Audits, Audit Effectiveness, and Post-audit Tax Compliance

Matthias Kasper[1] & James Alm[2]

Abstract

This study uses a laboratory experiment to investigate the effect of *audit effectiveness*, or the share of undeclared income that the tax agency detects in an audit, on post-audit tax compliance. We also study whether the effects of audits depend on a taxpayer's reporting behavior prior to the audit. Our findings show that tax audits have differential effects on post-audit compliance and that the effectiveness of audits determines these responses; that is, while effective audits increase post-audit tax compliance, ineffective audits have the opposite effect. We also find that relatively compliant taxpayers exhibit the strongest behavioral response to audits. Our results indicate that the specific deterrent effects of tax audits are more ambiguous than standard and behavioral models of tax compliance suggest, with these effects dependent on the effectiveness of audits and on the taxpayer's prior reporting behavior. Indeed, ineffective audits may well have counter-deterrent effects on some types of taxpayers.

Keywords: Tax compliance; Audit effectiveness; Specific deterrence; General deterrence; Laboratory experiments.

JEL Codes: C9; H26; H83.

[1] University of Vienna, Department of Psychology (matthias.kasper@univie.ac.at; Universitaetsstrasse 7, 1010 Wien, Austria) - corresponding author.
[2] Tulane University, Department of Economics (jalm@tulane.edu; Tulane University, New Orleans, Louisiana 70118 USA).

1. Introduction

Tax audits are an essential instrument in establishing and maintaining compliance, and increasing the number of audits has direct and indirect effects on taxpayer behavior. Audits have *direct* effects by raising revenue through the assessment of additional taxes, interest, and penalties on individuals who are audited. Additionally, tax audits have *indirect* effects by deterring future noncompliance among both audited taxpayers (*specific deterrence*) and unaudited taxpayers (*general deterrence*). A growing body of research analyzes these direct and indirect deterrent effects of tax audits, and generally shows that more audits lead to more compliance (Alm, 2019; Slemrod, 2019).

However, an important if often neglected feature of tax audits is that they do not always detect tax evasion when it is present and they may even find evasion when it is not in fact present. Early work by Feinstein (1991) suggests that the average detection rates of senior tax examiners are around 50 percent. Indeed, in fiscal year 2018 U.S. taxpayers challenged over $10 billion in additional taxes recommended by the Internal Revenue Service (IRS), while almost $4 billion of tax and penalties were under appeal in U.S. tax courts (Internal Revenue Service, 2019). This *audit effectiveness*, or the tax administration's capacity to detect noncompliance in an audit, seems likely to affect a taxpayer's behavioral response to enforcement. For example, Gemmell and Ratto (2012) and Beer et al. (2020) conclude that the specific deterrent effect of audits depends on the audit outcome. These studies find that tax audits increase subsequent, or post-audit, compliance among taxpayers who were found to be noncompliant, while they decrease compliance among those who were determined to be compliant.[3] These results raise important and unanswered questions about the effect of audit effectiveness on post-audit tax compliance.

Specifically, it remains unclear from existing research whether audit effectiveness – especially the lack of effectiveness – affects the specific deterrence effect of enforcement. A related question is whether the effect of audits depends on a taxpayer's reporting behavior prior to an audit; that is, it also remains unclear from existing research whether truly compliant and truly noncompliant taxpayers differ in their behavioral responses to enforcement.

This study addresses these questions by investigating the specific deterrent effect of audits on post-audit tax compliance when audit effectiveness varies. We run a preregistered

[3] Note that we use the terms *post-audit* and *subsequent* tax compliance interchangeably.

laboratory experiment with 333 participants in which we test how variation in the risk of detection affects post-audit tax compliance. The crucial feature of our experimental design is the addition of audit effectiveness to our audit mechanism, where we define audit effectiveness as the share of undeclared income that the tax agency detects in an audit (Rablen, 2014). This addition allows us to examine the effects of audit effectiveness on post-audit compliance. We also study whether the effect of audits depends on a taxpayer's reporting behavior prior to an audit; that is, we investigate whether prior reporting behavior affects behavioral responses to audits. Addressing these questions with field data is difficult, even problematic, because tax agencies typically do not know a taxpayer's true tax liability. In particular, the audit outcome is not a perfect measure of a taxpayer's true compliance, given the demonstrated inability of an audit to detect all evasion, so that the identification of audit effectiveness and its effects on truly compliant and noncompliant taxpayers is challenging with field data. In contrast to the use of field data, data generated from a laboratory experiment allows us to introduce changes in audit effectiveness, as well as in audit probability, and thereby allows clean identification of the causal effects of these changes on post-audit compliance of truly compliant and truly noncompliant individuals.

Our study differs from the previous literature by making contributions in three important dimensions. First, and most importantly, unlike most existing work, we account for the possibility that tax audits might not detect all undeclared income. This allows us to investigate whether ineffective audits reduce taxpayers' propensity to comply in the future. Second, we investigate whether behavioral responses to enforcement depend on taxpayers' prior reporting behavior. We do this by distinguishing between *relatively noncompliant* and *relatively compliant* individuals, where relatively noncompliant taxpayers are defined as the lower half of the compliance distribution in each reporting decision and relatively compliant taxpayers are defined as the upper half of the compliance distribution. Moreover, we distinguish between *honest* and *dishonest* individuals, where honest individuals report all income in all rounds prior to their first audit and dishonest individuals report zero income in these rounds. This overall design allows us to disentangle the possible mechanisms by which specific deterrence may drive post-audit tax compliance. Finally, our design allows us to investigate whether presenting taxpayers with a compound detection risk (where an audit does not result in certain detection) changes their willingness to comply compared to a decision where an audit results in certain detection.

Our results indicate that audit effectiveness is an important determinant of the specific deterrent effect of audits. Taxpayers declare a larger share of their income after experiencing an audit that detects all undeclared income, while ineffective audits decrease post-audit compliance. This effect does not diminish with audit frequency. Specifically, effective audits increase post-audit compliance, while ineffective audits decrease post-audit compliance even after taxpayers have experienced several audits. Moreover, we find that a taxpayer's prior reporting compliance affects these behavioral responses to audits. Relatively compliant taxpayers respond to audits by adjusting their post-audit compliance in response to the audit effectiveness. These taxpayers decrease their post-audit compliance after an ineffective audit, and they increase their post-audit compliance after an effective audit. In contrast, we do not find strong evidence that relatively noncompliant taxpayers alter their reporting behavior after "losing the audit lottery". Moreover, we find no evidence of a crowding-out effect of tax audits; that is, taxpayers who declared their entire income in all rounds prior to their first audit do not reduce their subsequent compliance in response to the audit. Finally, we find no evidence of a misperception of compound detection risk (where audits are ineffective).

Our study adds to the literature on behavioral responses to enforcement. In particular, our results suggest that ineffective audits contribute to the counter-deterrent effect of tax audits that has been found in some prior studies (Beer et al. 2020; Gemmel & Ratto, 2012). Moreover, we provide a new perspective on the tradeoff between audit frequency and audit effectiveness (Rablen, 2014) and the analysis of optimal tax administration (Keen & Slemrod, 2017). Our results suggest that a complete analysis of a revenue-maximizing audit strategy requires consideration of behavioral responses to audit effectiveness as well as recognition of differential responses of compliant and noncompliant taxpayers.

2. Related Literature

Prior work on the specific deterrent effect of tax audits does not account for audit effectiveness. These studies have typically used administrative data to analyze the aggregate response of those taxpayers who have been audited. Overall, these studies find that enforcement has a positive effect on post-audit tax compliance.[4] For example, Kleven et al. (2011) show that

[4] An exception is Erard (1992), who analyzes micro-level data from the U.S. Taxpayer Compliance Measurement Program (TCMP) of the IRS and who finds no significant effect of a prior tax audit on subsequent compliance. Similarly, a recent paper by Best, Shah, and Waseem (2021) finds no effect of audits on the subsequent VAT compliance of firms in Pakistan.

tax audits increase self-reported income among Danish taxpayers in the subsequent tax year. Similarly, Advani et al. (2017) find that reported income of self-employed UK taxpayers increases for at least 5 years after an audit, while DeBacker et al. (2018) show that compliance of U.S. taxpayers improves for three years after an audit before ultimately reverting to previous (and lower) levels. A more recent study of U.S. taxpayers by Beer et al. (2020) investigates whether the effect of audits on post-audit reporting behavior depends on the audit outcome, and they find that the specific deterrent effect of tax audits is positive in the aggregate but that post-audit compliance depends on the outcome of the examination. In particular, taxpayers who receive an additional tax assessment as a result of their audit report more income in subsequent years, while those who do not receive an additional assessment report less. This result is in line with a study by Gemmell and Ratto (2012) for the UK that finds that audited taxpayers who were found to be noncompliant report more income in their subsequent tax return than those who were not audited, while taxpayers who were found to be compliant show the opposite response. A study on the effects of audits on VAT compliance in Argentina and Chile by Bergman and Nevarez (2006) also finds that audits have a differential effect on post-audit compliance, although this study also finds that audits decrease compliance among those who were found to be cheating. More recently, some studies have investigated the effect of different audit types on post-audit tax compliance (Erard et al., 2019; Kotsogiannis et al., 2021). These studies suggest that more thorough face-to-face audits have a positive effect on post-audit tax compliance, while less thorough correspondence or desk audits tend to have a negative effect. Overall, these studies raise the question why enforcement appears sometimes to encourage rather than deter future noncompliance.[5]

Several behavioral explanations have been suggested for these results (Kirchler, 2007; Alm, 2019; Beer et al., 2020), but the underlying mechanisms remain unclear. One possible explanation relates to audit effectiveness, or the ability of the tax administration to detect evasion during an audit. An ineffective audit might stimulate a taxpayer's willingness to take risks; that is, if an audit fails to detect undeclared income, the taxpayer might infer that the agency is unable to discover cheating and thus underreport his or her income in subsequent years (Andreoni et al., 1998). Similarly, post-audit tax compliance might be affected by the fine for noncompliance that a taxpayer receives as a result of an audit: A higher fine might result in more post-audit tax compliance than a smaller fine. Indeed, prior work finds that unsanctioned

[5] There is some research in criminology that investigates the effect of punishment on an individual's future proclivity for crime (Dusek & Traxler, 2020). This work suggests mixed evidence for specific deterrence effects, and indeed there is some indication that the experience of punishment might increase, rather than decrease, future offending (Cullen et al., 2011; Nagin et al., 2009; Nagin 2013a, 2013b).

criminal offenses reduce perceived risk of detection and punishment (Matsueda et al., 2006). Such behaviors might be explained by the availability heuristic of Kahneman and Tversky who argue that individuals evaluate the risk of a decision by imagining the negative outcome. If the negative outcome is "vividly portrayed", then this event may "appear exceedingly dangerous, although the ease with which disasters are imagined need not reflect their actual likelihood. Conversely, the risk … may be grossly underestimated if some possible dangers are either difficult to conceive of, or simply do not come to mind" (Kahneman & Tversky 1974, p. 1128).

However, almost all prior work that estimates behavioral responses to tax enforcement assumes that tax audits always detect all undeclared income. The few exceptions employ laboratory experiments to investigate how variation in audit effectiveness affects the general population of taxpayers, rather than those taxpayers who experienced the audit. For example, Alm and McKee (2006) vary the fraction of undeclared income that the tax agency detects in an audit, and, surprisingly, they find higher compliance levels when audit effectiveness is low. However, the study also finds that audit effectiveness has a positive effect on compliance when taxpayers know that they will be audited with certainty. Similarly, Bernasconi and Bernhofer (2020) find some support for the hypothesis that ineffective tax audits increase compliance in the aggregate, attributing this result to taxpayers' misperception of compound detection lotteries where an audit does not result in a certain outcome. However, they also find that learning effects diminish the misperception of compound detection lotteries. In sum, these two studies provide mixed evidence for a general deterrent effect of audit effectiveness. However, prior work does not investigate the effect of audit effectiveness on post-audit tax compliance, so that the effects of ineffective audits on post-audit tax compliance remain unknown.

A second explanation for the unintended consequences of tax audits is the *bomb crater effect* (Guala & Mittone, 2005; Mittone, 2006). Contrary to the standard models of tax evasion that imply that audits affect subsequent compliance only when they provide the taxpayer with new information (Allingham & Sandmo, 1972; Srinivasan, 1973; Yitzhaki, 1974), it is common in laboratory experiments to find that participants declare a smaller share of their income after being audited. Such a response might result from an underestimation of the risk of future audits (Mittone, 2006; Mittone et al., 2017) or from loss-repair motivations (Andreoni et al., 1998; Maciejovsky et al., 2007). However, it remains unclear whether the perceived risk of future examinations is affected by the audit outcome or whether the tendency to make up for past losses pertains to individuals who have been found to be noncompliant. For example, some studies find that a decline in reported income after an audit cannot be explained by loss repair

motivations alone because individuals who were found to be compliant also report less income after experiencing an audit (Kastlunger et al., 2009; McKee et al., 2018; Bernasconi & Bernhofer, 2020). More recently, Kasper and Alm (2021) find that reporting compliance in the laboratory is volatile even absent any audits; that is, they find that individuals who reported all income correctly reduce their subsequent compliance irrespective of whether they have been audited or not, while audited and unaudited taxpayers who did not report any income increase their subsequent compliance.

A third potential explanation is that audits have differential effects on different types of taxpayers. Some scholars have suggested that taxpayers comply for different reasons (Erard & Feinstein, 1994; Torgler, 2003; Braithwaite, 2009). While some taxpayers are motivated entirely by the expected value of the evasion gamble, others comply regardless of any incentive to cheat (Braithwaite, 2003). However, such honest taxpayers may find being audited unfair, perceive the audit as a breach of trust, or experience negative emotions (Olsen et al., 2018; Enachescu et al., 2019). This experience might crowd out their *intrinsic motivation* to comply and reduce their propensity to comply in the future (Frey, 1997; Mendoza et al., 2017; Lederman, 2018; Hu & Ben-Ner, 2020). Therefore, a decline in post-audit compliance might also result from honest individuals who are less likely to comply after experiencing an audit. Dishonest taxpayers, on the other hand, might respond to an audit by increasing their post-audit compliance because the experience of being punished motivates them to comply more in the future (Braithwaite, 2003; Kirchler et al., 2008).

These prior studies suggest different behavioral explanations of responses to tax audits, but without resolving the actual mechanisms that drive these responses. Our study allows us to discern the potential explanations that have been proposed in the literature. To our knowledge, our study is also the first to investigate the effect of audit effectiveness on post-audit compliance.

3. Theoretical Foundations

Theories of deterrence distinguish between threat of punishment and experience of punishment (Chalfin & McCrary, 2007), and the literature in economics focuses mainly on the former. A taxpayer's compliance decision is typically analyzed within an expected utility framework that follows Becker's (1968) economics-of-crime approach, as first formalized by

Allingham and Sandmo (1972), Srinivasan (1973), and Yitzhaki (1974) as a decision under risk. These models explain the general deterrent effect of tax audits, predicting that an increase in the audit probability or the penalty rate leads to greater compliance.[6]

The standard models of tax evasion assume that an audit detects all undeclared income, but they can be easily adjusted to allow for ineffective audits. In this case a taxpayer receives income I and must decide how much to report to the tax agency. Reported income R is taxed at the rate t, and unreported income is not taxed. The taxpayer faces the risk of being audited with a probability p. In case of an audit, the agency detects a share e of undeclared income and imposes a fine f on the undeclared taxes that are detected. In case of no audit, the taxpayer simply pays taxes on reported income. The taxpayer chooses R to maximize the expected utility of the evasion gamble, or

$$(1) \qquad EU(I) = (1 - p)\, U(I - tR) + p(U(I - tR - etf(I - R)).$$

In this setting the risk of detection is the product of the audit probability p and the audit effectiveness e, where the compliance effect of a change in the audit probability is the same as the effect of an equivalent change in audit effectiveness.[7]

However, it is important to note that all of these models predict that audits do not affect a taxpayer's subsequent reporting decision because they assume that the audit does not provide the taxpayer with new information. As audit and penalty rates are fixed and known, experiencing an audit is merely a case of losing the evasion gamble, and this does not affect post-audit compliance. However, post-audit compliance depends on perceived rather than actual changes in the probability of detection. In fact, prior studies find that the experience of enforcement may change behavior, even absent any change in the underlying probability of detection (Simonsohn et al., 2008; Haselhuhn et al., 2012; Earnhart & Friesen, 2013). This effect is particularly well documented in laboratory experiments on tax compliance, where the relevant tax system parameters are typically public knowledge and unaffected by the audit outcome (Alm, 2019; Alm & Kasper, 2021).[8]

[6] There is ample empirical evidence that increasing p and f increases compliance. See Alm (2019) and Slemrod (2019) for comprehensive surveys of the literature.

[7] Simplifying equation (1) indicates that an x percentage point increase in p is offset by a $1/x$ percentage point increase in e and vice-versa. However, there is some evidence that decision makers misperceive such compound lotteries (Dillenberger, 2010; Harrison et al., 2015; Prokosheva, 2016; Bernasconi & Bernhofer, 2020).

[8] Field studies on the effect of tax audits attribute changes in post-audit tax compliance to new information that the audit provides (e.g., Kleven et al., 2011; Best et al., 2021; Kotsogiannis et al., 2021). In contrast to these studies,

These findings raise unresolved questions about the channels through which tax audits affect subsequent reporting compliance. In particular, the literature suggests three potential explanations for behavioral responses to tax audits. However, these theories make conflicting predictions on the effect of audits on subsequent compliance. First, audited taxpayers might assess the probability of a future audit by the ease of recalling their previous audit (Spicer & Hero, 1985). This *availability heuristic* (Kahneman & Tversky, 1973) predicts that the audit effectiveness (or the share of undeclared income that the tax agency detected in an audit) of a previous audit should affect a taxpayer's assessment of the risk of a future audit. As a result, the specific deterrent effect of an effective audit should be stronger than the specific deterrent effect of an ineffective audit. Similarly, audits that result in a higher fine should have a more positive effect of post-audit tax compliance than audits that result in a smaller fine. Second, taxpayers might falsely assume that a recent audit experience reduces the risk of a future audit (Mittone (2006). As a result, compliance levels would generally decrease after an audit. Third, the audit experience might change a taxpayer's motivation to comply. In particular, tax audits might crowd out the intrinsic motivation to comply and reduce post-audit compliance among honest taxpayers (Frey, 1997; Mendoza et al., 2017; Lederman, 2018). Conversely, dishonest taxpayers might respond to an audit by increasing their post-audit compliance because the experience of being punished motivates them to comply more in the future (Braithwaite, 2003; Kirchler et al., 2008). These considerations imply that pre-audit compliance levels determine the behavioral response to audits, but predict conflicting effects on the audit experience on post-audit tax compliance.

In sum, theoretical studies of tax compliance suggest that financial incentives determine a taxpayer's reporting decision and that increasing the audit probability and the fines for noncompliance deter tax evasion. However, the effect of the audit experience on post-audit compliance is not well understood, and the existing literature does not resolve crucial aspects. First, the effect of audit effectiveness on post-audit tax compliance remains unknown. Second, it remains unclear to what extent the experience of an audit affects the perceived risk of a subsequent audit. Third, it is also unknown whether taxpayers' previous history of compliance or noncompliance determines the behavioral response to audit; that is, the effects of audits on truly compliant and truly noncompliant taxpayers are unknown. The next section discusses our experimental design for examining these issues.

and in line with the standard models of tax evasion, we focus our analysis on a situation where all relevant tax system parameters are public knowledge and audits do not provide taxpayers with new information.

4. Experimental Setup: Design, Procedure, and Sample

Our experiment follows the standard procedure of tax compliance experiments (Alm and Jacobson, 2007), and it implements many of the basic elements of voluntary income tax reporting that are necessary to identify the effect of audits on post-audit tax compliance. In each reporting decision of the experiment, i.e. in each experimental *round*, participants receive a random amount of income that varies between 2,000 and 3,500 Experimental Currency Units (ECU).[9] They must decide how much income to report to the tax agency, and they may report any amount between 0 ECU and the amount they received. Reported income is taxed at a rate t of 25 percent ($t = 0.25$). Participants face the risk of being randomly selected for an audit, where audit selection is based on a lottery. The audit probability is announced at the beginning of each round of the experiment. After each reporting decision participants learn whether they have been selected for an audit, but they cannot observe the lottery draw itself.[10] Audit probabilities p range from 0.18 to 0.70, and tax audits differ in their effectiveness. While audits detect all undeclared income in some rounds, they detect only some fraction of undeclared income in others. Note that audit effectiveness e ranges between 0.30 and 1. Consequently, the overall detection risk (or the product of p and e) ranges from 0.18 to 0.49 (see Table 1 for details). The fine f for noncompliance is twice the evaded amount that has been detected. Once participants have reported their income, they learn whether they have been audited or not and the outcome of the audit. This process is repeated over 28 rounds in random order. Participants do not know the number of rounds.

[9] 1,000 ECU equals € 3.50.

[10] This approach is frequently used in laboratory experiments, although some studies use die rolls or similar observable mechanisms to simulate the audit selection mechanism. The experimental instructions emphasize that audit selection is random, but we acknowledge that some taxpayers might question the randomization of the audit mechanism.

Table 1: Experimental Parameters

Task	Audit type	Composition of detection risk	Parameter order	Audit probability	Audit effectiveness	Detection risk
1	Effective audit	Low *p*, *e* = 1	*p* first	0.18	1.00	0.18
2				0.21	1.00	0.21
3				0.24	1.00	0.24
4				0.28	1.00	0.28
5			*e* first	0.18	1.00	0.18
6				0.21	1.00	0.21
7				0.24	1.00	0.24
8				0.28	1.00	0.28
9	Medium effective audit	Low *p*, high *e*	*p* first	0.30	0.60	0.18
10				0.33	0.63	0.21
11				0.37	0.67	0.24
12				0.40	0.70	0.28
13			*e* first	0.30	0.60	0.18
14				0.33	0.63	0.21
15				0.37	0.67	0.24
16				0.40	0.70	0.28
17	Low effective audit	High *p*, low *e*	*p* first	0.60	0.30	0.18
18				0.63	0.33	0.21
19				0.67	0.37	0.24
20				0.70	0.40	0.28
21			*e* first	0.60	0.30	0.18
22				0.63	0.33	0.21
23				0.67	0.37	0.24
24				0.70	0.40	0.28
25	Medium effective audit	High *p*, high *e*	*p* first	0.60	0.60	0.36
26				0.63	0.63	0.40
27				0.67	0.67	0.44
28				0.70	0.70	0.49

Notes: Participants face all 28 tasks in random order. Parameters are presented to participants at the beginning of each round. *Parameter order* indicates how the *Audit probability* (*p)* and the *Audit effectiveness* (*e*) are presented to participants (*p* before *e* or vice versa).

Table 1 shows our experimental parameters. We calibrate these parameters such that a "reasonably" risk-averse taxpayer should not report any income to maximize expected profit.[11] By distinguishing between and introducing variation in the audit probability *p* and the audit effectiveness *e*, our design enables us to test whether effective versus ineffective audits differ in their capacity to deter noncompliance and it also allows us to investigate whether presenting taxpayers with a compound detection risk (where *p* and *e* are each less than 1) changes their

[11] An individual with "realistic" levels of constant relative risk aversion ($r \leq 1.5$) would optimally declare zero income for $d = 0.26$ (the average detection risk), $t = 0.25$, *and* $f = 2$. See Alm (2019) for details.

willingness to comply compared to a decision where an audit results in certain detection (where $e = 1$); see column *Audit type*. We also systematically vary the display of information on the audit probability p and the audit effectiveness e to rule out the possibility that order effects drive our results; see column *Parameter order*. Finally, our design allows us to study the effect of audit effectiveness on the future compliance behavior of audited taxpayers.

All parameters are known to the participants in each round. Also, to facilitate the compliance decision, we program a calculator that shows how declared income translates into after-tax income conditional on audit effectiveness. We provide the instructions and some screenshots of the experimental task in Appendix A.[12]

The experiment was conducted at the Vienna Center of Experimental Economics (VCEE) in December 2019 and January 2020. Participants were recruited via ORSEE (Greiner, 2015). We used a power analysis to determine the sample size, and we pre-registered our study at https://osf.io/uhpmw/.[13] The final sample (n = 333) comprises data from 13 experimental sessions.

At the beginning of the experiment participants learn that their information is private and that it is impossible to identify individual participants. The study starts with a few demographic questions. Subsequently, participants learn about the compensation mechanism. Each participant receives a show-up fee of € 5.00 and an additional compensation that is based on the after-tax income of a randomly selected round. After reading a detailed introduction to the experimental task and an example of the tax compliance decision, participants must answer two questions on the definition of audit probability and audit effectiveness correctly before they can proceed. Next, they play three practice rounds. One practice round is not audited, while the two other rounds result in one effective and one ineffective audit, respectively. Participants then proceed to the experiment which comprises 28 reporting decisions (rounds) in random order. In each experimental round participants receive a random amount of income (between 2,000 and 3,500 ECU). Participants learn the audit probability and the audit effectiveness (as depicted in Table 1) in every round before making their reporting decision. Once participants have reported their income, they learn whether they have been audited or not and the outcome of the

[12] The experiment was programmed in z-Tree (Fischbacher, 2007).
[13] Our target sample size estimate is based on a power analysis that indicated that a sample size of N = 327 is required to detect a difference between two means (mean compliance rate after an effective versus an ineffective audit) with the following parameters: power = 0.95, alpha = 0.05, Cohen's d = 0.2, t-test for two dependent means (two-tailed).

audit. After completing the 28th round, participants answer a few final questions. The experiment lasts approximately 45 minutes, and the mean payoff is € 12.66.

The participant pool has a slightly larger percentage of female participants (57 percent) than male participants, and the pool includes students and non-students. The mean age is 26 years (SD = 6.06) with a range from 18 to 59 years. Most participants hold at least a high-school degree (49 percent) and study business (19 percent). While 95 percent indicate that they participated in a laboratory experiment in the past, only 16 percent state that they participated in a study on tax compliance before. Moreover, 29 percent indicate that they self-prepared a tax return in the past.

Because our experiments are intended to help inform policy makers, they must satisfy Smith's (1982) precept of *parallelism*. Parallelism is satisfied when the experimental setting captures the essential elements of the decision problem faced in the naturally occurring setting. It is not necessary (and it is also not desirable) that the experimental setting implement all of the complexity of the naturally occurring setting (Plott, 1987). As implemented, our experimental design follows the main essential elements of laboratory experiments (Alm & Jacobson, 2007), but incorporates some additional features to improve parallelism with taxpayer's decision-making in the naturally occurring world. Most obviously, our design allows audits to vary in how effective they are in detecting unreported income, an essential feature that characterizes all audit systems in the naturally occurring world. In addition, subjects face values of the main fiscal parameters that parallel the values in the naturally occurring world. They face a reporting, auditing, and penalty process that also parallels the naturally occurring world, including the presence of multiple periods in which they must make decisions. Our experiments also utilize tax language in the instructions, as well as in the computer interface used to present information and to elicit income reporting behavior. Even so, our experimental design does not introduce some real-world features. For example, our study does not implement a public good because we are interested in isolating the specific deterrent effect of audits absent any social dynamics. The presence of a public good financed by total group tax payments introduces strategic considerations in the form of social interactions that complicate individual decisions and that might also induce perceptions of unfair or inefficient taxation. Our design also does not utilize endogenous audit selection rules, for similar reasons.[14] Overall, we believe that our

[14] Note that we also employ a within-subjects design. Such a design is frequently used in laboratory experiments because it provides the high levels of statistical power needed to identify the causal effect of interest. However, this design feature might also increase the salience of variation in the audit probability and audit effectiveness. Even so, behavioral responses to audits in our experiment should only be affected by variations in the audit

experimental design captures the essential elements of parallelism necessary for our results to generalize beyond the laboratory.

5. Results

5.1. Descriptive Statistics

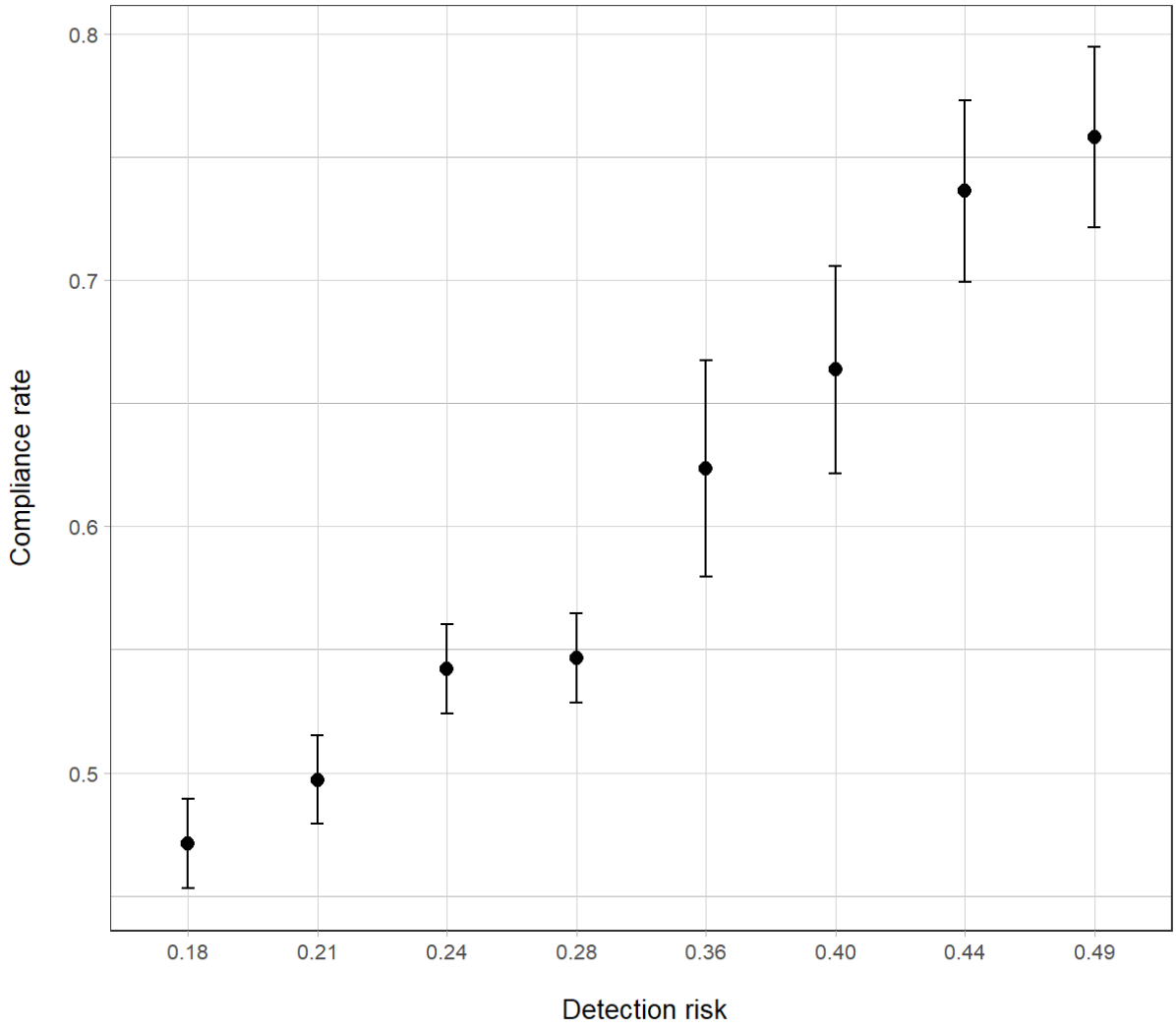
We observe 9,324 compliance decisions from 333 individuals, 4,131 of these decisions were randomly audited, and 5,193 decisions were not audited. The actual audit probability was 0.44, and the average audit effectiveness was 0.66. Our main dependent variable is the *Compliance rate*, defined as the share of received income that was reported to the tax agency. The mean compliance rate across all participants and all rounds was 0.54 (SD = 0.41), which indicates substantial underreporting in the aggregate. Panel A of Table B1 in Appendix B presents descriptive statistics in detail.

5.2. General Deterrent Effects of Audits

The purpose of this section is to establish that taxpayers' reporting behavior responds to an increase of the risk of detection. Moreover, we explore whether changes in the audit probability or the audit effectiveness have heterogenous effects on compliance. To investigate the general deterrent effect of audits, we show the relation between the detection risk and the compliance rate in Figure 1. Figure 1 shows a monotonous increase in compliance in the risk of detection, which demonstrates a strong general deterrent effect of tax audits. On average, taxpayers report less than half of their income when the risk of detection is below 20 percent, while they report around three quarters of their income when the risk of detection is close to 50 percent. Overall, reporting decisions are in line with the predictions of the standard models of tax compliance. This strengthens our confidence in the internal validity of our study.

probability and the audit effectiveness, so that we believe that the benefits of a within-subjects design outweigh its potential caveats.

Figure 1: General Deterrent Effect of Audits



Notes: *Detection risk* is the product of *Audit probability* (*p*) and *Audit effectiveness* (*e*) as described in Table 1. Error bars depict 95 percent confidence intervals.

We begin our analysis of the general deterrent effect of audits by estimating several variations of the following baseline model:

$$(2) \quad Compliance\ Rate_{i,t} = \beta_0 + \beta_1\ Received\ income_{i,t} + \beta_2\ Detection\ risk_{i,t} + \beta_3\ Z_i + \varepsilon_{it},$$

where i and t are individual and period indices. The traditional error term is denoted by w_{it}, and it is assumed to meet all of the usual requirements. The individual-specific random effect is denoted by u_i, which controls for individual level heterogeneity. Note that $\varepsilon_{it} = u_i + w_{it}$. The

variable Z_i represents individual-specific control variables as depicted in Panel A of Table B1 in Appendix B.[15]

We present regression results on the general deterrent effect of audits in Table 2. In line with Figure 1, we find that reporting decisions correspond to the predictions of the standard model of evasion. Specifically, we find that increasing the risk of detection has a strong effect on compliance, with a 1 percentage point increase in the risk of detection increasing compliance by 0.9 percentage points. This result is consistent with (if slightly larger than) the typical estimated response of reported income to changes in the detection risk (Alm, 2019).

Table 2 also provides insight into behavioral aspects. We find that neither the parameter order (audit probability presented before audit effectiveness or vice versa), nor the composition of the detection risk (as presented in Table 1) affects compliance. In particular, when accounting for the risk of detection, compliance is not affected relative to an audit that results in *Certain detection* (the reference category) by: decreasing the audit probability while increasing the audit effectiveness (*Low p, high e*), increasing the audit probability while decreasing the audit effectiveness (*Low p, high e*), or increasing both factors (*high p, high e*). This result indicates that taxpayers do not systematically misperceive the audit probability *p* relative to the audit effectiveness *e*.[16] The effects of demographic variables are discussed later.

[15] Since the tax rate *t* and the fine *f* for noncompliance are constant, they are not included in this model.
[16] To further investigate this issue, we depict the effect of the composition of the detection risk on compliance in Figure B1 in Appendix B. Figure B1 provides no indication of a systematic misperception of the audit probability or the audit effectiveness.

Table 2: General Deterrent Effect of Audits

Sample	**Entire Sample**	
Independent variable	(1)	(2)
Intercept	0.3199 ***	0.3505 ***
	(0.0244)	(0.0689)
Received income	-0.0198 ***	-0.0199 ***
	(0.0032)	(0.0032)
Detection risk	0.0086 ***	0.0086 ***
	(0.0008)	(0.0008)
Parameter order	0.0008	0.0008
	(0.0068)	(0.0068)
Low p, High e	-0.0049	-0.0049
	(0.0083)	(0.0083)
High p, Low e	-0.0060	-0.0060
	(0.0083)	(0.0083)
High p, High e	0.0093	0.0093
	(0.0189)	(0.0189)
Demographic variables?	No	Yes
N	333	333
Observations	9,324	9,324
Marginal R^2 / Conditional R^2	0.032 / 0.465	0.160 / 0.471

Notes: *, **, and *** indicate significance at the 10 percent, 5 percent, and 1 percent level. Received income and continuous demographic variables are scaled. Marginal R^2 provides the variance explained by fixed effects, and conditional R^2 provides the variance explained by fixed effects and individual-specific random effects.

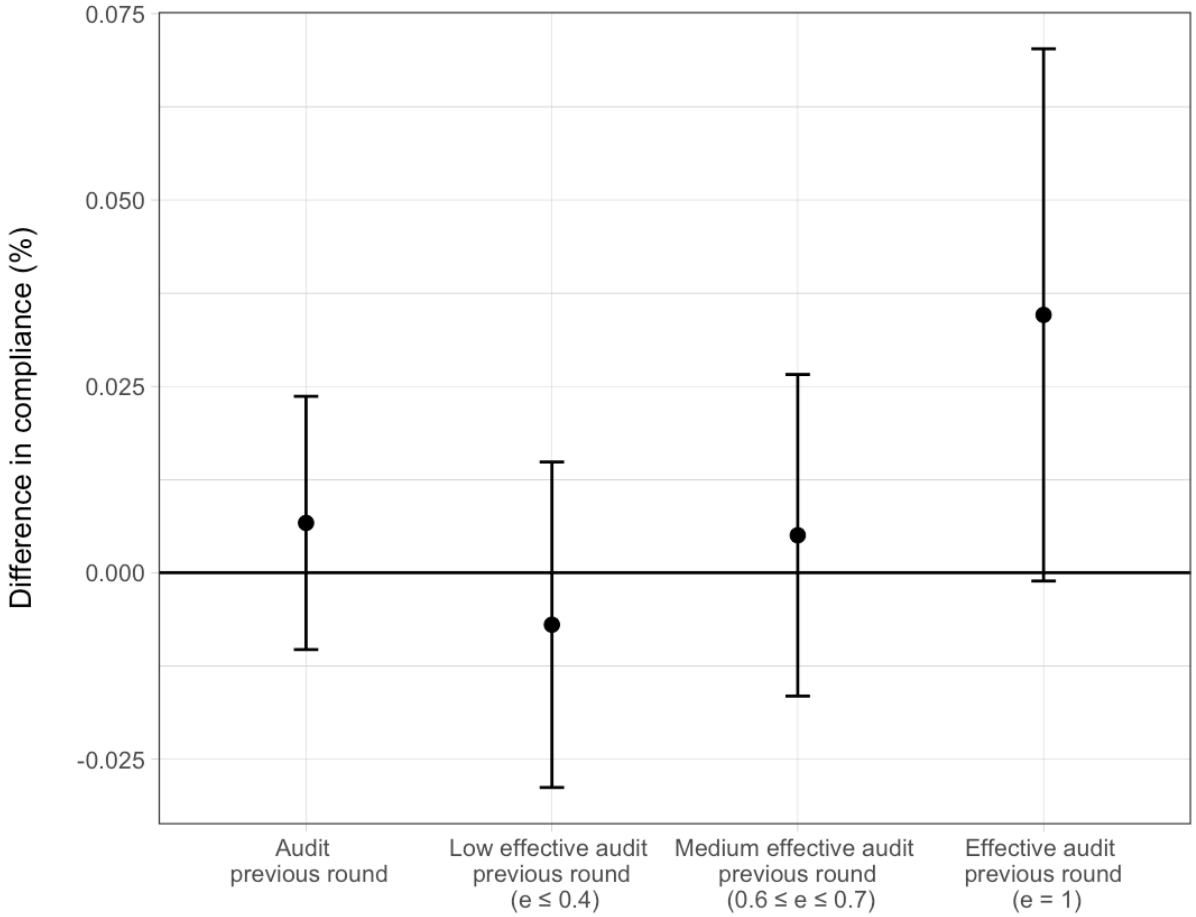
5.3. Specific Deterrent Effects of Audits

To investigate specific deterrent effects of audits, we add different indicator variables that reflect different types of past audits (see column *Audit type* in Table 1) to our baseline model. Specifically, we introduce the indicator variables *Audit previous round, Low effective audit previous round, Medium effective audit previous round,* and *Effective audit previous round*, which take respectively the value of 1 in rounds that directly follow a tax audit (*Audit previous round*), a relatively ineffective (*Low effective audit previous round*), a medium effective (*Medium effective audit previous round*), or an effective audit (*Effective audit previous*

round), and 0 otherwise. Therefore, these variables allow an initial assessment of the effect of audit effectiveness on post-audit tax compliance.

Figure 2 depicts mean differences in compliance between the two levels of these variables. For example, it shows the differences in compliance between all rounds that directly succeed audits (*Audit previous round*) and all other rounds. Figure 2 reveals that compliance rates in rounds after *Low effective audits* and also after *Medium effective audits* are not different from compliance in all other rounds. Importantly, however, Figure 2 provides some indication for a specific deterrence effect of effective audits because compliance is higher in rounds that follow an effective audit (*Effective audit previous round*).

Figure 2: Specific Deterrent Effect of Audits



Notes: Figure 2 depicts mean differences in compliance between the two levels of the indicator variables *Audit previous round, Low effective audit previous round, Medium effective audit previous round,* and *Effective audit previous round*. These variables take the value of 1 if the previous round was audited (*Audit previous round*), if the previous round was audited and the audit was low effective (*Low effective audit previous round*), if the previous round was audited and the audit was medium effective (*Medium effective audit previous round*), or if the previous round was audited and the audit was effective (*Effective audit previous round*) and 0 otherwise. Error bars depict 95 percent confidence intervals.

Table 3 presents regression results for these comparisons. In these analyses, we extend our baseline specification with the indicator variables introduced above (*Audit previous round*, *Low effective audit previous round*, *Medium effective audit previous round*, and *Effective audit previous round).* We also account for the fine that audited taxpayers might have had to pay in the previous round (*Paid fine*). Models 3 and 4 suggest that in the aggregate compliance does not change after an audit (*Audit previous round*, both $p>0.16$); that is, we find no evidence of a specific deterrent effect of audits, and we also find no evidence of a bomb crater effect. In line with Figure 2, these results also provide no indication of a specific deterrent effect of low or medium effective audits (all $p>0.18$). However, Models 9 and 10 reveal that compliance is approximately 4 percentage points higher in rounds that follow an effective audit that detected all undeclared income, as indicated by *Effective audit previous round* (both $p=0.01$). In sum, these results provide initial evidence that audit effectiveness affects post-audit tax compliance and that ineffective audits do not deter future noncompliance.

Table 3: Effects of Audits on Post-audit Tax Compliance

Dependent variable: Compliance rate								
Independent variable	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Intercept	0.3007 ***	0.3309 ***	0.3074 ***	0.3381 ***	0.3030 ***	0.3334 ***	0.3039 ***	0.3340 ***
	(0.0186)	(0.0667)	(0.0184)	(0.0669)	(0.0184)	(0.0667)	(0.0183)	(0.0666)
Audit previous round	0.0111	0.0109						
	(0.0079)	(0.0079)						
Low effective audit previous round			-0.0102	-0.0102				
			(0.0082)	(0.0082)				
Medium effective audit previous round					0.0117	0.0117		
					(0.0086)	(0.0086)		
Effective audit previous round							0.0383 **	0.0376 **
							(0.0151)	(0.0151)
Received income	-0.0199 ***	-0.0199 ***	-0.0199 ***	-0.0200 ***	-0.0199 ***	-0.0200 ***	-0.0199 ***	-0.0200 ***
	(0.0032)	(0.0032)	(0.0032)	(0.0032)	(0.0032)	(0.0032)	(0.0032)	(0.0032)
Detection risk	0.0092 ***	0.0092 ***	0.0091 ***	0.0091 ***	0.0092 ***	0.0092 ***	0.0091 ***	0.0091 ***
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Paid fine	-0.0064	-0.0062	-0.0023	-0.0023	-0.0046	-0.0045	-0.0071 *	-0.0070 *
	(0.0041)	(0.0041)	(0.0034)	(0.0034)	(0.0036)	(0.0036)	(0.0037)	(0.0037)
Demographic variables?	No	Yes	No	Yes	No	Yes	No	Yes
N	333	333	333	333	333	333	333	333
Observations	9,324	9,324	9,324	9,324	9,324	9,324	9,324	9,324
Marginal R^2 / Conditional R^2	0.033 / 0.462	0.161 / 0.468	0.033 / 0.464	0.160 / 0.470	0.033 / 0 .463	0.161 / 0.469	0.033 / 0.462	0.161 / 0.469

Notes: *, **, and *** indicate significance at the 10 percent, 5 percent, and 1 percent level. Received income, paid fine, and continuous demographic variables are scaled. Marginal R^2 provides the variance explained by fixed effects, and conditional R^2 provides the variance explained by fixed effects and individual-specific random effects.

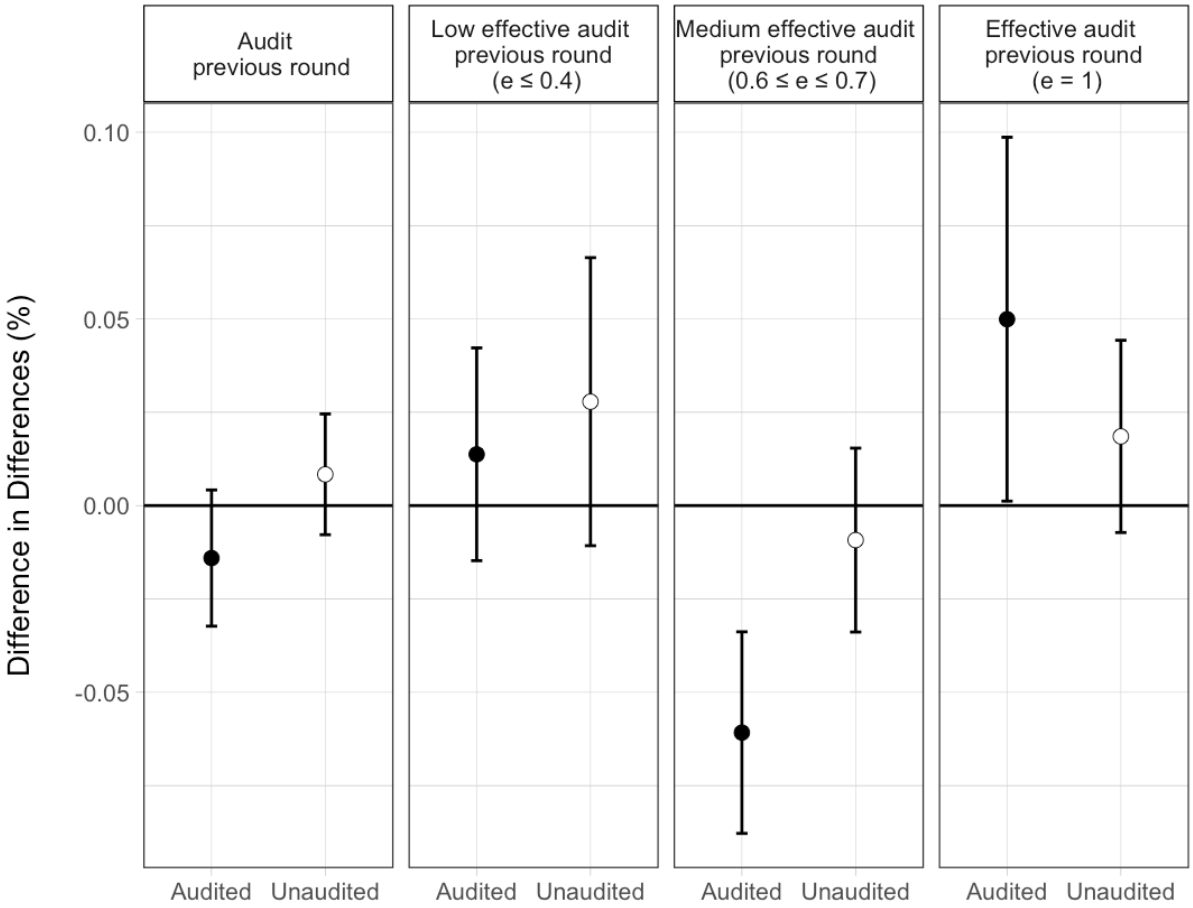
In addition, our analysis of the general deterrent effect of audits suggests that reporting compliance responds strongly to an increase in the risk of detection. However, while these results show that behavior is generally in line with the predictions of expected utility theory, behavioral responses to an audit are not always in line with the predictions of expected utility theory. In particular, the differential effect of audit effectiveness on post-audit tax compliance identified in Figure 2 and Table 3 indicate some behavioral dynamics.

In order to identify more precisely the causal effect of audit effectiveness on post-audit tax compliance, we compare changes in the reporting compliance of audited taxpayers to changes in the reporting compliance of unaudited taxpayers. This comparison accounts for changes in reporting compliance that do not result from the audit experience, and thus it provides a more accurate assessment of the effect of audit effectiveness on post-audit compliance. We use the sample described in Panel C of Table B1 in Appendix B, and we refer to the initial (un-)audited decisions as *initial decisions* and to the subsequent decisions as *subsequent decisions*. To identify changes in compliance between initial reporting decisions and subsequent reporting decisions for audited and unaudited taxpayers, we exclude observations where the initial reporting decision takes place in the last round of the experiment. This reduces our sample to 8,991 initial reporting decisions (4,016 audited reporting decisions and 4,975 unaudited reporting decisions), and thus allows us to analyze 8,991 changes in compliance between these 8,991 initial decisions and the 8,991 subsequent decisions.[17]

In Figure 3 we contrast changes in the reporting compliance of audited taxpayers with changes in the reporting compliance of unaudited taxpayers. We compare how taxpayers who experienced an audit change their reporting compliance in the subsequent round relative to unaudited taxpayers who faced the same risk of an effective or an ineffective audit, but they were *not* randomly selected for an audit. Figure 3 corroborates our initial finding that audit effectiveness has a strong effect on post-audit tax compliance. Specifically, Figure 3 reveals that on average compliance does not change much after an audit (*Audit previous round*). However, this aggregate result belies a more nuanced picture. While low effective audits (*Low effective audit previous round*) and medium effective audits (*Medium effective audit previous round*) have no, or even a counter-deterrent effect on subsequent compliance, effective audits increase compliance in the subsequent round (*Effective audit previous round*).

[17] In several robustness tests, discussed in section 5.3. below, we investigate the effects of different sample selection criteria as described in Table B2 in Appendix B.

Figure 3: Changes in Compliance after Audited and Unaudited Reporting Decisions



Notes: Figure 3 depicts mean changes in compliance from audited rounds to subsequent rounds for *All audits*, *Effective audits*, and *Ineffective audits*, as well as the corresponding changes in compliance from unaudited rounds to the subsequent rounds. Effective audits detected all undeclared income while ineffective audits did not. Sample sizes are shown in Panel C of Table B1 in Appendix B. Error bars depict 95 percent confidence intervals.

To investigate the effect of audit effectiveness on post-audit tax compliance in greater detail, we next analyze how audit effectiveness affects changes in the reporting compliance of audited taxpayers relative to changes in the reporting compliance of unaudited taxpayers, by estimating several variations of the following difference in differences model:

$$(3) \quad Compliance\ Rate_{i,t} = \beta_0 + \beta_1\, Audited_{i,t} + \beta_2\, Subsequent\ decision_{i,t} + \beta_3\, Audit\ effectiveness_{i,t} + \beta_4\, Audited_{i,t}\, x\ Subsequent\ decision_{i,t} + \beta_5\, Audited\ x\ Audit\ effectiveness_{i,t} + \beta_6\, Subsequent\ decision_{i,t}\, x\ Audit\ effectiveness_{i,t} + \beta_7\, Audited_{i,t}\ x\ Subsequent\ decision_{i,t}\, x\ Audit\ effectiveness_{i,t} + \beta_8\, Received\ income_{i,t} + \beta_9\, Detection\ risk_{i,t} + \beta_{10}\, Paid\ fine_{i,t} + \beta_{11}\, Z_i + \varepsilon_{it},$$

where *Audited* indicates whether a taxpayer has been audited or not, *Subsequent decision* indicates whether an observation stems from an (un-)audited initial round or the subsequent

round, and *Audit effectiveness* captures the effectiveness of the audit in the initial round. This specification allows us to estimate the effect of audits on audited taxpayers (relative to unaudited taxpayers), and it also allows us to estimate to what extent the effect of audits depends on audit effectiveness. Specifically, the interaction term *Audited* x *Subsequent decision* reflects the difference between the mean change in compliance of audited taxpayers between the initial (audited) decision and the subsequent decision and the mean change in compliance of unaudited taxpayers between the initial (unaudited) decision and the subsequent decision (i.e., the difference in differences). Accordingly, the interaction term *Audited* x *Subsequent decision* x *Audit effectiveness* reflects the extent to which the effect of audits depends on audit effectiveness. While the simple interaction *Audited* x *Subsequent decision* provides an estimate of the effect of ineffective audits that do not detect any income, the triple interaction *Audited* x *Subsequent decision* x *Audit effectiveness* indicates how a one percentage point increase in audit effectiveness affects the effect of audits. The control and demographic variables are analogous to our prior analysis (see Table B1 for details), and we account for income effects (*Received income*), for the risk of detection in the initial round and the subsequent round (*Detection risk*), and for the fine that taxpayers might have paid as a result of their audit in the initial round (*Paid fine*).

We report the results in Table 5. In line with our prior analyses, the regression results reveal three important results, which demonstrate overall that tax audits have differential effects on post-audit compliance, effects that vary by audit effectiveness and also by a taxpayer's prior reporting compliance. First, we find that audits have the potential to increase or to decrease post-audit tax compliance. Second, we find that effective audits have a more positive effect on post-audit tax compliance than ineffective audits. Third, we find that a taxpayer's reporting behavior prior to an audit determines the behavioral response to audits: audits have a strong effect on relatively compliant taxpayers (the upper 50 percent), while they have a weaker effect on the post-audit compliance of relatively noncompliant taxpayers (the lower 50 percent of the compliance distribution in each reporting decision).

Models 11 and 12 analyze how experiencing an audit affects compliance relative to not experiencing an audit. These models analyze whether changes in compliance after experiencing an audit differ from changes in compliance after not experiencing an audit, as depicted in Figure 3. The estimates again indicate that audit effectiveness has a strong effect on post-audit tax compliance in the aggregate. We estimate that an ineffective audit that does not detect any noncompliance reduces compliance in the subsequent reporting decision by approximately 5

percentage points (*Audited x Subsequent decision,* p=0.07). Conversely, a 1 percentage point increase in audit effectiveness increases compliance in the subsequent reporting decision by 0.09 percentage points (*Audited x Subsequent decision x Audit effectiveness,* p=0.01), so that an effective audit that detects all undeclared income increases compliance in the subsequent reporting decision by approximately 4 percentage points. This result is consistent with the regression results presented in Table 3 above.[18]

Models 13 to 16 complement these findings by analyzing the effect of audits among relatively compliant taxpayers (the upper 50 percent of the compliance distribution, who on average report 91 percent of the income they receive) and relatively noncompliant taxpayers (the lower 50 percent of the compliance distribution in each reporting decision, who on average report 18 percent of the income they receive). The coefficients on *Subsequent decision* indicate that reporting compliance is volatile, even absent any audits. In particular, participants decrease their compliance by 16 percentage points after a reporting decision in which they declared a relatively large share of their income, while they increase their compliance by 17 percentage points after a reporting decision in which they declared relatively little income (all $p < 0.001$).

Moreover, our estimates indicate that prior compliance has a strong effect on post-audit tax compliance, with the aggregate effect of audits on post-audit tax compliance mostly driven by relatively compliant taxpayers. Within this group, an ineffective audit decreases post-audit compliance by 7 percentage points (*Audited x Subsequent decision,* p=0.01), while a 1 percentage point increase in audit effectiveness increases post-audit compliance by 0.11 percentage points (*Audited x Subsequent decision x Audit effectiveness,* p=0.01). Thus, for relatively compliant taxpayers, our estimates suggest that an effective audit increases post-audit compliance by approximately 4 percentage points; for relatively noncompliant taxpayers, our estimates indicate that ineffective audits do not reduce subsequent compliance (*Audited x Subsequent decision,* p=0.78). However, we estimate that a 1 percentage point increase in audit effectiveness increases post-audit compliance by 0.10 percentage points (*Audited x Subsequent decision x Audit effectiveness,* p=0.08) within this group.

[18] We discuss the robustness of these results in section 5.3. below. Overall, we find that this result is robust to different sample selection criteria.

Table 4: Effects of Effective and Ineffective Audits

Dependent variable: Compliance rate						
Subsample	**Entire sample**		**Upper 50%**		**Lower 50%**	
Independent variable	(11)	(12)	(13)	(14)	(15)	(16)
Intercept	0.2835 ***	0.3191 ***	0.6614 ***	0.6862 ***	-0.0566 **	-0.0074
	(0.0215)	(0.0680)	(0.0194)	(0.0433)	(0.0223)	(0.0504)
Audited	0.0209	0.0211	0.0382 *	0.0388 *	-0.0201	-0.0194
	(0.0189)	(0.0189)	(0.0210)	(0.0210)	(0.0239)	(0.0239)
Subsequent decision	0.0174	0.0174	-0.1501 ***	-0.1501 ***	0.1728 ***	0.1728 ***
	(0.0190)	(0.0190)	(0.0212)	(0.0212)	(0.0237)	(0.0237)
Audit effectiveness	0.0003	0.0003	0.0000	0.0000	0.0003	0.0003
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Audited x Subsequent decision	-0.0477 *	-0.0477 *	-0.0746 **	-0.0748 **	0.0094	0.0095
	(0.0265)	(0.0265)	(0.0292)	(0.0292)	(0.0333)	(0.0333)
Audited x Audit effectiveness	-0.0003	-0.0003	-0.0006 **	-0.0006 **	0.0002	0.0002
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0004)	(0.0004)
Subsequent decision x Audit effectiveness	-0.0003	-0.0003	-0.0005 *	-0.0005 *	0.0000	0.0000
	(0.0002)	(0.0002)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Audited x Subsequent decision x Audit effectiveness	0.0010 **	0.0010 **	0.0011 **	0.0011 **	0.0010 *	0.0010 *
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0006)	(0.0006)
Received income	-0.0203 ***	-0.0204 ***	-0.0159 ***	-0.0161 ***	-0.0140 ***	-0.0142 ***
	(0.0023)	(0.0023)	(0.0025)	(0.0025)	(0.0029)	(0.0029)
Detection risk	0.0093 ***	0.0093 ***	0.0066 ***	0.0066 ***	0.0111 ***	0.0111 ***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0004)	(0.0004)
Paid fine	-0.0048	-0.0047	0.0070 **	0.0068 **	-0.0248 ***	-0.0245 ***
	(0.0030)	(0.0030)	(0.0029)	(0.0029)	(0.0076)	(0.0075)
Demographic variables?	No	Yes	No	Yes	No	Yes
N	333	333	320	320	307	307
Observations	17,982	17,982	8,972	8,972	9,010	9,010
Marginal R^2 / Conditional R^2	0.034 / 0.478	0.161 / 0.482	0.131 / 0.387	0.201 / 0.394	0.137 / 0.399	0.206 / 0.404

Notes: *, **, and *** indicate significance at the 10 percent, 5 percent, and 1 percent level. Received income, paid fine, and continuous demographic variables are scaled. Marginal R^2 provides the variance explained by fixed effects, and conditional R^2 provides the variance explained by fixed effects and individual-specific random effects.

To further investigate the effect of tax audits on individuals who differ in their propensity to comply, we investigate the effect of the first audit that taxpayers experience on *honest* and *dishonest* taxpayers, where honest individuals report all income in all rounds prior to their first audit and dishonest individuals report zero income in these rounds. Due to the small sample size, we do not distinguish between different levels of audit effectiveness. In Table 5 we show how (dis-)honest taxpayers who experience an audit as well as (dis-)honest taxpayers who do not experience an audit change their subsequent compliance. Table 5 suggests that the first audit does not have a strong effect on the post-audit tax compliance of honest and dishonest taxpayers in general, and it also does not indicate a crowding-out effect. More specifically, we find that audits do not affect the reporting compliance of taxpayers who report all income in all rounds prior to their first audit.

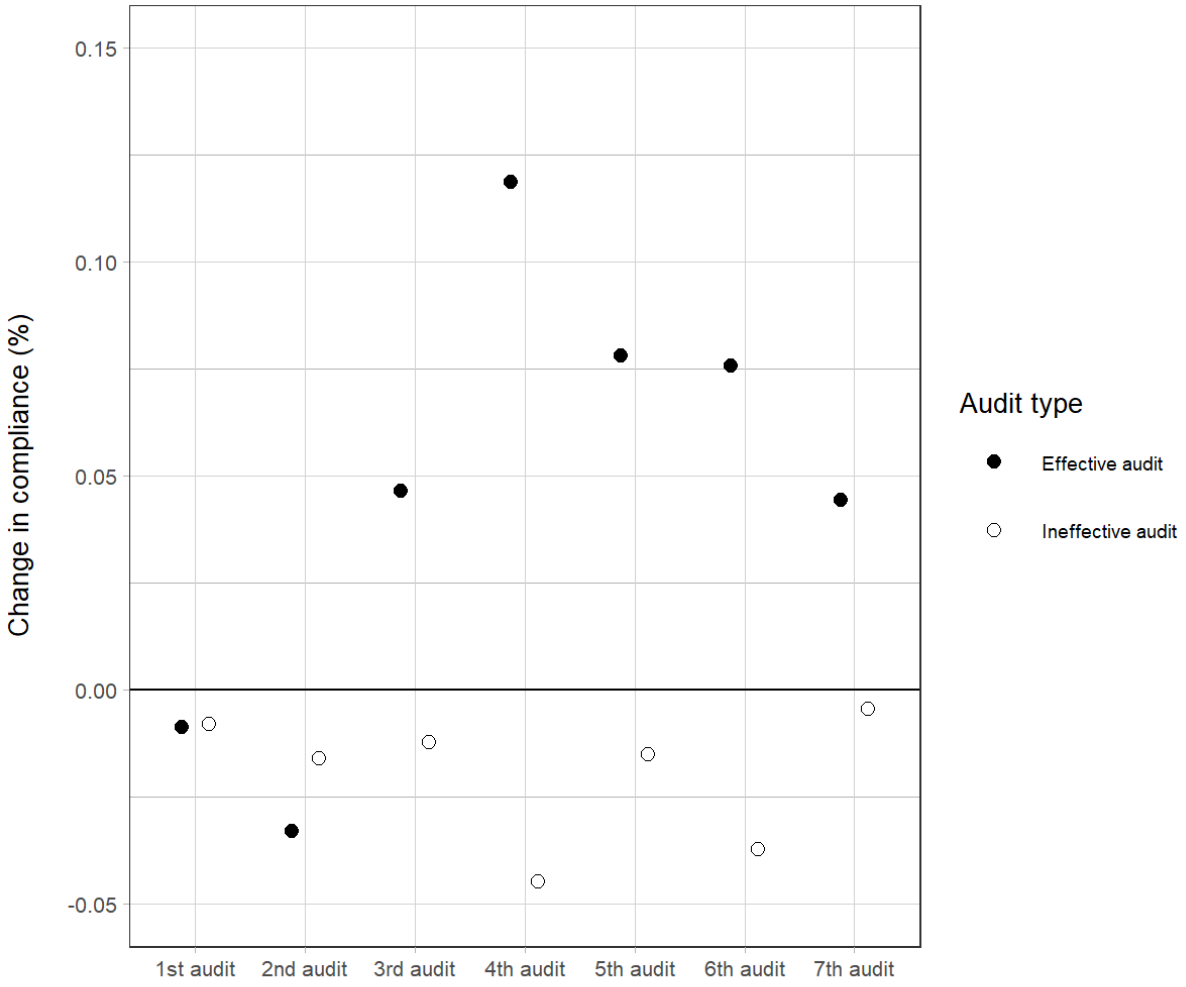
Table 5: Effect of Audits on Honest and Dishonest Taxpayers

Round	Honest taxpayers	Audited	Not Audited
1	Number of honest taxpayers	26	49
	Subsequent change in compliance	- 21%	- 24%
2	Number of honest taxpayers	8	15
	Subsequent change in compliance	- 10%	- 20%
3	Number of honest taxpayers	7	5
	Subsequent change in compliance	0%	0%
4	Number of honest taxpayers	3	2
	Subsequent change in compliance	0%	0%
5	Number of honest taxpayers	2	–
	Subsequent change in compliance	0%	–
Round	**Dishonest taxpayers**	**Audited**	**Not Audited**
1	Number of dishonest taxpayers	19	54
	Subsequent change in compliance	+ 32%	+ 36%
2	Number of dishonest taxpayers	9	15
	Subsequent change in compliance	+ 19%	+ 20%
3	Number of dishonest taxpayers	7	4
	Subsequent change in compliance	+ 7%	0%
4	Number of dishonest taxpayers	2	1
	Subsequent change in compliance	0%	0%

Notes: *Honest* taxpayers report all income in all rounds prior to their first audit. *Dishonest* taxpayers report zero income in these rounds

Finally, in Figure 4 we explore whether the effect of audit effectiveness on post-audit compliance diminishes over time. We depict the change in the compliance rate between the round that is audited and the subsequent round conditional on audit effectiveness. Figure 4 suggests that the effect of audit effectiveness on post-audit tax compliance does not diminish over time.

Figure 4: Changes in Compliance after Repeated Audits



Notes: Figure 4 depicts the aggregate change in compliance between the round that is audited and the subsequent round conditional on audit effectiveness for the first seven audits. Note that all taxpayers experience at least seven audits.

5.4. Robustness Tests and Supplemental Analysis

We have estimated numerous additional regression models to examine the robustness of our findings, as reported in Appendix B. First, we investigate whether the results presented in Table 5 are robust to different sample selection criteria (as described in Table B2). To this end we exclude cases where taxpayers experienced an audit in one (Models 17 and 18), two (Models 19 and 20), or three (Models 21 and 22) rounds prior to the audit under investigation. As reported in Table B3, we continue to find that ineffective audits reduce post-audit tax compliance while effective audits increase post-audit tax compliance.

Second, we investigate how audits affect compliance in subsequent reporting decisions. To this end, we analyze the effect of audits on the second (Models 23 and 24), third (Models 25 and 26), fourth (Models 27 and 28), and fifth (Models 29 and 30) subsequent reporting decision. Table B4 reveals that the specific deterrent effect of audits does not extend beyond the round that follows the audit. In line with experimental work on the bomb crater effect (Guala & Mittone, 2005; Mittone 2006), this result suggests that the audit experience does not have a sustained effect on post-audit tax compliance. The specific deterrent effect of audits might be particularly short-lived in our experiment because taxpayers were audited frequently and experienced at least 7 (and up to 19) audits throughout the experiment.

We also examine the effects of various control and demographic variables. We find a negative effect of received income, a common result in laboratory experiments (Alm, 2019). With regard to the demographic variables, we find that age and being female have positive effects on compliance, and that participants from German-speaking countries are less compliant than participants from other countries. We also find that individuals who indicated in the post-experimental survey that they tried to maximize their income reported smaller shares of their income, while individuals who indicated that they believe that tax evasion cannot be justified reported larger shares of their income. We present these results in detail in Table B5 in Appendix B.

Finally, in several unreported regressions, we test whether the results presented in Tables 3 and 4 are robust to changes in the dependent variable. Using *Evaded income* (i.e., received income minus reported income) as the dependent variable does not affect any of the results.

6. Conclusions

How do audits affect post-audit tax compliance? In this paper we study the specific deterrent effect of tax audits by analyzing two aspects of behavioral responses to enforcement. First, we investigate how ineffective audits that do not detect all undeclared income affect subsequent reporting behavior. Second, we examine the behavioral mechanisms that drive these responses and analyze differential responses of relatively noncompliant, relatively compliant, honest, and dishonest taxpayers. Moreover, our research design also allows us to investigate the general deterrent effect of audit effectiveness, and it allows us to test whether presenting the risk of detection as a compound risk with uncertain detection affects compliance decisions

relative to a decision with certain detection in case of an audit. We investigate these issues in a preregistered laboratory experiment in which taxpayers receive income and decide how much they declare to the tax agency and in which they also face the risk of being audited and fined for undeclared income that is detected on audit. We introduce variation in the audit probability and in the audit effectiveness in order to assess behavioral responses to changes in these factors.

Our results suggest that tax audits have different effects on post-audit compliance and that behavioral responses to enforcement are not always in line with the assumptions of the standard model of tax evasion. Our first main result relates to the role of audit effectiveness on the impact of audits. We find that tax audits do not have a positive effect on post-audit tax compliance in the aggregate. Instead, we find that the specific deterrent effect of tax audits depends strongly on audit effectiveness. While taxpayers increase their compliance after an effective audit that detected all undeclared income, they decrease their compliance after an ineffective audit that does not detect all undeclared income. This result suggests that ineffective tax audits stimulate risk-taking and that taxpayers whose underreporting was not detected during an audit contribute to the decline in post-audit compliance found in some prior studies (Gemmell & Ratto, 2012; Beer et al., 2020). In contrast, our results do not suggest that behavioral responses to audits are driven by the fine for noncompliance that has been detected in the past. We also find that increasing the effectiveness of audits increases the general deterrent effect of audits. Overall, participants responded strongly to an increase in the risk of detection. This strengthens our confidence in the internal validity of our results.

Our second main result relates to the ways in which a taxpayer's prior reporting behavior affects specific deterrence. We find consistent and robust evidence that post-audit compliance depends on taxpayers' prior reporting behavior. Specifically, relatively compliant taxpayers adjust their post-audit compliance in response to the audit effectiveness. These taxpayers appear to revise upwards their prior on the probability of a future audit after experiencing an audit that detected all their cheating, while they appear to revise downwards their prior on the probability of a future audit after experiencing an audit that did not detect all their cheating. This suggests that the availability heuristic informs a taxpayer's decision to increase or decrease his or her post-audit compliance. In contrast, relatively noncompliant taxpayers appear to be motivated strongly by the expected value of the evasion gamble and thus do not alter their reporting behavior substantially after "losing the audit lottery".

Third, we find no support for the hypothesis that audits affect taxpayers' motivations to comply. Specifically, we find no indication of a crowding-out effects of audits; that is, we find

that the audit experience does not reduce compliance among honest taxpayers, and similarly we do not find that dishonest taxpayers increase their compliance after experiencing an audit.

Fourth, in contrast to some prior work, we find no indication of a systematic misperception of compound detection risk (when audits are ineffective) relative to decisions where an audit leads to certain detection. Indeed, it is important to recognize that there was no uncertainty present in our design: participants knew the exact consequences of their reporting decisions, which reduces the margin for such bias. We also find that compliance choices are unaffected by the way in which the relevant factors are presented (e.g., showing the audit probability before the audit effectiveness and vice versa). In light of the results reported in Bernasconi and Bernhofer (2020), these findings suggest that transparency about the magnitude and the effects of these parameters as well as learning effects diminish the misperception of compound detection lotteries.

Taken together, our findings challenge the standard result – and common assumption – that more audits always lead to more compliance. Moreover, our results provide a more nuanced perspective on the finding that audited taxpayers generally tend to underestimate the risk of future examinations (Guala & Mittone, 2005; Mittone, 2006; Mittone et al., 2017) and indicate that loss-repair motivations alone do not explain behavioral responses to enforcement (Maciejovsky et al., 2007; McKee et al., 2018). This has important implications for tax administrations. Our study suggests that increasing the capacity of tax audits to detect noncompliance and improving the targeting of noncompliant taxpayers are both crucial in establishing and maintaining compliance. More specifically, our results indicate that tax agencies should publicize any increase in its capacity to conduct effective audits if they wish to increase aggregate compliance levels. Our findings also suggest that audits should be conducted thoroughly, for instance by focusing on the entire tax return rather than specific line items, in order to reduce the chances of missing noncompliance. Since the behavioral response to audit effectiveness is strongest among relatively compliant taxpayers who evade only a minor share of their income, audits that reveal seemingly small tax adjustments should be particularly thorough. Finally, our study suggests that lenient interpretation of the tax code in audits might stimulate future noncompliance.

Future work should investigate the effect of the audit selection mechanism on post-audit tax compliance. While in practice most audits target taxpayers with a relatively high likelihood of noncompliance, our study employs a random audit selection mechanism, common to many if not all laboratory experiments. A taxpayer, and particularly a relatively compliant taxpayer,

who has been randomly selected for audit might fall for the bomb crater fallacy, underestimate the risk of a future examination, and thus report less income after the audit. Conversely, taxpayers who have been targeted based on their prior reporting behavior might be less likely to exhibit such bias. Moreover, future work should investigate the role of loss-repair motivations in post-audit tax compliance. While prior studies find mixed evidence for taxpayers' tendency to make up for past losses, our results raise the question to what extent loss-repair motivations contribute to the decline in post-audit compliance after ineffective audits. Finally, future studies should investigate how uncertainty about the audit probability and the audit effectiveness affects post-audit tax compliance. In particular, taxpayers who experience a tax audit might only realize during the audit that the tax agency is unable to detect undeclared income. Conversely, taxpayers might suspect that tax audits sometimes fail to detect cheating, but they do not know how frequent such ineffective audits are. The effects of such uncertainty on post-audit tax compliance are unanswered questions, and they await future experimental investigations.

References

Advani, A., Elming, W., & Shaw, J. (2017). The dynamic effects of tax audits. No. W17/24. London, UK: The Institute for Fiscal Studies.

Allingham, M. G., & Sandmo, A. (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., & Jacobson, S. (2007). Using laboratory experiments in public economics. National Tax Journal 60 (1), 129-152.

Alm, J., & Kasper, M. (2021). Laboratory experiments. In B. Van Rooij & D. Sokol, eds., Cambridge Handbook of Compliance. Cambridge, UK: Cambridge University Press, 707-727.

Alm, J., McClelland, G. H., & Schulze, W. D. (1992). Why do people pay taxes? Journal of Public Economics 48, 21-38.

Alm, J., & McKee, M. (2006). Audit certainty, audit productivity, and taxpayer compliance. National Tax Journal 59 (4), 801-816.

Andreoni, J., Erard, B., & Feinstein, J. S. (1998). Tax compliance. The Journal of Economic Literature 36 (2), 818-860.

Becker, G. S. (1968). Crime and punishment – An economic approach. The Journal of Political Economy 76 (2), 169-217.

Beer, S., Kasper, M., Kirchler, E., & Erard, B. (2020). Do audits deter or provoke future tax noncompliance? Evidence on self-employed taxpayers. CESifo Economic Studies 66 (3), 248-264.

Bergman, M., & Nevarez, A. (2006). Do audits enhance compliance? An empirical assessment of VAT enforcement. National Tax Journal 59 (4), 817-832.

Bernasconi, M. (1998). Tax evasion and orders of risk aversion. Journal of Public Economics 67 (1), 123-134.

Bernasconi, M., & Bernhofer, J. (2020). Catch me if you can: Testing the reduction of compound lotteries axiom in a tax compliance experiment. Journal of Behavioral and Experimental Economics 84, Article 101479.

Bernasconi, M., and Zanardi, A. (2004). Tax evasion, tax rates, and reference dependence. FinanzArchiv 60 (3), 422-445.

Best, M., Shah, J., & Waseem, M. (2021). Detection without deterrence: Long-run effects of tax audits on firm behavior. Working Paper.

Braithwaite, V. (2003). Taxing democracy: Understanding tax avoidance and tax evasion. In V. Braithwaite, ed., Dancing with Tax Authorities: Motivational Postures and Non-compliant Actions. Aldershot, UK: Ashgate, 15-39.

Braithwaite, V. (2009) Defiance in Taxation and Governance – Resisting and Dismissing Authority in a Democracy. Cheltenham, UK and Northhampton, MA: Edward Elgar Publishing Limited.

Chalfin, A., & McCrary, J. (2017). Criminal deterrence: A review of the literature. The Journal of Economic Literature 55 (1), 5-48.

Cullen, F. T., Jonson, C. L., & Nagin, D. S. (2011). Prisons do not reduce recidivism: The high cost of ignoring science. The Prison Journal 91(3_suppl), 48S-65S.

DeBacker, J., Heim, B. T., Tran, A., & Yuskavage, A. (2018). Once bitten, twice shy? The lasting impact of IRS audits on individual tax reporting. The Journal of Law and Economics 61 (1), 1-35.

Dusek, L., & Traxler, C. (2020). Learning from law enforcement. CESifo Working Paper Series 8043. Munich, Germany.

Dhami, S., and al-Nowaihi, A. (2007). Why do people pay taxes? Prospect theory versus expected utility theory. Journal of Economic Behavior & Organization 64 (1), 171-192.

Dillenberger, D. (2010). Preferences for one-shot resolution of uncertainty and Allais-type behavior. Econometrica 78 (6), 1973-2004.

Earnhart, D., & Friesen, L. (2013). Can punishment generate specific deterrence without updating? Analysis of a stated choice scenario. Environmental and Resource Economics 56 (3), 379-397.

Enachescu, J., Olsen, J., Kogler, C., Zeelenberg, M., Breugelmans, S. M., & Kirchler, E. (2019). The role of emotions in tax compliance behavior: A mixed-methods approach. Journal of Economic Psychology 74, 102194.

Erard, B. (1992). The influence of tax audits on reporting behavior. In J. Slemrod, ed., Why People Pay Taxes: Tax Compliance and Enforcement. Ann Arbor, MI: The University of Michigan Press, 95-114.

Erard, B., & Feinstein, J. S. (1994). Honesty and evasion in the tax compliance game. The RAND Journal of Economics 25 (1), 1-19.

Erard, B., E. Kirchler, & Olsen, J. (2019). Audit impact study: The specific deterrence implications of increased reliance on correspondence audits. In Taxpayer Advocate

Service: Annual Report to Congress 2019. Washington, D.C.: Taxpayer Advocate Service, 257-268.

Feinstein, J. S. (1991). An econometric analysis of income tax evasion and its detection. The RAND Journal of Economics 22 (1), 14-35.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. Experimental Economics 10 (2), 171-178.

Frey, B. (1997). Not Just For the Money – An Economic Theory of Personal Motivation. Cheltenham, UK: Edward Elgar Publishing Limited.

Gemmell, N., & Ratto, M. (2012). Behavioral responses to taxpayer audits: Evidence from random taxpayer inquiries. National Tax Journal 65 (1), 33-58.

Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with orsee. Journal of the Economic Science Association 1 (1), 114-125.

Guala F., & Mittone, L. (2005). Experiments in economics: External validity and the robustness of phenomena. Journal of Economic Methodology 12, 495-515.

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

Harrison, G. W., Martinez-Correa, J., & Swarthout, J. T. (2015). Reduction of compound lotteries with objective probabilities: Theory and evidence. Journal of Economic Behavior and Organization 119, 32–55.

Haselhuhn, M. P., Pope, D. G., Schweitzer, M. E., & Fishman, P. (2012). The impact of personal experience on behavior: Evidence from video-rental fines. Management Science 58 (1), 52-61.

Hu, F., & Ben-Ner, L. (2020). The effects of feedback on lying behavior: Experimental evidence. Journal of Economic Behavior and Organization 171, 24-34.

Internal Revenue Service (2019). Data Book, 2018, Publication 55B. Washington, D.C.: Internal Revenue Service.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica 47 (2), 263-292.

Kastlunger, B., Kirchler, E., Mittone, L., & Pitters, J. (2009). Sequences of audits, tax compliance, and taxpaying strategies. Journal of Economic Psychology 30 (3), 405-418.

Keen, M., & Slemrod, J. (2017). Optimal tax administration. Journal of Public Economics 152, 133-142.

Kirchler, E. (2007). The Economic Psychology of Tax Behavior. Cambridge, UK: Cambridge University Press.

Kirchler, E, Hoelzl, E., & Wahl, I. (2008). Enforced versus voluntary tax compliance: The "slippery slope" framework. Journal of Economic Psychology 29 (2), 210-225.

Kleven, H. J., Knudsen, M. B., Kreiner, C. T., Pedersen, S., & Saez, E. (2011). Unwilling or unable to cheat? Evidence from a randomized tax audit experiment in Denmark. Econometrica 79 (3), 651-692.

Kotsogiannis, C., Salvadori, L., Karangwa, J., & Mukamana, T. (2021). Do tax audits have a dynamic impact? Evidence from corporate income tax administrative data." TARC Discussion Paper 035-21. Exeter, UK: University of Exeter.

Lederman, L. (2018). Does enforcement reduce voluntary tax compliance? Brigham Young University Law Review 3, 623-694.

Maciejovsky, B., Kirchler, E., & Schwarzenberger, H. (2007). Misperceptions of chance and loss repair: On the dynamics of tax compliance. Journal of Economic Psychology 28 (6), 678-691.

Matsueda, R. L., Kreager, D. A., & Huizinga, D. (2006). Deterring delinquents: A rational choice model of theft and violence. American Sociological Review 71(1), 95-122.

McKee, M., Siladke, C. A., & Vossler, C. A. (2018). Behavioral dynamics of tax compliance when taxpayer assistance services are available. International Tax and Public Finance 25 (3), 722-756.

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

Mittone, L. (2006). Dynamic behaviour in tax evasion: An experimental approach. Journal of Socio-Economics 35 (5), 813-835.

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

Nagin, D. S. (2013a). Deterrence in the twenty-first century. Crime and Justice 42 (1), 199-263.

Nagin, D. S. (2013b). Deterrence: A review of the evidence by a criminologist for economists. Annual Review of Economics 5 (1), 83-105.

Nagin, D. S., Cullen, F. T., & Jonson, C. L. (2009). Imprisonment and reoffending. Crime and Justice 38 (1), 115-200.

Olsen, J., Kasper, M., Enachescu, J., Benk, S., Budak, T., & Kirchler, E. (2018). Emotions and tax compliance among small business owners: An experimental survey. International Review of Law and Economics, 56, 42-52.

Plott, C. R. (1987). Dimensions of parallelism: Some policy applications of experimental methods. In A. E. Roth, ed., Laboratory Experimentation in Economics: Six Points of View. New York, NY: Cambridge University Press, 193-217.

Prokosheva, S. (2016). Comparing decisions under compound risk and ambiguity: The importance of cognitive skills. Journal of Behavioral and Experimental Economics 64, 94-105.

Rablen, M. D. (2014). Audit probability versus effectiveness: The Beckerian approach revisited. Journal of Public Economic Theory 16 (2), 322-342.

Simonsohn, U., Karlsson, N., Loewenstein, G., & Ariely, D. (2008). The tree of experience in the forest of information: Overweighing experienced relative to observed information. Games and Economic Behavior 62 (1), 263-286.

Smith, V. L. (1982). Microeconomic systems as an experimental science. The American Economic Review 72 (4), 923-955.

Snow, A., & Warren, R. S. (2007). Audit uncertainty, Bayesian updating, and tax evasion. Public Finance Review 35 (5), 555-571.

Skinner, J., & Slemrod, J. (1985). An economic perspective on tax evasion. National Tax Journal 38 (3), 345-353.

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

Spicer, M. W., & Hero, R. E. (1985). Tax evasion and heuristics: A research note. Journal of Public Economics 26 (2), 263-267.

Srinivasan, T. N. (1973). Tax evasion: A model. Journal of Public Economics 2 (4), 339–346.

Torgler, B. (2003) Tax morale, rule-governed behaviour, and trust. Constitutional Political Economy 14 (2), 119-140.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. Cognitive Psychology 5 (2), 207-232.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. Science 185 (4157), 1124-1131.

Yaniv, G. (1999). Tax compliance and advance tax payments: A prospect theory analysis. National Tax Journal 53 (4), 753-764.

Yitzhaki, S. (1974). A note on income tax evasion: A theoretical analysis. Journal of Public Economics 3 (2), 201-202.

Appendix A

Instructions & Screen Shots

Attached in a separate document ("Experimental Instructions")

Appendix B

Table B1: Data Description (Panel A)

Variable	Description	Mean	SD
Dependent Variables			
Compliance rate	Reported income divided by received income	0.54	0.41
Experimental Treatment Variables			
Received income	Income received (in ECU)	2700.16	430.04
Detection risk	Audit probability x Audit effectiveness	0.26	0.08
Demographic Variables			
Female	= 1 if participant is female and 0 if male	0.57	0.50
Age	Participant's age in years	25.94	6.06
Higher education	= 1 if completed Bachelor Studies or higher and 0 if otherwise	0.51	0.49
Economics major	= 1 if Major in Economics and 0 if otherwise	0.08	0.27
German speaking	= 1 if Austrian or German and 0 if otherwise	0.48	0.50
Prior experiments	= 1 if prior participation in laboratory experiments and 0 if otherwise	0.95	0.23
Prior tax experiments	= 1 if prior participation in tax experiments and 0 if otherwise	0.16	0.37
Self-preparation	= 1 if self-prepared tax return in the past and 0 if otherwise	0.29	0.46
Risk seeking[#]	Do you like to gamble? (0 to 9)	4.36	2.36
Income maximization[#]	To what extent did you try to maximize your income? (0 to 9)	6.27	2.34
Tax morale[#]	Do you think cheating on tax if you have a chance can be justified? (0 to 9)	6.05	2.68
Grouping Variables			
Parameter order	= 1 if audit probability presented before audit effectiveness and 0 if audit effectiveness presented before audit probability		
Audit previous round	= 1 if previous round was audited and 0 if previous round was not audited		
Low effective audit previous round	= 1 if previous round was audited and audit effectiveness was smaller than 0.6 and 0 in all other rounds		
Medium effective audit previous round	= 1 if previous round was audited and audit effectiveness was smaller than 1 but greater than 0.4 and 0 in all other rounds		
Effective audit previous round	= 1 if previous round was audited and audit was effective and 0 if previous round was not audited or audit was ineffective		
Ineffective audit previous round	= 1 if previous round was audited and audit was ineffective and 0 if previous round was not audited or audit was effective		
Audited	= 1 if reporting decision was audited and 0 if reporting decision was not audited		
Audit effectiveness	Effectiveness in round that was (un-)audited		
Subsequent decision	= 1 if reporting decision directly follows an (un-)audited round and 0 if reporting decision is (un-)audited		

Notes: [#] denotes a scale from 0 to 9, where higher values indicate more risk-seeking, more income maximization, and higher tax morale.

Table B1: Sample Description (Panel B)

Sample	Time	Audit type	Audited	Unaudited	Total
All observations					
	All	Effective audit	580	2,084	2,664
		Ineffective audit	3,551	3,109	6,600
		All	4,131	5,193	9,324

Table B1: Sample Description (Panel C)

Sample	Time	Audit type	Audited	Unaudited	Total
Initial decisions (Rounds 1 – 27) and subsequent decisions					
	Initial decision	Effective audit	547	1,936	2,483
		Ineffective audit	3,469	3,039	6,508
	Subsequent decision	Effective audit	547	1,936	2,483
		Ineffective audit	3,469	3,039	6,508
		All	8,032	9,950	17,982

Notes: *Audited* reporting decisions were randomly selected for an audit, while *Unaudited* reporting decisions were not. *Effective audits* detect all undeclared income, while *Ineffective audits* do not detect all undeclared income. *Initial decision* refers to the round that is (un-)audited, and *Subsequent decision* refers to the subsequent round.

Table B2: Subsample Description

Sample Selection	Time	Audit type	Audited	Unaudited	Total
Subsample I: No audit in round prior to initial decision					
	Initial decision	Effective audit	275	997	1,272
		Ineffective audit	1,872	1,639	3,511
	Subsequent decision	Effective audit	275	997	1,272
		Ineffective audit	1,872	1,639	3,511
		All	4,294	5,272	9,566
Subsample II: No audit in two rounds prior to initial decision					
	Initial decision	Effective audit	144	495	639
		Ineffective audit	1,008	877	1,885
	Subsequent decision	Effective audit	144	495	369
		Ineffective audit	1,008	877	1,885
		All	2,304	2,744	5,048
Subsample III: No audit in three rounds prior to initial decision					
	Initial decision	Effective audit	78	243	321
		Ineffective audit	530	460	990
	Subsequent decision	Effective audit	75	243	321
		Ineffective audit	530	460	990
		All	1,216	1,446	2,662
Subsample IV: Initial decision and second subsequent decision					
	Initial decision	Effective audit	287	1,032	1,319
		Ineffective audit	1,826	1,604	3,430
	Second subsequent decision	Effective audit	287	1,032	1,319
		Ineffective audit	1,826	1,604	3,430
		All	4,226	5,272	9,498
Subsample V: Initial decision and third subsequent decision					
	Initial decision	Effective audit	150	507	657
		Ineffective audit	974	865	1,839
	Third subsequent decision	Effective audit	150	507	657
		Ineffective audit	974	865	1,839
		All	2,248	2,744	4,992
Subsample VI: Initial decision and fourth subsequent decision					
	Initial decision	Effective audit	80	264	344
		Ineffective audit	512	439	951
	Fourth subsequent decision	Effective audit	80	264	344
		Ineffective audit	512	439	951
		All	1,184	1,406	2,590
Subsample VII: Initial decision and fifth subsequent decision					
	Initial decision	Effective audit	45	133	178
		Ineffective audit	267	217	484
	Fifth subsequent decision	Effective audit	45	133	178
		Ineffective audit	267	217	484
		All	624	700	1,324

Notes: To investigate the specific deterrent effect of tax audits over time, we analyze (un-)audited reporting decisions and the 2nd to 5th subsequent reporting decision. We exclude taxpayers from these analyses who were audited more than once (or audited at all in case of unaudited taxpayers) in the respective time frame. For example, audited taxpayers in subsample VII were audited and they were subsequently not audited for at least five rounds. Unaudited

taxpayers in this subsample were not audited for at least 6 rounds. *Effective audits* detect all undeclared income, while *Ineffective audits* do not detect all undeclared income. *Initial decision* refers to the round that is (un-)audited, and *Subsequent decision* refers to the n^{th} subsequent round.

Table B3: Effect of Audits on Post-audit Tax Compliance with Different Inclusion Criteria (Not Audited for One (I), Two (II), or Three (III) Rounds Prior to the Initial Decision)

Dependent variable: Compliance rate

Subsample	**I**		**II**		**III**	
Independent variable	(17)	(18)	(19)	(20)	(21)	(22)
Intercept	0.2720 ***	0.3106 ***	0.2456 ***	0.2753 ***	0.1937 ***	0.2457 ***
	(0.0259)	(0.0711)	(0.0327)	(0.0778)	(0.0431)	(0.0905)
Audited	0.0288	0.0288	0.0451	0.0447	0.1397 ***	0.1423 ***
	(0.0260)	(0.0260)	(0.0358)	(0.0357)	(0.0501)	(0.0499)
Subsequent decision	0.0262	0.0262	-0.0152	-0.0151	-0.0259	-0.0259
	(0.0258)	(0.0258)	(0.0351)	(0.0351)	(0.0481)	(0.0481)
Audit effectiveness	0.0003	0.0003	0.0002	0.0002	0.0003	0.0003
	(0.0002)	(0.0002)	(0.0003)	(0.0003)	(0.0005)	(0.0005)
Audited x Subsequent decision	-0.0616 *	-0.0617 *	-0.0230	-0.0229	-0.1113 *	-0.1112 *
	(0.0362)	(0.0362)	(0.0491)	(0.0491)	(0.0671)	(0.0671)
Audited x Audit effectiveness	-0.0004	-0.0004	-0.0005	-0.0005	-0.0021 ***	-0.0022 ***
	(0.0004)	(0.0004)	(0.0005)	(0.0005)	(0.0007)	(0.0007)
Subsequent decision x Audit effectiveness	-0.0004	-0.0004	0.0002	0.0002	0.0005	0.0005
	(0.0003)	(0.0003)	(0.0005)	(0.0005)	(0.0006)	(0.0006)
Audited x Subsequent decision x Audit effectiveness	0.0013 **	0.0013 **	0.0007	0.0006	0.0021 **	0.0020 **
	(0.0005)	(0.0005)	(0.0007)	(0.0007)	(0.0010)	(0.0010)
Received income	-0.0200 ***	-0.0202 ***	-0.0181 ***	-0.0184 ***	-0.0125 **	-0.0128 **
	(0.0031)	(0.0031)	(0.0043)	(0.0043)	(0.0062)	(0.0061)
Detection risk	0.0092 ***	0.0092 ***	0.0103 ***	0.0102 ***	0.0119 ***	0.0119 ***
	(0.0004)	(0.0004)	(0.0005)	(0.0005)	(0.0007)	(0.0007)
Paid fine	-0.0096 **	-0.0096 **	-0.0071	-0.0066	-0.0074	-0.0068
	(0.0042)	(0.0042)	(0.0057)	(0.0057)	(0.0078)	(0.0078)
Demographic variables?	No	Yes	No	Yes	No	Yes
N	333	333	333	333	314	314
Observations	9,566	9,566	5,048	5,048	2,622	2,622
Marginal R^2 / Conditional R^2	0.035 / 0.482	0.161 / 0.488	0.044 / 0.502	0.168 / 0.511	0.061 / 0.515	0.176 / 0.523

Notes: *, **, and *** indicate significance at the 10 percent, 5 percent, and 1 percent level. Received income, paid fine, and continuous demographic variables are scaled. Marginal R^2 provides the variance explained by fixed effects, and conditional R^2 provides the variance explained by fixed effects and individual-specific random effects.

Table B4: Effect of Audits on Post-audit Tax Compliance (Two (IV), Three (V), Four (VI), or Five (VII) Rounds After the Audit)

Dependent variable: Compliance rate								
Subsample	**IV**		**V**		**VI**		**VII**	
Rounds after audit	**2**		**3**		**4**		**5**	
Independent variable	(23)	(24)	(25)	(26)	(27)	(28)	(29)	(30)
Intercept	0.3118 ***	0.3318 ***	0.3085 ***	0.3120 ***	0.2844 ***	0.2577 ***	0.1624 **	0.2113 *
	(0.0261)	(0.0713)	(0.0334)	(0.0762)	(0.0456)	(0.0895)	(0.0646)	(0.1100)
Audited	-0.0005	0.0001	-0.0295	-0.0281	-0.0570	-0.0563	0.0626	0.0672
	(0.0264)	(0.0264)	(0.0367)	(0.0367)	(0.0521)	(0.0519)	(0.0749)	(0.0743)
Subsequent decision	-0.0206	-0.0206	-0.0430	-0.0430	-0.0177	-0.0177	0.0600	0.0600
	(0.0261)	(0.0261)	(0.0362)	(0.0362)	(0.0507)	(0.0507)	(0.0720)	(0.0717)
Audit effectiveness	-0.0000	-0.0000	-0.0002	-0.0002	-0.0002	-0.0002	0.0011	0.0010
	(0.0002)	(0.0002)	(0.0003)	(0.0003)	(0.0005)	(0.0005)	(0.0007)	(0.0007)
Audited x Subsequent decision	-0.0068	-0.0068	0.0177	0.0176	0.0090	0.0088	0.0343	0.0328
	(0.0366)	(0.0366)	(0.0503)	(0.0503)	(0.0697)	(0.0696)	(0.0981)	(0.0978)
Audited x Audit effectiveness	0.0001	0.0001	0.0003	0.0003	0.0007	0.0007	-0.0009	-0.0010
	(0.0004)	(0.0004)	(0.0005)	(0.0005)	(0.0008)	(0.0008)	(0.0011)	(0.0011)
Subsequent decision x Audit effectiveness	0.0002	0.0002	0.0004	0.0004	0.0000	0.0000	-0.0009	-0.0009
	(0.0003)	(0.0003)	(0.0005)	(0.0005)	(0.0007)	(0.0007)	(0.0009)	(0.0009)
Audited x Subsequent decision x Audit effectiveness	0.0001	0.0001	0.0002	0.0002	0.0000	0.0000	-0.0004	-0.0006
	(0.0006)	(0.0006)	(0.0008)	(0.0008)	(0.0010)	(0.0010)	(0.0015)	(0.0015)
Received income	-0.0190 ***	-0.0191 ***	-0.0199 ***	-0.0201 ***	-0.0192 ***	-0.0194 ***	-0.0196 **	-0.0185 **
	(0.0032)	(0.0032)	(0.0044)	(0.0044)	(0.0061)	(0.0061)	(0.0091)	(0.0090)
Detection risk	0.0089 ***	0.0089 ***	0.0098 ***	0.0098 ***	0.0111 ***	0.0111 ***	0.0119 ***	0.0119 ***
	(0.0004)	(0.0004)	(0.0006)	(0.0006)	(0.0008)	(0.0008)	(0.0011)	(0.0011)
Paid fine	0.0017	0.0018	-0.0034	-0.0032	-0.0042	-0.0038	-0.0068 *	-0.0051
	(0.0030)	(0.0030)	(0.0032)	(0.0032)	(0.0034)	(0.0034)	(0.0036)	(0.0036)
Demographic variables?	No	Yes	No	Yes	No	Yes	No	Yes
N	333	333	333	333	312	312	240	240

Observations		9,498	9,498	4,992	4,992	2,590	2,590	1,324	1,324
Marginal R^2 / R^2	Conditional	0.032 / 0.478	0.157 / 0.484	0.037 / 0.477	0.163 / 0.484	0.049 / 0.482	0.175 / 0.489	0.062 / 0.436	0.185 / 0.453

Notes: *, **, and *** indicate significance at the 10 percent, 5 percent, and 1 percent level. Received income, paid fine, and continuous demographic variables are scaled. Marginal R^2 provides the variance explained by fixed effects, and conditional R^2 provides the variance explained by fixed effects and individual-specific random effects.

Table B5: Regressions with Control Variables[19]

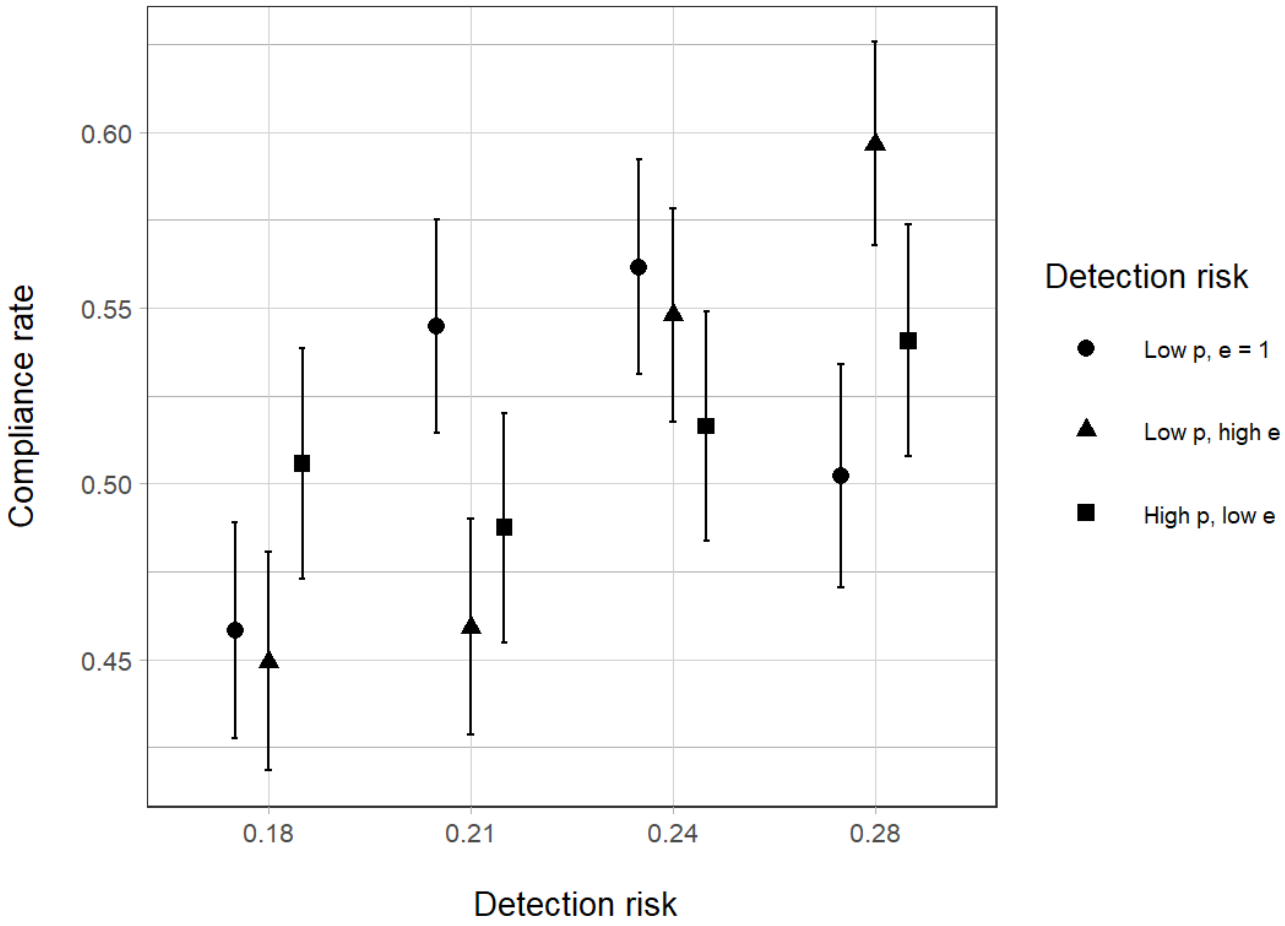
Dependent variable: Compliance rate					
	(31)	(32)	(33)	(34)	(35)
Intercept	0.3505 ***	0.3309 ***	0.3381 ***	0.3334 ***	0.3340 ***
	(0.0689)	(0.0667)	(0.0669)	(0.0667)	(0.0666)
Parameter order	0.0008				
	(0.0068)				
Low p, High e	-0.0049				
	(0.0083)				
High p, Low e	-0.0060				
	(0.0083)				
High p, High e	0.0093				
	(0.0189)				
Paid fine		-0.0062	-0.0023	-0.0045	-0.0070 *
		(0.0041)	(0.0034)	(0.0036)	(0.0037)
Audit previous round		0.0109			
		(0.0079)			
Low effective audit previous round			-0.0102		
			(0.0082)		
Medium effective audit previous round				0.0117	
				(0.0086)	
Effective audit previous round					0.0376 **
					(0.0151)
Received income	-0.0199 ***	-0.0199 ***	-0.0200 ***	-0.0200 ***	-0.0200 ***
	(0.0032)	(0.0032)	(0.0032)	(0.0032)	(0.0032)
Detection risk	0.0086 ***	0.0092 ***	0.0091 ***	0.0092 ***	0.0091 ***
	(0.0008)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Female	0.0896 ***	0.0890 ***	0.0893 ***	0.0892 ***	0.0889 ***
	(0.0283)	(0.0281)	(0.0282)	(0.0281)	(0.0280)
Age	0.0447 ***	0.0444 ***	0.0446 ***	0.0445 ***	0.0443 ***
	(0.0157)	(0.0156)	(0.0157)	(0.0156)	(0.0156)
Higher education	-0.0380	-0.0378	-0.0379	-0.0379	-0.0375
	(0.0311)	(0.0309)	(0.0310)	(0.0310)	(0.0309)
Economics major	0.0711	0.0706	0.0709	0.0707	0.0705
	(0.0483)	(0.0480)	(0.0482)	(0.0481)	(0.0480)

[19] The effects of demographic variables are very similar in all other analysis and thus not presented here.

German speaking	-0.1036 ***	-0.1029 ***	-0.1032 ***	-0.1030 ***	-0.1027 ***
	(0.0273)	(0.0271)	(0.0272)	(0.0272)	(0.0271)
Prior experiments	-0.0241	-0.0236	-0.0241	-0.0238	-0.0237
	(0.0593)	(0.0588)	(0.0591)	(0.0590)	(0.0588)
Prior tax experiments	0.0352	0.0347	0.0350	0.0347	0.0348
	(0.0370)	(0.0367)	(0.0369)	(0.0368)	(0.0367)
Self-preparation	-0.0035	-0.0034	-0.0035	-0.0035	-0.0035
	(0.0304)	(0.0302)	(0.0303)	(0.0302)	(0.0302)
Risk seeking	-0.0040	-0.0038	-0.0040	-0.0039	-0.0040
	(0.0140)	(0.0139)	(0.0140)	(0.0139)	(0.0139)
Income maximization	-0.1026 ***	-0.1020 ***	-0.1024 ***	-0.1022 ***	-0.1018 ***
	(0.0139)	(0.0138)	(0.0139)	(0.0138)	(0.0138)
Tax morale	0.0318 **	0.0317 **	0.0318 **	0.0317 **	0.0316 **
	(0.0135)	(0.0134)	(0.0134)	(0.0134)	(0.0134)
N	333	333	333	312	240
Observations	9,324	9,324	9,324	9,324	9,324
Marginal R^2 / Conditional R^2	0.160 / 0.471	0.161 / 0.468	0.160 / 0.470	0.161 / 0.469	0.161 / 0.469

Notes: *, **, and *** indicate significance at the 10 percent, 5 percent, and 1 percent level. Received income, Paid fine, and continuous demographic variables are scaled. Marginal R^2 provides the variance explained by fixed effects, and conditional R^2 provides the variance explained by fixed effects and random subject effects.

Figure B1: Effects of the Composition of the Detection Risk on Compliance



Notes: *Detection risk* results from different combinations of *Audit probability* (*p*) and *Audit effectiveness* (*e*) as described in Table 1 (Tasks 1 to 24). Error bars depict 95 percent confidence intervals.