CESifo
CONFERENCES 2019

11th Norwegian-German Seminar on Public Economics

Munich, 6 – 7 December 2019

Behavioral Responses to Tax Audits: Between Deterrence and Approval

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

20 YEARS CESifo

NHH

OFS Oslo Fiscal Studies

Behavioral Responses to Tax Audits: Between Deterrence and Approval[*]

Shafik Hebous[¤], Zhiyang Jia[§], Knut Løyland[†], Thor O. Thoresen[‡], and Arnstein Øvrum[**]

Preliminary, 28 August 2019

Abstract

Auditing the income tax return not only ensures that the correct amount is paid in the year of the intervention, it is expected to affect the tax compliance in the subsequent years too. A random audit selection scheme operated by the Norwegian tax administration is used to identify magnitude and duration of post-audit deterrence effects. Moreover, we explore to what extent there is a counteracting "approval effect" too, among the taxpayers found to be compliant by the audit. We find estimates in accordance with a modest deterrence effect, however statistically significant only in the first year after the audit. Behind this we see substantially larger effects in the non-compliant group, lasting for five years after the audit. The compliant taxpayers, however, show no signs of behavioral adjustments.

Keywords: Tax compliance, deterrence of audits, administrative data

JEL codes: H26, C23

[*]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 Edvin Leuwen, Monique de Haan, participants of "Modelling tax policy and compliance workshop" (TARC, University of Exeter, 14–15 March 2019) and Skatteforum 2019 (Yearly meeting of Norwegian tax economists, Asker, 12–13 June), Annual Congress of the International Institute of Public Finance (Glasgow, 21–23 August 2019) for comments to an earlier version of the paper.

[¤]International Monetary Fund and Oslo Fiscal Studies (Department of Economics, University of Oslo), email: Shebous@imf.org

[§]Statistics Norway, email: zhiyang.jia@ssb.no

[†]Norwegian Tax Administration, email: knut.loyland@skatteetaten.no

[‡]Statistics Norway and Oslo Fiscal Studies (Department of Economics, University of Oslo), email: thor.olav.thoresen@ssb.no

[**]Norwegian Tax Administration, email: arnstein.ovrum@skatteetaten.no

1. Introduction

Tax audits hold the promise of enhancing tax compliance. Although effects of audits on tax compliance have been extensively studied, see Andreoni, Erard and Feinstein (1998), Slemrod (2007; forthcoming), and Alm (2019), there is less information about how effects develop over time. From a perspective of efficient tax administration, it is crucial to obtain information on to what extent audits deter taxpayers from non-compliance also in the years after being undergoing inspection. In the present study we have had access to an administrative dataset of the Norwegian Tax Administration, which implies that we follow a large number of individual taxpayers before and after being audited. Thus, this analysis adds to the relatively scarce literature that discusses effects of audits in a longer time perspective.

A clear advantage of the present study is that data are generated by a purely random audit procedure. The schedule is as follows. Most Norwegian taxpayers receive a complete prefilled income tax return based on an extensive third-party reporting system. This means that many taxpayers do not need to make amendments, just confirm that they accept.[1] However, although the information base is wide-ranging, not all items are third party reported, and taxpayers have scope for entering information missing from the form. With respect to deductions, there is an item on the income tax return form named "Other deductions", used by both wage earners and self-employed for claiming deductions not already recorded. This particular item has been subject to further inspection by the tax administration, which has generated the data set of the present study. Critically, and as already noted, both the selection process and the assignment into treated and non-treated are random. Around 10 percent of taxpayers with deductions under "Other deductions" in the range from 5,000 to 50,000 Norwegian kroner (NOK) were assigned to further control. The data set generated by these audits includes information on all taxpayers eligible for the inspection, not only the selected, and all are observed in up to six years after the control.

The present analysis uses data for approximately 30,000 individuals from these administrative inspections in 2009, 2010 and 2011, adopting an experimental empirical design by letting the year of the audit represent a demarcation line. The treatment refers to the attention given to taxpayers in terms of being notified and informed that their deduction claim has been further investigated. Given the random assignment into the treated and non-treated groups, our main estimate corresponds to an average treatment effect on the treated (ATT), comparing the difference in outcomes with and without treatment in a randomly selected group of taxpayers reporting "Other deductions" within a specific range. Obviously, and as we soon will return to, the auditing generates two groups among the audited, the "compliant" and the "non-compliant", where the latter group consists of taxpayers whose income tax return have been examined and dismissed due to irregularities on the item "Other deductions".

[1] In fact, if the person (for some reason) does not review it, it is regarded as filed and accepted. The extensive third-party reporting scheme means that information goes directly from, for example, employers (wage income), banks (interest income and wealth), and charitable organizations (donations are deductible in income).

These taxpayers may get fined, but, in practice, because of the magnitude of the criminal act, they are not. This also points to the fact that it can be questioned to what extent the behavior results from deliberate criminal behavior or simply follow from misunderstandings of the tax code. In any case, the intervention from the tax authorities comes in the form of "attention" more than inflicting penalties. In the "compliant group" we find those who have been checked and cleared. Importantly, most compliant taxpayers are also informed about the inspection, as they normally are asked about further documentation of their claims.[2]

Firstly, this study presents estimates of the average effect of audits, with estimates derived from a simple treatment-control empirical design, utilizing the panel dimension of the data.[3] It is the random assignment of taxpayers into audit that defines the key to obtain non-biased estimates. It follows that an investigation along these lines holds the promise of contributing to the understanding of one of the most central question of the literature on effects of auditing, namely to what extent attention by the taxman deters or encourages non-compliance in the years after inspection. The informational content of an audit can be referred to a type of Bayesian updating, see Snow and Warren (2007), i.e., the experience leads to a revision of beliefs. The literature refers to both a "target effect" (Hashimzade, Myles and Tran-Nam, 2013), which means that audits deter under-reporting because the agents perceive that chances for another inspection is high, and "a bomb-crater" effect (Maciejovsky, Kirchler and Schwarzenberger 2007; Mittone, Panebianco and Santoro, 2017), a "bomb" would not strike exactly the same place again.

Further, the auditing itself reveals information about the treated, in that it divides them into compliant and non-compliant taxpayers, which we shall utilize to further add to the understanding of how people react to the attention from the tax authorities. Thus, in the second part of the empirical investigation we discuss the behavioral responses to audits by expanding on separate effects of "negative" and "positive" attention. Although audits may discourage future illegal activities for taxpayers who have been caught evading, the experience of being checked and cleared may generate other reactions. The key to obtain credible estimates of subgroup behavior is that the deduction trend prior to the audit in both groups is close to what we observe in the non-audited group. We shall use a difference-in-differences (DID) estimation technique to obtain subgroup estimates.

With respect to the second part of our study, the focus on subgroups of the treated, two studies are very close to the empirical design of the present study, as they address the behavior of both compliants and non-compliants – Gemmell and Ratto (2012) and Beer et al. (2015). Both studies find that the non-compliant taxpayers increase their subsequent compliance after an audit. But notably,

[2] We are aware that there are some taxpayers in the compliant group who are not informed about their assignment into auditing. These taxpayers essentially behave like taxpayers in the non-treated group and therefore represents a bias towards the mean.

[3] As a matter of terminology, note that the "treatment" of the present study is a "control".

both analyses find evidence suggesting that the compliant taxpayers react in the opposite way, as they become less compliant after being audited.

The paper is organized as follows. In Section 2 we survey the literature on responses to audits, whereas Section 3 describes in more detail the institutional setting that has produced the audit data exploited here. In Section 4 we present the empirical framework and estimation results for the average overall effect of audit, whereas Section 5 presents regressions results when the audited are divided into compliant and non-compliant taxpayers. Several robustness tests are included. Section 6 concludes.

2. Responses to attention from audits

Audits influence the tax collection directly, as it means that additional revenue is collected from people not abiding by the rules. Here, the attention is on the effects of audits in a longer time perspective – to what extent the taxpayers are deferred by the tax audit, and thereby change their subsequent compliance behavior, often referred as the indirect effects of audits.[4] As the previous literature finds that the indirect effects outweigh the direct effects (Dubin, 2007; Ratto, Thomas and Ulph, 2013), from an efficiency of tax administration point of view it is imperative to enhance the knowledge about magnitudes.

In the Allingham-Sandmo model (Allingham and Sandmo, 1972), as summarized in Sandmo (2005), we have that net income, *Y*, of the taxpayer is defined by $Y = W - t\left(W - E\right)$, where *W* is gross income, *E* is the amount of underreporting, and *t* is the proportional tax rate. Given that there is a penalty paid on the evaded tax (Yitzhaki, 1974), the net income, *Z*, can be seen as $Z = \left(1-t\right)W - \left(\theta - 1\right)tE$, where $\theta(>1)$ is the penalty rate. As the taxpayer's subjective probability of detection is *p*, and he maximizes $V = \left(1-p\right)U\left(Y\right) + pU\left(Z\right)$, the first order condition becomes $\frac{U'\left(Z\right)}{U'\left(Y\right)} = \frac{1-p}{p\left(\theta-1\right)}$.[5] The main subject of the present analysis is to discuss how the amount of underreporting, *E*, is determined in Period 2, when the taxpayers have received attention in Period 1. Given that there are no penalties involved in our case, the main line of reasoning builds on the attention of the tax administration influences the perception of *p* in Period 2. Snow and Warren (2007) refer to this process as Bayesian updating, since the detection probabilities are updated based on the experience from Period 1.

[4] This effect may also be characterized as "corrective" (Gemmell and Ratto, 2012) or a "direct deterrent effect" (Alm, Jackson and McKee, 2009). The "indirect deterrent effect" then refers to spillover effects on the non-audited, see discussion of spillover effects in for example Fortin, Lacroix and Villeval (2007).
[5] It should be noted that this model confronted with values of fines and audit probabilities, in most cases, overshoots the level of tax evasion observed. It is a general understanding that other explanations must be added to the framework (Hashimzade, Myles and Tran-Nam, 2013).

Given this theoretical set-up, there are two important characteristics of the present case that represent deviations from the reasoning. Firstly, in general, the tax administration will not issue fines to the non-compliant taxpayers, since the gravity of the criminal act must be characterized as modest. However, as it is generally acknowledged that tax evasion involves more than amoral cost-benefit calculations of the agents, we believe that there are other (intrinsic) motivations for being compliant. The basic model has been extended in several directions, including accounting for moral sentiments of guilt and shame (Erard and Feinstein 1994) and social conformity effects (Myles and Naylor 1996; Fortin, Lacroix, and Villeval 2007). Thus, we believe that the mechanism of deterrence (if there is one) works even without issuing actual fines. Secondly, a key characteristic of the updating of the present study is that we have two types of taxpayers in Period 2: taxpayers who in Period 1 have been found to comply and not to comply, respectively. As we soon shall return to, these two groups are not expected to show the same response to the attention given to them from the tax administration.

The literature is dominated by studies which do not separate between compliant and non-compliant taxpayers. With respect to the behavior of the audited in general, there are mixed results on post-audit behavior (Alm, 2019). Although overview studies report that there is a tendency to reduce tax evasion, compatible with deterrence, effects are small or even negligible (Andreoni et al., 1998; Gangl et al., 2014). One potential reason is that tax enforcement may crowd out the intrinsic motivation of paying taxes (Luttmer and Singhal, 2014; Dwenger et al., 2015; Mendoza, Wielhouver and Kirchler, 2017). However, Slemrod, Blumenthal and Christian (2001) find evidence in accordance with the most intuitive mechanism – a random sample of taxpayers report higher income after being warned about future close examination of their income returns. This can be characterized as a target effect (Hashimzade, Myles and Tran-Nam, 2013), simply stating that agents perceive that an audit is likely to be followed by another in the future. Along the line of Tversky and Kahneman (1974), a possible explanation is that individuals' increase their awareness of audits, the decision-makers are inclined to believe that an event is more likely to happen if the event (attention of the tax administration) is easily retrieved from the memory. In contrast, some studies refer to a "bomb-crater effect" (Maciejovsky, Kirchler and Schwarzenberger, 2007; Mittone, Panebianco and Santoro, 2017), referring to individuals perceiving that the risk of being audited again falls immediately after an audit, because "a bomb" would not strike exactly the same place again (hence soldiers in WW1 hide in bomb craters). This can be explained by misperception of chance – in this case it is expected that a random event (audit) is less likely to happen again if it just happened.

With respect to the distinction between compliant and non-compliant taxpayers among the audited, the post-audit behavior of the compliant may be particularly hard to predict, see Beer et al. (2015). One may reasonably expect that the taxpayers in this group just simply continue their deduction behavior, as before. However, there are reasons for finding more illegal behavior in this group after audit. For example, the previously compliant taxpayers now infer that the risk of future checks is low and thereby exploit the information provided by the tax administration to decrease the subsequent tax burden. Further, as already discussed, given that it is widely established that taxpayers are motivated by intrinsic motivation, auditing level may signal distrust in taxpayers and lead to the perception that the tax authority and its enforcement actions are excessive and unfair, which in turn lead to reduced compliance. Also, if the compliant in our data in reality are non-detected non-compliant taxpayers, they may become even more motivated to continue their illegal activities by the lack of detection.

Effects of audits in a longer time perspective are both discussed in studies based on laboratory experiments and analyses of actual data. With respect to the former type of studies, Alm, Jackson and McKee (2009) use laboratory experiments to examine the compliance impact of types of information dissemination regarding audit frequency and find results conditional on whether the taxpayer is well informed about the audit rate prior to filing. Results suggest that it would be advantageous to pre-announce audit rates. Turning to studies that discuss compliance behavior by analysis of tax data, both Dubin, Graetz and Wilde (1990) and Dubin (2007) show strong positive effects on compliance after the IRS (Internal Revenue Service) has shown interest in the filing of taxpayers. Newer studies confirm that audits or other types of interventions seem to deter taxpayers from non-compliance.[6] However, the results of both DeBacker et al. (2018) and Advani, Elming and Shaw (2019) point to relatively moderate effects. Like as in Kleven et al. (2011), both studies report that effects are stronger with respect to self-reported income components than third-party reported income. Moreover, both analyses find that effects last for about 5–6 years.

There are two empirical investigations, Gemmell and Ratto (2012) and Beer et al. (2015), that are particularly close to the present study in that they distinguish between effects of compliant and non-compliant taxpayers. Gemmell and Ratto investigate the response of U.K. of a sample of taxpayers to randomly selected to audits in the year 2000, consisting of both business owners and "personal" taxpayers. If the direct yield (additional tax plus penalties) is positive, the taxpayer is classified as non-compliant, whereas no change defines compliance. The responses are measured in terms of comparing declared tax in three years before the audit to three years after the audit, using difference-in-differences regression analysis to identify effects. Whereas all the involved are informed about the audits, in some cases the closure of the inquiry is well into the post-audit observation period, which is found to influence results. Moreover, and important with respect to the present study, two specifications are estimated: one measuring the overall effect on the audited, and another allowing for separate effects of the non-compliant and the compliant taxpayers. In the latter case, the effects of the two groups are estimated jointly, which implies that the same group of taxpayers (the non-audited) are used as control group for the two treatment groups.[7] Controlling for individual fixed effects, observable and unobservable characteristics independent of time, is claimed to account for pre-audit

[6] However, DeBacker et al. (2015) find that U.S. firms reduce tax payment immediately after an audit and then increase it gradually to the pre-audit level.
[7] As we soon shall return to, this empirical strategy is also followed here.

differences between the three groups. Results suggest that taxpayers are deferred by audits, as audited taxpayers who were found to be non-compliant increased their subsequent compliance. But there are indications that approved behavior may give reduced compliance after the audit. Thus, the authors write (p. 55): "These results serve to highlight the importance of testing for the responses of the so-called compliant and noncompliant subgroups separately to avoid conflating their different responses."

Interestingly, Beer et al. (2015) find that effects of audits differ with respect to compliance and non-compliance in the same manner as reported by Gemmell and Ratto (2012). Beer et al. use the impact of enforcement activity and subsequent compliance income reporting of sole proprietors in the U.S. in their analysis, comparing a random sample of filers who were audited after filing in 2007 to a control sample of non-audited under the same schedule (Schedule C filers). Thus, this means that the selection of audited taxpayers is operational rather than random, which is important when interpreting the magnitudes of the effects. The identification strategy is close to the approach of Gemmell and Ratto (2012), in that they employ a difference-in-difference empirical design. Three years after the audit a positive reporting effect is seen among the non-compliant taxpayers, in accordance with a deterrent effect, whereas audits have a detrimental impact on the reporting behavior of taxpayers who do not experience an additional tax assessment – see above for possible explanations. Moreover, the sensitivity tests include both specifications with selection on both observables and unobservables, employing matching techniques and the selection model of Heckman (1978), respectively.

3. Third-party reporting, random audits and filters

Before we present the empirical approach and results, we shall briefly refer to the institutional setting from which the data are derived. The Norwegian third-party information schedule has been developed to the extent that most taxpayers are not in contact with the tax administration before receiving the form for approval: most income tax returns are prefilled based on third-party information from employers, the financial sector and others. For example, as donations are deductible up to a threshold in the taxation of ordinary income in Norway, the recipients of donations (as the Red Cross) report the individual donations directly to the tax authorities. The taxpayer is therefore usually presented to a complete digital prefilled income tax return and is asked to approve before filing. Of course, if he finds errors or incomes/deductions not reported, he makes amendments. As the Norwegian tax administration is well underway to make the tax filing system fully digitalized, the income tax return is typically electronically filed, and, accordingly, most people make amendments by addressing their income tax return directly through the internet. Thus, there is usually no administrative staff of the tax authorities involved.

In the present analysis we focus on an item on the income tax return which is frequently used to report additional deductions, the "Other deductions", used for claiming deductions not already recorded through other items of the third-party reporting system. For example, judicial help to be able to earn an income in the labor market or assistance by a stockbroker (for capital income) can be deducted here. A problem from a tax administrative point of view is that any taxpayer can make substantiated or unsubstantiated claims, in the latter case using this item to illegally reduce the tax burden. Thus, the claimants are exposed to audits.

In general, the tax administration employs a whole range of filters to select individuals for audits. With respect to this particular item, the selection is based on establishing a pool of taxpayers who have claimed "Other deductions" in the range from 5,000 to 50,000 NOK. From this sample, a subsample of approximately 10 percent is randomly assigned for further investigation. In the present study we utilize information from this type of auditing for the years 2009, 2010 and 2011.[8]

From an empirical identification perspective, it is important to note that not only the non-compliant are informed about the attention form the tax administration, but also most of the compliant taxpayers.[9] Of course, the non-compliant taxpayers are informed since the tax authorities adjust their income tax return.[10] Some of the compliant taxpayers may go through the process without notification, if they already have provided all the necessary documentation needed. However, most of them would have been asked to provide additional information, which implies that they are aware of the attention of the tax administration.

4. Average effect of audit

4.1. Assignment to treatment is random

The key identifying feature of the analysis is that the audit is completely random. This means that identification of effects can be obtained by techniques associated with natural experiments. Since we have observation of taxpayers belonging to the pool of taxpayers in danger of being exposed to audit, independent of being controlled or not, we overcome the main problem of individual *i* not being both treated and non-treated, by letting the non-treated represent the counterfactual. We define a binary variable, D_i, as

$$D_i = \begin{cases} \dfrac{1 \text{ if individual } i \text{ receives treatment}}{0 \text{ otherwise}} \end{cases}. \quad (1)$$

Further, let the deduction behavior of the treated and non-treated be symbolized by y_{0i} and y_{1i}, respectively. The effect of the audit can then be seen as an average treatment effect of the treated (ATT),

[8] After 2011 the tax administration changed the audit procedures.
[9] Obviously, effects may depend on how much information that is conveyed to the agents. For example, taxpayers in the U.S and Denmark, randomly selected for audits, are informed about the investigation being based on random selection.
[10] In some grave cases the taxpayers may be fined too, but such incidences would be rare given the amounts involved here.

(2) $$E\left[y_{1i}\,\middle|\,D_i=1\right]-E\left[y_{0i}\,\middle|\,D_i=0\right]=E\left[y_{1i}-y_{0i}\,\middle|\,D_i=1\right].$$

This simple framework relies on the behavior of the non-audited representing the counterfactuals for the audited, i.e., how the treated would have behaved if not being audited. Given the random assignment to auditing, we find this assumption justifiable. Note that so far we have not distinguished between compliant and non-compliant members of the audited group; we return to this in Section 5.

4.2. Data descriptions

In Table 1 we present the number of observations we have had available for this study, allocated on different groups. As the data are collected from three different audits, in 2009, 2010, and 2011, and the observation period starts in 2008 and lasts until 2015, we observe the taxpayers in minimum three years and in maximum six years after the audit. The table shows that approximately 30,000 taxpayers qualified for audits in the three years, from which approximately 3,500 have been assigned to audits by the random selection. As the demarcation line is the year of the audit, the number of observations in each year varies: if the individual is audited in 2011, we observe he/she in three years before the audit and in four years after the audit, whereas a person audited in 2009 is observed in six post-audit years, but only in one year before the audit.

Table 1. Number of observations in samples

	Number of individuals in panel	Number of observations
Non-treated	26,775	197,396
Audited	3,476	22,646
Compliant	2,238	14,338
Non-compliant	1,238	8,308

Table 2 presents descriptive statistics for the two groups. As expected, given the random assignment, there are no clear differences in characteristics between the two groups. We note that the average claimed deduction is around 23,000 NOK in both groups.[11] Figure A1 and Figure A2 in Appendix A shows how the deductions are distributed among the audited and the non-audited, both in the year of the audit and across all years used empirical investigation. After the audit, the tax authorities have decided to reduce deductions by approximately 9,000 NOK on average among the audited taxpayers, due to non-verified claims.

[11] Use exchange rates for 2015 to convert to euros and US dollars: 1€=NOK 8.95 and 1$=NOK 8.07.

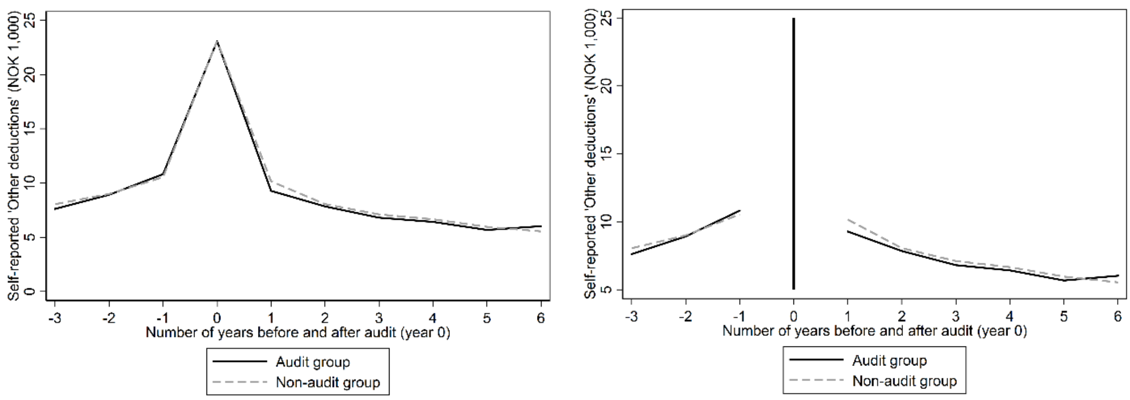
Table 2. Descriptive statistics. Average figures for audited and non-audited, 2008–2015

	Audited	Non-audited
Claimed "Other deductions"	23,055	23,101
	(11,502)	(11,290)
Direct correction in "Other deduction" due to audit	8,993	–
	(20,162)	
Self-employed	0.210	0.208
	(0.408)	(0.406)
Female	0.285	0.290
	(0.451)	(0.454)
Age	48.9	48.4
	(13.0)	(13.0)
Temporary work migrant	0.048	0.060
	(0.213)	(0.237)
Total deductions	216,280	216,652
	(209,417)	(205,798)
Third-party reported deductions	143,635	141,245
	(113 261)	(108,897)
Total gross income	774,454	772,372
	(747,502)	(736,564)
Total third-party rep. gross income	608,880	607,771
	(483,927)	(484,428)
Observations	3,476	26,775

Note: Standard deviations in parentheses

Next, in Figure 1, the trend in the deduction behavior is shown. Recall that as the data are from the period 2008–2015 and that there were audits in 2009–2011 (year 0), we observe the taxpayers up to three years before the audit and a maximum of six years after. Note that the sample selection rule applied by the tax authority implies that only individuals whose deduction values lie in a given interval are selected into our analysis. Thus, the distribution of deduction values in year of the audit (t_0) is substantially different from the rest of the years, as clearly depicted by the left-hand diagram of Figure 1.

Figure 1. Mean deductions before and after audit, audited and non-audited



Note: 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.

Including data from the year of the audit or not in the empirical analysis would not likely influence results, as this year represents a similar deduction outlier for both the audited and the non-audited. We have decided to leave out data of the year of the audit in the empirical investigation, which means that the right-hand diagram is representative for the data employed in the study.

Moreover, as already discussed in Section 4.1, in order to obtain estimate of the average treatment effects we do not need to employ data before the audit. Given the random assignment, one could simply compare the sample means of the audited and non-audited group to obtain an estimate of the average treatment effect. However, as we soon shall return to, we employ a DID approach also for the estimation of the average treatment effect of the treated.

4.3. Estimates of the average treatment effect

The overall post-audit effect of the audit on the deduction behavior, y_{it}, given that we observe the taxpayers in up till six years after the audit, can be obtained by the following DID set-up

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

where s measures the distance in years to the year of the audit, $s = 0$. Recall that the year of the audit is excluded from the data set used in the estimation. Thus, δ_s measures the treatment effect of audit at year s after the audit. As in Equation (1), the binary regressor, D_i, takes the value 1 for the audited. α_i measures the individual fixed effect and λ_t represents the calendar year effect.

When we estimate Equation (3) for all post-audit years together, we obtain an estimate of -750 NOK.[12] Thus, on average the audit resulted in a reduction in claimed deductions of 750 NOK. As a tentative estimate of the overall effect of this type of audit, one may use this estimate as a yearly average effect per year for the first six years after the audit, obtaining a measure for the increased tax payments resulting from it. Recall that approximately 3,500 taxpayers have been checked in the present study, see Table 1.[13]

Next, we estimate Equation (3) separately for each post-audit year ($s = j \in \{1, 2, ..., 6\}$); estimation results are reported in Table 3. Although all point estimates suggest that taxpayers reduce their reporting of deductions because of the audit, effects are only statistically significant[14] in the first year after the audit. The estimate of the first post-audit year picks up the difference between the audited and

---

[12] The estimate is statistically significant at the 0.1 level.

[13] Of course, an audit generates revenue for the year of the audit too, due to the so-called direct effect, see Table 2. The main contribution of the present study is to provide estimates of the after-audit effects. A very simple "back-of-the-envelope" calculation suggests that due to this particular audit (with filters as described in Section 3), tax revenue increased by 4.4 million NOK. This follows from the tax base (after deductions) being taxed by a flat rate of 28 percent: $(750 \times 6 \times 3500)0.28$. In addition there could be spillover effects on other items, which should be taken into account; we will return to this in Section 5. There are also possible "network effects" influencing the overall deterrence effect, obviously hard to measure.

[14] Recall that due to our empirical design, there are less observations in the last years of the period, which likely contributes to the lack of statistical significance. We have also carried out tests in terms of tests of differences between mean values in the two groups, and find large *t* score for the first year after the audit and small scores in the other years.

the non-audited after year, as seen in Figure 1. Although estimates for $s>1$ are not statistically significant, the point estimates depict a pattern which is in accordance with a diminishing deterrence effects after the audit, which is accordance with findings of other studies, see for example Advani, Elming and Shaw (2019). In Section 5 we shall see that this average treatment effect is composed of differing behavioral changes from taxpayers who been found compliant and non-compliant by the audit.

Table 3. Effects of 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.
$*p < 0.1, **p < 0.05, *** p < 0.01$

5. Distinguishing between compliant and non-compliant taxpayers

5.1. Extended information on the audited

The auditing process generates two distinctively different groups among the treated: those who have been caught not reporting correctly, the non-compliant, and, at the other side, those who can substantiate that their claims are correct and therefore get cleared (compliant). Thus, the reactions we see in the treatment group after the audit is made up by reactions in two different subsets. Simply measuring the average effect, as seen so far, would therefore conflate informative evidence about differences between the two groups, as also discussed by Gemmell and Ratto (2012) and Beer et al. (2015).

Table 4 shows descriptive statistics for the audited, when distinguishing between compliant and non-compliant taxpayers, and reveals some notable differences between the two subgroups of the audited. For example, we see that the share of females seems to be larger among the non-compliant than among the compliant. Given that most evidence, both results come from lab experiments (Alm, Jackson and McKee, 2009) and from sample surveys (Torgler and Valev, 2010; Nygård, Slemrod and Thoresen, 2019), suggest that females are more compliant than males, this is not expected. However, it may indicate that the deduction behavior among the non-compliant in reality is explained by misunderstandings of the use of the item rather than plain criminal activity. The higher ratio of self-employed among the compliant taxpayers may point in the same direction: one could argue that the self-employed, due to their more complicated income tax returns, have a higher level of understanding.[15]

Table 4. Descriptive statistics. Average figures for compliant, non-compliant and non-audited, 2008–2015

	Compliant	Non-compliant	Non-audited
Claimed "Other deductions"	22,275	23,153	23,104
	(8,914)	(10,463)	(11,290)
Direct correction in "Other deductions" due to audit	–	25,948	–
		(22,182)	
Self-employed	0.25	0.14	0.21
	(0.293)	(0.243)	(0.406)
Female	0.25	0.34	0.29
	(0.458)	(0.490)	(0.454)
Age	52	43	48
	(12.5)	(12.9)	(13.0)
Temporary work migrant	0.034	0.073	0.060
	(0.199)	(0.243)	(0.237)
Total deductions	232,364	191,417	217,410
	(79,232)	(70,882)	(205,798)
Third-party rep. deductions	153,390	128,443	141,596
	(52,507)	(46,071)	(108,897)
Total gross income	859,141	617,731	770,830
	(355,130)	(243,154)	(736,564)
Total third-party rep. gross income	655,057 (296,625)	526,447 (235,434)	607,691 (484,428)
Observations	2,238	1,238	26,775

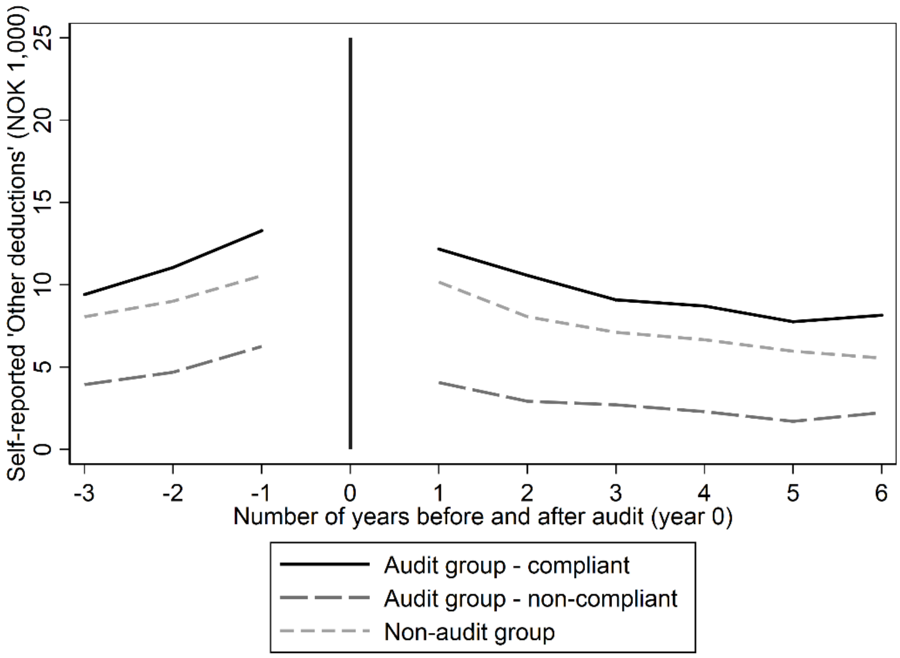
Note: Standard deviations in parentheses

Figure 2 confirms that the post-audit behavior of the two groups is different, but it also shows that this pattern is relatively parallel to the deduction behavior before the audit. Figure 2 demonstrates that the non-compliant taxpayers move into the group exposed to audit from a lower average deduction level, compared to the compliant, and this pattern is repeated after the audit.[16] Does this represent a challenge for the identification of sub-group effects? The problem boils down to a concern about the non-audited representing a valid the counterfactual for both groups. In this perspective Ratto and Gemmell (2012), Beer et al. (2015) and Advani, Elming and Shaw (2019) raise the question to what extent one can use the whole group of non-audited as defining the counterfactual behavior when obtaining results for the compliant group and the non-compliant group, respectively. Instead one would like to see post-audit deduction behavior for the compliant and non-compliant when the behavior in the two groups is compared to compliant and non-compliant types among the non-audited, respectively. But, of course, these two latent groups are not easily identified, as allocation primarily

[15] Although it is not clear whether high knowledge increases or decreases tax compliance (Alm, 2019).
[16] One may speculate how this pattern develops. One possibility is that there are people in the two groups with different "risk patterns". For example, could the non-compliant taxpayers have lower "permanent claims", but increase their claims more when they want to take advantage of a non-verified claim. [this can be further developed]

will be based on unobservables. As a response to this, Advani, Elming and Shaw (2019) suggest estimating effects based on the treated only, as they let the different groups identified by the audit be compared to themselves prior to the audit. This method overcomes a possible endogeneity problem, but the method only allows them to identify the effect for the compliants, not that of the non-compliants.

Figure 2. Deductions before and after audit. Compliant and non-compliant taxpayers compared to the non-treated



Notes: Deductions of the year of the audit not reported in the figure

In the following we further discuss to what extent employing a "mixed" control group, consisting of both (hypothetical) "compliant" and "non-compliant" taxpayers, leads to inconsistent estimates of the type-specific audit effects. To do this, we first introduce notations. Let $Q_i = 1$ denote that individual i is of type non-compliant and $Q_i = 0$ if the person is the compliant 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 non-compliant and compliant taxpayers, respectively, using all the non-audited as the control group. Denote $\Delta Y_i(Q_i, D_i)$ as the difference between the post-audit and the pre-audit deduction of individual i the auditing. Then the DID estimator for the non-compliant group can be written as:

$$\begin{aligned}
\delta_1 &= E[\Delta Y_i(Q_i = 1, D_i = 1) - \Delta Y_i(D_i = 0)] \\
&= E[\Delta Y_i(Q_i = 1, D_i = 1)] - E[\Delta Y_i(D_i = 0)] \qquad (4) \\
&= E[\Delta Y_i(Q_i = 1, D_i = 1)] - pE[\Delta Y_i(Q_i = 1, D_i = 0)] - (1-p)E[\Delta Y_i(Q_i = 0, D_i = 0)],
\end{aligned}$$

where p is the probability for individual i being a non-compliant taxpayer. Given the random assignment of audit, it can be consistently estimated. It is easy to see that 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. We shall return to the use of these expressions when establishing bounds of the group-specific effects in Section 5.4.

The reason for us, at least as a start, to let the non-audited represent the counterfactual for both groups is that the common trend assumption seems to be fulfilled for both, see Figure 2.[17] Following Autor (2003), we have checked this more formally 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 non-compliant taxpayer, denoted by Q_i (as established after the audit), [this and other equations must be checked]

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

where α_i measures the individual fixed effect and λ_t represents the calendar year effect. In practice, we estimate Equation (5) by letting alternate years, $s \in \{-1, -2, -3\}$, be omitted. As we obtain clearly non-significant estimates of κ, π and ξ for all pre-audit years, we conclude that there is no strong support for the rejection of the common trend. Thus, we present evidence for a common control first.

5.2. Effect of audit on subgroups

It follows that we employ the same specification as seen in Equation (3) to estimate average effects when the treatment group is divided into two subgroups. The specification is simply an extension of Equation (3) as we introduce a further distinction in the post-treatment years, dependent on the taxpayers being able to verify his/her claim for deductions. Thus, we let subscript $j \in \{0,1\}$ denote that the treated belong to the subgroups 0 and 1, compliant and non-compliant taxpayers, for which the regressions have been done separately,

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

It follows that estimates of effects of audits, δ_{sj}, now measure average effects of audits in the compliant and non-compliant subgroups.

The results described in Table 5 suggest that there are large differences between taxpayers being told to adjust their claims and those experiencing that they have been cleared. Whereas the compliant taxpayers do not alter their deduction behavior after the audit, the non-compliant reduce their deductions substantially. Thus, this gives support for audits having no approval effect but a clear deterrence effect on the non-compliant is seen. The latter materializes despite, as noted above (see

[17] Although, it must be admitted that the pre-reform period is short.

Section 2), these taxpayers have not been fined because of their unverified claims. Moreover, the estimates of Table 5 give support to the deterrence effect diminishing over time, and after six years the effect is no longer significant.

As noted in Section 3, we are not confident that all taxpayers belonging to the compliant-group have been aware of the audit. To the extent that this fact represents a substantial contribution to bias, these taxpayers will likely behave as belonging to the non-audited, and therefore potentially weaken the observed effects. With respect to the results for the compliant in Table 5, we see that point estimates shift between positive and negative values, which means that it is not clear to which direction a potential bias contributes.

Table 5. Effects of audit on post-audit deduction behavior. Compliant and non-compliant 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
Non-compliant	First	δ_{11}	-2,876***(589)	-4.88
	Second	δ_{12}	-1,858***(589)	-3.15
	Third	δ_{13}	-1,091*(622)	-1.75
	Fourth	δ_{14}	-1,045*(602)	-1.74
	Fifth	δ_{15}	-1,219**(611)	-1.99
	Sixth	δ_{16}	-405(740)	-0.55

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

5.3. Spillover effects on other items

As noted in Section 4.3 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 some extent explore this issue further, we have estimated Equation (6) when 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. The average figures reported in Table 4 may indicate such effects. However, we find no indications of spread to the gross income reporting.[18]

[18] Results are not reported here, but are available upon request.

One may object that the third-party reporting schedule of Norway (see Section 3) prevent us from observing any behavioral effects along this channel, simply because most taxpayers do not control their income reporting. Given that the self-employed have wider scope for such behavioral adjustments we have estimated Equation (6) (with gross income as the dependent variable) only for the self-employed part of the sample. No statistically significant estimates are obtained when restricting to the self-employed. Most noteworthy, we see a clearer approval effect among the self-employed, in terms of the point estimates, than as reported (for all) in Table 5.

5.4. Employing 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. In the case when the common trend fails, we could use Equation (4) and some additional assumptions to derive bounds for the true audit effects for the compliants (γ_0) and non-compliants (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 non-compliant 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 on would expect results to move when not relying on a subgroup common trend. Then, one should be aware that the first assumption, that the non-compliant 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 non-compliant, δ_{1s} in Table 5. Moreover, the second assumption rules out the so-called bomb-crater effect, see Section 2.

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 the non-compliants:

(7) $$\gamma_1 = \delta_1 - (1-p)(E[\Delta Y_i(Q_i = 1, D_i = 0)] - E[\Delta Y_i(Q_i = 0, D_i = 0)]) \leq \delta_1.$$

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

Thus, we bound the true treatment effect for the non-compliants as

$$\frac{\delta_1}{p} \leq \gamma_1 \leq \delta_1.$$

For the compliants we correspondingly have,

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

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

$$p\gamma_1 + (1-p)\gamma_0 = \text{ATT},$$

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

$$\gamma_0 = \frac{\text{ATT} - p\gamma_1}{1-p} \leq \frac{\text{ATT} - \delta_1}{1-p},$$

which follows from $\gamma_1 \geq \frac{\delta_1}{p}$.

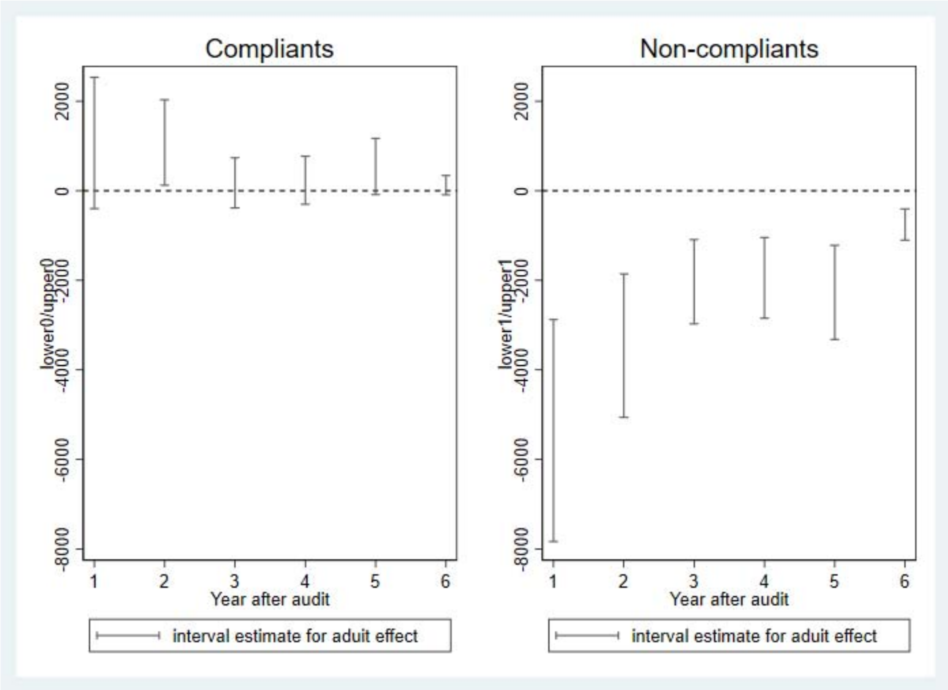
Under these assumptions the bounds the type-specific audit effects can be seen as

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

Hence, based on estimates reported in Section 4.3 and in Section 5.2, 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 5, represent the lower and upper bound for the compliant and non-compliant, respectively. Intuitively, the tightness of the bounds for γ_1 is an increasing function of the share of non-compliants (p). If there is no non-complaints in the population, that is when $p = 0$, there is no information in the data to identify γ_1, while the exact identification is obtained when $p = 1$. In this case the interval is reduced to a single point.

The bounds are reported in Figure 3 (without standard errors), showing that bounds are relatively wide for the non-compliant 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, the non-compliants may reduce their deduction claims more after being audited, whereas the upper bounds of the compliant signify a possibility for approval.

Figure 3. Bounds for the effects of auditing



5.5. Robustness test using matching method

Another way to deal with the problem of latent types in the control group is to use matching methods based on observed individual characteristics. By doing this we can reduce the potential bias generated by the fact that the control groups comprised of both compliants and non-complaints. However, this method will never fully solve this problem unless one can perfectly predict the latent type of a given individual. Nevertheless, we consider this as a useful robustness check.

There are many different matching methods. Here we apply the Coarsened Exact Matching algorithm (CEM), see Iacus, King and Porro (2011) and use pre-audit control variables to obtain better balance between the treated and the control groups.[19] Note that around 10% of audited individuals were not matched to anyone in the control group so they were excluded from the matched regression analysis.

Table 6 presents the results using only the matched sample. Compared with the non-matched sample, the estimated effects audits for the non-compliant 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.

[19] We have also obtained results for propensity score matching, basically similar to the results reported.

Table 6. Effects of audit on post-audit deduction behavior. Compliant and non-compliant 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
Non-compliant	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. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$

6. A simple model of tax deduction claiming

In the following, we set up a simple model where we try to explain the observed differences in deduction claiming behaviors before and after the audit.

In our model, the individuals will not cheat intentionally. The non-compliant behavior is simply a result of individuals mistakenly claim illegitimate deductions which they are unsure about themselves – i.e. due to a lack of understanding of the tax rules.

6.1. Model setup

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.

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.

The "riskfree" claim consist 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 $E[\varepsilon_{it}] = 0$. The "risky" claims $u_{it} > 0$ is also independent over time and uncorrelated with ε_{it}. This is a subject belief probability $0 < p_i < 1$ that the 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 "riskfree" deductions. There is, however, a positive cost if the tax payer's claimed deduction is audited and found to be illegitimate. So, she will only claim the risky potential 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.

6.2. Model Implications

Let's consider a three-period model, t=-1,0,1 where t=0 is the year of audit and t=-1 and t=1 corresponds to the year before and after the audit respectively. At any given year t, tax payer i's claimed deduction is denote 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: the non-compliance is temporary

When a tax audit is performed, we can group tax payers into compliant group and non-compliant groups. We see immediately that the non-compliant group consists of only type II individuals who claimed "risky" deductions, while the compliant group consists of both types. Since u_{it} is uncorrelated over time, then D_{it} is uncorrelated over time. This implies that the compliance behavior is not correlated over time. In other words, being a non-compliant at a given audit gives no indication on her compliant behavior in years prior to the audit. Note that this is not true for the behavior after audit, since the audit will change the key parameters governing the model, as discussed later.

Implication 2: the non-compliants have on average a lower claimed deduction level in the years prior to the audit

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

$$E[y_{it}|G_{i0} = 0] < E[y_{it}|G_{i0} = 1]$$

According to the model, only when the total deduction level is above a certain level, the taxpayer is eligible to audit. This implies that

$$\begin{aligned} &E[\lambda_i|G_{i0} = 0, eligible\ to\ audit] \\ &= E[\lambda_i|G_{i0} = 0, y_{it} > C] \\ &> E[\lambda_i|G_{i0} = 1, y_{it} > C] \\ &= E[\lambda_i|G_{i0} = 1, eligible\ to\ audit] \end{aligned}$$

So we have

$$E[\lambda_i|compliants\ at\ 0] > E[\lambda_i|non-compliants\ at\ 0]$$

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

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

The intuition is rather straightforward: Suppose the there are two individuals who claimed the same amount deduction. One is complaint and the other is non-compliant. Since individuals only claim "risky" deduction when the amount is high, the non-compliant will have lower time invariant risk-free claim than the compliant.

Implication 3: the non-compliants will adjust downwards their claimed deduction level in the years after the audit

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

After being found that their "risky" deductions are illegitimate, the tax payers would likely to adjust downwards 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 downwards q_i. This will result in a decrease in the threshold and reduce the claimed deduction after the audit.

Implication 4: the compliants may adjust their claimed deduction level either way in the years after the audit

For the compliants, 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 quite likely that they will adjust upwards the subject belief p_i. On the other hand, they may also adjust upwards the audit probability q_i where the overall effect could go either way.

The above model is quite simple, and assumptions are strong. However, we think similar implications should still hold if relax the assumptions and adding in additional features, for example, individual heterogeneity in the probability that the risky claim being legitimate.

These implications from 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 tax payers evade when they have the chances. What we found points to another possible sources of tax non-compliance 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.

7. Conclusion

It is crucial for the tax administration to know how much tax revenue that is generated by audits. Such calculations should not only account for the corrections made in the year of the control, but must take into account that taxpayers being exposed to audits most likely adjust their behavior in the years after the audit. In the present study we use data established by a random audit schedule operated by the Norwegian tax administration, carried out with respect to the item "Other deductions".

We find evidence in support of audited taxpayers reacting by reducing their claimed income deductions in the post-audit years, suggesting that they have been deterred by the audit. This is in accordance with a target effect, simply stating that agents perceive that an audit is likely to be followed by another in the future. Although we obtain a strictly significant estimate of the indirect effect in the first year after the audit, the other point estimates appear to be plausible too, in that they show diminishing effect over time.

Moreover, the study benefits from being able to further explore the dynamics of tax audit by distinguishing between behavior of the compliant and non-compliant taxpayers. Some recent studies (Gemmell and Ratto, 2012; Beer et al., 2015) find that the compliant taxpayers show opposite reactions after an audit, suggesting a counteracting approval effect also being in operation after an audit. This can be explained by previously compliant taxpayers now inferring that the risk of future checks is low and thereby exploit the information provided by the tax administration to decrease the subsequent tax burden. However, we do not find strong evidence of an approval effect among the compliant taxpayers; they basically show the same deduction behavior as before the audit. The point estimates obtained when using matching methods are all at the positive side, but statistically significant only in the sixth year after the audit.

However, the distinction between compliant and non-compliant taxpayers generates more markedly identified deterrence effects among the non-compliant: we obtain effects that are statistically significant for five years after the audit, with a decreasing deterrence effects over time. The results for the matched sample are larger and last for all six post-audit years.

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 (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.

Beer, S., M. Kasper, E. Kirchler, and B. Erard (2015). Audit Impact Study, 2015 Annual Report to Congress, TAS Research and Related Studies.

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

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.

Dubin, J. A. (2007). Criminal investigation enforcement activities and taxpayer noncompliance. *Public Finance Review*, 35(4), 500–29.

Dubin, J. A., M.J. Graetz, and L. L. Wilde (1990). The effect of audit rates on the federal individual income tax, 1977–1986. *National Tax Journal*, 43(4), 395–109.

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.

Gangl, K., S. Muehlbacher, M. de Groot, S. Goslinga, E. Hofmann, C. Kogler, G. Antonides, E. Kirchler, (2013). ‘‘How can I help you?’’ Perceived service orientation of tax authorities and tax compliance. *FinanzArchiv: Public Finance Analysis*, 69(4), 487–510.

Gangl, K., B. Torgler, E. Kirchler, and E. Hofmann (2014). Effects of supervision on tax compliance: Evidence from a field experiment in Austria. *Economics Letters*, 123, Issue 3, 378–382.

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.

Heckman, J. J. (1978). Dummy endogenous variables in a simultaneous equation system. *Econometrica*, 46(4), 931–959.

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.

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.

Nygård, O. E., J. Slemrod, and T. O. Thoresen (2019). Distributional Implications of Joint Tax Evasion. *Economic Journal*, 129 (620), 1894–1923.

Ratto, M., R. Thomas, and D. Ulph (2013). The indirect effects of auditing taxpayers. *Public Finance Review*, 41(3), 317–333.

Sandmo, A. (2005). The theory of tax evasion: A retrospective view. *National Tax Journal*, 58(4), 643–63.

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

Slemrod, J. (forthcoming). Tax compliance and enforcement, *Journal of Economic Literature*, forthcoming.

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.

Tversky, A. and D. Kahneman (1974). Judgement under uncertainty: Heuristics and biases. *Science*, 185, 1124–1131.

Yitzhaki, S. (1974) A note on: “Income tax evasion: a theoretical analysis”. *Journal of Public Economics*, 3, 201–202.

Appendix A. Distribution of “Other deductions”

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

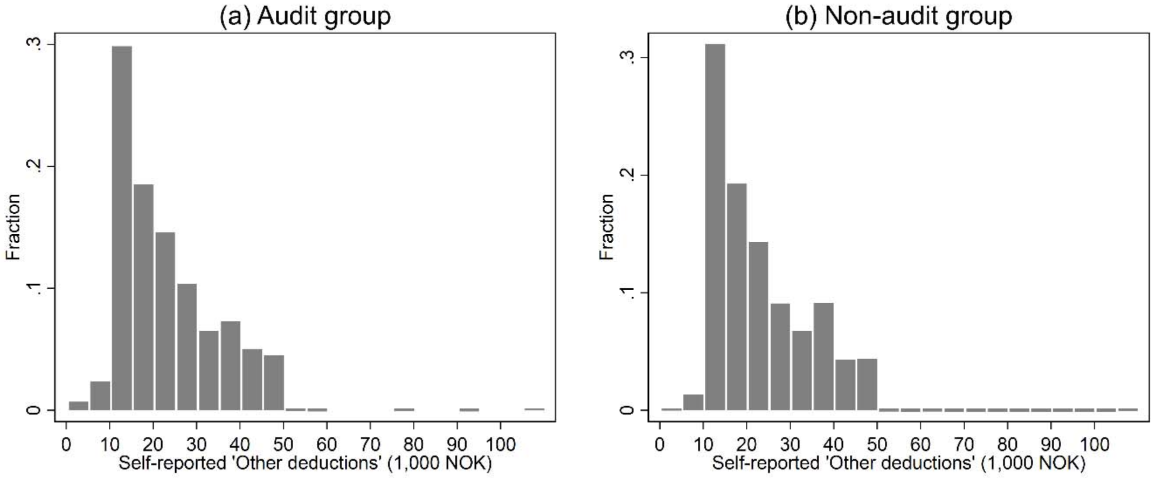
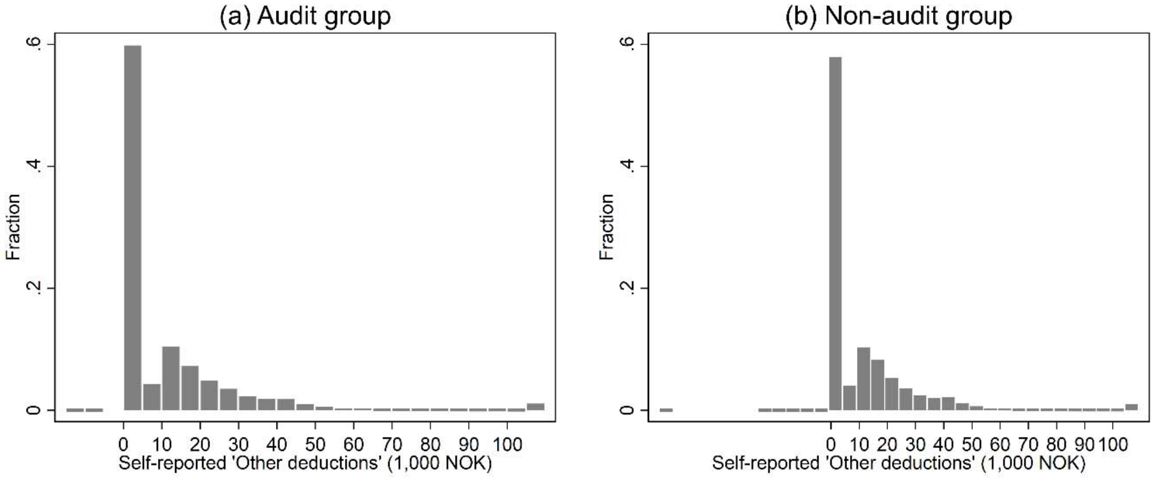


Figure A2. Distribution of “Other deductions” among audited and non-audited. Averages over all years used in the empirical analysis



Appendix B. Distribution of treatment effect [Jia rewrites, reorganizes]

As the results so far show that there is a negative shift in the mean deduction after the audit, it is interesting to examine to what extent the shape of the deduction distribution has been changed. Here we follow Hernæs and Jia (2013) and Brinch, Hernæs and Jia (2017) and look at the Complementary Cumulative Distribution Functions (CCDF), $\bar{F}(y \mid X) = \Pr(Y > y \mid X)$, before and after audit. To be more precise, 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 $t \leq 1$

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

where X_{it} denotes individual characteristics and F represents the logit function. We estimate the logit specification letting y_k vary 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 2 shows the estimated marginal effects with 95% 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 4.

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

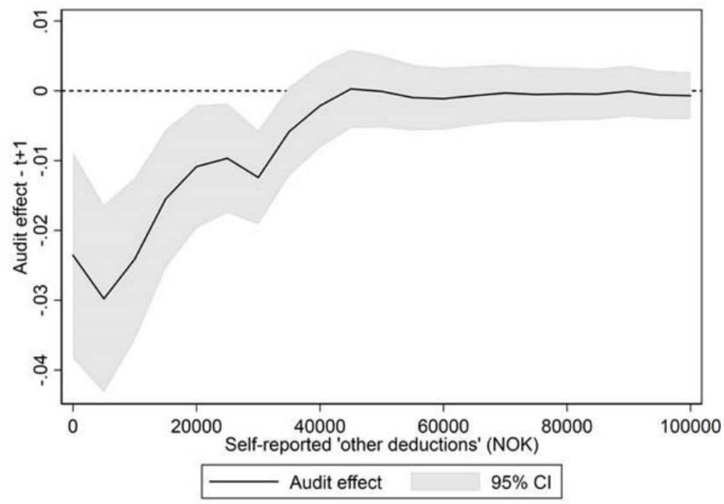


Table B1. Estimate of shift into deduction interval after audit

	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