Tax Compliance and Whistleblowing – The Role of Incentives

Ludger Breuer

Introduction

With sovereign debt levels reaching record heights, governments face a growing pressure to effectively combat tax evasion. In recent years, tax authorities have turned their hands to a novel method for detecting tax evaders: the purchase of CDs with incriminating bank information about large-scale tax evasion. Most notably, German tax authorities have been avid for such CDs that were offered for sale by employees of Swiss, Liechtenstein, and Luxembourgian banks\(^1\). At present, legality, morality, and effectiveness of the actions taken by the German

\(^*\) Ludger Breuer received his degree in Economics (Diplom, M. Sc. equivalent) from the University of Bonn in 2012. The present article refers to his diploma thesis submitted in August 2012

\(^1\) In January 2006, a former employee of Liechtenstein’s LGT Bank was the first to offer data on several hundred affluent tax evaders. The German Government paid roughly €4.2 million for the CD and used the data to start a tax evasion probe in February 2008. The probe received much media attention, mainly because the well-known CEO of German postal giant Deutsche Post AG was detained during the first hours of the police raid. According to an interim result published by the public prosecutor’s office in March 2010, 244 out of 596 preliminary proceedings have been completed, generating proceeds of €181 million. Moreover, the case triggered also voluntary declarations of tax liabilities amounting to €626 million, €222 million thereof related to LGT Bank (Leyendecker, Hans 2010. "Liechtenstein-CD bringt 626 Millionen Euro." Süddeutsche Zeitung, April 7). In 2010, German authorities bought three other CDs with i.a. data on customers of Swiss banks Credit Suisse and Julius Baer. Another CD with data on customers of Luxembourgian banks was purchased in 2011. In July 2012, North Rhine-Westphalia confirmed that it had bought, yet again, CDs containing information about tax evasion by customers of Swiss banks. It is estimated that the acquisition of the above-mentioned CDs will generate proceeds of several billion Euro in total.
state are subject of a vigorous debate. The highly-charged atmosphere of the
debate can be attributed to the fact that the purchase of "tax CDs" conflates two
tricky subjects: tax evasion and whistleblowing. In this paper, we do not immerse
into the broader issues of law and justice but focus on a preliminary question:
Are monetary incentives for whistleblowing effective in fostering tax compliance?
In particular, we ask the following three questions: How do monetary incentives
for whistleblowing

   I. influence the decision to report tax evasion?

   II. influence taxpayers’ decision to declare taxable income?

   III. influence the revenues of the state?

We find that monetary rewards for whistleblowing lead to a significant increase
in the reporting of tax evasion: the larger the reward the more pronounced the
increase in whistleblowing. This result contradicts previous research reporting
that external whistleblowing rewards may cause crowding-out effects. However,
since some subjects do not follow the incentives, the results are also contrary to
standard economic theory. Interestingly, this supposedly surprising whistleblow-
ing behavior is, by and large, correctly anticipated by taxpayers. Thus, even
without monetary incentives, whistleblowing proves to be fairly effective in curbing
tax evasion. But incentivizing whistleblowing makes whistleblowing an even
more powerful tool against tax evasion.

The remainder of the paper is organized as follows: Providing the background for
our study, the following section discusses the complex phenomenon of whistle-
blowing. Then we will give a review of the related literature. After we have
presented the design of the experiment, we will present our hypotheses. The
results of the experiment will be reported in several sections and the last section
concludes.
Whistleblowing – Loyalty and Betrayal

Though there is no generally accepted definition of whistleblowing, it is usually described as the deliberate disclosure of information about illegal, immoral, or illegitimate practices by current or former organization members to persons or organizations that may be able to effect action (Near and Miceli 1985). The term ”whistleblower” became popular in the early 1970s since it allowed to avoid existing expressions (”informer”, ”snitch”, ”denouncer”) that all have more or less negative connotations. The struggle about words and definitions points to the profound ambiguity in the way whistleblowing is ethically perceived by the public. Essentially, this ambiguity can be attributed to the fact that social enforcement raises conflicts of loyalties which are often hard to resolve. On the one hand, it is sometimes suggested that all citizens have an ethical obligation to aid in the arrest and prosecution of criminal actors (Asbury 2010). This notion corresponds to the ideas of loyalty to the state, loyalty to principles, or loyalty to actual or potential victims of the observed wrongdoing. However, the whistleblower also belongs to the group whose members are – at least partially – involved in the observed wrongdoing. Since the afore-mentioned loyalties collide with group loyalty, particularly the group loyalty to coworkers or employers, potential whistleblowers face a complex ethical dilemma (Bok 1983; Hersh 2002). In view of the intricate conflict of loyalties, it does not come as a surprise that whistleblowing is sometimes condemned as a treacherous breach of trust, while others glorify it as a heroic act of civic duty. The social acceptance of whistleblowing apparently depends on several factors. First, not every legal norm is backed by a corresponding social norm, and the intensity of underlying social norms may vary, too. Broadly speaking, the weaker the underlying social norm, the stronger the social rejection of reporting behavior. Second, whistleblowing always involves cutting some ties of loyalty, but the strength of these ties varies: The stronger
the social expectation that some form of trust should not be breached, the more pronounced the social rejection of whistleblowing (Koch, 2007). Third, the acceptance of whistleblowing hinges on the perceived motives of the whistleblower: The more selfish the motives, the lower the acceptance of whistleblowing. Considering the purchase of "tax CDs", social acceptance especially depends on the following questions:

I. How strong is the social norm that backs the legal prohibition of tax evasion?

II. To what degree are employees of banks expected to be loyal to their employers and to the customers of their banks?

III. Do financial rewards for whistleblowers reduce the social acceptance of whistleblowing?

Among the many strategies used to foster whistleblowing behavior the offer of monetary rewards is particularly controversial. In most countries, it is a well-established practice for public authorities to offer monetary rewards for information that will help to solve crimes or enable the police to locate and apprehend criminals. In contrast, there are considerable differences between legal orders when it comes to rewards that transfer a part of the revenue collected by the state to the whistleblower.

---


3 Under current German law, no provisions allow for such a profit-sharing between public authorities and whistleblowers. But historically, some examples for such provisions can be found in German law. Moreover, other legal orders often resort to this instrument. Under US American law, employees of fraudulent government contractors can file qui tam suits on behalf of the government and receive a compensation of up to 30% of the recovery if they win the suit. Likewise, according to Section 7623(b) of the US American Internal Revenue Code, whistleblowers receive as an award at least 15% but not more than 30% of the collected proceeds. Another recent example is Dodd-Frank Act Section 922: under this provision, whistleblowers who provide the Securities and Exchange Commission (SEC) with original information about violations of the securities laws are entitled to rewards in the range of 10% to 30% of the monetary sanctions.
The mixed legal views on monetary rewards for whistleblowing correspond to an ongoing lively debate among behavioral economists about the relationship between intrinsic and extrinsic aspects of motivation. While extrinsic motivation "refers to the performance of an activity in order to attain some separable outcome" (Ryan and Deci 2000), such as monetary rewards, intrinsic motivation means that people behave in a certain way because of the inherent satisfaction of the activity itself (Ryan and Deci 2000). Some studies suggest that, for a range of activities, the introduction of monetary rewards can undermine intrinsic motivation (Frey and Jegen 2011). This finding runs counter to classic economic predictions. However, monetary rewards do not always cause such "crowding-out effects" (for a meta-study see: Deci, Koestner, and Ryan 1999) and sometimes outside intervention in form of rewards can even increase intrinsic motivation (Osterloh, Frost, and Frey 2002).

Hence, different attitudes towards whistleblower rewards in various legal orders may not only reflect different moral concepts, but also point to uncertainty about the effects of such rewards. The present study aims to provide valuable insight into the impact of whistleblower rewards on reporting of illegal activities, such as tax evasion. Thus, it contributes to the existing experimental literature on the interplay between intrinsic and extrinsic motivation.\footnote{Rewards for whistleblowers can be viewed through the lens of expressive law, too. Once monetary incentives for whistleblowing are implemented through legislation, the legal provisions may send out a signal about moral values. In recent years, legal scholars put forward the idea that individuals do not only react to rewards and sanctions laid down in the law, but also respond to expressive signals embodied within our legal system. Monetary rewards for cooperation may not just offer material gains but also signal that whistleblowing is a socially desired behavior. If this signal is able to shape people’s attitude towards whistleblowing in that direction, monetary rewards and expressive signals may mutually reinforce one another. However, the signal can also backfire. For example, people may get the impression that blowing the whistle is generally considered an abhorrent breach of loyalty, since substantial monetary rewards are needed to induce at least some level of reporting. Thus, in the case of whistleblowing rewards, it is entirely unclear how the expressive function of law will work. The present study aims to shed some light on this topic, too.}
Related Literature

The related experimental research is basically divided in two pillars: tax evasion experiments and whistleblowing experiments. To our knowledge, so far there are no laboratory experiments that directly combine both types of experiments. Beginning with the first study of Friedland, Maital, and Rutenberg (1978), most tax evasion experiments rely on the standard model of tax evasion first proposed by Allingham and Sandmo (1972) and Yitzhaki (1974), in which the taxpayer acts as an isolated expected utility maximizer facing a gamble with detection probability and punishment. The most important finding of the early tax evasion experiments is that the level of income reporting is much higher than the expected utility model would predict (Torgler, 2002). In order to unravel the puzzle of tax compliance, some experiments investigate the effects of various additional aspects of tax administration. On the other hand, a growing number of experiments examine how various societal institutions and processes influence tax compliance. Basically, our experiment belongs to this second branch of recent tax evasion experiments.

In particular, audit rules can be designed so that a subject’s chance of audit depends upon the behavior of other subjects. Alm and McKee (2004) conducted an experiment to test an endogenous audit rule based upon how far a taxpayer’s tax report drops below the average of tax reports in the relevant cohort. Since audit selection is based upon relative reporting behavior, there is a coordination game among taxpayers with multiple equilibria. Those equilibria that involve low reporting are clearly preferable for the group, but they are also vulnerable to defection. The results presented by Alm and McKee show that subjects are unable to coordinate on the low compliance equilibrium if they are not permitted to discuss their strategies among themselves prior to reporting their income.
While tax evasion experiments flourished during the last decades, only recently economists have begun to include whistleblowing options in at least two well established types of laboratory experiments: antitrust experiments and experimental research on corruption.

A fast-growing body of antitrust experiments examines the effects of leniency programs. Leniency in antitrust enforcement can be defined as the reduction of penalties for firms or individuals that first confess to involvement in a cartel. The experimental research on leniency in antitrust enforcement has provided ambiguous results. If a leniency program is in place, subjects make more often use of the whistleblowing option and, thus, more cartels are detected (Bigoni, Fridolfsson, Le Coq, and Spagnolo, 2011). But often, cartel formation, cartel stability and market prices are not reduced, and the prices charged by cartel members may even increase (Apesteguia, Dufwenberg, and Selten, 2007; Bigoni, Fridolfsson, Le Coq, and Spagnolo, 2011; Dijkstra, Haan, and Schoonbeek, 2011). The most prominent explanation argues that the threat of whistleblowing deters cartel members from underbidding the agreed upon prices. This has motivated the investigation of whistleblower rewards. The theoretical prediction (Spagnolo, 2004; Aubert, Kovacic, and Rey, 2000) that these rewards would strengthen the effectiveness of leniency rules has received mixed support in recent experimental studies. Whistleblowing rewards did not perform well in a one-shot game (Apesteguia, Dufwenberg, and Selten, 2007), but, in a repeated game, they turned out to be the only welfare enhancing leniency scheme (Bigoni, Fridolfsson, Le Coq, and

To some extent, this experiment neighbors on our study, since both whistleblowing-triggered audits and the endogenous audit rule described above can lead to coordination problems. But there are major differences. In particular, subjects facing an endogenous audit rule have to guess the tax compliance behavior of other subjects, whereas subjects facing a whistleblowing-triggered audit have to guess the whistleblowing behavior of other subjects. Since the social norms of tax compliance may substantially differ from the social norms of whistleblowing, the results of the experiments are not directly comparable.
Corruption experiments have recently been combined with whistleblowing mechanisms, too. [Schikora (2011)] and [Engel, Goerg, and Yu (2012)] investigate whether whistleblowing mechanisms with symmetric or asymmetric punishment are more effective in curbing corruption. Both studies find that whistleblowing mechanisms may serve as a threat that forces officials to reciprocate, thus stabilizing the corrupt relationship[6]

However, there is a substantial difference between the whistleblowing behavior of cartel members, bribers or corrupt officials on the one hand and whistleblowing in the context of taxation on the other hand. The first type of whistleblowing serves to break up conspiracies from the inside and may be called "traitorous whistleblowing". The second type of whistleblowing covers the reporting of wrongdoing by a third person that observed the wrongdoing but did not take part in it. It seems fitting to label this as "watchdog whistleblowing"[7] Since the role of incentives for watchdog whistleblowing has not yet been investigated in laboratory experiments, our study adds a novel facet to the existing experimental research on whistleblowing.

This facet has recently received some attention in the theoretical literature and in vignette studies. [Yaniv (2001)] was the first to model the individual’s decision to blow the whistle on tax evaders. But since Yaniv takes tax evasion as given, his model is unable to provide insight into the impact of whistleblowing.

---

6 In particular, punishing the briber less harshly than the bribee leads to more corruption [Engel, Goerg, and Yu (2012)]. On the other hand, asymmetric leniency programs for officials who blow the whistle significantly reduce the level of corruption, since the official is protected from retaliation [Schikora (2011)].

7 Traitorous whistleblowing is tantamount to self-indictment causing punishment for the whistleblower himself. However, the watchdog whistleblower can abide by the law and still be able to observe and report the wrongdoing of others. Thus, the problem that, without special mechanisms like leniency, whistleblowers are automatically punished for their own wrongdoing does not arise. Nevertheless, whistleblowers might face some risk of retaliation so that whistleblowing remains a costly activity. It should also be noted that due to these differences between the two types of whistleblowing the relevant social norms might vary, too.
on tax compliance. To our knowledge, so far, only Mealem, Tobol, and Yaniv (2010) model the taxpayer’s decision to evade taxes under the threat of whistleblowing and the tax authority’s optimal design of audit policy in the presence of whistleblowers. The authors’ main findings can be summarized as follows: If audit costs are too high to audit so many taxpayers that everybody prefers to report his income truthfully, the tax authority might be better off running a second, whistleblowing-triggered audit round although all reports about tax evasion will stem from dishonest whistleblowers who falsely denounce taxpayers to harass them.

A large part of the research on the characteristics of whistleblowers and the contextual variables of whistleblowing uses scenario-based approaches to ascertain when an observer of wrongdoing will be likely to blow the whistle. Most notably, Feldman and Lobel (2010) use a large vignette study to compare the effect of different regulatory mechanisms on individual whistleblowing motivation and behavior. The study produced several intriguing results. First, participants predict that they themselves would be more likely to report than others (“holier than thou effect”). Second, participants who have a low perception of misconduct severity and, thus, a low internal motivation to report, are less willing to blow the whistle if they are offered low monetary rewards compared to legal mechanisms that do not offer any whistleblowing rewards. Third, compared to the low reward scenario, participants with low internal motivation are much more likely to report if they are offered high monetary rewards. Forth, the reporting behavior of participants with high internal motivation is very similar across the various

---

8 In the second round the tax authority threatens to audit a sufficiently high fraction of denounced taxpayers so that all taxpayers who fear a whistleblower refrain from evading taxes. Hence, no honest whistleblower will report a taxpayer. The results derived from the model are somewhat counterintuitive. Since the motivations that underpin the behavior of whistleblowers are very complex, it is questionable to focus exclusively on the desire to take revenge. In terms of external validity, one may also question the complete lack of penalties for false denunciations. Therefore, our experiment is not primarily designed to test the model developed by Mealem, Tobol, and Yaniv (2010).
legal mechanisms used to incentivize whistleblowing. The authors conclude that these findings "indicate that often offering monetary rewards to whistleblowers will lead to less, rather than more, reporting of illegality" and hint at "a type of crowding effect in which the introduction of an external reward interferes with the moral dimension of reporting". However, fifth, participants think that the size of the whistleblowing reward is far more influential when it comes to others in comparison to themselves. Considering the general limitations of vignette studies, the last finding casts some doubts on the alleged crowding out effect. Therefore, we feel it desirable to experimentally test the results presented by Feldman and Lobel (2010).

The Experiment

The experiment was programmed and implemented using z-Tree (Fischbacher, 2007) version 3.3.11 and is subdivided into four treatments. We shall refer to these treatments as "BASE", "NO INC", "SM INC" and "LA INC". The experiment consisted of two sessions for each of the four treatments. All eight sessions were conducted at the Laboratory for Experimental Economics, at the University of Bonn, Germany. We had six sessions with 24 subjects and two sessions with 21 subjects. Hence, a total of 186 subjects, mostly students from various undergraduate courses at the University of Bonn, including law, economics, political science, etc., took part in the experiment. When participants arrived at the laboratory they were seated in a lecture room where the experimenter thanked them for coming and informed them that they take part in an experiment that allows them to earn money. The participants were informed that the money earned in the experiment would be paid anonymously at the end of the experiment, so that no participant could learn about the money earned by other participants. They

9 The four legal mechanisms were: anti-retaliation protections, affirmative reporting duties, fines for the failure to report, and monetary rewards.
were also told not to communicate with fellow participants. Finally, the participants were randomly assigned to visually isolated cubicles equipped with computer terminals. The cubicles were numbered from 1 to 24. Each session lasted less than two hours including the time for payments. Earnings were recorded in an experimental currency named "Taler". Talers were convertible to Euros at the rate of 7000 Talers per Euro. Average earnings in the experiment were 10.75 Euro.

Each session comprises two parts. The first part is once more divided into two stages.

At the beginning of the experiment, the participants receive only the instructions for the first stage of the first part, a real effort task based on sliders [Gill and Prowse 2011a,b]. This stage consists of three rounds. The first round is a practice round and the two remaining rounds are paying rounds. In every round, the slider task consists of a single screen displaying 48 sliders. The screen does not vary across subjects or across rounds. Initially, all 48 sliders are positioned at 0. The subjects can use the mouse to position each slider at any integer location between 0 and 100 inclusive. The current position of each slider is displayed to the right of the slider. The subjects’ task is to position as many sliders as possible exactly at 50 within the allotted time of 120 seconds. For every slider that is correctly positioned at 50 at the end of the allotted time, the subjects receive 1000 points. As the task proceeds, the subject’s current points score is displayed above the 48 sliders and the remaining time is displayed in the right upper corner of the screen. At the end of the first stage, for each subject, the points scored in the first and the second paying round are added up to a total score. The subjects receive a payoff of 1 Taler for every point of their total score. Additionally, each subject gets a fixed payoff of 4000 Taler.

After the first stage is completed, the instructions for the second stage of the first
part are handed out to the participants. At the beginning of the second stage, the subjects receive the payoff that they have earned in the previous stage and are asked to “pay taxes” on their income. For all subjects, the tax rate is set at 30% and the taxes paid by the subjects are donated to a charity with the DZI donation seal. This is common knowledge, but the name of the charity is only revealed to the participants at the end of the experiment. The subject must decide how much income to report, and must pay taxes on all reported income. The subject pays no taxes on underreported income. To simplify decision-making, the subject can choose the percentage of income that he wants to report with the help of a slider on the computer screen. This slider can be positioned at any integer location between 0 and 100 inclusive. For each possible position of the slider, the corresponding percentage of income reported, the value of income reported (in Taler), the resulting tax payment (in Taler), and the remaining income after taxes (in Taler) are displayed to the subject. Before the subject decides how much income to report, he is informed about the following steps.

The subject is told that his “tax declaration” will be randomly audited with a probability of \( \frac{1}{24} \). To make the randomness of the audit more salient, an old fashioned hand turning bingo cage containing 24 balls numbered from 1 to 24 is put up in the middle of the room and the subjects are allowed to watch the drawing. Once the balls are mixed, one ball is drawn from the bingo cage, and the number on the ball is displayed to all subjects. Afterwards, the subject that is seated in the cubicle labeled with the number on the drawn ball is audited. Based on a comparison between the actual and the reported income of the subject, the audit entails the payment of any unpaid taxes and the additional payment of a "fine" if the reported income is less than the actual income. The fine is fixed

---

10 In all eight sessions, the taxes paid by the subjects were donated to UNICEF.
11 If less than 24 subjects take part in one session, and the ball drawn from the bingo cage is labeled with one of those numbers that belong to the unoccupied cubicles, no subject is selected for audit. This was made clear to all subjects taking part in sessions with less than 24 subjects.
at 200% of unpaid taxes. The paid tax arrears are also donated to the charity whereas the fine is not donated. To facilitate the decision-making, a small table depicting the subject’s payoff depending on the percentage of income reported and the occurrence of an audit is presented to the subject. In ”BASE”, stage 2 of part 1 ends after the random audit. In the other treatments, stage 2 of part 1 continues with the third step.

Once the random audit has been conducted, every subject observes the tax declaration of one other subject. Accordingly, the tax declaration of every subject is observed by exactly one other subject. It is common knowledge that two subjects cannot mutually observe each other. In order to have as many independent observations as possible, in every session, there were eight random matching groups consisting of three subjects that observe each other in circles: A observes B, B observes C, and C observes A.  

Since it is well possible that many subjects will declare the same percentage of their income, for many other tax declaration values there may not be enough observations. To circumvent this problem, we adopt the strategy method (Selten 1967). The subject is presented with six different tax declarations: five fictive tax declarations and the actual tax declaration of the observed subject. The subject is told that five of the six tax declarations are fictive but he cannot figure out which of the six tax declaration is the actual tax declaration. For all six tax declarations, the subject has to decide whether or not to report the tax declaration. In the further course of the experiment, only the decision to or not to report the actual tax declaration leads to consequences for the subjects.

If the actual tax declaration is reported, the reported tax declaration is audited. If the declared income is less than the actual income (i.e. the percentage of income reported ranges between 0% and 99%), the audited subject must pay any

\[12\] The subjects were not explicitly informed about these random matching groups. The participants did not ask any questions with regard to the exact observation pattern.
unpaid taxes and an additional fine fixed at 200% of unpaid taxes. Again, the fine is not donated but the paid tax arrears are donated to the charity. If the declared income equals the actual income (i.e. the percentage of income reported amounts to 100%), the audit does not entail consequences for the audited subject. If a subject that did not fully declare his income has already been randomly audited, the actual tax declaration observed by the other subject is set at 100% and, thus, a possible second audit would not entail further consequences for the audited subject.

Reporting the actual tax declaration costs the whistleblower 2000 Taler. In "NO INC", the whistleblower is not entitled to any financial rewards. In contrast, "SM INC" provides an exiguous financial incentive for reporting: The whistleblower gets a bonus of 2200 Taler if the audit leads to the detection of some level of tax evasion, i.e. if the actual tax declaration ranges between 0% and 99%. In "LA INC", the bonus for whistleblowing is raised to 12000 Taler, thus strengthening the incentives for reporting. The conditions for the award of the bonus do not vary between the treatments. To prevent the dissemination of information about the actual tax compliance of other subjects, the subjects that report the tax declaration in at least one of the six decision situations learn only at the end of the experiment (i.e. after part 2) whether they are eligible to get the whistleblowing bonus. Stage 2 of part 1 ends after the whistleblowing procedure and the experiment continues with part 2. It is common knowledge that the decisions taken in part 2 cannot influence the subject’s payoff from the first part of the experiment.

After the first part of the experiment is completed, the instructions for the second part are handed out to the participants. The second part is based on a "surprise restart", i.e. all stages and steps of the first part are repeated without any substantial changes to the rules of the game. We only omitted the practice round of

---

13 Since only the decision to or not to report the actual tax declaration leads to consequences for the subjects, it is not costly to report one of the five fictive tax declarations.
the slider task since the subjects were already able to sufficiently practice the task during the first part of the experiment. The matching groups remain unchanged and the subjects are explicitly told that the observation mechanism stays the same.

Hypotheses

Standard economic theory is known for poorly predicting the behavior of subjects taking part in tax evasion experiments. Hence in order to formulate our hypotheses, we mainly use the results provided by a vignette study of Feldman and Lobel (2010) to predict subjects’ tax compliance and whistleblowing behavior. According to this study, some subjects predict that they would likely report wrongdoing even if there are no whistleblowing rewards and no fines for failure to report. This behavior can be attributed to intrinsic motivation. Some antitrust experiments using one-shot games also find whistleblowing in treatments without incentives for reporting. Given this evidence, we formulate our first hypothesis about subject’s whistleblowing behavior:

**Hypothesis 1.** In "NO INC", some subjects will blow the whistle on tax evaders.

With regard to subjects with low internal motivation, Feldman and Lobel (2010) find that low rewards for whistleblowing reduce the willingness to report compared to situations where no incentive is present. This reaction to the introduction of external rewards may be interpreted as a crowding-out effect. Along these lines, we formulate our second hypothesis:

**Hypothesis 2.** In "SM INC", the probability that subjects blow the whistle on tax evaders will be smaller than in "NO INC".

However, the results presented by Feldman and Lobel (2010) suggest that high whistleblowing rewards are able to overcome the crowding-out effect. Thus, we
expect that an increase in whistleblowing rewards leads to more reporting:

**Hypothesis 3.** In "LA INC", the probability that subjects blow the whistle on tax evaders will be higher than in "SM INC".

The following predictions concern the tax evasion behavior. Since the tax evasion behavior in "BASE" does not depend on any whistleblowing mechanism, we derive our forth hypothesis from the numerous tax evasion experiments that have been used to test the standard model of tax evasion. In line with these experiments, we expect to observe a substantial amount of tax evasion, but we also assume that the level of income reporting will be higher than the expected utility model would predict for risk-neutral or moderately risk-averse subjects:

**Hypothesis 4.** In "BASE", some subjects will declare more than 0% of their income, but only few subjects will fully declare their income.

With regard to the three other treatments, we assume that tax evasion behavior depends on what subjects expect other subjects to do under the various incentive schemes for whistleblowing. Data from vignette studies demonstrates that the perceived effect of whistleblowing mechanisms differs from the effect that subjects report for their own actions. In general, individuals predict that others are less likely to blow the whistle than they themselves are (Feldman and Lobel, 2010). Despite "this holier than thou" effect, subjects predict that others will to some degree report wrongdoing even if there are no whistleblowing rewards and no fines for failure to report. Considering these results, we expect subjects to believe that some other subjects with high intrinsic motivation will make use of costly whistleblowing options even if there are no monetary rewards. Put another way, we assume that, in "NO INC", subjects expect a higher audit probability than in "BASE". Thus, we formulate our fifth hypothesis as follows:

**Hypothesis 5.** In "NO INC", subjects will declare a larger share of their income
compared to "BASE".

Feldman and Lobel (2010) also report that subjects believe that the average person’s whistleblowing behavior is much more externally motivated than their own actions. Therefore, we expect that subjects do not to anticipate the crowding-out effect described in hypothesis 2 but rather believe that the introduction of small rewards does not have a significant impact on whistleblowing behavior. This expectation results in our sixth hypothesis:

**Hypothesis 6.** In "SM INC", subjects will declare a share of income that does not significantly differ from the share of income declared in "NO INC".

Moreover, following the above-mentioned results, we expect that subjects believe that the introduction of large rewards will substantially increase the willingness of others to blow the whistle on tax evaders. Hence we assume that, in "LA INC", subjects expect a higher audit probability than in all other treatments. As a result, we formulate our seventh hypothesis as follows:

**Hypothesis 7.** In "LA INC", subjects will declare a larger share of their income compared to "NO INC" and "SM INC".

**Experimental Results**

We present our results in three parts. At first we start with the investigation of whistleblowing behavior under the three different incentive schemes. After that we will turn to the differences in tax compliance between the four treatments. Finally, we take a closer look at the efficiency of whistleblowing incentives by evaluating the differences in tax payments and state revenues between the treatments.
Experimental Results – Whistleblowing

Figure 1 gives the frequencies of whistleblowing over all decisions for the two rounds in each of the three treatments that include a whistleblowing mechanism. The whistleblowing frequencies are calculated as follows: In the whistleblowing stage each subject observes six different tax declarations, one of them always being 100%. Since reporting an honest taxpayer is obviously not a sound decision\footnote{Reporting an honest taxpayer does not entail monetary consequences for the audited taxpayer, but for the reporting subject it is a costly activity that does not yield any monetary profits. Moreover, it seems implausible that subjects may be intrinsically motivated to report honest taxpayers. The inclusion of the "100%"-scenario mainly allows to control whether subjects understood the rules of the experiment.}, we excluded reactions to observed full tax compliance from the calculation\footnote{With regard to the remaining five observed tax declarations, for every subject the number of decisions to blow the whistle is divided by the number of observed tax declarations, i.e. 5.}

Whistleblowing in the Various Treatments

We start with a quick look at the whistleblowing behavior in "NO INC". In the first round, in 27.1% of all cases, subjects decided to blow the whistle on tax evaders. In the second round, the whistleblowing frequency slightly rises to 31.6%, but this increase is not significant (p = 0.269, two-sided Wilcoxon signed-rank test).

The introduction of small rewards changes the picture significantly. In the first round of "SM INC", subjects blew the whistle on tax evaders in 48.4% of all cases. This whistleblowing frequency is significantly higher than the whistleblowing frequency of 27.1% observed in the first round of "NO INC" (p = 0.013, two sided Mann-Whitney u-test). For the second round of "SM INC", whistleblowing frequency even amounts to 62.2%, again resulting in a significant increase compared to the whistleblowing frequency of 31.6% in the second round of "NO INC" (p = 0.001, two sided Mann-Whitney u-test). It is also worth noting that the whistleblowing frequency in the second round of "SM INC" is significantly higher than
the whistleblowing frequency in the first round of "SM INC" ($p = 0.017$, two-sided Wilcoxon signed-rank test).

Increasing the size of rewards has a significant impact on whistleblowing behavior, too. In "LA INC", subjects were even more inclined to blow the whistle on tax evaders. In the first round, tax evasion was reported in 79.2% of all cases, and, in the second round, subjects even decided to blow the whistle in 85.8% of all cases. The increase from the first to the second round is only weakly significant ($p = 0.054$, two-sided Wilcoxon signed-rank test), but in both rounds the whistle was blown much more often than in the other treatments: In the first round of "LA INC", the whistleblowing frequency is significantly higher than in the first round of "NO INC" ($p < 0.001$, two sided Mann-Whitney u-test) and in the first round of "SM INC" ($p < 0.001$, two sided Mann-Whitney u-test). Likewise, in the second round of "LA INC", tax evasion was reported significantly more often than in the second round of "NO INC" ($p < 0.001$, two sided Mann-Whitney u-test) and in the second round of "SM INC" ($p = 0.001$, two sided Mann-Whitney u-test).

Explaining Whistleblowing Behavior

In this subsection, we investigate the effects of monetary rewards for whistleblowing on the willingness to report tax evaders in greater detail, i.e. we account for individual characteristics when comparing treatments. The variable that we seek to explain in the analysis is the whistleblowing frequency for each of the two rounds. Whistleblowing frequency is a left- and right-censored dependent variable: Whistleblowing frequency cannot be smaller than 0 and it cannot exceed 1. Therefore, we run a Tobit regression that censors the observations at both minimum and maximum whistleblowing frequency.
Table 1 gives the results of the Tobit regression models used to explain whistleblowing frequency in the first round. Model 1 of Table 1 basically repeats the nonparametric tests. The difference between "NO INC" and the other treatments is striking: Both "SM INC" and "LA INC" have a significantly positive effect ("SM INC": $p < 0.025$; "LA INC": $p < 0.01$) on the whistleblowing frequency. The picture does not change if we use subjects’ effort in round 1 as an additional explanatory variable (Model 2 of Table 1). Because subjects earn 1000 Taler for every correctly positioned slider, this additional variable does not just show how much effort subjects put in\(^\text{[16]}\), but it also depicts the income level of subjects. Thus, no statistically significant dependence of whistleblowing frequency on either effort or income level was detected. Adding subjects’ attitude towards charities in general (with higher values indicating a more positive attitude) as an additional explanatory variable (Model 3 of Table 1) yields a similar result. The coefficient of attitude towards charities is positive but not statistically significant. Finally, Model 4 of Table 1 extends the explanatory variables by including age and gender (gender is set to 1 for females) of subjects. Again, the coefficients for both variables are not statistically significant. In contrast, the coefficients for the treatments remain virtually unchanged compared to Model 1, Model 2 and Model 3: In all four models they are positive and statistically significant ("SM INC": $p < 0.025$; "LA INC": $p < 0.01$).

The results of the Tobit regression models used to explain whistleblowing frequency in the second round are shown in Table 2. The overall picture is virtually unchanged compared to the results for the first round: In all four models, both "SM INC" and "LA INC" have a significantly positive effect ($p < 0.01$) on the whistleblowing frequency, while the coefficients of all further explanatory vari-

\(^{16}\) The variable "Effort (round1)" shows the number of correctly positioned sliders plus 4 (taking into account that subjects received a fixed payoff equal to the earnings for four correctly positioned sliders). Thus, strictly speaking, the variable depicts the income level rather than the effort level.
ables are not statistically significant.

**Summing up – Results on Whistleblowing**

Given the results of the Mann-Whitney u-tests and the regression analyses, we do not reject Hypothesis 1 and Hypothesis 3, but we reject Hypothesis 2:

**Result 1.** *Even if, as in "NO INC", whistleblowing is costly and no monetary incentives for whistleblowing are provided, some subjects still blow the whistle on tax evaders.*

**Result 2.** *If, as in "SM INC", small monetary incentives for whistleblowing are introduced, subjects decide significantly more often to blow the whistle on tax evaders.*

**Result 3.** *If, as in "LA INC", substantial monetary incentives for whistleblowing are provided, the whistleblowing frequency is significantly higher compared to treatments where only small rewards or even no rewards are provided.*

Summing up, monetary rewards lead to a significant increase in whistleblowing frequency, and the larger the reward the more pronounced the increase in whistleblowing and the resulting detection probability of tax evasion. These findings do not lend support to the hypothesized crowding-out effect of external rewards. The results also run counter to predictions based on standard economic theory: Though, in general, incentives for whistleblowing do not backfire, there are some subjects who report tax evasion even if whistleblowing is costly, while other subjects refuse to blow the whistle even if substantial rewards are provided. Thus, the whistleblowing behavior does not go from one extreme to the other but rather steadily shifts, thereby cushioning the effects of incentives.
Experimental Results – Tax Compliance

Figure 2 gives the arithmetic mean of tax declarations for the two rounds in each of the four treatments. Figure 3 provides a supplementary box plot diagram for tax declaration. The term "tax declaration" is here used to designate the percentage of income reported.

Tax Compliance in the Various Treatments

We observe a substantial amount of tax evasion in "BASE". In the first round subjects declare on average only 36.38% of their income, and in the second round tax declaration even declines to a meager 25.21% of income, significantly less than in the first round (p = 0.039, two-sided Wilcoxon signed-rank test). Only few subjects fully declare their income (16.67% in round 1, and 6.25% in round 2).

When a non-incentivized whistleblowing mechanism is introduced, a completely different picture emerges. In the first round of "NO INC", tax declaration amounts to 69.82% of income. This value significantly drops to 59.62% in the second round (p = 0.042, two-sided Wilcoxon signed-rank test). A comparison between treatments shows that, in both rounds, tax declaration in "NO INC" is significantly higher than in "BASE" (for round 1: p < 0.001, two sided Mann-Whitney u-test; for round 2: p < 0.001, two sided Mann-Whitney u-test).

Compared to "NO INC", introducing small whistleblowing rewards does not have a major impact on tax declaration. In "SM INC", tax declaration amounts to 68.49% in the first round, and 73.47% in the second round. The small increase from the first to the second round is not statistically significant (p = 0.271, two-

\[17\] The boxplot also includes subjects’ expectations about average tax declaration of all subjects in the respective treatment.
\[18\] Since a flat tax is applied, the percentage of income reported is naturally equal to the share of taxes due voluntarily paid by a subject.
sided Wilcoxon signed-rank test). There is also no significant difference in tax declaration between "NO INC" and "SM INC" (for round 1: p = 0.99, two sided Mann-Whitney u-test; for round 2: p = 0.109, two sided Mann-Whitney u-test). Consequentially, we find that tax declaration in "SM INC" is significantly higher than in "BASE" (for round 1: p < 0.001, two sided Mann-Whitney u-test; for round 2: p < 0.001, two sided Mann-Whitney u-test).

However, providing large whistleblowing rewards strongly affects tax declaration. In the first round of "LA INC", subjects declare on average 85.67% of their income, and in the second round tax declaration even increases to 90.21%. The increase in tax declaration from round 1 to round 2 is only weakly significant (p = 0.086, two-sided Wilcoxon signed-rank test), but in both rounds tax declaration is much higher than in the other treatments: In the first round of "LA INC", tax declaration is significantly higher than in the first round of "BASE" (p < 0.001, two sided Mann-Whitney u-test), "NO INC" (p = 0.006, two sided Mann-Whitney u-test), and "SM INC" (p = 0.02, two sided Mann-Whitney u-test). Likewise, in the second round of "LA INC", subjects declared a significantly higher share of income than in the second round of "BASE" (p < 0.001, two sided Mann-Whitney u-test), "NO INC" (p < 0.001, two sided Mann-Whitney u-test), and "SM INC" (p = 0.007, two sided Mann-Whitney u-test).

### Explaining Tax Compliance in Round 1

In order to explore the effects of whistleblowing rewards on tax compliance more thoroughly, the following analysis takes account of individual characteristics when comparing treatments. The variable that we seek to explain is the tax declaration for round 1. Since tax declaration is a left- and right-censored dependent variable\(^{19}\), we run a Tobit regression.

\(^{19}\) Tax declaration is expressed as a percentage of income declared and, thus, ranges from 0 to 100.
Table 3 shows the results of the Tobit regression models used to explain tax compliance behavior in the first round. In line with the results of the nonparametric tests, Model 1 of Table 3 reveals clear-cut treatment effects. Compared to "BASE", all treatments with whistleblowing mechanisms have a significantly positive effect (p < 0.01) on tax declaration. Model 2 of Table 5 adds subjects’ effort – and thereby subjects’ income – as explanatory variable. The coefficient of effort and income on tax declaration is negative but not statistically significant. Model 3 of Table 5 also uses subjects’ attitude towards charities in general as an explanatory variable. The attitude towards charities has a significantly positive effect (p < 0.025) on tax declaration, i.e. subjects who have a positive view of charities declare a higher percentage of their income. However, the treatment effects in Model 3 do not substantially differ from the treatment effects in Model 1 and 2. The same applies to Model 4 of Table 5. Here, age and gender are added as further explanatory variables. Neither of them has a significant effect on tax declaration.

Put simply, it seems that subjects react to a perceived higher risk of whistleblowing with a higher degree of tax compliance. But this interpretation might be premature: Figure 4 gives subjects’ predictions of the whistleblowing behavior of other subjects depending on the observed percentage of income declared (0%, 20%, 40%, 60%, 80%, and 100%). These predictions are remarkably similar for "SM INC" and "LA INC", casting some doubt on the notion that the higher level of tax compliance observed in "LA INC" can be explained with a higher expected whistleblowing frequency. Therefore, we run further linear and Tobit regression

\[ \text{20} \]

20 However, the data on the expected whistleblowing frequency should be treated with caution. Subjects’ predictions of other subjects’ whistleblowing behavior were only made after round 2 and may, thus, be shaped by subjects’ experiences with whistleblowing in the two previous rounds of the experiment. Since, in "LA INC", only few subjects attempted to evade taxes at all, most subjects had no actual information about other subjects’ willingness to report tax evaders. In "SM INC", tax compliance was lower and therefore more subjects got actual feedback on the whistleblowing behavior of other subjects. If we assume that tax evaders believe that other subjects are not very likely to blow the whistle, and if we also take into ac-
models that use the average whistleblowing expectation as an additional explanatory variable for tax compliance in round 1. Table 4 shows the results of the linear regression models. "SM INC" and "LA INC" are compared to "NO INC", since there are no whistleblowing expectations for "BASE". Model 1 of Table 4 shows that, compared to "NO INC", "LA INC" has a significantly positive effect (p < 0.05) on tax declaration, whereas Model 2 and 3 of Table 4 demonstrate that effort and attitude towards charities do not have a significant effect on tax declaration. The significant effect of "LA INC" remains largely unchanged (p < 0.05 and p < 0.025, respectively). However, the picture changes with regard to Model 4 of Table 4. Once the whistleblowing expectation is added as an explanatory variable for tax compliance, "LA INC" has no longer a significantly positive effect on tax declaration, whereas we now observe that the coefficient of whistleblowing expectation is positive and statistically significant. This is strong evidence that it is indeed the difference in the whistleblowing expectation that causes the different levels of tax compliance in the treatments. However, we must also admit that the corresponding Tobit regression models (Table 5) yield slightly different results, thus calling for future statistical in-depth analysis.

Explaining Tax Compliance – Adoptions in Round 2

Finally, in this subsection we take a closer look at tax compliance in the second round. In order to find out how subjects adapt their tax declarations to the experiences from the first round, we run another linear regression, seeking to explain tax declaration in round 2. We used five linear regression models, the results of which are shown in Table 6. Model 1 of Table 6 uses only one explanatory variable: "Controlled and Punished" is a dummy variable that takes the value count that whistleblowing frequency in "SM INC" is rather high. Many tax evaders experienced whistleblowing-triggered audits and may have adjusted upwards their beliefs about whistleblowing frequency. This would explain why the clear-cut treatment effect of higher whistleblowing rewards on tax declaration is not mirrored by a much higher expected whistleblowing frequency.
1 for every subject that was audited\footnote{Audits include random audits as well as whistleblowing-triggered audits.} and punished for tax evasion in the first round. The coefficient of this variable is positive and statistically significant ($p < 0.01$). However, it would be premature to conclude that tax declarations in round 2 are mainly driven by previous experiences of audits and punishments. As we have seen, whistleblowing frequency strongly depends on the treatments and, thus, the number of subjects that were audited and punished in round 1 depends on the treatments, too. Therefore, Model 2 of Table\ref{table:models} includes the three treatments with whistleblowing mechanisms as additional explanatory variables. Once again the treatments have a significantly positive effect ($p < 0.01$) on tax declaration. The coefficient of ”Controlled and Punished” remains positive but it is not longer statistically significant. Models 3, 4, and 5 of Table\ref{table:models} successively add effort, attitude towards charities, and, finally, age and gender as further explanatory variables. None of these variables has a statistically significant effect on tax declarations in round 2. It is noteworthy that the attitude towards charities had a significantly positive effect ($p < 0.025$) on tax declaration in round 1, but in the second round this effect is no longer detectable. This finding suggests that tax compliance in round 2 is so predominantly driven by the various whistleblowing mechanisms that the influence of other, rather intrinsic motivations is harder to identify.

**Summing up – Results on Tax Compliance**

Based on the results of the Mann-Whitney u-tests and the regression analyses, we do not reject Hypotheses 4, 5, 6, and 7:

**Result 4.** *Even in ”BASE”, subjects declare on average more than 0% of their income, but there is a high level of tax evasion and full tax compliance is rare.*
Result 5. If, as in "NO INC", taxpayers face a whistleblowing mechanism without monetary incentives for whistleblowing, subjects declare a significantly larger share of their income compared to "BASE".

Result 6. If, as in "SM INC", taxpayers face a whistleblowing mechanism with small monetary incentives for whistleblowing, subjects declare a share of their income that does not significantly differ from the share of income declared in "NO INC".

Result 7. If, as in "LA INC", taxpayers face a whistleblowing mechanism with large monetary incentives for whistleblowing, subjects declare a significantly larger share of their income compared to "NO INC" and "SM INC".

In sum, these results nicely fit the hypotheses that were based on the perceived effect of whistleblowing reported by Feldman and Lobel (2010). Broadly speaking, it seems that subjects react to a perceived higher risk of whistleblowing with a higher degree of tax compliance.

Efficiency – Tax Payments and State Revenues

Finally, we take a closer look at the efficiency of tax regimes under the various whistleblowing mechanisms. The efficiency of a tax regime can be defined in a number of different ways. Here we focus on three distinct yardsticks: voluntary tax compliance ($C_V$), total tax compliance ($C_T$), and state revenues ($R$). Denoting all taxes due as $T_D$, all taxes paid voluntarily as $T_V$, all taxes paid due to audits as $T_A$, all fines collected as $F$, and all rewards paid to whistleblowers as $B$, we define:

$$C_V = \frac{T_V}{T_D}, \quad C_T = \frac{T_V + T_A}{T_D} \quad \text{and} \quad R = \frac{T_V + T_A + F - B}{T_D}$$
Table 7 gives the values of $C_V$, $C_T$, and $R$ for both rounds in each of the four treatments. The values for $C_V$ are almost identical to the values for tax declaration reported in Section 6.2. The values for $C_T$ show that, in "BASE", the tax authority is only able to collect 38.4% of taxes due in the first round, while all treatments with whistleblowing mechanisms perform much better ("NO INC": $C_T = 77.7\%$, "SM INC": $C_T = 85.1\%$, "LA INC": $C_T = 96.9\%$). With regard to this yardstick for efficiency, we can also conclude: The larger the whistleblowing reward the higher the efficiency of the tax regime. However, the values for $R$ tell a different story. While "BASE" still performs poorly ($R = 42.2\%$), "SM INC" is now the treatment with the highest efficiency score: Since $R$ equals $114.6\%$, the bottom line is that the state actually profits from tax evasion attempts. As we have seen, in "SM INC", medium-scale tax evasion collides with medium-scale whistleblowing frequency, leading to a comparably high number of whistleblowing triggered audits (Figure 5) and correspondingly high fines. Therefore, the fines overcompensate the losses from undetected tax evasion and the costs of paying whistleblowing rewards. Compared to "SM INC", the number of reported and punished tax evaders in "NO INC" is too small to generate a similar-sized effect ($R = 97.2\%$). In "LA INC", the number of reported and punished tax evaders is also very high, but the amount of evaded taxes and, consequently, fines is smaller. Moreover, the costs of paying whistleblowing rewards are much higher. As a result, the "overcompensation effect" is less pronounced than in "SM INC" ($R = 103.1\%$). In the second round, the difference between "SM INC" ($R = 121.6\%$) and "LA INC" ($R = 97.4\%$) becomes even larger.

**Conclusion**

In this paper, the effect of monetary incentives on whistleblowing behavior and the effect of different whistleblowing regimes on tax compliance were experimen-
tally investigated. A tax regime without any whistleblowing mechanism was compared to three tax regimes that allow subjects to blow the whistle on tax evaders. Those three tax regimes differ from each other in terms of the incentives for whistleblowing: In one treatment, whistleblowing is costly, while in the other two treatments whistleblowers receive a reward that exceeds the costs in one case marginally and in the other case substantially.

We observe that monetary rewards for whistleblowing lead to a significant increase in the reporting of tax evasion. Moreover, we find a distinct pattern: the larger the reward the more pronounced the increase in whistleblowing. These findings do not lend support to a crowding-out effect of external whistleblowing rewards that was hypothesized in the literature. In the present experiment, subjects’ behavior seems to be much more motivated by external rewards than the actions that participants of vignette studies predicted to take. However, our results also contradict predictions based on standard economic theory, since subjects do not always follow the incentives. While some subjects blow the whistle even if this behavior results in certain financial losses, other subjects refuse to report tax evasion even if they have to forego substantial whistleblowing rewards. Thus, the results suggest that some whistleblowers as well as some people who refuse to blow the whistle are motivated by intrinsic ethical concerns. The whistleblowing behavior is, by and large, correctly anticipated by taxpayers. Only the increase in whistleblowing behavior due to the introduction of small monetary incentives is underestimated by subjects. These results nicely fit the perceived general whistleblowing behavior predicted by participants of vignette studies. It thus seems that subjects have a fairly realistic view of others when it comes to assessing the influence of external rewards on whistleblowing behavior. From an efficiency point of view, especially large monetary incentives for whistleblowers are an effective way to maximize tax compliance. This makes whistleblowing
rewards a very powerful tool against tax evasion. But even without monetary incentives, whistleblowing proves to be fairly effective in curbing tax evasion.
References


Appendix

Figures

Figure 1: Whistleblowing (Probability of Whistleblowing)

Figure 2: Tax Declaration (Percentage of Income Reported)
Figure 3: Box Plot for Tax Declaration

Figure 4: Whistleblowing Expectations
Figure 5: Share of Controlled and Punished Subjects (Round 1)
### Tables

<table>
<thead>
<tr>
<th>Whistleblowing (round 1)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small Incentives</td>
<td>0.522**</td>
<td>0.537**</td>
<td>0.514**</td>
<td>0.518**</td>
</tr>
<tr>
<td></td>
<td>(0.2098)</td>
<td>(0.2112)</td>
<td>(0.2102)</td>
<td>(0.2069)</td>
</tr>
<tr>
<td>Large Incentives</td>
<td>1.394***</td>
<td>1.394***</td>
<td>1.564***</td>
<td>1.588***</td>
</tr>
<tr>
<td></td>
<td>(0.2463)</td>
<td>(0.2465)</td>
<td>(0.3232)</td>
<td>(0.3224)</td>
</tr>
<tr>
<td>Effort (round 1)</td>
<td>-0.0045</td>
<td>0.0001</td>
<td>-0.0029</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0068)</td>
<td>(0.0076)</td>
<td>(0.0079)</td>
<td></td>
</tr>
<tr>
<td>Charity</td>
<td>0.0255</td>
<td>0.0332</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0486)</td>
<td>(0.0483)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.0110</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0203)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td>-0.2660</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.1973)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-0.0546</td>
<td>0.1684</td>
<td>-0.2314</td>
<td>-0.2661</td>
</tr>
<tr>
<td></td>
<td>(0.1539)</td>
<td>(0.3662)</td>
<td>(0.5358)</td>
<td>(0.7371)</td>
</tr>
<tr>
<td>Observations</td>
<td>138</td>
<td>138</td>
<td>114</td>
<td>114</td>
</tr>
<tr>
<td>Prob &gt; Chi^2</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

*Standard errors in parentheses*

*** p < 0.01, ** p < 0.025, * p < 0.05

Table 1: Whistleblowing (Round 1), Tobit Regression Models
### Table 2: Whistleblowing (Round 2), Tobit Regression Models

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small Incentives</td>
<td>0.956***</td>
<td>0.956***</td>
<td>0.935***</td>
<td>0.950***</td>
</tr>
<tr>
<td></td>
<td>(0.3139)</td>
<td>(0.3143)</td>
<td>(0.3053)</td>
<td>(0.3056)</td>
</tr>
<tr>
<td>Large Incentives</td>
<td>2.047***</td>
<td>2.047***</td>
<td>2.038***</td>
<td>2.045***</td>
</tr>
<tr>
<td></td>
<td>(0.4071)</td>
<td>(0.4081)</td>
<td>(0.4843)</td>
<td>(0.4833)</td>
</tr>
<tr>
<td>Effort (round 1)</td>
<td>0.0002</td>
<td>-0.0016</td>
<td>-0.0050</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0091)</td>
<td>(0.0098)</td>
<td>(0.0104)</td>
<td></td>
</tr>
<tr>
<td>Charity</td>
<td>0.0090</td>
<td>0.0006</td>
<td>(0.0683)</td>
<td>(0.0687)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.0979)</td>
</tr>
<tr>
<td>Age</td>
<td>-0.0124</td>
<td></td>
<td>(0.0288)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.0286)</td>
</tr>
<tr>
<td>Gender</td>
<td>-0.2375</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.2826)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.0636</td>
<td>-0.0734</td>
<td>0.1007</td>
<td>0.6250</td>
</tr>
<tr>
<td></td>
<td>(0.2158)</td>
<td>(0.5554)</td>
<td>(0.7719)</td>
<td>(1.0825)</td>
</tr>
<tr>
<td>Observations</td>
<td>138</td>
<td>138</td>
<td>114</td>
<td>114</td>
</tr>
<tr>
<td>Prob &gt; Chi$^2$</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Standard errors in parentheses  
*** $p < 0.01$, ** $p < 0.025$, * $p < 0.05$

### Table 3: Tax Compliance (Round 1), Tobit Regression Models

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Incentives</td>
<td>65.97***</td>
<td>67.82***</td>
<td>59.31***</td>
<td>62.48***</td>
</tr>
<tr>
<td></td>
<td>(16.87)</td>
<td>(16.98)</td>
<td>(15.78)</td>
<td>(15.89)</td>
</tr>
<tr>
<td>Small Incentives</td>
<td>66.08***</td>
<td>69.48***</td>
<td>57.86***</td>
<td>59.74***</td>
</tr>
<tr>
<td></td>
<td>(17.09)</td>
<td>(17.44)</td>
<td>(16.27)</td>
<td>(16.33)</td>
</tr>
<tr>
<td>Large Incentives</td>
<td>112.3***</td>
<td>113.8***</td>
<td>123.5***</td>
<td>125.3***</td>
</tr>
<tr>
<td></td>
<td>(18.78)</td>
<td>(18.90)</td>
<td>(24.49)</td>
<td>(24.48)</td>
</tr>
<tr>
<td>Effort (round 1)</td>
<td>-0.553</td>
<td>-0.346</td>
<td>-0.0925</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.510)</td>
<td>(0.525)</td>
<td>(0.534)</td>
<td></td>
</tr>
<tr>
<td>Charity</td>
<td>7.730**</td>
<td>7.229**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.985)</td>
<td>(2.989)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>1.504</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.497)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>16.66</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(12.70)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>27.29**</td>
<td>52.96*</td>
<td>-3.680</td>
<td>-58.80</td>
</tr>
<tr>
<td></td>
<td>(11.37)</td>
<td>(26.17)</td>
<td>(32.45)</td>
<td>(53.79)</td>
</tr>
<tr>
<td>Observations</td>
<td>186</td>
<td>186</td>
<td>162</td>
<td>162</td>
</tr>
<tr>
<td>Prob &gt; Chi$^2$</td>
<td>3.20e-09</td>
<td>7.61e-09</td>
<td>4.28e-10</td>
<td>1.47e-09</td>
</tr>
</tbody>
</table>

Standard errors in parentheses  
*** $p < 0.01$, ** $p < 0.025$, * $p < 0.05$
### Table 4: Tax Compliance and Whistleblowing Expectation (Round 1), Linear Regression Models

<table>
<thead>
<tr>
<th>Tax Declaration (round 1)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small Incentives</td>
<td>-1.333</td>
<td>-0.750</td>
<td>-1.498</td>
<td>-6.708</td>
</tr>
<tr>
<td>(7.219)</td>
<td>(7.256)</td>
<td>(6.999)</td>
<td>(7.301)</td>
<td></td>
</tr>
<tr>
<td>Large Incentives</td>
<td>15.84*</td>
<td>15.58*</td>
<td>20.32**</td>
<td>15.74</td>
</tr>
<tr>
<td>(7.105)</td>
<td>(7.118)</td>
<td>(8.600)</td>
<td>(8.726)</td>
<td></td>
</tr>
<tr>
<td>Effort (round 1)</td>
<td>-0.201</td>
<td>-0.171</td>
<td>-0.237</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.229)</td>
<td>(0.250)</td>
<td>(0.248)</td>
<td></td>
</tr>
<tr>
<td>Charity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.976</td>
<td>2.568</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.592)</td>
<td>(1.578)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whistleblowing Expectation</td>
<td></td>
<td></td>
<td></td>
<td>0.623*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.122)</td>
</tr>
<tr>
<td>Constant</td>
<td>69.82**</td>
<td>79.83***</td>
<td>57.64***</td>
<td>53.56***</td>
</tr>
<tr>
<td></td>
<td>(5.104)</td>
<td>(12.52)</td>
<td>(17.44)</td>
<td>(17.27)</td>
</tr>
<tr>
<td>Observations</td>
<td>138</td>
<td>138</td>
<td>114</td>
<td>114</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.051</td>
<td>0.057</td>
<td>0.121</td>
<td>0.157</td>
</tr>
<tr>
<td>$Prob &gt; Chi^2$</td>
<td>0.0284</td>
<td>0.0490</td>
<td>0.00659</td>
<td>0.00215</td>
</tr>
</tbody>
</table>

*Standard errors in parentheses*

*** $p < 0.01$, ** $p < 0.025$, * $p < 0.05$
<table>
<thead>
<tr>
<th>Tax Declaration (round 2)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controlled and punished</td>
<td>24.57***</td>
<td>8.835</td>
<td>8.616</td>
<td>12.78</td>
<td>11.88</td>
</tr>
<tr>
<td>No Incentives</td>
<td>(8.680)</td>
<td>(7.613)</td>
<td>(7.723)</td>
<td>(9.042)</td>
<td>(9.068)</td>
</tr>
<tr>
<td>Small Incentives</td>
<td>33.62***</td>
<td>33.54***</td>
<td>32.54***</td>
<td>31.10***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(7.477)</td>
<td>(7.507)</td>
<td>(7.739)</td>
<td>(7.802)</td>
<td></td>
</tr>
<tr>
<td>Large Incentives</td>
<td>46.28***</td>
<td>46.16***</td>
<td>44.45***</td>
<td>43.59***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(7.638)</td>
<td>(7.687)</td>
<td>(8.074)</td>
<td>(8.117)</td>
<td></td>
</tr>
<tr>
<td>Effort (round 1)</td>
<td>63.16***</td>
<td>63.20***</td>
<td>66.81***</td>
<td>66.14***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(7.494)</td>
<td>(7.517)</td>
<td>(9.760)</td>
<td>(9.769)</td>
<td></td>
</tr>
<tr>
<td>Charity</td>
<td>0.0394</td>
<td>0.00728</td>
<td>(0.211)</td>
<td>(0.239)</td>
<td>(0.253)</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td>(1.437)</td>
<td>(1.451)</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-8.575</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(6.188)</td>
</tr>
<tr>
<td>Constant</td>
<td>58.28***</td>
<td>25.02***</td>
<td>22.93</td>
<td>18.51</td>
<td>37.38</td>
</tr>
<tr>
<td></td>
<td>(3.368)</td>
<td>(5.182)</td>
<td>(12.36)</td>
<td>(16.42)</td>
<td>(26.27)</td>
</tr>
</tbody>
</table>

Observations 186 186 162 162 162

$R^2$ 0.042 0.323 0.323 0.325 0.334

$Prob > Chi^2$ 0.00516 0 0 0 1.04e-10

*Standard errors in parentheses*

*** $p < 0.01$, ** $p < 0.025$, * $p < 0.05$

Table 6: Tax Compliance (Round 2), Linear Regression Models