Deterrence Effects of Auditing Rules: An Experimental Study

Tan, Fangfang and Yim, Andrew

Tilburg University, CentER, Cass Business School, City University, London

14 December 2010

Deterrence Effect of Auditing Rules: An Experimental Study*

Fangfang Tan † Andrew Yim ‡

This version: December 14, 2010

Abstract

This paper experimentally examines a new auditing rule termed the bounded rule, which takes into account the budget constraint of the auditor (e.g., a tax authority). Compared to a traditional rule that audits income reports with a constant probability, the bounded rule can induce the same deterrence effect with a smaller budget. The basic setting follows a classic tax-compliance game in which each taxpayer receives either high or low income with certain probability. On knowing an auditing rule, the taxpayers have to decide simultaneously and independently whether to report their income truthfully to the auditor. The traditional rule audits every low-income report with a constant probability. The bounded rule audits a randomly selected sample of low-income reports whenever the number of these reports exceeds the maximum number of audits allowed by the budget, or otherwise all of the low-income reports. The experimental evidence suggests that, as predicted, the two auditing rules have the same deterrence effect. The bounded rule needs a smaller budget ex-ante, and conducts fewer audits ex-post. The results provide support for the bounded rule as a more cost-effective alternative to the traditional rule.

JEL Classification numbers: H26, M42, C9, C72

Keywords: Audit sampling plan, tax audit, tax compliance, tax evasion, experimental economics.

*We are particularly grateful for the valuable suggestions by Ralph-C Bayer, Luc Bissonnette, Eddy Cardinaels, Vilen Lipatov, Wieland Mueller and Charles Noussair. We thank James Alm, Cedric Argenton, Rudolf Avenhaus, Jochen Bigus, Kim Bloomquist, Folkert Botma, Liesbeth Bruynseels, Douglas DeJong, Lata Gangadharan, Karla Johnstone, Steven Kachelmeier, William Kinney, Tobias Klein, Ben Li, Ksenia Panidi, Jan Potters, Ernesto Reuben, Jason Shachat, Joel Slemrod, Bin Srinidhi, Konrad Stahl, Jeroen Suijs, Stefan Trautmann, Peter van Oudheusden, Zongxin Qian, Arthur van Soest, Gijs van de Kuilen, Ann Vanstraelen, Eyal Winter, Shlomo Yitzhaki and seminar participants of Tilburg University, the GSS interdisciplinary workshop at CentER, 2010 Behavioral Decision Research in Management (BDRM) Conference held at Carnegie Mellon University and 66th Annual Congress of the International Institute of Public Finance (IIPF) for helpful comments. We thank Chung Kim and Michael McKee for sharing their experimental instructions, and CentER lab for financial support. Siqi Pan and Huojun Sun provided excellent research assistance. All remaining errors are ours.

†Department of Economics, Tilburg University, Tilburg, the Netherlands. E-mail: F.tan@uvt.nl.

‡Cass Business School, City University London, 106 Bunhill Row, London EC1Y 8TZ, UK. E-mail: a.yim@city.ac.uk.

1 Introduction

Tax compliance has been a central topic for government tax agencies. According to the US Internal Revenue Service (IRS), the most recent estimated tax gap (i.e. the difference between what taxpayers should pay and what they actually pay) ranges from \$312 billion to \$353 billion. About 50 percent of the gap can be attributed to individual income underreporting. Due to the enormous size of the gap, each 1 percent reduction of the gap would likely yield more than \$2.5 billion annually (US Department of the Treasury (2005)). Hence, in the last five years, roughly half of the IRS's annual budget (\$4.7 to \$5.5 billion) was allocated for enforcement purposes, including increasing tax compliance, improving data-matching technology and promoting the effectiveness of evasion detection (US Government Accountability Office (2009)). In light of the huge budgets devoted to reduce the tax gap, this paper explores the possibility of conducting tax audits more efficiently. In particular, it examines an auditing rule that induces the target deterrence level with a lower budget.

The existing tax-compliance literature widely analyzes a simple random auditing rule. That is, each taxpayer is independently selected for audit with a constant probability (see, for example, Moser et al. (1995), Zimbelman and Waller (1999), Boylan and Sprinkle (2001), Kim et al. (2005), Kim and Waller (2005), Alm et al. (2009) and Kleven et al. (2010)). For simplicity, we term this the *traditional* rule.

One undesirable feature of the traditional rule is that the auditing budget can be used inefficiently. Due to uncertainty arising from the random auditing decisions or other sources[1], the actual auditing expenditure could vary substantially across years. This leads to difficulties in planning resources ex-ante. Consequently, the IRS needs a large budget to support the traditional rule. In reality, however, budgets set aside for organizational activities including auditing have limited flexibility for other purposes unless under extraordinary circumstances (see, e.g. US Department of the Treasury (2006)). If the IRS sets aside a large budget for auditing purposes but only uses a fraction of the budget, then it inefficiently ties up resources that could have been better used elsewhere for an entire fiscal year. In sum, the

[1] For instance, the proportion of the "red-flagged" (suspicious) reports filed could vary significantly across years. If the tax-evasion rate fluctuates along with the condition of the general economy, the IRS cannot distinguish whether the cause is tax evasion or other reasons, upon observing many red-flagged reports.

IRS might have a greater incentive to formulate auditing rules that plan and use budgets more efficiently.

This paper analyzes an auditing rule named the *bounded* rule. When applying the bounded rule, the auditor needs to set up an audit capacity (i.e., a maximum number of audits allowed by the budget). The bounded rule chooses an audit sample from the population of "red-flagged" (suspicious) tax reports, given the audit capacity. It audits a randomly selected sample of these suspicious reports whenever its total number exceeds the audit capacity, or otherwise all of these reports. Unlike the traditional rule, the audit probability of a taxpayer under the bounded rule is no longer exogenously given. Instead, it hinges on the proportion of "red-flagged" reports submitted. The feature of endogenous audit probability induces strategic interactions among taxpayers. Nevertheless, for any given number of players and income distribution, the bounded rule can induce the same deterrence effect as the traditional rule with a targeted level of audit probability. Put differently, the induced tax-compliance level is the same under both rules. However, the bounded rule requires a lower budget and uses the budget more efficiently.[2]

Our laboratory experiment supports the bounded rule as a more cost-effective alternative to the traditional rule. The laboratory offers a controlled environment to test the deterrence effect of auditing rules directly. Such control can isolate many factors that confound behavior. It is also helpful in examining factors that have been omitted in the theory, such as bounded rationality and risk attitudes.

The laboratory setting in this paper is as follows. Every taxpayer has a certain probability of receiving high or low income. Knowing a certain auditing rule, they have to decide simultaneously and independently whether to report their income truthfully to the tax authority. The tax authority implements either the traditional or the bounded rule after deducting taxes according to players' reported income. Parameters are selected such that the two auditing rules induce the same level of compliance.

[2]The idea of the bounded rule is inspired by Yim (2009). A key assumption distinguishing this paper from Yim (2009) is the ability of the auditor to commit to an auditing strategy. In Yim (2009), the auditor interacts strategically with taxpayers, and chooses the audit probability on observing the behavior of taxpayers. In this study, the auditor commits to an auditing strategy so that the focus is on the reactions of the taxpayers. Consequently, the properties of the bounded rule explored in this study are fundamentally different from Yim (2009).

The main experimental results are the following. The deterrence effect of the bounded rule is as strong as that of a traditional rule. However, the bounded rule is more cost-effective for two reasons. First, it conducts fewer audits to attain the same level of deterrence. Second, it uses the budget more efficiently. The budget-usage ratio, which is defined as the percentage of resources actually used in auditing for a given budget, is higher for the bounded rule.

The data also show that theory underpredicts the level of deterrence. In other words, the compliance level for both rules is higher than theoretical predictions. The reason is that subjects' decisions are highly stochastic, as a large proportion of them switch their decisions across periods. In order to account for behavioral anomalies, this paper develops and compares several structural models on choices under risk and uncertainty. The behavior in our data is consistent with a model of loss aversion with stochastic decision errors.

So far, the bounded rule has generated deterrence solely based on behavioral instead of theoretical reasons. This paper also examines the bounded rule in another parameter domain where theory could predict full compliance. In this new treatment, all parameters remain the same except that the ex-ante probability of receiving high income increases. The game induced by the bounded rule has both a payoff-dominant equilibrium in which all strategic taxpayers underreport, and a risk-dominant equilibrium in which all strategic taxpayers report truthfully. The experimental evidence suggests that in the presence of multiple equilibria, the bounded rule generates even higher deterrence, as players' behavior tends to converge to the full-compliance equilibrium. The parameters chosen in this treatment resemble a rich neighborhood where every taxpayer is likely to earn high income. The results provide further support for the bounded rule as a more cost-effective auditing rule when the income distribution is left-skewed.

This study makes three contributions. First, it offers experimental evidence on a new cost-effective auditing rule. Several studies discuss alternative auditing rules opposed to the simple random audit.[3] Although these papers suggest that rules contingent on strategic interactions among players might prove to be more deterrent, little is known about the actual responses of taxpayers to these rules. This study empirically shows that the bounded rule is

[3]See, for instance, Reinganum and Wilde (1985), Harrington (1988) and Bayer and Cowell (2009). Slemrod and Yitzhaki (2002) provide a detailed discussion of these alternative auditing rules.

more cost-effective than the traditional rule.

Second, the paper presents an example of using more realistic assumptions in the experimental tax-compliance literature. The traditional rule and its variants have been widely studied in this literature (see the literature review by Alm and McKee (1998), Torgler (2002)). A meta study by Blackwell (2007) based on twenty laboratory experimental studies finds that an increase in audit probability or fine rate leads to higher compliance, but an increase in the tax rate has no significant effect. The previous studies, by assuming a simple random audit, neglect the resource constraint in reality.[4] This paper proposes a relatively easy way of modeling budget constraints in constructing auditing rules, and shows a first piece of evidence that the deterrence effect induced is indeed the same.

Third, this paper is the first to estimate structural models of utility using data from a tax-compliance experiment. These models require less strict assumptions regarding cognitive reasoning or the ability to form correct beliefs, and hence offer a much more satisfactory account of behavior in our data. Moreover, the exercise of structural estimation allows the comparison of alternative behavioral models.

The paper is organized as follows. Section 2 describes the tax-compliance model and auditing rules that are examined in the experiment. Section 3 constructs an experimental design in which the two auditing rules induce the same level of compliance, and then analyses the experimental data with both nonparametric and parametric methods. Section 4 examines the bounded rule in another parameter domain where the interactions among players have multiple equilibria. Section 5 displays the effect of learning and social demographics on compliance behavior. Section 6 concludes and discusses directions for future research.

[4] Cason and Gangadharan (2006) test the Harrington (1988) model in which the audit probability negatively correlate with the agents' previous compliance performance. Alm et al. (1993) examine a cut-off rule by combining a sure audit below a threshold on reported income and a small, random audit above the threshold. Alm and McKee (2004) consider another cut-off rule with audit probability depending on the deviation of an individual's reported income from the average of the incomes reported by all other players. This paper differs from theirs in that the bounded rule induces interactions among taxpayers, but such interaction does not always need to be a coordination game.

2 Model Description

The model of this paper builds upon some basic features of a classic tax-compliance game developed by Graetz et al. (1986). Consider a player population of size N. For simplicity, the model assumes two income classes: high and low, denoted I_H and I_L, respectively, where $I_L < I_H$. Each player has a probability q of being a high-income taxpayer (H-type) and $1-q$ of being a low-income taxpayer (L-type), where $0 < q < 1$. Players know the type distribution as well as their own types, but they do not know the exact types of the other players. Each player has to decide simultaneously and privately whether to report high income (I_H) or low income (I_L) to the tax authority. Let T_H and T_L for the tax payment by high- and low-income taxpayers, respectively, where $T_H < I_H$, $T_L < I_L$, and $T_L < T_H$. If cheaters are audited, then a fine F is imposed on top of the tax they should have paid ($F > 0$). However, taxpayers who report truthfully are never fined and incur no cost if they are audited. The following analysis assumes that players are homogeneous, rational, risk-neutral profit maximizers.

The *traditional rule* can be presented easily. Any taxpayer who has filed a "low-income" report will face a flat probability a_{TR} of being audited independently. Since reporting truthfully does not incur any cost when being audited, L-type players always report their income truthfully. If they report high income, they will be taxed T_H, which is strictly larger than the tax T_L they need to pay if they honestly state income. For H-type players, the honest-reporting payoff is $I_H - T_H$. If they underreport, the payoff is $I_H - T_L$ if they are not audited, and $I_H - T_H - F$ if they are audited. Therefore, they choose to underreport if and only if the expected profit is strictly larger:

$$(1 - a_{TR})(I_H - T_L) + a_{TR}(I_H - T_H - F) > (I_H - T_H).$$

If the audit probability is less than the threshold $\bar{a}$ defined by

$$\bar{a} = \frac{T_H - T_L}{F + T_H - T_L},$$

the H-type players will underreport. Otherwise, if the audit probability is larger than $\bar{a}$, they choose to report truthfully.

Note, however, that the traditional rule does not model the budget constraint explicitly. In fact, due to its "coin-flipping" nature, the application of the traditional rule implicitly assumes that the tax agency has the budget to carry out a full audit of N files. Even in terms of expected number of audits, the larger the number of "low-income" reports L turned in, the larger is the expected number of audits needed. The following paragraph presents an alternative auditing rule taking into account the resources of the tax agency. It allows the tax agency to induce the same level of compliance with a lower budget.

The *bounded rule* requires the auditor to set up a maximum number of K audits allowed, given the budget. It then constructs an audit sample size contingent on the number of "low-income" reports L. If L is smaller than or equal to the audit capacity K, the auditor will audit all L reports. However, if L is strictly larger than K, then the auditor will randomly audit K reports. Expressed more formally, every "low-income" taxpayer under the bounded rule faces the following audit probability:

$$a_{BD} = \begin{cases} 1 & \text{if} \quad L \leq K \\ K/L & \text{if} \quad L > K \end{cases}$$

for $L = 0, 1, ...N$.

The key feature of the bounded rule is that the audit probability a_{BD} is no longer exogenously given. Instead, it depends on the audit capacity K and the number of reported "low-income" files L. The latter is a function of population size N and the ex-ante probability q of being an H-type. The following proposition characterizes a property of the bounded rule.

Proposition 1 *For any given N and q, the auditor can always choose an audit capacity K for the bounded rule such that it induces the same compliance level as the traditional rule.*

Proof: See appendix A. ∎

The intuition of Proposition 1 is as follows. Any audit probability a_{TR} under the traditional rule induces all-or-none compliance behavior. If the maximum number of K is so high that all "low-income" reports will always be audited for sure, H-type players will have no incentive to underreport. On the other hand, if K is zero (meaning that no audit is conducted regardless of the number of "low-income" reports submitted), then H-type players will underreport with certainty. Between these two extreme cases there exists a threshold $\bar{K}$ such that any $K > \bar{K}$ sustains compliance behavior regardless of the actual income-realization parameter q. That is, even in the scenario which all taxpayers claim low income, the audit probability is still high enough to deter tax evasion.

To induce full compliance, however, the committed budget K does not always need to be larger than $\bar{K}$. Put it differently, even when $K < \bar{K}$, the bounded rule is still able to induce full compliance. Depending on the parameters, the interactions among taxpayers induced by the bounded rule could either be a dominance-solvable game with one unique equilibrium, or a coordination game with multiple equilibria. Section 3 examines the former case focusing on comparing its deterrence effect to that of the traditional rule. Then it examines the latter case in Section 4.

The above analysis shows a desirable feature of the bounded rule. That is, the maximum audit number needed to sustain the full compliance equilibrium, K, is always less than N, which is the maximum number of audits needed for a traditional rule. Consequently, without sacrificing the induced-compliance effect, the tax agency can always commit to a lower budget to support the implementation of the bounded rule.

3 Dominance

3.1 Experimental Design and Procedure

This section presents the experimental procedures and parameters that induce the same deterrence effect for both auditing rules. Based on the capacity constraint in the lab, the size of the taxpayer population is fixed to be $N = 8$.

The tax-compliance game in both treatments has three stages: (i) income reporting and tax deduction, (ii) audit and fine deduction, and (iii) feedback. Subjects receive either high income (I_H) €25 or low income (I_L) €10 with probability (q) 0.5. Subjects are informed about the group size N and the probability q. During the income-reporting stage, they have to decide simultaneously and independently the type of income to report to an auditor, which is simulated by a computer. The computer automatically deducts taxes according to the reported income. The tax for subjects reporting "high income" (T_H) is €12.5, whereas the tax for subjects reporting "low income" (T_L) is €2.5.[5] Subjects are told that taxes are deducted based on their reported income instead of true income. For instance, H-type players receive €22.5, instead of €12.5, if they submit "low-income" reports. Similarly, L-type players receive -€2.5, instead of €7.5, if they submit "high-income" reports.[6] In the audit stage, the computer implements either a traditional rule or a bounded rule to audit "low-income" taxpayers.

Traditional: In the traditional rule sessions, subjects filing "low-income" reports face an independent audit probability of 0.4. This audit probability induces the same compliance rate to the bounded rule.[7] If they indeed report honestly, nothing will happen to their final payoffs. However, if cheaters are caught by the auditor, then they need to pay back the €10 of taxes evaded plus a fine (F) of €10.

Bounded: In the bounded rule sessions, the audit probability depends on the total number of "low-income" reports received. The maximum number of audits to be conducted is $K = 2$. This means that if the number of low-income reports does not exceed two, then all of them will be audited with probability 1. Otherwise, the audit probability decreases monotonically with the number of "low-income" reports L. In particular, the probability is 0.67 for $L = 3$; 0.5 for $L = 4$; 0.4 for $L = 5$; 0.33 for $L = 6$; 0.29 for $L = 7$; and 0.25 for $L = 8$. This parameter K guarantees a unique Nash equilibrium based on non-cooperative game theory (see analysis below). The fine for cheaters is exactly the same as in the *Traditional* treatment.

The experiments are conducted at the CentER Lab in Tilburg University from October

[5] Experimental parameters concerning taxation are chosen to be in line with reality. For instance, the real-world tax rates for high-income and low-income taxpayers are usually dependent on the levels of their incomes. In particular, many countries such as Britain, the Netherlands, Germany, Italy and the USA use a progressive tax system instead of a proportional one. Hence, this experiment adopts a progressive tax system for the sake of facilitating subjects' understanding.

[6] Even when a subject with low income makes a loss by submitting "high income" reports and that decision is selected for payment, the potential loss is covered by a show-up fee of €3. During the experiment sessions, this situation never actually happens.

[7] Due to the fact that the traditional rule induces all-or-none behavior in compliance, any audit probability $a < 0.5$ is theoretically equivalent to the bounded rule. Nevertheless, this statement only holds for perfectly rational, risk- neutral players. To what extent this holds is an empirical question to be tested.

to December 2009. Tilburg University students, mostly major in economics or business, participate as subjects in the experiment. Each treatment consists of four sessions of 16 subjects each. The duration of a session is about 1 hour (including the initial instruction and final payment to subjects). The average earnings are €16.23 (including the €3 show-up fee). The experiments are programmed and conducted in Z-Tree software (Fischbacher (2007)).

The instructions of the tax-compliance game are modified from instructions in prior studies of the literature, namely Alm et al. (2009), Kim et al. (2005), and Kim and Waller (2005) (see Appendix B.2). At the beginning of each session, subjects are randomly assigned to the computer terminals. Before the experiment starts, subjects have to complete an exercise to make sure they understand the rules of the game.

The game consists of 30 periods. At the beginning of each period, 16 subjects are randomly allocated into two groups of eight. The random rematching protocol minimizes the chances that subjects encounter the same group of participants again. It simulates a one-shot scenario but allows the subjects to be familiar with the game environment. At the end of each period, a summary screen is presented to subjects with feedback information including the subject's true and reported income, and the final payoff for the period. Subjects are not informed of others' payoffs.

Upon completing the tax-compliance experiment, subjects are asked to complete a risk elicitation task similar to the one used by Holt and Laury (2002).The instructions for the risk elicitation task are handed out only after the tax-compliance game. Hence, the subjects are not aware of its existence beforehand. In this task, subjects have to make selections of a set of 21 lottery pairs. Each lottery pair consists of a safe and a risky lottery. The expected payoff of the risky lottery compared to the safe one is the lowest in the first pair, and the highest in the last pair. The switching point from the safe to the risky lottery reflects subjects' risk tolerance level. These data is used to explain behavior in the tax-compliance game.

At the end of the experiment, subjects are asked to complete two questionnaires. The first one concerns social background information such as gender, nationality, and years of studying economics. The second one elicits subjects' Machiavelli scores by means of the

Machiavellian scale personality test (see Christie and Geis (1970)).[8]

During the payment stage, one period of the tax game and the realization of one lottery are randomly selected to determine the final payment of a subject. This random payment scheme mitigates the potential income effect that subjects carry across games and over different periods within a game.

3.2 Hypotheses

This section presents hypotheses regarding the deterrence effect of both rules. The deterrence effect is indicated by the underreporting rate in the population: namely, the proportion of high-income taxpayers filing "low-income" reports in a certain period. As discussed in Section 2, the analysis focuses on the H-type players, as the L-type players have a dominant strategy of reporting honestly, regardless of the auditing rules.[9]

In the following, let h be the honestly reporting strategy for H-type players, and u be the underreporting strategy. As the audit probability a_{TR} is set to be 0.4 for the traditional rule, an underreporting decision is equivalent to selecting a lottery of €22.5 with probability 0.6 and €2.5 with probability 0.4. The expected payoff is therefore: $E(\pi_u) =$ €$22.5 \times 0.6 +$ €$2.5 \times 0.4 =$ €14.5. As it is strictly larger than the sure payoff €12.5 from an honest report, H-type players are expected to underreport.

Under the bounded rule, the H-type players again face the tax-evasion gamble of choosing a sure payoff of €12.5, or a high payoff of €22.5 if they are not audited but a low payoff €2.5 otherwise. Unlike the traditional rule, however, the audit probability a_{BD} is not exogenously given. Instead, it depends on the players' perception of the actions of others. In particular, it depends on player i's subjective belief on the likelihood of the proportion of "low-income" reports turned in by another player, denoted by B_i.

A "low-income" report could come from two sources. The first source is from a truth-telling L-type player with probability $1-q$. Alternatively, it could come from H-type players who dishonestly report that they have received low income. If a player thinks that the

[8]This test measures a person's predisposition to be opportunistic and manipulative; with higher scores indicating that these properties are more pronounced.

[9]The actual percentage of honest reports among L-type taxpayers are 99.68% and 99.28% across treatments, suggesting that they do play the dominant strategy.

underreporting probability of H-type players i is b_i, this scenario will occur with probability qb_i. Hence, the overall probability of observing a "low-income" report B_i for player i is the sum of the probabilities in these two situations: $B_i = 1 - q + qb_i$.

The Nash equilibrium of this game can be reached by iterated elimination of dominated strategies. The intuition proceeds as follows. Reporting high income is a dominated strategy for L-type players, since they have to pay a high tax and incur a lower payoff than they would otherwise. If the H-type players believe that the L-type obey dominance, then the strategy of reporting truthfully (h) is dominated. That is, even when a H-type player believes that no other players evade taxes, the expected payoff of underreporting is still higher than that of honest reporting. Such a high expected payoff is caused by a low audit probability strictly less than 0.5, which stems from the fact that all of the L-type players (about half of the population) state low income truthfully. The calculation also gurantees that evading taxes is always a best response for a H-type player when L-type players obey dominance. Proposition 2 derives the equilibrium underreporting decisions.

Proposition 2 *The game introduced by the bounded rule is dominance solvable. In the equilibrium, both the L-type and H-type players report "low income".*

Proof: See appendix A. ∎

Note that the above hypothesis holds for strategic, self-regarding profit maximizers. Now suppose that some players are intrinsically honest: they report their income truthfully, regardless of their type. This assumption does not change the direction in terms of treatment differences. Recall that in the *Bounded* treatment, the optimal strategy of the H-type players does not depend on their beliefs towards other H-type players. As long as they believe that L-types will not play dominated strategy (i.e. reporting high income), they can form expectations on the proportion of "low-income" reports filed in each realized income distribution. Given that the ex-ante probability of being a L-type player is sufficiently high ($q = 0.5$), the sure payoff for a H-type player to report honestly is lower than the expected payoff from underreporting, even when s/he does not expect any other H-types to underreport. This ensures that all H-type players will continue to underreport with or without honest players.

The analysis in the *Traditional* treatment is simpler. As player decisions are independent, the audit probability facing self-regarding profit maximizers is unaffected by honest players. In sum, if the percentage of intrinsically honest players is assumed to be the same in both treatments, the compliance rate is the same. For the mathematical formulation, see Appendix A.

Let b^{TR} be the underreporting rate in the *Traditional* treatment, and b^{BD} be the underreporting rate in the *Bounded* treatment. The first hypothesis is built upon Proposition 2:

Hypothesis 1 *The underreporting rate is the same under both rules:* $b^{TR} = b^{BD}$.

The expected number of audits under the *Traditional* treatment, L^{TR}, depends on the number of "low-income" reports. Let p^{TR} denote the percentage of players that submitted "low-income" among N players. In our setting, if $p^{TR} > \frac{5}{8}$, then L^{TR} will be larger than the two audits committed in the *Bounded* treatment. Assume the cost of an audit to be the same in both treatments. Since the *Bounded* treatment needs fewer audits, it has a lower implementation cost.

Hypothesis 2 *If the percentage of "low-income" reports submitted is larger than 62.5% , the number of audits is smaller in the Bounded treatment than in the Traditional treatment.*

3.3 Average Treatment Effect

Table 1 summarizes the descriptive results of non-compliance behavior and profits across experimental treatments. All statistics reported in this table are on the session level. Columns 2 and 3 contain averages over all 30 periods of play, and columns 4 and 5 contain the results for the last 10 periods, where the behavioral pattern is more stable.

Table 1: Summary statistics across treatments (standard errors in parentheses)

	All 30 Periods		Last 10 Periods	
	Traditional	*Bounded*	*Traditional*	*Bounded*
All subjects				
High-income probability	0.514	0.491	0.527	0.519
	(0.007)	(0.039)	(0.042)	(0.038)
Percentage of "low-income" reports	79.741%	78.853%	77.969%	75.935%
	(0.074)	(0.015)	(0.066)	(0.018)
H-type subjects				
Underreport frequency	60.829%	57.114%	58.163%	53.321%
	(0.144)	(0.049)	(0.143)	(0.052)
Average underreport profit	14.513	16.446	13.272	15.781
	(0.650)	(0.285)	(0.967)	(1.979)
Auditing statistics				
Total audit number	153.751	120	53.75	40
	(18.140)	(0.000)	(8.098)	(0.000)
Average audit number	2.563	2	2.692	2
(per group per period)	(0.300)	(0.000)	(0.414)	(0.000)
Audit frequency	40.161%	31.712%	42.958%	32.943%
	(0.030)	(0.006)	(0.038)	(0.007)
Budget usage ratio	32.033%	100%	32.091%	100%
	(0.181)	(0.000)	(0.181)	(0.000)
Cheater detection rate	38.762%	33.134%	42.082%	31.876%
	(0.065)	(0.043)	(0.107)	(0.125)

The two rows on top of the table report statistics concerning all subjects. The first row indicates that the actual probability of being an H-type in both treatments is very close to their pre-specified levels with repeated drawing. The second row displays the percentage of low-income reports among all reports (i.e. reports from L-type players and the untruthful ones by H-type players). This number is around 80% in both treatments, which satisfies the condition in Hypothesis 2 allowing the comparison of the implementation costs of the two rules.

The next two rows focus on H-type players. The third row reports the overall underreport frequency, which is 60.83% in the *Traditional* treatment and 57.11% in the *Bounded* treatment. A two-sided Wilcoxon rank-sum test cannot reject the null hypothesis that the underreport frequencies of the two treatments are the same ($p = 0.386$). The profit for the cheaters in the *Bounded* treatment is €2 higher ($p < 0.05$), which may be due to the fact that the detection rate in this treatment is lower than that in the traditional rule (see the last row).

The final four rows concern audit statistics. Two pieces of evidence support the bounded rule to be more cost-effective. To begin with, it can sustain the same level of compliance with a lower cost. Due to the fact that the auditor effectively commits to fewer audit resources under the bounded rule, both the total audit number and the audit frequency are significantly lower ($p < 0.05$). This result is robust even when comparing the average number of audits per group per period ($p < 0.05$).

Apart from a lower implementation cost, the bounded rule has a higher budget-usage ratio. The budget usage ratio is defined as the percentage of resources actually used for a given budget. This figure is 100% in the *Bounded* treatment, which means that all resources committed are used at their full capacity in each period (i.e. two audits). Under the traditional rule, the budget-usage ratio is only 32.03%. The inefficiency comes from the fact that while the auditor has to prepare resources to do all eight audits in each period, only a small fraction of audits are actually carried out.

The last column shows the effectiveness of cheater detection. The success rate is higher in the *Traditional* than in the *Bounded* treatment, though not statistically significant ($p = 0.113$).

In both treatments, the underreporting rates decrease over time. Due to fewer "low-income" reports, the relative audit frequencies increase, and cheaters earn less. Nevertheless, the results of cross-treatment comparisons remain the same. Results 1 and 2 summarize the main findings in this section.

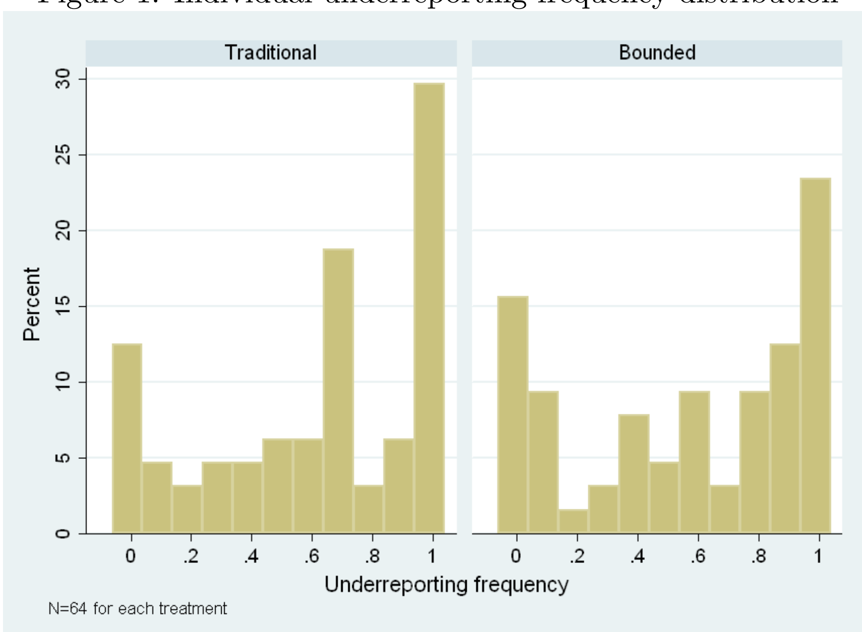
Result 1 *Hypothesis 1 is supported. The observed underreporting rates are not statistically different between the two treatments, although the absolute levels are significantly lower.*

Result 2 *Hypothesis 2 is supported. The bounded rule uses resources more efficiently in that 1) The average number of audits in the Bounded treatment is significantly smaller than* *that in the Traditional treatment, and 2) The budget-usage ratio is higher.*

3.4 Individual Behavior in the Game

Figure 1 displays the distribution of the underreporting rate for H-type players across treatments. The horizontal axis represents subjects' underreporting frequency throughout the game (i.e. the percentage of times when they receive high income and decide to underreport). The vertical axis represents the proportion of players having similar underreporting frequency in each treatment.

Figure 1: Individual underreporting frequency distribution



The main message conveyed by Figure 1 is that theory has limited explanatory power over the individual-level data: Only 29.13% of the subjects in the *Traditional* treatment and 23.43% of subjects in the *Bounded* treatment behave exactly in accordance with theory. That is, they underreport whenever they receive high income throughout the experiment. The percentage of intrinsically honest subjects who always report their income truthfully is 12.5% and 15.63%, respectively. Even corrected for the presence of honest players, theory underpredicts the deterrence effect of both auditing rules. According to Figure 1, around 60 percent of the subjects switch between the two options with various levels of frequency.

This pattern is the same in both treatments (Mann-Whitney test, $p = 0.322$).

3.4.1 Choice models under uncertainty

This section attempts to develop alternative models that explain the stochastic component of behavior. Theory based on individual profit maximization makes two unrealistic assumptions regarding behavior. The first is the assumption of perfect rationality. In reality, people are usually bounded by the cognitive limitation of their minds, given the amount of time they have to make decisions. The second assumption is of risk neutrality. The experimental literature documents mounting evidence that subjects are not risk-neutral profit maximizers, but rather risk-averse utility maximizers.

The discrete-choice model is a framework to relax the perfect rationality assumption and to accommodate boundedly rational behavior (McFadden (2001)). Models in this framework are motivated by empirical studies in which observed decisions exhibit some noise (see, e.g., Fischbacher and Stefani (2007), Loomes (2005), Rieskamp (2008) and Wilcox (2010)). Such noise could come from observed sources like decision errors, but could also come from other unobserved or unmodeled channels such as individual perceptions of the game, or sensitivity to payoff changes. Due to the presence of such noise, people make decision errors and hence do not behave consistently with their choices. Our *Baseline* treatment is essentially a non-strategic choice-under-uncertainty problem for H-type players. Therefore, the classic individual discrete-choice model is a natural setting to explore behavioral anomalies. The bounded treatment introduces interactions of players. A general way to incorporate decision error is the quantal response equilibrium first proposed by McKelvey and Palfrey (1995), which is based on the random utility-maximization model of McFadden (1973).

According to the discrete-choice framework, H-type players will choose to underreport if and only if the difference in the expected utilities is sufficiently large to exceed a stochastic error denoted by ε; i.e.,

$$EU(\pi_u) - \pi_h > \varepsilon.$$

In the expression, π_u and π_h denote the expected profits from underreporting and reporting honestly, respectively. The parameter ε is commonly assumed to be independently and identically distributed across players and actions with a Type 1 extreme value ("logit") distribution. The error can arise come from many sources, including the inability to calculate the expected payoff or trembling hands during decision making. A standard result of the discrete-choice model framework is that under the above error distributional assumptions, the underreporting probability $\widehat{b}$ is given by the relation below:

$$\begin{aligned} \widehat{b} &= \Pr\{EU(\pi_u) - \pi_h > \mu\varepsilon\} \\ &= \frac{1}{1 + \exp\left[-\frac{EU(\pi_u)-\pi_h}{\mu}\right]} . \end{aligned} \tag{1}$$

The parameter $\mu > 0$ captures the sensitivity of subjects' choices to the relative payoffs of the two choices. When μ approaches infinity, players choose underreporting and honest-reporting with equal probability, independent of the relative expected payoffs. When μ decreases, on the other hand, players put less probability weight on choices that yield suboptimal payoffs, and the probability that they make the optimal choice converges to 1 when μ approaches 0. Put differently, μ is an index of the measurement error when subjects calculate expected utility from underreporting.

Within this framework, this paper further relaxes the assumption of risk neutrality. In particular, three behavioral models are estimated and compared: risk-aversion, and loss aversion with- and without combining probability weighting. In the risk-aversion model, subjects are assumed to have a CRRA-form utility function $u(\pi) = \frac{\pi^{1-r}}{1-r}$.[10,11] This model offers the possibility of explicitly testing the assumption of risk neutrality. If the estimated r is significantly different from zero, then the null hypothesis that subjects are risk neutral can be rejected.

While the observed compliance behavior can be explained by risk attitude, it is also consistent with the notion of loss aversion. Recent research has shown that that loss aversion

[10] Alternative utility forms such as CARA and power-expo utility do not change the fit of the data.

[11] Data from the tax-compliance game alone do not have any identification power to jointly estimate three parameters, since they only contain two moments (i.e., the fraction of subjects selecting the "risky" lottery in the traditional rule and that in the bounded rule) given a fixed payoff structure. To gain enough identification power, we pool data from both the risk elicitation task and the tax-compliance game.

provides a much better account of tax evasion both in the lab and in the field (see, e.g.,Elffers and Hessing (1997), Yaniv (1999), King and Sheffrin (2002), Dhami and Al-Nowaihi (2007) and Dhami and Al-Nowaihi (2010)). The loss-aversion model characterizes individuals as loss averse in terms of reference income, denoted by R. For a given amount of money, $x > 0$, and the value function $v(x)$ (specified below), losses are weighted more than gains ($|-v(-x)| > v(x)$). This study follows Dhami and Al-Nowaihi (2007) and Dhami and Al-Nowaihi (2010) by taking the honest post-tax income as the reference point: $R = I_H - T_H$. The rationale for this reference point is as follows. If the reference point is selected differently say, the initial income or the income after cheating detection, then taxpayers are always in the domain of losses or gains. In those cases, the asymmetry of gains and losses disappears, and the analysis completely falls back to an expected-utility framework.[12] The income relative to the reference point is as follows:

$$\pi_i = \begin{cases} I_H - T_H - F - R & \text{for} \quad i \text{ is caught.} \\ I_H - T_L - R & \text{for} \quad i \text{ is not caught.} \end{cases}$$

The form of the utility function follows Tversky and Kahneman (1992). It is defined separately over gains and losses: $U(\pi_i) = \pi_i^\alpha$ if $\pi_i \geq 0$, and $U(\pi_i) = -\lambda(-\pi_i)^\beta$ if $\pi_i < 0$. The α and β are the parameters controlling for the curvature of the utility functions, and λ is the coefficient of loss aversion. Subjects are considered loss-averse if $\lambda > 1$.

Besides value functions, subjects could also have a nonlinear transformation of the probability scale (i.e. they overestimate low probabilities and underestimate high probabilities (see, e.g. Kahneman and Tversky (1979))). In order to examine the effect of subjective probability weight, this paper estimates a third model combining the loss-averse utility form with a probability-weighting function. In particular, this paper adopts a popular form of the one-parameter probability-weighting function: $w(p) = \frac{p^\delta}{(p^\delta + (1-p)^\delta)}$, where $\delta \geq 0$. Note that if $\delta < 1$, the weighting function has an inverted "S" shape, which is concave for low probabilities and convex for high probabilities, and crosses the diagonal at the probability of 1/3.

[12] More specifically, such a framework is called Rank dependent expected utility theory (RDEU), which can be considered as expected utility theory applied with a transformed cumulative probability distribution. See Dhami and al-Nowaihi (2007) for more detail.

Recall that H-type players are choosing between a safe lottery and a risky one with fixed probabilities in the traditional rule, but endogenous probabilities under the bounded rule. In the following, denote parameter "a" as the perceived audit probability in the *Bounded* treatment. The estimated parameter a answers the following question: If a bounded rule is transformed into the context of a traditional rule, which exogenous audit probability "a" best justifies behavior? Moreover, how do risk attitude, probability weighting, or loss aversion influence subjects' perception of the audit probability? The conditional log-likelihood is the following:

$$
\begin{aligned}
\ln L(\mu, a_{it}) &= \sum_{i,t} \left\{ y_{it} \cdot \ln \left(\frac{1}{1+\exp[\frac{\pi_h - E(\pi_u)}{\mu}]} \right) + (1-y_{it}) \cdot \ln \left(\frac{\exp[\frac{\pi_h - E(\pi_u)}{\mu}]}{1+\exp[\frac{\pi_h - E(\pi_u)}{\mu}]} \right) \right\} \quad (2) \\
E(\pi_l) &= \begin{cases} 0.6 \times 22.5 + 0.4 \times 2.5 & \text{for } \; i \in \textit{Traditional} \\ (1-a) \times 22.5 + a \times 2.5 & \text{for } \; i \in \textit{Bounded} \end{cases}
\end{aligned}
$$

where $y_{i,t} = 1(0)$ denotes that subject i underreports (reports honestly) in the tax-compliance game in period t. Table 2 reports the estimation results of various behavioral models.

Table 2: Comparison of behavioral models

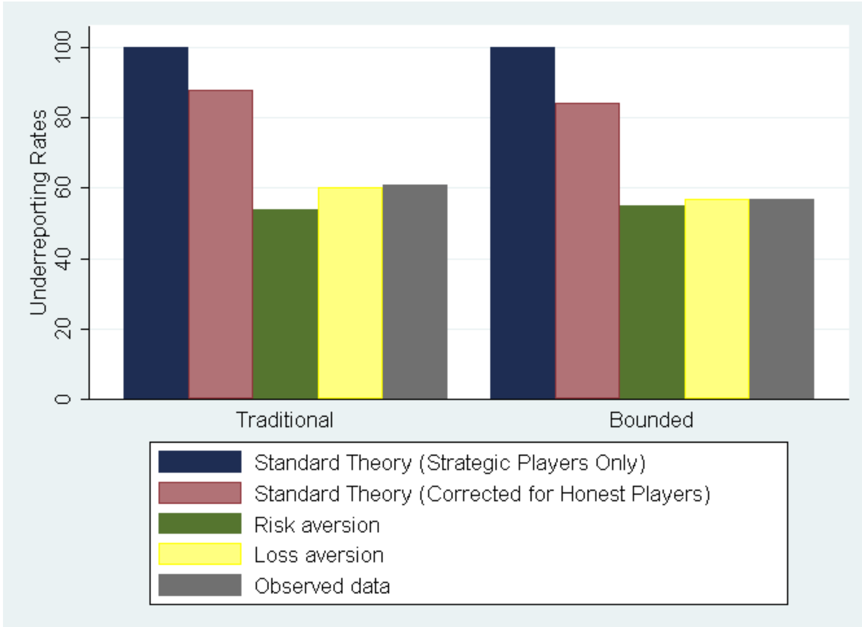
	Risk aversion		**Loss aversion**		**Loss aversion & Probability Weighting**	
	Traditional	*Bounded*	*Traditional*	*Bounded*	*Traditional*	*Bounded*
Risk magnitude r	0.366***	0.594***				
	(0.350)	(0.055)				
Gain domain curvature α			0.445***	0.428***	0.640***	0.533***
			(0.034)	(0.038)	(0.459)	(0.075)
Loss domain curvature β			0.548***	0.708***	0.586***	0.858***
			(0.052)	(0.030)	(0.068)	(0.073)
Loss-aversion coefficient λ			1.100***	1.148***	1.674***	1.283***
			(0.802)	(0.030)	(0.123)	(0.171)
Weighting parameter δ					1.150***	0.899***
					(0.193)	(0.120)
Perceived audit prob. a		0.336***		0.305***		0.240***
		(0.017)		(0.007)		(0.023)
Noise μ	0.667***	0.618***	0.266***	0.256***	0.430***	0.289***
	(0.067)	(0.098)	(0.019)	(0.424)	(0.042)	(0.040)
Log-likelihood	-1163.773	-1087.292	-1141.710	-1082.473	-1141.353	-1082.111
Observations	2331	2287	2331	2287	2331	2287

Notes: *10% significance; **5% significance, ***1% significance. The standard errors are clustered on subjects.

Upon first glance, all parameters in these models are significant, suggesting that the alternative behavioral models help to explain the compliance behavior in our study. For instance, the risk-aversion specification suggests that subjects are risk averse in both treatments, as the CRRA coefficient r is significantly larger than zero. It indicates that risk aversion helps to explain our data. The perceived audit probability for a risk-averse subject in the *Bounded* treatment is about 0.34. The explanation is straightforward: To induce a similar compliance pattern among subjects who are risk-averse, the audit capacity of the bounded rule can be set smaller, such that it induces the same deterrence effect compared to a traditional rule with audit probability $a = 0.336$. In other words, fewer resources are needed to achieve the same level of deterrence for risk-averse subjects for risk-neutral ones.

In the loss-aversion specification, subjects in both treatments exhibit loss aversion: The

Figure 2: Observed and predicted underreporting rates



coefficients of the loss-aversion parameter λ are larger than 1 in both treatments, which means that subjects are more sensitive to loss than to the equivalent magnitude of gain. The slopes of the value function indicate concavity in the gain domain (α) and convexity in the loss domain (β). Moreover, a Vuong test on non-nested models favors the loss-aversion model over the risk-aversion model ($p < 0.05$). If subjects are loss-averse, the bounded rule is even cheaper to implement, as the induced deterrence rate only needs to be the same as a traditional rule with audit probability $a = 0.306$.

The third specification combines loss-aversion utility and probability weighting. However, the likelihood of this specification does not improve significantly. Moreover, the probability-weighting parameter δ is not significantly different from 1 for both treatments ($p = 0.438$ and 0.397 respectively). This means that the average subjective probability of the subjects is pretty much in line with the objective audit probability. Overall, the results seem to indicate that the driving force for the observed compliance frequency is more likely to be found in the way they view losses and gains, rather than in how they assess probabilities.

Figure 2 displays the observed and predicted underreporting rates based on risk- and loss-aversion models. Since estimation results suggest that probability weighting does not explain the data well, parameters are taken from the second specification of loss aversion without probability weighting. Among the three models, the one using loss aversion fits our data the best. Result 3 summarizes the section.

Result 3 *The proportion of compliance behavior in both treatments is consistent with the presence of loss aversion together with some stochastic decision errors, although not in probability weighting.*

4 Coordination under the Bounded Rule

So far, the game introduced by the bounded rule is dominance solvable. In fact, it is not difficult to show that as long as the ex-ante probability of receiving high income (q) is lower than 0.5, the H-type players always underreport, given the dominant strategy of L-type. Essentially, the more L-type players in the population who honestly state their type with certainty, the easier it is for the H-type players to pretend to be L-type.

This subsection examines the bounded rule in another parameter domain where the game has multiple equilibria.[13] In this new bounded-rule treatment called *Bounded-Hq*, everything remains the same as the *Bounded* treatments, except that the ex-ante probability of receiving high-income q becomes 0.9 instead of 0.5. This parameter q determines an important property of the bounded rule: A high q resembles a rich neighborhood where each inhabitant is very likely to be wealthy. Hence, this new treatment explores the performance of the bounded rule when the income distribution is left-skewed. In particular, it examines whether the bounded rule loses its deterrence effect in the presence of multiple equilibria.

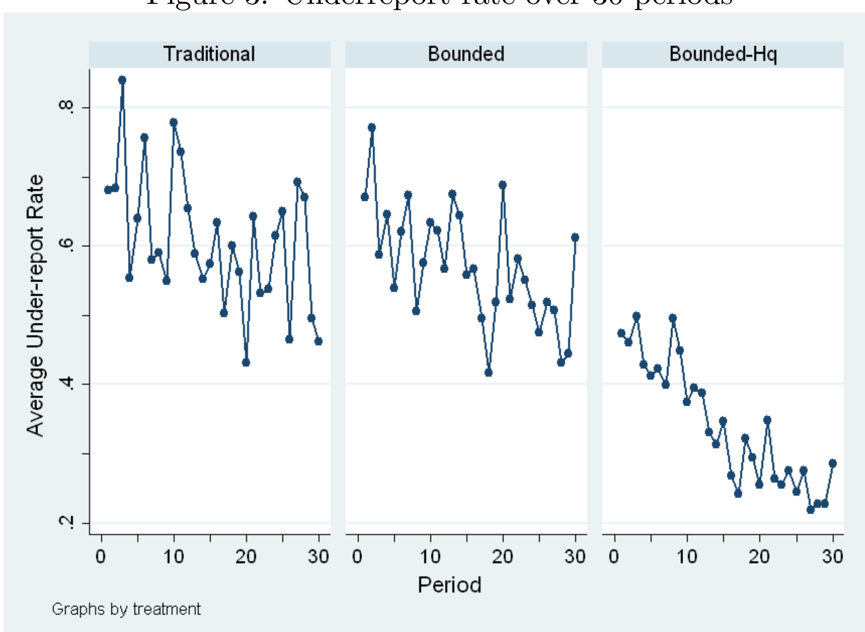
To derive predictions for this treatment, first assume that every taxpayer is a strategic, expected-profit maximizer. According to non-cooperative game theory, the introduction of the bounded rule with the same audit capacity changes the interaction of players into a coordination game with incomplete information. There are two pure-strategy Nash equilibria and one mixed-strategy Nash equilibrium in the game. In the pure-strategy equilibria, L-type players play their dominant strategy of reporting truthfully. All H-type players opt for underreporting (truth-reporting) if they believe other H-type players are going to cheat

[13] A formal illustration is in Appendix A.

with probability higher(lower) than 0.432. There is also a symmetric mixed-strategy Nash equilibrium, in which H-type players underreport with probability 0.432.[14]

This prediction can be modified when introducing intrinsically honest players. Specifically, let each taxpayer have a probability ρ of being an honest player. If ρ is sufficiently large, strategic players will find underreporting too risky to be worth the attempt. If that is the case, this modification could be considered as a refinement of the coordination game. However, if ρ is small, the payoff-dominant Nash equilibrium still exists, if a strategic player has a strong belief in the noncompliance behavior of the other strategic players. The proportion of honest players is about 15% in the other two treatments. Assuming strategic players correctly anticipate that $\rho = 0.15$, the threshold beliefs inducing underreporting behavior increases to 0.508. Nonetheless, the two pure-strategy equilibria remain the same.

Figure 3: Underreport rate over 30 periods



The overall underreport rate in the *Bounded-Hq* treatment is 33.95% over all 30 periods, and 26.16% in the last 10 periods. According to Figure 3, it is clear that the deterrence effect is the strongest, as the non-compliance frequency is significantly lower compared to

[14]For the proof, please refer to Appendix A. Note that there are other asymmetric equilibria in the game. However, we ignore them in a symmetric setting, since these equilibria require unrealistic coordination among symmetric players.

the other two treatments (a two-sided Mann-Whitney ranksum test with $p < 0.05$). This difference is already salient in the first period, and remains highly significant throughout the game. Note that the drastic change in behavior compared to the *Bounded* treatment cannot be explained by subjects' risk attitude: The estimated CRRA coefficient r, based only on risk elicitation data, does not indicate any significant differences between treatments.

Regarding auditing statistics, despite the fact that the total number of audits is smaller in this treatment (even compared to the *Bounded* treatment ($p < 0.05$)), the audit frequency turns out to be significantly higher ($p < 0.05$), due to the fact that fewer "low-income" reports need to be audited. The audit success rate is remarkably higher as well ($p < 0.05$), leading to the dishonest H-type players receiving a significantly lower payoff than under the *Traditional* treatment ($p < 0.05$). The average budget-usage ratio is 95.63%, which is again significantly higher than that under the traditional rule (32.03%).

Result 4 *The non-compliance rate in the Bounded-Hq treatment is significantly lower than it is in both the Traditional and the Bounded treatments. This high deterrence rate is achieved with significantly lower implementation costs, and a higher budget-usage ratio.*

The previous subsection shows that behavior in both treatments is consistent with a loss-aversion model with stochastic decision errors. To examine how they explain the pattern in this treatment, we perform the following exercise. We first estimate the perceived audit probability of the *Bounded-Hq* treatment given the α, β and λ parameters obtained in the *Bounded* treatment. Then we use all of the information to calculate the predicted underreporting rate of the treatment $\widehat{b}^{SE}$. It turns out that $\widehat{b}^{SE}$ is 33.80%, which is again very close to the actual prediction 33.95%. This is an indication that behavior in the *Bounded-Hq* treatment is again consistent with the loss-aversion model with decision errors.

Interestingly, the perceived probability in the *Bounded-Hq* treatment, 0.344, is only mildly larger than that in the *Bounded* treatment, 0.305. That means a five percentage increase in audit probability perception leads to a 23 percent increase in compliance level. This asymmetry stems from the fact that subjects value gains and losses differently with respect to the reference point. When the perceived audit probability is 0.305 in the *Bounded* treatment, the value of underreporting is in the gain domain, and is marginally larger than 0, the value of the reference point. When this probability increases to 0.344, however, the value of underreporting falls into the loss domain with a larger distance to the reference point. Since subjects are loss-averse, they weight losses more than gains, which lowers underreporting frequency more drastically.

Loss-averse players face a higher degree of strategic uncertainty and hence are more likely to fail the coordination. If taxpayers are expected-profit maximizers, they will underreport as long as they think the probability that others are going to underreport is larger than 0.432. It would be much harder, on the other hand, for loss-averse players to choose to underreport. Given that they are more sensitive to losses than they are to gains, they will choose to underreport only when they think the other H-types are going to underreport with probability higher than 0.774. With 15% honest players, this probability even goes up to 0.911. This threshold requires more coordination among taxpayers, and involves a substantially higher degree of strategic uncertainty.

According to Brandenburger (1996)'s definition, strategic uncertainty arises when there is "uncertainty concerning the purposeful behavior of players in an interactive decision situation", as opposed to a game against nature. Strategic uncertainty is widely documented in many experimental studies such as coordination games (e.g. Huyck et al. (1990), Huyck et al. (1991)), market entry games (e.g. Sundali et al. (1995); Erev and Rapoport (1998)) and bank runs (e.g. Garratt and Keister (2009); Schotter and Yorulmazer (2009)). Recently, Heinemann et al. (2009) propose a method to measure strategic uncertainty by eliciting certainty equivalents analogous to measuring risk attitudes in lotteries. In their experiment, N subjects have to choose simultaneously between a series of lottery pairs. In each pair, lottery A always yields a sure fixed payoff, while lottery B yields a payoff if the minimum number of players selected is k. They find that the number of B-choices in coordination games decreases with an increasing coordination requirement k. Holding k constant, N has a strong positive effect on coordination, since a large N reduces the relative hurdle to coordination. These behavioral patterns indicate that subjects are strategic uncertainty averse. Applying the study by Heinemann et al. (2009) reveals that the risk-dominant equilibrium is more likely to be chosen by the loss-averse subjects than by expected-profit maximizers.

5 Learning and Social demographics

In the post-experiment questionnaire, subjects are asked to provide their social background information such as gender and nationality. This information allows us to study how subjects form and adjust their underreporting decisions under different rules. The first specification concerns compliance behavior. We use the following random-effect probit model specification:

$$y_{it} = \gamma x_{it} + u_i + \varepsilon_{it} \ .$$

The variable y equals 1 if subjects decide to underreport, and is equal to 0 otherwise. Furthermore, x is a vector of explanatory variables, the u_i represent individual random effects and γ is a vector of parameters. The explanatory variables include subjects' social backgrounds such as gender, nationality and experience of economics. They also contain a history of play such as underreporting performance in the previous period, period number and its square term.

Apart from compliance behavior, this subsection also investigates how individual characteristics and previous performance influence players' perceived audit probability of the bounded rules. The loss-aversion model without probability weighting is estimated, allowing the perceived a parameter to depend on the social background information vector η: $a = \eta x + \varepsilon_{it}$. The results of the two specifications appear in Table 3.

Table 3: The influences of social background and learning on compliance behavior and audit probability perception

	Compliance behavior			Audit Probability Perception	
	Traditional	*Bounded*	*Bounded-Hq*	*Bounded*	*Bounded-Hq*
Underreport Detection Experience	-0.102	-0.498***	0.010	0.075**	0.030***
	(0.186)	(0.195)	(0.105)	(0.038)	(0.006)
Period	-0.039	-0.040	-0.082***	0.008	0.002***
	(0.026)	(0.032)	(0.023)	(0.007)	(0.006)
Period^2	0.0006	0.0004	0.001	-0.0001	-0.0002
	(0.0008)	(0.0009)	(0.0006)	(0.0002)	(0.0002)
Gender (1 for men)	0.756*	0.960*	0.459	-0.020	0.003
	(0.413)	(0.544)	(0.478)	(0.017)	(0.011)
Years of learning economics	0.052	0.797***	0.278	-0.015	-0.001
	(0.186)	(0.287)	(0.214)	(0.014)	(0.005)
Econ experience × Game theory	0.083	-0.574**	-0.066	0.012	-0.0009
	(0.144)	(0.274)	(0.187)	(0.013)	(0.004)
Dummy for Eastern Europeans	0.015	-0.409	0.125	0.043**	-0.020
	(0.899)	(0.922)	(0.925)	(0.022)	(0.201)
Dummy for Dutch	-0.102	1.190	0.620	0.004	-0.025
	(0.691)	(0.777)	(0.677)	(0.019)	(0.018)
Dummy for Chinese	-0.234	1.336	-0.009	-0.003	-0.014
	(0.652)	(0.798)	(0.778)	(0.019)	(0.023)
Dummy for other Asian	-0.878	0.159	0.940	0.022	-0.026
	(0.784)	(0.971)	(0.968)	(0.025)	(0.023)
Mach-IV Scale	0.025	-0.019	0.018	0.003	-0.0003
	(0.016)	(0.018)	(0.015)	(0.003)	(0.004)
Tax Filing Experience	-0.005	-0.304	-0.339	0.011	0.001
	(0.458)	(0.492)	(0.575)	(0.012)	(0.011)
Constant	-0.508	5.318**	0.918	0.209***	0.305***
	(2.091)	(2.308)	(1.800)	(0.042)	(0.038)
Log-likelihood	-444.868	-359.410	-640.430	-997.726	-1437.232
Observations	957	912	1670	2287	3014

Notes: *10% significance; **5% significance, ***1% significance. We only include observations which players receive high income. Standard errors are clustered on individuals.

Various effects are found in the regressions. The first one concerns the effect of learning. In the *Bounded* treatment, detection experience in the previous round decreases non-compliance propensity. Interestingly, players with a background in economics are more likely to underreport, which seems to suggest that training in economics results in behavior more in line with homo-economicus. These effects, however, do not exist in the other two treatments. In both the *Traditional* and *Bounded* treatments, men are more likely to underreport than women are. However, this is not the case in the *Bounded-Hq* treatment. In fact, no other social demographic information affects behavior except for time. When perceived audit probability depends on social demographic information under the bounded rule treatments, the only variable that is significant is the lag audited experience. That is, if a cheater was caught in the previous period, the perceived audit probability increases. This pattern is largely in line with reinforcement learning. The invariant influences for social background information might indicate that the effect of bounded rules is robust. Subjects with a different social background seem to react to the bounded rule in a similar manner, especially in the *Bounded-Hq* treatment.

6 Conclusion

This paper examines the bounded rule as a new method in constructing an audit sample more cost-effectively. It audits all "red-flagged" reports whenever the total number of these reports is no more than the maximum number of audits allowed by the budget, and merely the maximum otherwise. Compared to a traditional rule which audits files with a constant probability, the bounded rule can induce the same deterrence effect with a lower budget.

The paper then tests the model's predictions in a controlled laboratory experiment. Subjects receive either high- or low income with a predetermined probability. On knowing a certain auditing rule (traditional or bounded), they report income to the tax agency. Individual profit maximization and non-cooperative game theory suggest that the two rules have the same deterrence effect. The experimental results indicate that the compliance rate in the bounded rule is the same as that in the traditional rule. Given the same compliance level induced, the bounded rule is more cost-effective in terms of both the implementation costs (i.e., the average number of audits conducted) and the budget-usage ratio (i.e., the percentage of the budget used in actual audits). The deterrence effect of a bounded rule is even stronger in another parameter domain where it introduces multiple equilibria. It deters subjects from coordinating on the payoff-dominant equilibrium with a low implementation cost. These results strongly favor the bounded rule to be a superior auditing rule.

In all treatments, the compliance rates are higher than the prediction, even taking into account the fraction of intrinsically honest players. About 60 percent of the subjects switch their decisions alternatively. To explain behavioral anomalies, several choice models under uncertainty are estimated and compared. Among these specifications, loss aversion combined with stochastic errors are more successful at tracking observed data patterns. History of play also affects their perception toward audit probability under the bounded rule. Incorporating this evidence would better help tax administrations to adjust their policies to encourage people to pay their taxes in a more cost-effective way.

This study is just a first step into the investigation of the bounded rule. In our current setup, taxpayers only can decide whether to underreport or honestly report. In future studies, the model could be extended to allow choices on the extent of underreporting. Another possible extension might involve introducing a human auditor to further examine the strategic interactions. This would be useful, due to the fact that taxpayers can communicate with each other in reality. Alm and McKee (2004) show that such cheap-talk communication could help taxpayers to coordinate on zero-compliance (payoff-dominant) equilibrium. However, if a strategic auditor could observe this, s/he would be able to adjust the audit capacity accordingly to combat collusion among taxpayers.

References

Alm, J., M. B. Cronshaw, and M. McKee (1993). Tax compliance with endogenous audit selection rules. *Kyklos 46*(1), 27–45.

Alm, J., B. R. Jackson, and M. McKee (2009). Getting the word out: Enforcement information dissemination and compliance behavior. *Journal of Public Economics 93*(3-4), 392–402.

Alm, J. and M. McKee (1998). Extending the lessons of laboratory experiments on tax compliance to managerial and decision economics. *Managerial and Decision Economics 19*(4-5), 259–275.

Alm, J. and M. McKee (2004). Tax compliance as a coordination game. *Journal of Economic Behavior & Organization 54*(3), 297–312.

Bayer, R. and F. Cowell (2009). Tax compliance and firms' strategic interdependence. *Journal of Public Economics 93*(11-12), 1131–1143.

Blackwell, C. (2007). A meta-analysis of tax compliance experiments. *International Studies Program Working Paper 07-24*.

Boylan, S. J. and G. B. Sprinkle (2001). Experimental evidence on the relation between tax rates and compliance: The effect of earned vs. endowed income. *Journal of the American Taxation Association 23*(1), 75–90.

Brandenburger, A. (1996). Strategic and structural uncertainty in games. In R. J. Zeckhauser, R. L. Keeney, and J. K. Sebenius (Eds.), *Wise Choices: Decisions, Games, and Negotiations*, pp. 221–232. Boston: Harvard Business Press.

Cason, T. N. and L. Gangadharan (2006). An experimental study of compliance and leverage in auditing and regulatory enforcement. *Economic Inquiry 44*(2), 352–366.

Christie, R. and F. Geis (1970). *Studies in Machiavellianism.* Academic Press, New York.

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

Dhami, S. and A. Al-Nowaihi (2010). Optimal taxation in the presence of tax evasion: Expected utility versus prospect theory. *Journal of Economic Behavior & Organization 75*(2), 313–337.

Elffers, H. and D. J. Hessing (1997). Influencing the prospects of tax evasion. *Journal of Economic Psychology 18*(2-3), 289–304.

Erev, I. and A. Rapoport (1998). Coordination, magic, and reinforcement learning in a market entry game. *Games and Economic Behavior 23*(2), 146–175.

Fischbacher, U. (2007). Z-tree: Zurich toolbox for readymade economic experiments. *Experimental Economics 10(2)*, 171–178.

Fischbacher, U. and U. Stefani (2007). Strategic errors and audit quality: An experimental investigation. *Accounting Review 82*(3), 679.

Garratt, R. and T. Keister (2009). Bank runs as coordination failures: An experimental study. *Journal of Economic Behavior & Organization 71*(2), 300 – 317.

Graetz, M., J. Reinganum, and L. Wilde (1986). The tax compliance game: Toward an interactive theory of tax enforcement. *Journal of Law, Economics, and Organization 2*, 1–32.

Harrington, W. (1988). Enforcement leverage when penalties are restricted. *Journal of Public Economics 37*(1), 29–53.

Heinemann, F., R. Nagel, and P. Ockenfels (2009). Measuring strategic uncertainty in coordination games. *Review of Economic Studies 76*(1), 181–221.

Holt, C. A. and S. K. Laury (2002). Risk aversion and incentive effects. *American Economic Review 92*(5), 1644–1655.

Huyck, J. B. V., R. C. Battalio, and R. O. Beil (1990). Tacit coordination games, strategic uncertainty, and coordination failure. *American Economic Review 80*(1), 234–248.

Huyck, J. B. V., R. C. Battalio, and R. O. Beil (1991). Strategic uncertainty, equilibrium selection, and coordination failure in average opinion games. *Quarterly Journal of Economics 106*(3), 885–910.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica 47*(2), 263–91.

Kim, C. K., J. H. Evans, and D. V. Moser (2005). Economic and equity effects on tax reporting decisions. *Accounting, Organizations and Society 30*(7-8), 609–625.

Kim, C. K. and W. S. Waller (2005). A behavioral accounting study of strategic interaction in a tax compliance game. *Experimental Business Research*, 35–56.

King, S. and S. Sheffrin (2002). Tax evasion and equity theory: An investigative approach. *International Tax and Public Finance 9*(4), 505–521.

Kleven, H. J., M. Knudsen, C. Kreiner, S. Pedersen, and E. Saez (2010). Unwilling or unable to cheat? evidence from a tax audit experiment in denmark. *Econometrica*, forthcoming.

Loomes, G. (2005). Modelling the stochastic component of behaviour in experiments: Some issues for the interpretation of data. *Experimental Economics 8*(4), 301.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. *Frontiers of Econometrics New York: Academic press.*

McFadden, D. (2001). Economic choices. *American Economic Review 91*(3), 351–378.

McKelvey, R. D. and T. R. Palfrey (1995). Quantal response equilibria for normal form games. *Games and Economic Behavior 10*(1), 6–38.

Moser, D. V., J. H. Evans, and C. K. Kim (1995). The effects of horizontal and exchange inequity on tax reporting decisions. *Accounting Review 70*(4), 619–634.

Reinganum, J. F. and L. L. Wilde (1985). Income tax compliance in a principal-agent framework. *Journal of Public Economics 26*(1), 1–18.

Rieskamp, J. (2008). The probabilistic nature of preferential choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 1446–1465.

Schotter, A. and T. Yorulmazer (2009). On the dynamics and severity of bank runs: An experimental study. *Journal of Financial Intermediation 18*(2), 217–241.

Slemrod, J. and S. Yitzhaki (2002). Tax avoidance, evasion, and administration. In A. J. Auerbach and M. Feldstein (Eds.), *Handbook of Public Economics*, Volume 3, pp. 1423–1470. Elsevier.

Sundali, J. A., A. Rapoport, and D. A. Seale (1995). Coordination in market entry games with symmetric players. *Organizational Behavior and Human Decision Processes 64*(2), 203 – 218.

Torgler, B. (2002). Speaking to theorists and searching for facts: Tax morale and tax compliance in experiments. *Journal of Economic Surveys 16*(5), 657–83.

Tversky, A. and D. Kahneman (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty 5*(4), 297–323.

US Department of the Treasury, . (2005). Fy 2006 budget in brief. Washington, DC.

US Department of the Treasury, . (2006). Fy 2007 budget in brief. Washington, DC.

US Government Accountability Office, . (2009). Review of the fiscal year 2010 budget request. Washington, DC.

Wilcox, N. T. (2010). Stochastically more risk averse: A contextual theory of stochastic discrete choice under risk. *Journal of Econometrics*, forthcoming.

Yaniv, G. (1999). Tax compliance and advance tax payments: A prospect theory analysis. *National Tax Journal 52*(4), 753–764.

Yim, A. (2009). Efficient committed budget for implementing target audit probability for many inspectees. *Management Science 55*(12), 2000–2018.

Zimbelman, M. F. and W. S. Waller (1999). An experimental investigation of Auditor-Auditee interaction under ambiguity. *Journal of Accounting Research 37*, 135–155.

Appendix

A Proofs

A.1 Proof of Proposition 1

Let $\bar{K} = [\bar{a}N]$, as K needs to be an integer. Thus, $a_{BD} = \min\{1, \overline{K}/L\} = \min\{1, [\bar{a}N]/L\}$. Since $L \leqslant N$, $a_{BD} \geqslant \bar{a}$. That means, in the scenario where all players declare low income, the audit probability a_{BD} is equal to $\bar{a}$. The H-type players are indifferent between the decisions of underreporting and reporting honestly. If $K > \bar{K}$, that means the lowest probability of being audited is strictly larger than $\bar{a}$. Hence, any $K > \bar{K}$ is sufficient to support full compliance.

The simplest case to induce zero compliance is to set $K = 0$. Because of zero audit, self-regarding, profit-maximizing H-type players always report low income, regardless of their beliefs towards other H-types. More generally, if $K < [\bar{a}]$, the bounded rule cannot induce any compliance for strategic players regardless of the income distribution. In other words, in the worst-case scenario in which only one H-type player claims low income, the audit probability he or she faces is lower than $[\bar{a}]$. Hence, strategic H-type players will underreport.

A.2 Proof of Proposition 2

This subsection contains two parts. The first part proves that given all players are rational, strategic expected profit maximizers, the game introduced by the bounded rule is dominance solvable. The second part shows that this claim still holds by introducing intrinsically honest players.

The proof is trivial that reporting high income is a dominated strategy for the L-type players. To prove that the best response of H-type players is underreporting given that L-type players comply dominance, the expected payoff from underreporting should be strictly larger than the sure payoff from reporting truthfully. Moreover, this holds regardless of the beliefs that H-type players hold towards the other H-types.

First assume that a H-type player anticipates that no body other than him or her will underreport. That is, $\overline{b_0} = (b_1, b_2, ..., b_{N-1}) = (0, 0., , , 0)$. In this situation, "low-income" reports are submitted by L-type. Since the probability of being a L-type is $q = 0.5$ for every other player, the probability that exactly n out of $N-1$ players submit "low-income" reports follows the binomial distribution $\mathbf{Bin}\ (n, N-1; q) = \mathbf{Bin}\ (n, 7; 0.5)$. The expected payoff from underreporting is therefore:

$$\begin{aligned} E(\pi_l|\overline{b_0}) &= \sum_{n=0}^{N-1} \mathbf{Bin}(n; N-1, q) \times \{\min(\frac{2}{n+1}, 1) \times \pi_F + [1 - \min(\frac{2}{n+1}, 1)] \times \pi_S\} \\ &= \pi_S - (\pi_S - \pi_F) \times \sum_{n=0}^{N-1} \mathbf{Bin}(n; N-1, B_i) \times \min(\frac{2}{n+1}, 1) \\ &= 22.5 - 20 \times \sum_{n=0}^{7} \mathbf{Bin}(n; 7, 0.5) \times \min(\frac{2}{n+1}, 1) \\ &= 12.698 \end{aligned}$$

The sure payoff of reporting truthfully is 12.5. Hence, a self-interest, risk neutral H-type player will underreport.

The remaining proof shows that for any given set of beliefs held by a H-type player, the expected payoff from underreporting is always not less than $E(\pi_l|\overline{b_0})$. Assume that player N thinks the first $N-1$ players underreport with probability $\overline{b} = (b_1, b_2, ..., b_{N-1})$. The probability that player i submit "low-income" is $B_i = 1 - q + qb_i = \frac{1}{2}(1 + b_i)$. Note that $B_i \in [\frac{1}{2}, 1]$. To facilitate notation, define an index vector $\mathbf{I} = (i_1, i_2, ...i_7)$, with $i_1 \neq i_2 \neq ...i_7$. Each index takes a value from the set $\{1, 2, ..., 7\}$. The probability that n out of 7 other players submit "low-income" reports is:

$$\Pr(n|\overline{b}) = \sum_{s=1}^{C_7^s} \prod_{j=1}^{s} B_{i_j} \prod_{k=s+1}^{i_7} (1 - B_{i_k})$$

The expected payoff from underreporting is therefore:

$$E(\pi_l|\overline{b}) = \sum_{n=0}^{N-1} \Pr(n|\overline{b}) \times \{\min(\frac{2}{n+1}, 1) \times \pi_F + [1 - \min(\frac{2}{n+1}, 1)] \times \pi_S\}$$

It turns out that for any given b_i, $\partial E(\pi_l)/\partial b_i = (\partial E(\pi_l)/\partial B_i) \cdot (\partial B_i/\partial b_i) > 0$.[15] This

[15] Calculation is available upon request.

means that the expected payoff from underreporting is increasing in the (subjective) propensity to evade taxes. Hence, given any set of belief $\overline{b} = (b_1, b_2, ..., b_{N-1})$, $E(\pi_l|b) \geq E(\pi_l|\overline{b_0})$. Hence, the best response of the H-type players is to underreport. ∎

The second part of this subsection proves that the introduction of intrinsically honest players does not change the directions of treatment difference. Let ρ be the probability that a player is intrinsically honest, and $1-\rho$ be the probability that a player is a strategic, self-regarding profit maximizer, where $0 \leq \rho < 1$. We do not allow $\rho = 1$, since at least one strategic player is thinking of this problem. In our setting, in particular, the number of honest players ρN can be any number from 0 to 7 out of 8 players. We further assume that the ρ is the same in both treatments.

To prove the statement, we only need to show that the inclusion of honest players does not affect the strategy of the profit maximizers. When the strategic players are assigned to be L-types, they gain a higher payoff by reporting truthfully, regardless of the auditing rule implemented. In the *Traditional* treatment, H-type profit maximizers only compare a sure payoff of reporting truthfully and the expected payoff from the tax evasion gamble if they underreport. Hence, the existence of honest players will not affect their choices. In the *Bounded* treatment, the subjective beliefs of strategic, H-type players of the number of "low-income" reports now become: $B_i = (1-q) + q(1-\rho)b$. Given that $q = 0.5$, $0 \leq \rho < 1$, B still lies in the interval $[\frac{1}{2}, 1]$. Therefore, Proposition 2 still holds.

In the presence of honest players, the non-compliance rate of both treatments becomes:

$$\sum \mathbf{Bin}(n; N, q)(1-\rho) = (1-\rho) \ .$$

A.3 The Existence of Coordination

If this game is a coordination game, there exists an $b \in [0,1]$ such that the payoff from underreporting is equal to the honest payoff:

$$
\begin{aligned}
E(\pi_u; N, q, K, b_i) &= \sum_{n=0}^{N-1} \mathbf{Bin}(n, N-1; B_i)\,[(1-a_{BD}) \times (I_H - T_L) + a_{BD} \times (I_H - T_H - F)] \\
&= I_H - T_H \quad .
\end{aligned}
$$

Due to the discrete nature of the distribution, a direct proof is difficult. However, just for illustration purposes, if N is large, the expected number of "low-income" reports is $B_i N = [(1-q) + qb_i]N$. The expected profit from underreporting could be simplified as

$$
\begin{aligned}
E(\pi_u) &= \frac{K}{B_i N}(I_H - T_H - F) + (1 - \frac{K}{B_i N})(I_H - T_L) \\
&= I_H - T_H \quad .
\end{aligned}
$$

Solving the equation yields $B_i = K(T_H + F - T_L)/N(T_H - T_L)$. Hence, there exists a set of parameters K,T_H,F,T_L, N and q such that $B_i \in (0, 1)$. Thus, in certain parameter domains, the H-type players under the bounded rule find themselves indifferent between underreporting and honestly-reporting if $b_i = \bar{b} = \frac{B_i-(1-q)}{q}$. If $b_i > \bar{b}$, then the H-types all underreport; if $b_i < \bar{b}$, then the H-types all report honestly.

A.4 Equilibrium Analysis for *Bounded-Hq* Treatment

Let $\sigma_i(j)$ be the probability that type i player (H-type or L-type) will use strategy j (u or h). There are two pure Nash equilibria and one mixed-strategy equilibrium in this treatment:

$$
\{(\sigma_H(u) = 1, \sigma_L(h) = 1), (\sigma_H(h) = 1, \sigma_L(h) = 1), (\sigma_H(u) = 0.432, \sigma_L(h) = 1)\}.
$$

In words, the two pure Nash equilibria are 1) all H-type players underreport and 2) all H-type players honestly report. L-type players always honestly report.

Let us examine the former case. Given that a H-type player thinks that all other H-types choose strategy u, s/he will have an expected payoff of 17.5 by playing strategy l.

By deviating to h, the payoff decreases to 12.5. Since we assume symmetry among players, no one has the incentive to deviate from underreporting, which constitutes a NE. A highly similar analysis applies to the latter case. Given that all other H-type players play strategy h, a strategy deviation from h to l will yield a lower expected payoff for H-type players (from 12.5 to 3.59). Hence, no one has an incentive to deviate.

On top of the two pure equilibria, the game has also a mixed-strategy equilibrium in which each H-type player is indifferent between the strategy of honest-reporting and underreporting. Given the game parameters, the underreporting probability b that induces utility indifference is $b^*_{SE} = 0.432$. ∎

Note that the game has other asymmetric equilibria. However, we ignore them in a symmetric setting, since these equilibria require unrealistic coordination among symmetric players.

B Instructions

B.1 Instructions Comparison

The instructions given in the next subsection are for the *Bounded* treatment. These instructions differ from those given for the other treatments as follows:

- *Traditional* treatment

  1. The second bullet (concerning matching protocol) of the list under "Task Description" in the instructions for the "Tax Compliance Game" is absent.
  2. The "Audit Probability Table" is absent.
  3. The phrase "see audit prob. table" in the "Payoff Table" becomes 0.4.

- *Bounded-Hq* treatment

  1. In the third bullet of the list under "Task Description" in the instructions for the "Tax Compliance Game", the probability of receiving €25 becomes 0.9, and accordingly the probability of receiving 10 becomes 0.1.
  2. In the "Payoff Table" (immediately before "Payment Method" in the instructions for the "Tax Compliance Game"), the probabilities in the second column become 0.9 and 0.1, respectively.

B.2 Instructions for *Bounded* Treatment

- Please read these instructions carefully!
- Please do not talk to your neighbours and remain quiet during the entire experiment.
- If you have a question, please raise your hand. We will come to you to answer it.
- You will receive a show-up fee of €3 for completing all tasks in the experiment, independent of your performance.

Task Description

- This session consists of 30 periods of play; each period is completely independent of the others.

- Of the participants in the room, two groups of 8 participants will be randomly formed at the beginning of each period. You will not know the identity of the other players in your group in any period.

- At the beginning of each period, you will receive a taxable income of either €25 or €10. The probability of receiving €25 is 0.5; the probability of receiving €10 is 0.5.

- Your task is to report your income to the auditor, which is played by a computer. The amount that you report is your decision. You can report either €25 or €10, regardless of your received income.

After-tax Income Determination

Your after-tax income in this period is determined by the following two steps: tax payment and an audit.

Step One: Tax payment

The tax rate is 50% for those who reported €25 and 25% for those who reported €10. Suppose the income you received is €25:

- If you report €25 to the auditor, the auditor will charge €12.5 (50% of €25) as tax. So your after-tax income in this period equals to €25 – €12.5 = €12.5.

- If you report €10 to the auditor, the auditor will charge €2.5 (25% of €10) as tax. So your after-tax income in this period equals to €25 – €2.5 = €22.5.

Suppose the income you received is €10:

- If you report €10 to the auditor, the auditor will charge €2.5 (25% of €10) as tax. So your after-tax income in this period equals to €10 – €2.5 = €7.5.

- If you report €25 to the auditor, the auditor will charge €12.5 (50% of €25) as tax. So your after-tax income in this period equals to €10 – €12.5 = -€2.5.

- In sum, the auditor charges tax based on your reported income, instead of your received income.

Step Two: Audit

The auditor does not know your received income unless your report is audited later .

Auditing procedure:

- If your reported income is €25, it will not be audited. That means what you have earned in step one (€12.5 or -€2.5) will be your after-tax income (if your received income is €25 and €10, respectively).

- Regardless of your received income, if your reported income is €10, there is a chance that your report will be audited. The outcome is as follows:

  – Suppose your reported income is €10 AND your received income is also €10. Then what you have earned in step one (€7.5) will be your after-tax income, no matter whether your report is audited or not.

  – Suppose your reported income is €10 AND your received income is €25. If your report is not audited, you will keep the €22.5 earned in step one; if audited, you will get €2.5.

Auditing probability:

The number of reports the auditor will audit depends on the number of players reporting an income of €10 in a group.

- If the number of €10 income reports is equal to two or less, the auditor will audit all of the €10 reports.

- If the number of €10 income reports is three or more, then two out of such reports will be randomly selected for audit.

- The "Audit Probability Table" below shows the audit probabilities for a player who reported an income of €10.

Audit Probability Table

Number of €10 reports	0	1	2	3	4	5	6	7	8
Audit Probability	100%	100%	100%	66.7%	50%	40%	33.3%	28.6%	25%

- The "Payoff Table" below summarizes all of the possible scenarios you may encounter in one period and the related payoffs:

Payoff Table

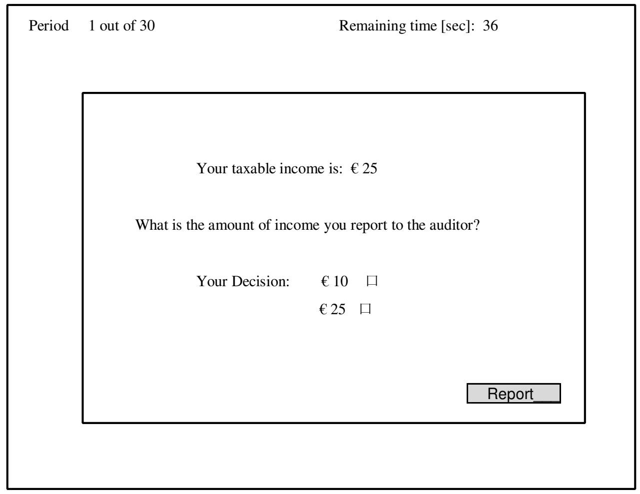
Received Income	**Probability**	**Reported Income**	**Audit Probability**	**After-tax Income if audited**	**After-tax Income if NOT audited**
€25	0.5	€25	0	€12.5	€12.5
		€10	see audit prob. table	€2.5	€22.5
€10	0.5	€10	see audit prob. table	€7.5	€7.5
		€25	0	−€2.5	−€2.5

Payment Method

- At the end of this experiment, one out of 30 periods will be selected to determine your payoff for this task. The computer program will generate a random number from 1 to 30. This number will determine one of the 30 periods. Your performance in that period determines your payoff.
- You will be paid based on your after-tax income for the randomly selected period.
- Because each period is equally likely to be selected for payment determination, you should make your decision in each period as if that period would be selected for payment.
- Your payoff will be paid out in cash at the end of the experiment along with your earnings in the other task(s).

We will now show you what the computer screens look like.

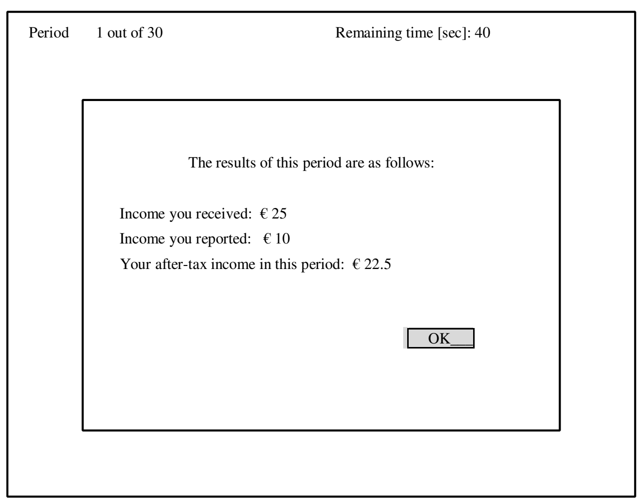
SCREEN 1



In “Screen 1”, you can decide the amount of income to report to the auditor. Please select either “€10” or “€25”, and confirm your choice by pressing the “Report” button.

Warning: Before pressing the button, make sure your choice is correct. You cannot change your decision after you have pressed OK.

SCREEN 2



“Screen 2” is the feedback table you will receive regarding your after-tax income. Your will find information on the initial taxable income you received, the income you reported and your after-tax income in this period.

Click on OK when you finish checking the information.

Note that the purpose of the screen shots is to clarify the procedure, rather than provide advice about how to act. You should make the decisions that are best for you.

B.2.1 Risk Elicitation Task[16]

Task Description

In this task, you are asked to make decisions related to 21 choice pairs. In each choice pair, you need to select between two lotteries labeled "Lottery A" and "Lottery B". Please, take your time and read each choice pair carefully. An example of a typical choice pair is given below:

Choice No.1	Lottery A	€5.5 with probability 0.5 or €3.5 with probability 0.5	Your choice:	Lottery A	☐
	Lottery B	€9 with probability 0.5 or €0.5 with probability 0.5		Lottery B	☐

Payment Method

- You need to make choices for all 21 choice pairs. However, only one of the 21 choices you have made will be chosen for the payoff determination of this task. First, the computer program will generate a random number from 1 to 21. This number will determine a choice pair. Then, the computer program will simulate the lottery you have chosen and reveal the outcome on your screen. The outcome of this lottery will determine your payoff.

- For example, suppose that the computer program has generated a random number 2. It will then check what you have selected in choice pair number 2. Suppose that you have chosen Lottery A in that choice pair. Then the computer program will simulate Lottery A and reveal your payoff (either €5.5 or €3.5). Your payoff will be paid out in cash at the end of the experiment along with your earnings for the other task.

It is important that you fully understand the lottery selection task. Please raise your hand if you have any questions at this moment.

[16] The risk elicitation task is conducted after the tax-compliance game. However, the subjects do not know the existence of this task when they were playing the tax-compliance game.

B.2.2 Post-experimental Questions

Questions on Treatment Manipulation

Please evaluate the following statements with respect to the tax reporting task:[17]

1=strongly disagree, 2=somewhat disagree, 3=slightly disagree, 4=no opinion, 5=slightly agree, 6=somewhat agree, 7=strongly agree

1. The instructions were clearly formulated.
2. I felt that I performed well on the task.
3. I received plenty of time to carry out the task.
4. I was motivated to do well on the task.
5. The task was fun to perform, motivating me to achieve a payoff as high as possible.
6. I considered the tax reporting task as fairly complex.
7. My payoff is determined not only by my own decision, but also by the decisions of the other players.
8. When making my decision, I thought about what other players might do.
9. I feel obliged to report the received income in each period.
10. The chance I have received €25 is about 50%.[18]

Questions on Background Information

Please answer the following survey questions. Your answers will be used for this study only. Individual data will not be exposed.

1. What is your gender?

[17] The first five questions are used to understand the subjects' perception about the experimental setup and instructions in general. We do not expect to find differences across treatments. The last five questions focus on capturing different types of manipulations of the treatments; therefore, we expect to see differences across manipulations.

[18] In the *Bounded-Hq* treatment, the chance should be 90%, instead of 50%.

2. What is your nationality?

3. How many years have you already studied in economics?

4. Have you ever had a course related to game theory?

5. Have you ever had a part-time job?

Questions on Mach IV Scale[19]

In the following you will find a list of statements. Please read them carefully and answer them to what extent you agree or disagree. Even if in some cases you would like to say that your answers depend on the circumstances, you should only choose one of the answers. Since all responses are anonymous you can answer freely. There is nobody on whom you need to make a good impression. Only if you answer very honestly can the results be used.

1=strongly disagree, 2=somewhat disagree, 3=slightly disagree, 4=no opinion, 5=slightly agree, 6=somewhat agree, 7=strongly disagree

1. Never tell anyone the real reason you did something unless it is useful to do so.

2. The best way to handle people is to tell them what they want to hear.

3. One should take action only when sure it is morally right.

4. Most people are basically good and kind.

5. It is safest to assume that all people have a vicious streak and it will come out when they are given a chance.

6. Honesty is the best policy in all cases.

7. There is no excuse for lying to someone else.

8. Generally speaking, people won't work hard unless they're forced to do so.

9. All in all, it is better to be humble and honest than to be important and dishonest.

[19] Questions 3, 4, 6, 7, 9, 10, 11, 14, 16 and 17 are reverse coded.

10. When you ask someone to do something for you, it is best to give the real reasons for wanting it rather than giving reasons which carry more weight.
11. Most people who get ahead in the world lead clean, moral lives.
12. Anyone who completely trusts anyone else is asking for trouble.
13. The biggest difference between most criminals and other people is that the criminals are stupid enough to get caught.
14. Most people are brave.
15. It is wise to flatter important people.
16. It is possible to be good in all respects.
17. Barnum was wrong when he said that there's a sucker born every minute.
18. It is hard to get ahead without cutting corners here and there.
19. People suffering from incurable diseases should have the choice of being put painlessly to death.
20. Most people forget more easily the death of their parents than the loss of their property.