. Despite the simplicity of the model, the general implication would hold under less restrictive assumptions such as individual heterogeneity in the probability of legitimacy of risky claims.

Overall, the above-mentioned implications of the model are consistent with what we found in empirical analysis. While we cannot really test the basic assumptions of our model directly against data, the empirical results do show some inconsistencies with the theory that taxpayers evade when they have the chances. What we found points to another possible sources of tax noncompliance behavior, namely the complicated tax rules. Similar problems have been found in other cases where economic policies induce unintended outcomes, see for example Brinch, Hernaes and Jia (2017) for an example in the pension policy.