

Social Design Engineering Series

SDES-2015-9

Who is audited? Experimental study on rule-based and human tax auditing schemes

Yoshio Kamijo Kochi University of Technology Research Center for Social Design Engineering, Kochi University of Technology

Takehito Masuda Research Center for Social Design Engineering, Kochi University of Technology Japan Society for the Promotion of Science

Hiroshi Uemura Kochi University of Technology Research Center for Social Design Engineering, Kochi University of Technology

15th January, 2015

School of Economics and Management Research Center for Social Design Engineering Kochi University of Technology

KUT-SDE working papers are preliminary research documents published by the School of Economics and Management jointly with the Research Center for Social Design Engineering at Kochi University of Technology. To facilitate prompt distribution, they have not been formally reviewed and edited. They are circulated in order to stimulate discussion and critical comment and may be revised. The views and interpretations expressed in these papers are those of the author(s). It is expected that most working papers will be published in some other form. Who is audited? Experimental study on rule-based and human tax auditing schemes

January 2015

Yoshio Kamijo^{a, b}, Takehito Masuda^{b, c, 1}, Hiroshi Uemura^{a, b}

^a School of Management, Kochi University of Technology, Tosayamada, Kami, Kochi 782-8502, Japan

^b Research Center for Social Design Engineering, Research Institute, Kochi University of Technology, Tosayamada, Kami, Kochi 782-8502, Japan

^c Japan Society for the Promotion of Science, Kojimachi Business Center Building, 5-3-1 Kojimachi, Chiyoda-ku, Tokyo 102-0083, Japan

Abstract

In this study, we employ a game theoretic framework to formulate and analyze tax audit schemes; we test the theoretical predictions in a laboratory experiment. We compare five audit schemes including three rule-based audits: random audit rule, cut-off audit rule, and lowest income reporter audited rule. The cut-off audit rule is theoretically optimal but, to the best of our knowledge, it has not been experimentally examined. We also employ a novel experimental design for two schemes involving the human auditor conditions. The rule-based audits experimentally enhance tax compliance as predicted, and cut-off yields the highest tax revenue among the three rule-based audits in the lab. Moreover, beyond our prediction, the human auditor conditions maximized tax revenue among the five schemes in the lab. This suggests that auditors' strategic ambiguity is another route to enhance tax compliance. We also show that subjects' social norms regarding tax payment influence the choice of tax evasion, in accordance with the experimental literature.

JEL Classification: C91; D81; H26

Keywords: audit schemes; tax evasion; laboratory experiment; cut-off rule; lowest income reporter audited rule; ambiguity

¹ Corresponding author. Postal address: Institute of Social and Economic Research, Osaka University, 6-1 Mihogaoka, Ibaraki, Osaka, 567-0047, Japan. Phone: +81-6-6879-8559; Fax: +81-6-6879-8583. E-mail addresses: yoshio.kamijo@gmail.com (Y. Kamijo), takehitomasuda@gmail.com (T. Masuda), and hiroshiuemura42@gmail.com (H. Uemura).

1. Introduction

Securing government tax revenues is a persistent and fundamental problem for all nations (Webber and Wildavsky, 1986). The incentive for individuals and companies to avoid excessive tax payments is high, which leads to tax avoidance, tax evasion, and payment delays. The results of a well-known audit program—the National Research Program, conducted by the US Internal Revenue Service (IRS)—estimated the tax gap (i.e., tax that is due but not paid in a voluntary or timely manner) in 2001 to be 345 billion dollars; this amount represented approximately 3.2% of the nominal GDP for that year (Slemrod, 2007). Although the analyses of the tax gaps in other countries are limited for several reasons (such as resource constraints and non-publication of survey results), the gaps are estimated or speculated to be considerable (see Slemrod, 2007 for details). Thus, the research on policy devices to enhance tax compliance has become increasingly significant. Therefore, this study intends to analyze auditing schemes.

Previous research examined tax auditing schemes that enhance tax compliance by analyzing a game theoretic situation involving taxpayers and the tax authority (IRS in the US, the National Tax Agency in Japan, HM revenue and Customs in the UK, etc.). One strand of the theoretical research assumes that the tax authority can use an apriori determined rule for investigation. For instance, Sánchez and Sobel (1993) analyze a dynamic game between an auditor and taxpayers where the auditor chooses the auditing strategy that determines the probability of auditing for each reported income; knowing this, the taxpayers determine the reported income, possibly untruthfully. They show that the optimal rule that maximizes the expected net tax revenue (tax + penalty - auditing cost) is the cut-off rule. According to the cut-off rule, the income range is classified into two or three classes, and there is strict audit for the lower-income class; however, there is no audit for the higher-income class.² Another strand of the literature assumes that the tax authority cannot decide a rule for audit investigations a priori. According to this line of research, a taxpayer reports his/her income; subsequently, based on the reported income, an auditor determines the effort required to reveal the true income in order to maximize the net tax revenue. Using a model of one auditor and one taxpayer, Reinganum and Wilde (1986) show that an equilibrium exists where almost all the income ranges of taxpayers underreport their true incomes. A common assumption in these two lines of research is that the true income of a particular taxpayer is private information; i.e., the auditor and other taxpayers do not know the true income of a taxpayer.

 $^{^2\,}$ According to Andreoni et al. (1998), the tax authorities in the US actually use the cut-off rule based on their prior work experience.

The prior experimental studies can be separated into two research lines: one deals with the behavioral aspect of tax-evasion decision-making, which deviates from rational decision-making;³ the other deals with the strategic interaction between auditors and taxpayers under various tax schemes. The latter approach is close to what we adopt in our study. Our experimental study focuses on the audit rule where out of a finite number of taxpayers (participants in the lab), the taxpayer whose reported income is the lowest is investigated. We name this the lowest income reporter audited (LIRA) rule. There were prior experimental studies of tax compliance and tax evasion. For instance, Collins and Plumlee (1991) consider a model wherein an individual must choose a labor supply decision and a tax evasion decision; they experimentally verify that a LIRA rule enhances tax compliance when compared to completely random auditing. Additionally, Alm and McKee (2004) conduct a tax compliance experiment involving two tax audit schemes similar to the LIRA (in fact, one rule is the LIRA); in the case of a tie, no one is inspected, and in the case of a tie, the taxpayers are inspected at random. However, they adopt a complete information setting where the taxpayers have identical income, and this is common knowledge. Thus, these prior experimental studies focus on a rather simplified setting compared to the theoretical works that use an incomplete information setting; further, these studies examined rules that are different from those suggested by theoretical studies.

Our study aims to address this gap between the theoretical and experimental studies on auditing schemes. We employ a game theoretic framework to formulate and analyze tax audit schemes, and we test the theoretical predictions with a laboratory experiment. In our setting, four players with different taxable incomes simultaneously report their incomes; based on their reports, tax proportional to the reported incomes is levied. The true income of a particular taxpayer is private information. Each player has an incentive to underreport the true income in order to reduce the tax burden. Following the taxpayers' decisions, some of them may be inspected by the auditor. If an inspected taxpayer is found to have underreported his/her income, the tax for this

³ Baldry (1986) compares the subjects' behaviors in a tax-evasion task and the equivalent gambling task (without a frame of tax) and shows that subjects choose a safer option (no-evasion) in the tax-evasion task than in the gambling task because of the moral cost incurred by the person who evades tax. Coricelli et al. (2010) investigate the relationships between emotions and rational decision-making by means of an experiment on tax evasion, where emotions are measured by skin conductance responses and self-reports. In the experiment of Gërxhani and Schram (2006), the subjects choose a source of income, where one type enables subsequent tax evasion; they show that the subjects take into account the possibility of evasion when deciding on the income source. Kastlunger et al. (2009) focus on the effect of different audit patterns on future compliance; using a 60-times repeated design experiment, they show that that early audit experiences in a "taxpaying life span" lead to increased compliance.

concealed income is levied, multiplied by the penalty rate.

In this study, we compare five audit schemes: the random audit rule (henceforward, "Random"), the cut-off audit rule (Cut-off), the lowest income reporter audited rule (LIRA), and two types of human auditor conditions (Human_1 and Human 2). In the Random scheme, a taxpayer is randomly chosen and inspected, irrespective of the reported incomes of the taxpayers. This rule is adopted quite often; it is the most common rule used in the experiment to examine the canonical tax-evasion model of Allingham and Sandmo (1972) and Yitzhaki (1974). In the Cut-off scheme, the probability is high that taxpayers whose income is less than some threshold will be inspected, and those whose income is more than the threshold are never inspected. The resources for audits are limited, and the lower-income class should be inspected more frequently to prevent a taxpayer belonging to the high-income class from imitating a taxpayer from the low-income class. Therefore, theoretical studies (e.g., Sánchez and Sobel, 1993) show that the cut-off rule with appropriate parameters becomes the optimal (revenue-maximizing) audit scheme. In the LIRA, the auditor investigates the taxpayer/reporter whose reported income is the lowest among the four reported incomes. An audit scheme similar to the LIRA was examined using laboratory experiments by Collins and Plumlee (1991) and Coricelli et al. (2010) without much theoretical analysis. In contrast, we make theoretical predictions related to the LIRA, and we run experiments to test these predictions.

In the three audit schemes discussed so far, which taxpayer is to be inspected is determined by the rules; thus, these schemes are rule-based. In contrast, in the Human_1 and Human_2 schemes, there is no *a priori* determined rule. After the taxpayers make their decisions, the subject who plays the role of the auditor must choose one of the four taxpayers as the target based on his/her own discretion. The difference between the two human conditions is the ambiguity in the auditor's choices. While the auditor chooses exactly one taxpayer in Human_1, the auditor in Human_2 can decide to inspect multiple taxpayers (however, inspecting more than one taxpayer is irrational or does not pay). However, the auditor's right to audit may itself affect the reporting behavior of the taxpayers.

From the theoretical analysis, we find that the cut-off rule with an optimal choice of threshold dominates the other three schemes in terms of truth-telling rate, minimizing the evaded incomes, and maximizing the tax revenue. The LIRA is better than the Random and Human schemes in terms of truth-telling rate and minimizing the evaded incomes; however, the Random and Human schemes are better than the LIRA in terms of penalty. The performance of Random is equal to that of Human_1 and

Human_2.

Our findings are summarized as follows. First, as the theory predicted, the rules for choosing the target (such as Cut-off and LIRA) enhance tax compliance. We observe that although the LIRA ranks first in terms of the average tax revenue, the Cut-off rule has the highest performance among the three rule-based audit schemes in terms of the average total revenue, including revenue from tax and penalty. In terms of the frequency of truthful reporting, Cut-off ranks highest when we combine all data. This observation is qualitatively consistent with the theoretical prediction. However, in terms of the percentage of reported income vs. true income, LIRA ranks first when we combine all the data. This is consistent with the first observation that LIRA ranks highest in terms of tax revenue (without the revenue from penalty). Moreover, the reporting strategies in LIRA and Cut-off have kinks, as was predicted.

Second, although the theory made the extremely negative prediction of zero reporting under Human audit conditions, the total revenue was found to be highest in the Human_2 condition among the five conditions. This observation suggests that non-game theoretic factors such as the reporters' innate norms (e.g., tax awareness) and/or ambiguity in the auditors' strategy need to be taken into account. Therefore, the auditors' strategy is investigated in this study. In the experiment, some of the auditors behaved in a manner similar to what is expected in LIRA—they investigated the low-income reporters.

Third, we conducted the regression analysis by controlling the subjects' identity and characteristics measured by the questionnaire (related to tax-payment awareness, risk attitude, etc.). The results confirmed that the earlier observations hold true even after controlling for these factors. Further, we found that tax awareness and the subjects' need for tax audit are negatively correlated with the rate of underreported income. On the other hand, the aggressiveness toward tax evasion has a positive correlation with the frequency of evading. The results suggest that the social norms of subjects related to tax payment influence the choice of tax evasion.

The rest of this paper is organized as follows. In the following section, we present a basic theory of tax evasion decision-making; subsequently, we present our theoretical predictions related to several tax audit schemes. Section 3 describes our experimental design and procedure. In Section 4, we report the results of our experiment and statistically analyze them. Section 5 concludes the paper.

2. Theory of tax audit schemes

2.1. Basic model

This section summarizes a canonical model of taxpayer decision-making proposed by Allingham and Sandmo (1972) and Yitzhaki (1974). A taxpayer decides whether and to what extent to evade taxes in the same way that an individual would weigh a risky gambling decision. The taxpayer (an individual or a firm) has a true taxable income of Y, where Y > 0 which is private information. Let t be the basic tax rate. The taxpayer pays tY as tax if (s)he reports his/her true income. However, if the income is under-reported, the taxpayer should pay tR, where R represents the under-reported income $(R \leq Y)$, and Y - R represents the amount of evaded or concealed income.⁴ However, detailed auditing is randomly executed in probability p, and the tax evasion is detected. In our model, tax evasion is revealed if the tax authority inspects the under-reporting taxpayer. In the case of inspection, the individual must pay tq(Y - R), as penalty for the tax evasion, where q represents the penalty rate for the illegal activity (q > 1). Thus, the penalty is proportional to the concealed income.

The expected utility for an individual reporting his/her income as R (where $0 \le R \le Y$) is EU = (1-p)U(Y-tR) + pU(Y-tR-tq(Y-R)), where U is a utility

function with U(Y) > 0 and U'(Y) > 0 for any Y > 0. By differentiating EU by R and

evaluating it at R = Y, we obtain $\frac{\partial EU}{\partial R}|_{R=Y} = t(pq-1)U'((1-t)Y)$. Thus, tax evasion occurs when pq < 1 or p < 1/q.

While the evasion decision depends on neither the basic tax rate t nor the true income Y, the extent of the evasion may depend on these variables.⁵ However, if we assume risk neutrality, the taxpayer fully evades tax liability (i.e., reports 0 income) whenever (s)he decides to evade taxes. In the discussion that follows, we assume risk neutrality for the taxpayers. A comprehensive review of the theory is presented in Andreoni et al. (1998).

The canonical model does not deal with how the detection probability (p) is determined. Later studies (such as Alm and McKee, 2004) pointed out that the

⁴ There are other types of reporting decisions such as non-filing and late payment of taxes owed. However, according to the 2001 IRS estimate of the tax gap, under-reporting represents approximately 82% of the gap, and non-filing and late payment represent 8% and 10% of the gap, respectively (see Slemrod, 2007). Thus, the major source of the tax gap is under-reporting.

⁵ Yitzhaki (1974) showed that under the assumption of decreasing absolute risk aversion, the extent of evasion decreases as the basic tax rate increases, and the extent of evasion increases as income increases.

determination of p is the result of the strategic interdependence between auditors and taxpayers. Thus, the detection probability seems to vary with the reported incomes (Reinganum and Wilde, 1986; Sánchez and Sobel, 1993), the past experience of cheating or auditing (Clark et al., 2004; Friesen, 2003; Greenberg, 1984; Harrington, 1988), the relative positions of the reported income (Alm and McKee, 2004; Collins and Plumlee, 1991), etc. In order to ensure strategic interdependence among taxpayers, we assume that there are n taxpayers. In the following subsections (Sections 2.2–2.5), we describe four audit schemes (random audit rule, cut-off rule, lowest income reporter audit rule, human), and we theoretically show how the taxpayer decisions are different in these four schemes.

We explain the four audit schemes using the following parameters: n = 4, t = 0.2, and q = 3. This simplification facilitates the understanding of the rules, and this is the setting we adopt in our experiment. To fairly compare these four schemes, we propose the condition that the (expected) number of investigated taxpayers is one due to the resource constraints of the audit authority. We assume that the true income of each player is selected independently from an identical uniform distribution on [0, 1000]. For each taxpayer i ($i \in \{1, 2, 3, 4\}$), Y_i and R_i denote i's true income and reported income, respectively.

2.2. Random audit scheme

In the random audit rule, the auditor chooses one of the four taxpayers at random, irrespective of their reported incomes; the chosen taxpayer is inspected. Under our setting, the probability of detection (p) is 1/n = 1/4, and the penalty rate q is 3. Thus p < 1/q holds true, indicating that the optimal strategy for each taxpayer is to report 0 income. Thus, the random audit rule does not incentivize the taxpayers to report their true income.

2.3. Cut-off audit scheme

In the cut-off audit scheme, the detection probability varies according to the reported incomes. In particular, we choose the cut-off audit where the reported income in the income class [0, 750] is inspected with probability 1/3, and the reported income of the class with income more than 750 is never inspected. According to our selected parameters, the detection probability of 1/3 is the smallest probability for a taxpayer to truthfully report his/her income. The range of [0, 750] is determined by the restriction that the expected number of inspections is one out of four taxpayers $((1/3) \times (750/1000) = 1/4)$. The optimal strategy of a taxpayer in the cut-off audit scheme

is to report his/her income truthfully when the income is less than 750 and to report the threshold when the income is more than 750. Thus, in this audit scheme, a taxpayer with higher income evades the tax burden. It is theoretically known that the cut-off rule discussed here is the audit rule that maximizes tax revenue (Sánchez and Sobel, 1993).

2.4. Lowest income reporter audited rule

In the lowest income reporter audited (LIRA) rule, the auditor investigates the income that is the lowest among the four reported incomes. Thus, a strategic interdependence exists among the taxpayers. Since the true income of each taxpayer is private information, this scheme involves incomplete information (Harsanyi, 1967).

Under the LIRA rule, the smaller the reported income is, the more likely the income is to be inspected. Therefore, the optimal strategy for the taxpayers is to report the income truthfully if the true income is less than some critical value c, and to cheat otherwise. The critical value c is calculated as follows. Assume that the four players follow the same strategy; thus, they report the true income when their income is less than c. Consider a taxpayer whose true income is c. The probability of detection when (s)he reports c is $(1-c/1000)^3$, and this probability decreases in c. According to our selected parameters, a detection probability greater than or equal to 1/3 (= 1/q) is needed for truthfully reporting income (see Section 2.1). Since the income c is the marginal value between the true income and the income when cheating, $(1-c/1000)^3$ must be equal to 1/q = 1/3. Thus, we have $c^* = 1000 \times (1-(1/q)^{1/3}) \approx 306$. In fact (as shown in Appendix A), in the LIRA rule, the equilibrium strategy of each i becomes the one where (s)he truthfully reports the income ($R_i = Y_i$) if $Y_i < c^*$, and (s)he cheats by $e(Y_i)$ ($R_i = Y_i - e(Y_i)$) if $Y_i \ge c^*$, where e represents the extend of cheating with $e(c^*) = 0$, $e(Y_i) > 0$ for $Y_i \ge c^*$, and $e'(Y_i) > 0$ for $Y_i \ge c^*$.

Comparing the equilibrium strategies in the cut-off and LIRA rules, the income range of those who truthfully report income is larger in the cut-off scheme than in the LIRA. Moreover, for any income $Y > c^*$, the taxpayer of type Y reports more income in the cut-off rule than in the LIRA (see Figure 1). Thus, the cut-off rule theoretically dominates the LIRA.

2.5. Human audit condition

In the three audit schemes discussed in Sections 2.2–2.4, the auditor follows the rule that is specified for a particular audit scheme. In contrast, in the human audit condition, the auditor himself/herself is seen as a player in the tax-reporting and auditing game; this role is played by a subject in our experiment. Thus, in this condition, after the

taxpayers' decisions are made, the auditor must choose one from the four taxpayers as the target based on his/her own discretion.

There is some difficulty in predicting the consequences of this human audit condition, even though the incentive of the auditor is to maximize the tax revenue in this one-shot game. The problem arises from the inability of the auditor to determine an audit rule. For instance, assume that the auditor announces to the taxpayers in advance that the audit rule will be the cut-off rule (explained in Section 2.2), and all the taxpayers believe it. Further assume that the income profiles reported to the auditor are (200, 400, 600, 750). In this scenario, whom should the auditor inspect? Given that the taxpayers follow the equilibrium strategy in the cut-off rule, the reported income of 750 is the only one with the possibility of cheating: therefore, the auditor should inspect this taxpayer. However, this contradicts the pre-announcement; if (s)he follows the cut-off rule, (s)he must randomly choose one taxpayer from those who reported 200, 400, and 600 and inspect that individual. This demonstrates that the pre-announcement is not credible, and the taxpayers do not believe it. A similar situation holds for other audit rules such as other versions of the cut-off rule and the LIRA rule.

A theoretical prediction with regard to this human audit condition gives some unintuitive results. In Appendix B, we show that there is a perfect Bayesian equilibrium for the dynamic incomplete information game such that all the taxpayers fully cheat irrespective of their true incomes, and the auditor randomly chooses an inspected taxpayer. Thus, there is a pooling equilibrium in this game, and the performance of the human audit condition is equal to that of the random audit rule.

2.6. Summary of theoretical predictions under our parameter selection

Figure 1 summarizes the equilibrium tax-reporting behaviors under the four audit schemes. The cut-off rule dominates the other three rules, and the LIRA rule dominates the random and human audit rules. The predicted strategies under LIRA and Cut-off have kinks at Y = 306.6 and Y = 750, respectively. Table 1 shows the expected tax revenue per taxpayer under these four schemes, with revenues from the ordinal tax and the penalty. Consistent with the reporting behaviors, the tax revenue from reported income is highest in the cut-off scheme, second highest in the LIRA scheme, and worst in the random and human audit schemes. However, for the revenue from the penalty, this order is reversed. Overall, the total tax revenue is highest in the cut-off scheme, second highest in the total tax revenue is highest in the LIRA scheme, second highest in the cut-off scheme, second highest in the random and the human audit schemes, and worst in the LIRA scheme.



Figure 1. Theoretical predictions of reported incomes under four audit schemes

 Table 1. Theoretical predictions of tax revenues (including penalty) per taxpayer under
 four audit schemes

		Audit Scheme	
	Random / Human	Cut-off	LIRA
Tax Rev.	0	93.6	68.2
Penalty Rev.	75	0	2.7
Total	75	93.6	70.9

3. Experimental design

We experimentally compare the audit rules discussed in Section 2. We have five treatments: Random, Cut-off, LIRA, Human_1, and Human_2; each of these has two sessions. Since we explained Random, Cut-off, and LIRA in the Sections 2.2–2.4, we begin with two treatments where the subjects play the role of the auditor, referred to as Human_1 and Human_2. In Human_1, given the reported incomes, an auditor must choose one reporter to audit. On the other hand, in Human_2, the auditor can choose all four of the reporters in a group; however, the auditor must pay 600 yen per reporter after the second audited target. We ran the Human_2 treatment to explore the effect of the auditor's right to audit on income reporting, although the argument is not based on economic theory. Note that under our parameter setting, a "rational" auditor does not

choose two or more reporters since the additional audit is covered only if the investigated reporter with an income of 1000 reports 0 income. Thus, the additional inspection is almost meaningless. Further, our matching protocol does not allow the auditors to create a reputation for severe auditing because the auditor and the four taxpayers were rematched in every period, and their IDs were shuffled.

We conducted all the sessions at Kochi University of Technology's Experimental Social Design Lab in July 2014. Each session lasted one and a half hours. We used the experimental software z-Tree (Fischbacher, 2007). We recruited 170 student subjects from Kochi University of Technology through campus-wide advertisements. The number of subjects for each treatment is 24 for Random, 36 for Cut-off, 40 for LIRA, 30 for Human_1, and 40 for Human_2. No subject participated in more than one session. Moreover, none of them had prior experience in a similar type of experiment. The subjects were seated at individually partitioned computer terminals assigned by lottery. We did not allow any communication among the subjects.

Each subject received a copy of the instructions. Moreover, the instructions were read aloud by an experimenter. Subsequently, the subjects answered a quiz about the audit rule in which they participated. Following the quiz, an experimenter publicly announced the answers of the quiz. The subjects then proceeded to twenty payment periods. In every session, we employed the stranger matching protocol so that every group in every period included four reporters (plus one auditor in Human_1 and Human_2). The subjects were informed that they would be randomly rematched in every period.

We first explain the process followed in one period of the Random, Cut-off, and LIRA treatments. Once a group was formed, every reporter faces the reporting screen. At the reporting screen, (s)he privately receives and confirm his/her income, which is drawn independently from the uniform distribution on [0, 1000] (yen), with an increment of 10. Every reporter can confirm t = 0.2 and q = 3. Given this information, the reporter determines how much income to report, and (s)he inputs a number that is between 0 and his/her income, with an increment of 10. Once every subject inputs the reported income and clicks the OK button, the subjects proceed to the results screen. The results screen displays (from the top) one's own income (Y), one's reported income (Y - R), tax on reported income (tR), penalty (tq(Y - R)) if any, and one's payoff in the period. In every period after the second period, the history box appears, where the subjects can confirm the information contained in the results screen in all of the previous periods. Once all the subjects click the Next button, the subjects proceed to the next period.

In the Human_1 and Human_2 treatments, the following process was added. At the beginning of the first period of Human_1 and Human_2, a lottery assigns the role of auditor to 1/5 of the subjects; the remaining 4/5 of the subjects are assigned the role of reporter. This role does not change throughout a session, and the role is displayed at the beginning of the first period. After the reporting stage, every auditor faces the decision screen that displays (on the left side) the four incomes reported in his/her group. In Human_1, every auditor inputs a reporter ID from one to four; the reported IDs are randomly assigned after the group is formed. Hence, the reporter ID is anonymous. In Human_2, on the other hand, every auditor checks the radio button for each reporter ID (from one to four) to select whether or not (s)he audits that reporter ID. The payoff for the auditor in a period is the sum of the tax and penalty in his/her group.

After participating in twenty payment periods, the subjects completed two sets of the questionnaire. The first set is related to taxpayer awareness; the questions are broadly categorized into six groups: tax awareness; acceptable tax rate; aggressiveness; need for audit; satisfaction with public services; and tax compliance. The questionnaire on taxpayer awareness is adapted from the one that is widely used in the literature (see for instance, Gërxhani, 2004; Lefebvre et al., 2014). The second set included questions on lottery vs. safe cash choice to elicit the risk preferences of the subjects. After the questionnaires were completed, the subjects were immediately paid in cash (privately). Each subject was paid the show-up fee of 800 yen plus the total earnings over three periods, which are randomly decided by a lottery.

4. Experimental results

4.1. Comparison of audit rules

We begin this section with a comparison of the amount of tax collected, the amount of penalty collected, and the total revenue.

4.1.1. Revenue

Result 1. With regard to tax and penalty revenue, the following hold true:

(i) LIRA generates the highest average tax revenue, beyond the theoretical prediction.

(ii) Random generates the highest average penalty revenue, consistent with the theoretical prediction.

(iii) Human_2 generates the highest average total tax and penalty revenue, exhibiting the highest positive deviation from the prediction.

Support. Table 2 summarizes the revenues generated by each audit scheme. The second,

third, fourth, and fifth columns indicate the number of observed reports, the average tax revenue, the average penalty revenue, and the total revenue, respectively. The last column indicates the average total revenue minus the predicted total revenue (see Figure 1).

Scheme	Number of observations	Avr. tax rev. (A)	Avr. pentalty rev. (B)	Total rev. (A) + (B)	Predicted total rev. (C)	Difference (A) + (B) - (C)
Random	480	52.3*	40.2	92.5	75.0	17.5
Cut-off	720	61.4	32.7	94.1	93.6	0.5
LIRA	800	62.6	24.0***	86.6*	70.9	15.7
Human_1	480	56	34.1	90.1	75.0	15.1
_Human_2	640	61.4	36.7	98.1	75.0	23.1

Table 2. Revenue from tax and penalty

Notes: a) The unit is yen. b) Bold indicates the highest ranked scheme in that column. c)

* p < 0.10, ** p < 0.05, *** p < 0.01, relative to the highest ranked scheme in the column.

Notably, LIRA yields the highest average tax revenue, while the theory predicts that Cut-off would do so. Moreover, the deviation in tax revenue in LIRA is 62.6-68.2 = -5.6, which is quite smaller than that in Cut-off, 61.4-93.6 = -32.2. This suggests that the subjects in Cut-off tend to under-report their income. However, a pairwise comparison by Wilcoxon rank-sum test with the average tax per subject as one data point shows that only one pair of audit schemes, LIRA and Random, has statistically significant differences in tax revenue (z = 1.796, p=0.073). Second, Random generates the highest average penalty revenue, which is significantly higher than what LIRA does (z = 2.795, p = 0.005). The third finding related to the total revenue from tax and penalty exceeds the theoretical prediction. Interestingly, Human_2 ranks the highest (98.1), which is significantly higher than LIRA (z = 1.842, p = 0.066).

If we compare the realized total revenue to the predicted one, among the five audit schemes, the revenues of Human_2 positively deviate from the predicted revenues the most. It is noteworthy that the increase in Human_2's revenues seems to be attributed to the auditor's option to audit two or more reporters, because the data show that the auditors in Human_2 rarely audit two or more reporters in a period (7 out of 160 games in total). In fact, even if we exclude the data where the auditor audits two or more reporters, the average total revenue becomes 96.5; hence, result (iii) still holds.⁶

The remarkable thing is that the data in Human_1 and Human_2 exceed what was predicted by the theory: even though a perfect Bayesian Nash equilibrium predicts that zero income would be reported under these schemes, the subjects consistently report a certain amount of their income. This could be attributed to their norm-like tax awareness, which will be discussed in the Section 4.2. Another factor in Human_1 and Human_2 is the ambiguity in the auditor's strategy, which means that the reporters (unlike in a Bayesian scenario) cannot figure out the distribution of the auditor's strategy; subsequently, they may overestimate the probability of being audited. For the theoretical foundation that ambiguity mitigates fraud, see Lang and Wambach (2013). Moreover, this ambiguity is greater in Human_2 compared to Human_1, which is consistent with the highest performance of Human_2 in levying total tax revenue.

4.1.2. Frequency of truthful reporting

In this subsection, we examine whether or not the subjects truthfully report their income. We refer to the relative frequency of truthful reporting (R = Y) divided by the total frequency of reporting as the "frequency of truthful reporting." Although the tendency to report truthfully does not become clear if we combine all the data, it becomes clearer if we classify the data into three parts using income range. In what follows, we classify $10 \le Y \le 330$, $340 \le Y \le 660$, and $670 \le Y \le 1000$ as Low, Middle, and High, respectively. ⁷ We employed this classification because the predicted reporting strategies are nonlinear, such that cheating occurs for $Y \ge 310$ under the LIRA rule.

Result 2. With regard to the frequency of truthful reporting, we have the following results.

(i) When we combine all the data, Cut-off generates the highest frequency of truthful reporting, which is qualitatively consistent with the theoretical prediction.

(ii) In the Low range, LIRA generates the highest frequency of truthful reporting.

(iii) In the Middle and High ranges, Cut-off generates the highest frequency of truthful

⁶ On the other hand, of the five audit schemes, Cut-off generates the total revenue closest to the predicted one. This could be because the unpredicted under-reporting leads to unpredicted penalties; hence, the increase in penalty revenue offsets the decrease in tax revenue. Such an offset is consistent with the theoretical prediction that sincere reporting is not different from insincere reporting under our parameter settings.

⁷ We excluded the data for Y=0 since only truthful reporting $(Y \neq 0)$ is allowed.

reporting.

Support. Table 3 summarizes the frequency of truthful reporting categorized by audit schemes and income classification.

		Income classification				
Scheme	All data	Low	Middle	High		
Random	21.8	29.6	19.5**	17.1		
Cut-off	31.7	40.4	30.8	24.3		
LIRA	18.3*	41.8	8.5***	7.3***		
Human_1	22	37.1	17.9^{**}	11.8***		
Human_2	23.9	39.3	18.7*	14.3*		

Table 3. Ranking of audit schemes according to frequency of truthful reporting

Notes: a) Bold indicates the highest ranked scheme in the column. b) p < 0.10, ** p < 0.05, *** p < 0.01, relative to the highest ranked scheme in the column.

First, when considering the overall data shown in the second column, Cut-off outperforms the other four audit schemes in terms of frequency of truthful reporting, achieving 31.7%. However, a pairwise comparison by Wilcoxon rank-sum test with the frequency of truthful reporting per subject as one data point⁸ shows that no pair of audit rules (except Cut-off and LIRA) has statistically significant differences in the frequency of truthful reporting (z = 1.768, p = 0.077). Second, in the Low range shown in the third column in Table 3, LIRA generates the highest frequency of truthful reporting (41.8%). Further, since no pair of audit schemes has statistically significant differences in the Low range, the theoretical prediction that both LIRA as well as Cut-off (40.4%)induces truthful reporting at the same degree in the Low range is qualitatively supported. However, since LIRA generates a frequency of truthful reporting far from 100 in the Low range, the theoretical prediction that both LIRA as well as Cut-off induces fully truthful reporting in the Low range is not supported. On the other hand, in the Middle and High ranges, Cut-off outperforms the other four audit schemes. A Wilcoxon rank-sum test rejects the equality of the frequency of truthful reporting if we compare Cut-off with each of other schemes in the Middle range (z=2.447, p=0.014 vs. Random; z = 3.714, p < 0.001 vs. LIRA; z = 2.114, p = 0.035 vs. Human_1; z = 1.802, p = 1.0020.072 vs. Human_2). A similar result holds in High range (z = 2.816, p = 0.005 vs. LIRA;

⁸ Andreoni and Miller (1993) use this method to eliminate cross-period correlation.

z = 2.657, p = 0.008 vs. Human_1; z = 1.685, p = 0.092 vs. Human_2). Similar results hold if we classify income range into four (see Appendix B).

4.1.3. Reporting percentages

Note that Result 2 (Section 4.1.2) says nothing about the extent of reporting. If the reported incomes under an audit scheme are sufficiently close to the actual income on average, we can say that the scheme works. According to this viewpoint, the percentage of income reported by the subjects can be called "reporting percentages." The ranking of the schemes in terms of the reporting percentages is different from the one based on the frequency of truthful reporting in Result 2.

Result 3. Based on the reporting percentages, we have the following results.
(i) When we combine all the data, LIRA generates the highest reporting percentages.
(ii) In the Low and Middle ranges, LIRA generates the highest reporting percentages.
(iii) In the High range, Cut-off generates the highest reporting percentages, which is qualitatively consistent with the theoretical prediction.

(iv) The slope of the reporting strategies in LIRA and Cut-off decreases at the predicted points.

Support. Table 4 shows the reporting percentages of the various audit schemes.

		Income classification				
Scheme	All data	Low	Middle	High		
Random	51.0***	50.0***	53.1**	50.0		
Cut-off	59.3	54.9***	57.0	65.7		
LIRA	66.3	78.8	65.1	56.9		
Human_1	58.4	60.4*	61.2	53.0*		
Human_2	62.6	66.1*	61.2	60.5		

Table 4. Ranking of audit schemes according to reporting percentages

Notes: a) Bold indicates the highest ranked scheme in the column. b) p < 0.10, p < 0.05, p < 0.01, relative to the highest ranked scheme in the column.

Based on the overall data, LIRA yields the highest reporting percentages among the five audit schemes. However, a pairwise comparison by Wilcoxon rank-sum test with the average reporting percentage per subject as one data point shows that just one pair of audit schemes, i.e., LIRA and Random, has statistically significant differences in the reporting percentage (z = 2.760, p = 0.006). If we focus on the Low income range, the outperformance of LIRA is clearer, as the pairwise comparison with any other scheme shows that LIRA yields higher reporting percentage that is statistically significant (z = 3.212, p = 0.001 vs. Random; z = 2.844, p = 0.004 vs. Cut-off; z = 1.936, p = 0.053 vs. Human_1; z = 1.905, p = 0.059 vs. Human_2). In the Middle range, however, the significant difference appears only in the comparison of LIRA with Random (z = 2.351, p = 0.019). In the High range, Cut-off outperforms the other schemes, but the significant difference appears only in the comparison of Cut-off with Human_1 (z = 1.751, p = 0.080).

To validate Result 3 (iv), we regress the reported income (*R*) on the true income (*Y*) using ordinary least squares (OLS) to test the predicted reporting strategy with one kink for LIRA and Cut-off data. As for LIRA, we estimate $R = \beta_0 + \beta_1 Y + \beta_2 D_1 + \beta_3 D_1 Y$, where D_1 is a dummy variable that takes the value 1 if $Y \ge 300$, and 0 otherwise. Note that the theory predicts $\beta_1 + \beta_3 < \beta_1$, i.e., $\beta_3 < 0$. We get $R = 2.046 + 0.776 Y + 93.393 D_1 - 0.329 D_1 Y$, where the number in parentheses is the

t-value. Moreover, Chow test rejects the null hypothesis of no structural break: $\beta_2 = \beta_3 = 0$ (F = 5.52, p < 0.01). These findings indicate that the subjects tend to evade a greater proportion of their income when $Y \ge 300$. Similarly, for the Cut-off data, we get $R = -18.027 + 0.638Y + 395.171D_2 - 0.424D_2Y$, where D_2 is a dummy variable $(-0.93)^{-10.027} + (-0.000)^{-$

that takes the value 1 if $Y \ge 750$, and 0 otherwise. Moreover, there is a kink at Y = 750 (Chow test; F = 2.58, p = 0.076).⁹

4.1.4. Auditors' choices in Human treatments

In this subsection, we examine how subjects assigned to the role of auditor behaved. Since we employed the random matching and the auditor cannot see reporters' identities, we use only the reported income profile to explain audit decisions. We have the following results.

Result 4. The auditors in Human_1 as well as Human_2 tend to pick up reporters with low reported income for audit. Moreover, the auditors in Human_2 are more likely to

⁹ We get $R = \underset{(0.39)}{8.649} + \underset{(12.97)}{0.494} Y$ for Random; $R = \underset{(1.634)}{27.899} + \underset{(17.09)}{0.507} Y$ for Human_1; and $R = \underset{(0.85)}{13.703} + \underset{(21.00)}{0.578} Y$ for Human_2.

pick up reporters with the minimum reported income in the group.

Support. First, we made a contingency table with the row indicating whether or not a reporter is audited, and the column indicating whether the reporter has the minimum reported income in the group, by treatment. A chi-square test rejects the null hypothesis of no association in Human_2 (χ^2 =17.844, *p* < 0.001), but does not do so for Human_1 (χ^2 =0.511, *p* = 0.475). On the other hand, a chi-square test replacing the column with the three reported income ranges (Low, Middle, and High) rejects the null hypothesis of no association in both Human_1 (χ^2 =10.892, *p* = 0.004) as well as Human_2 (χ^2 =27.541, *p* < 0.001).

4.2. Regression results

In this section, we confirm through multiple regression analyses that audit scheme, amount of income, and awareness about tax payment affect the decision of tax evasion. The regression analysis involving tax payment awareness enriches our understanding of why Human_1 and Human_2 work beyond the theory (as stated in Result 1). Moreover, the analysis is in line with the claims put forward in the extant literature that the motivation to comply depends on subjective constructs of tax phenomena and collective sense-making of subjective tax knowledge, on myths and legends about taxation and others' tax behavior, on subjective constructs and evaluations of perceived and internalized norms, perceived opportunities not to comply, and perceptions of fairness (Braithwaite, 2003). The aggregation of these variables results in the motivation and socially shared beliefs and evaluations is related to motivational postures (Braithwaite, 2003).

Each subject replied to the question about tax awareness (tax-payment awareness, acceptable tax rate, aggressiveness against tax evasion, needs for audit, satisfaction for public service, and tax compliance) after the experiment. Moreover, all participants replied to the questionnaires for measurement of risk preference. In order to measure the subjects' degree of risk aversion, we set 11 lotteries vs. safe constant cash questions, varying the winning probability of the lottery from 0% to 100% with an increment of 10%. We measured each subject's switching point where (s)he begin to prefer the lottery to safe constant cash. We included the tax awareness and risk appetite of the participants (each answer to the questions) in the regression model as independent variables.

We present the descriptive statistics before considering the determinants of

cheating behavior. The definitions of the variables for the descriptive statistics and multiple regressions are presented in Table 5. The total number of observations is 3,400 (170 subjects \times 20 times). From these, 300 observations were excluded from the analyses since the subjects did not reply to the questions after the experiment. Further, 280 observations were excluded because the subjects played the role of auditors; 17 observations for which we were unable to calculate the decision time were excluded. As a result, 2,803 observations were used as the sample in our statistical analysis.

Variable	Definition
Low (reference) ¹	equal to 1 if 10≤Income≤330, and 0 otherwise
Middle ¹	equal to 1 if 340≤Income≤660, and 0 otherwise
High ¹	equal to 1 if 670≤Income≤1000, and 0 otherwise
Random(reference) ¹	equal to 1 if an audit scheme is Random, and 0 otherwise
$Cut-Off^{1}$	equal to 1 if an audit scheme is Cut-Off, and 0 otherwise
LIRA ¹	equal to 1 if an audit scheme is LIRA, and 0 otherwise
Human_1 ¹	equal to 1 if an audit scheme is Human_1, and 0 otherwise
Human_2 ¹	equal to 1 if an audit scheme is Human_2, and 0 otherwise
tax awareness ²	the tax-payment awareness of a subject
acceptable tax rate	the acceptable tax rate of a subject on 10,000
aggressiveness ²	the aggressiveness of a subject against tax evasion
needs for audit ²	the needs for tax audit that a subject feels
satisfaction for public survices ²	the satisfaction for public survices of a subject
tax compliance ²	the tax compliance of a subject
Decision Time	time spent in order that a subject reports income
Male	male of a subject
rate of evaded income (t-1)	the ratio of evasion income to the truth income at t-1
audit (t-1)	an indicator variable equal to 1 if a subject was audited at t-1, and 0 otherwise
<pre>rate evade(t-1) *audit(t-1)</pre>	the interaction between rate of evaded income (t-1) and audit (t-1)
Risk appetite ³	the risk preference of a subject

Table 5. Definition of variables

Notes: a) The variables are indicator variables. b) Each question is categorized into one of six questionnaire items: tax awareness, aggressiveness toward tax aversion, need for tax audit satisfaction with public services, and tax compliance. 10 means that each answer for questionnaires item of a subject is low, and 100 means that each item is high. For correspondence between each question and questionnaire item, see the online supplementary material. c) We used the answers for question 3 where every outcome is positive. A larger number means that the subject is risk averse.

4.2.1. Comparison between evading decision and non-evading decision

Table 6 presents the descriptive statistics for the evading decision and the non-evading decision. First, the decision to evade taxes is more likely to increase with an increase in the true income (Low: approximate t = -11.36, p < 0.001; Middle: approximate t = 3.84, p < 0.001; High: approximate t = 8.87, p < 0.001)¹⁰. This result is consistent with the results presented in Coricelli et al. (2010).

¹⁰ Low: Pearson $\chi^2 = 148.77$, p < 0.001; Middle: Pearson $\chi^2 = 13.88$, p < 0.001; High: Pearson $\chi^2 = 67.88$, p < 0.001.

	evading decision		non-evadeir	differences	
	mean	std.dev	mean	std.dev	t or χ^2
Low	0.26	0.43	0.5	0.50	-11.4 ***
Middle	0.35	0.47	0.28	0.48	3.84 ***
High	0.39	0.48	0.22	0.41	8.87 ***
Tax awareness	82.24	1 4.9	84.29	14.7	-3.17 **
Acceptable tax rate	10.49	8.02	9.66	7.21	2.42 **
Aggressiveness	40.26	2.21	30.48	2.32	7.95 ***
Needs for audit	64.55	15.5	69.09	16.5	-6.63 ***
Satisfaction for public services	42.67	11.8	42.32	12.9	0.64
Tax compliance	76.70	1 6	76.87	17.9	-0.24
Risk appetite	57.70	27.3	55.39	33.3	2.23 **
Decision time	21.97	6.98	23.11	6.47	-3.99 ***
Male	0.66	0.47	0.68	0.46	0.39
Number of observations	2.095		708		

Table 6. Summary of descriptive statistics for evading vs. non-evading

Notes: p < 0.10, p < 0.05, p < 0.01.

Subjects with low awareness regarding tax payment are more likely to cheat about their income (tax awareness: t = -3.17, p = 0.002). Moreover, subjects with high acceptable tax rate and subjects with motivation to evade tax are more likely to cheat about their income (acceptable tax rate: t = 2.42, p = 0.061; aggressiveness: t = 7.95, p < 0.001). However, subjects who do not feel the necessity for tax audits are more likely to evade tax payments (need for audit: t = -6.63, p < 0.001). Our results support the results of previous studies (e.g., Braithwaite, 2003).

Moreover, the risk attitude of individuals affects their decision to evade tax. Subjects who dislike risky choices are less likely to report their income truthfully (risk appetite: t = 2.23, p = 0.030). Finally, when the decision time to report income is short, the probability that the subjects cheat about their income becomes high (decision time: t = -3.99, p < 0.001).

4.2.2. Multiple comparisons of audit schemes

Table 7 shows the frequency of cheating (=1-the frequency of truthful reporting) in each audit scheme and the results of the analysis of the differences (mean) using the Tamhane multiple comparison procedure.¹¹ The

 $^{^{11}\,}$ We consider true reporting frequency to be an interested variable. However, since previous studies

dependent variable is an indicator variable equal to 1 if a subject underreports her/his income, and 0 otherwise. The results indicate that the frequency of cheating in Cut-Off is the lowest (Table 7: Panel A). The differences in the frequency of cheating in Random, LIRA, and Human_1 are not significant. On the other hand, the frequency of cheating in LIRA is higher than that in Human_2 (Table 7: Panel B). Eventually, Cut-Off is the most effective audit scheme for the prevention of decisions to evade taxes. This result is consistent with Result 2 (Section 4.1.2).

Panel A. Descriptive Statistics for audit schemes								
audit scheme	frequency	mean	std.dev	std.err				
Random	417	0.760	0.428	0.020				
Cut-Off	674	0.699	0.471	0.018				
LIRA	695	0.809	0.394	0.015				
Human_1	437	0.769	0.422	0.020				
Human_2	580	0.740	0.439	0.018				

Table 7. Summary of frequency of cheating categorized by audit scheme and the results of the analysis of differences

Panel B. One-way analysis of variance							
audit scheme		differences	std.err	p-value			
Random	Cut-Off	0.091	0.027	0.010 **			
	LIRA	-0.048	0.025	0.461			
	Human_1	-0.008	0.029	0.999			
	Human_2	0.020	0.027	0.998			
Cut-Off	LIRA	-0.135	0.023	<0.001 ***			
	Human_1	-0.099	0.027	0.003 **			
	Human_2	-0.070	0.257	0.060 *			
LIRA	Human_1	0.398	0.251	0.701			
	Human_2	0.292	0.027	0.964			
Human_1	Human_2	0.029	0.027	0.964			

Notes: p < 0.10, p < 0.05, p < 0.01.

4.2.3. Multiple regression analyses

We assume that the decision to evade taxes and the amount of income evaded can be explained by different factors and by the differential effects of the same factors that affect both the decisions (Coricelli et al., 2010). We deal with this situation by referring to a two-step Heckman model.

⁽e.g., Coricelli et al., 2010) include the frequency of cheating as a dependent variable in the regression analysis, we follow this precedent. Levene-value is 34.959, and *F*-value is 9.430; p < 0.001. Therefore, we use a non-parametric method (Tamhane) to compare the frequency of cheating of the audit schemes.

Result 5. Logistic regression shows the following results.

(i) Subjects are more likely to decide to cheat about their income with increasing income.

(ii) Cut-off and Human_2 affect the decision to evade taxes.

(iii) The tax payment awareness (need for tax audit, aggressiveness toward tax evasion) of the subjects affects their decision to evade taxes.

Support. Table 8 presents the results of the regression analysis.

	DV=Proba	bility to e	vade		DV=Propo	ortion of eva	ded income		
	Random-e	ffects Logi	stic regressi	on	Random-e	ffects GLS 1	nodels with	robust Std.Err	
	А			в			С		
Variables	Coef.	Std.Err,	z	Coef.	Std.Err,	z	Coef.	Std.Err,	z
Low (reference)	-			-					
Middle	0.593	0.064	9.21 ***	26.157	19.430	1.35	25.617	17.776	1.44
High	0.808	0.066	12.17 ***	36.201	24.354	1.81 *	34.773	22.241	1.77 *
Random(reference)	-								
Cut-off	-3.490	0.087	-3.99 ***	-24.291	13.048	-1.86 *	-20.47	11.891	-1.72 *
LIRA	0.076	0.092	0.83	-20.358	8.610	-2.36 **	-12.95	7.912	-1.64
Human_1	-0.032	0.099	-0.33	-13.059	9.348	-1.40	-10.44	8.563	-1.22
Human_2	-0.152	0.091	-1.69 *	-24.201	9.669	-2.50 **	-18.73	8.887	-2.11 **
Tax awareness	-0.003	0.002	-1.82 *	-0.310	0.220	-1.41	-0.230	0.205	-1.13
Acceptable tax rate	0.004	0.003	1.10	0.339	0.363	0.93	0.267	0.329	0.81
Aggressiveness	0.093	0.012	7.57 ***	5.307	2.883	1.87 *	4.788	2.597	1.84 *
Needs for audit	-0.007	0.001	-4.12 ***	-0.434	0.294	-1.47	-0.413	0.270	-1.53
Satisfaction for pulic service	0.006	0.014	0.44	0.018	0.199	1.22	0.014	0.182	0.08
Tax compliance	-0.001	0.010	-0.02	-0.019	0.136	-0.14	-0.013	0.128	-0.11
Male	-0.068	0.058	-1.18	7.499	5.949	1.26	4.482	5.479	0.82
Prop evade(t-1)							0.424	0.061	6.93 ***
Audit (t-1)							5.465	7.954	0.69
Prob evade(t-1) *andit(t-1)							-0.033	0.128	-0.28
Risk appetite							0.047	0.104	0.46
Decision time							-0.203	0.290	-0.70
Inverse Mill's ratio				4.889	3.490	1.49	3.544	5.129	1.22
Constant	0.804	0.194	4.13 ***	20.321	31.392	0.65	2.123	13.53	0.06
Number of Observations	2803			2803				2803	
Log Likelihood	-1142.19			-				-	
Wald χ^2	223.31			146.30				864.48	
Prob> χ^2	0.000			0.000				0.000	
ρ	0.459 **	tek		0.329				0.127	
\mathbb{R}^2	-			0.161				0.073	

Table 8. Determinants of the proportion of evaded income (Heckman two-step model)

Notes: p < 0.10, p < 0.05, p < 0.01.

First, we estimate the determinants of the decision to evade taxes using a random-effects logistic model for the panel data (Column A). We include individual

random effects for the subjects to control for the lack of independence between each trial because each subject repeats the task 20 times. The dependent variable is an indicator variable that equals 1 if a subject underreports her/his income, and 0 otherwise. The independent variables include each income category (the minimum income category being the omitted reference category), each audit scheme (Random being the omitted reference category), and the tax payment awareness of each subject. We also include gender (1 if male and 0 if female).

The results of the regression analysis indicate that the Middle and High income categories have a significant positive correlation with the decision to evade tax (z = 9.21, p < 0.001; z = 12.17, p < 0.001, respectively). This is consistent with the findings reported in previous studies (e.g., Coricelli et al., 2010). Second, Cut-off and Human_2 are negatively correlated with the decision to evade tax (Cut-off: z = -3.99, p < 0.001; Human_2: z = -1.69, p = 0.089). Third, the tax payment awareness of the subjects affects their decision to evade tax (tax awareness: z = -1.82, p = 0.073; aggressiveness: z = 7.57, p < 0.001). While aggressiveness toward tax evasion has a positive correlation with evaded income, the need for tax audit has a negative correlation with the decision to evade taxes.

By including the percentage of cheating (=1 - the reporting percentage) we get the next result, which is a counterpart of Result 3 (Section 4.1.3).

Result 6. The generalized least squares (GLS) model (Heckman two-step model) presents the following results.

(i) Three audit schemes (Cut-off, LIRA, and Human_2) have a negative correlation with the percentage of cheating.

(ii) The aggressiveness toward tax evasion of a subject has a positive correlation with the percentage of cheating.

(iii) The previous percentage of cheating has a positive correlation with the subsequent percentage of cheating.

Support. In the second step, we estimate two models of the rate of evaded income using random-effects GLS models with robust standard errors. In Panel B (Table 8), we add the inverse of the Mill's ratio extracted from the first-step estimation to control for a potential correlation among the error terms of the two equations. In Panel C (Table 8), we include the risk preference of each subject and the decision time to report their income (Alm, 1988; Coricelli et al., 2010).

First, some of the audit schemes correlate significantly and negatively with the rate of the percentage of cheating (Cut-off: z = -1.86, p = 0.074; LIRA: z = -2.36, p = 0.029; Human_2: z = -2.50, p = 0.021). This result indicates that the effect of these audit schemes in reducing the percentage of cheating is higher than that of a random audit scheme. This is in accordance with our theoretical prediction that Cut-off and LIRA are more effective audit schemes than Random. Further, Human_2 has a mitigating effect on the rate of evaded income. Second, the higher the aggressiveness toward tax evasion, the higher is the percentage of cheating (aggressiveness: z = 1.84, p = 0.071). Finally, the previous rate of evaded income and auditing have a positive correlation with the subsequent rate of evaded income (rate evade (t-1): z = 6.93, p < 0.001).¹²

4.2.4. Additional analysis

An important concern that remains is whether the amount of income in each audit scheme effects the decision to cheat about the reported income. In order to deal with this issue, we included the interaction between each income category and each audit scheme as an independent variable in the regression model (Table 9).

Result 7. The regression model that includes the interaction between each income category and each audit scheme shows the following results.
(i) Even if true income increases, Cut-off moderates the percentage of cheating.
(ii) The percentage of cheating in LIRA increases with increasing true income.

Support. The result indicates that even when the true income increases, Cut-off continues to have a moderating effect on the increase in the rate of evaded income (see Panel B in Table 9). On the other hand, LIRA loses its moderating effect on the rate of under-reporting with increasing true income. Consequently, Cut-Off is the most effective audit scheme for the prevention of tax evasion. Although this conflicts a little with the results in the single regression analysis (Result 3 in Section 4.1.3), the result is consistent with our theoretical prediction.

Table 9. Decision to cheat about income and interaction between true income and audit scheme

 $^{^{12}}$ This result differs from what Coricelli et al. (2010) reported; according to them, an audit in the previous period has a negative correlation with the rate of evaded income.

	DV=Probability to evade		DV=Proportion of evaded income			
	Random-eff	ects Logistic	c regression	Random-effects GLS models with robus		
	А			В		
Variables	Coef.	Std.Err,	z	Coef.	Std.Err,	z
Low (reference)	_			_		
Middle	0.592	0.064	9.20 ***	8.998	12.742	0.71
High	0.807	0.066	12.17 ***	12.468	15.151	0.82
Random(reference)	_					
Cutoff	-0.346	0.087	-3.97 ***	-8.878	9.388	-1.82 *
LIRA	0.076	0.092	0.83	-26.716	7.404	-3.61 ***
Human_1	-0.031	0.099	-0.32	-10.175	7.665	-1.33
Human_2	-0.151	0.091	-1.65 *	-19.549	7.484	-2.61 **
Tax awareness	-0.003	0.002	-1.81 *	-0.142	0.118	-1.21
Acceptable tax rate	0.004	0.003	1.11	0.150	0.193	0.78
Aggressiveness	0.093	0.012	7.57 ***	2.832	1.591	1.78 *
Needs for audit	-0.008	0.001	-4.25 ***	-0.263	0.161	-1.63
Satisfaction for pulic service	0.008	0.017	0.38	0.015	0.192	0.12
Tax compliance	-0.005	0.009	-0.10	-0.129	0.215	-0.44
Male	-0.068	0.058	-1.17	6.271	3.166	1.98 *
Prop evade(t-1)				0.428	0.035	12.13 ***
Audit (t-1)				4.276	4.239	1.01
Prob evade(t-1) *andit(t-1)				-0.046	0.068	-0.68
Risk appetite				0.033	0.041	0.82
Decision time				-0.024	0.166	-1.45
Cutoff*Middle				-0.037	8.744	-0.12
Cutoff*High				-10.723	8.768	-1.79 *
LIRA*Middle				10.989	9.013	1.22
LIRA*High				20.797	8.986	2.31 **
- Human_1*Middle				-2.798	9.632	-0.29
Human_1*High				3.041	9.907	0.31
- Human_2*Middle				4.359	8.968	0.49
Human_2*High				5.886	8.983	0.66
Inverse Mill's ratio				13.013	12.132	1.63
Constant	0.816	0.193	4.21 ***	12.963	18.651	1.39
Number of Observations	2803			2803		
Log Likelihood	-2829.66					
Wald γ^2	368.44			279.65		
$Prob > \gamma^2$	0.000			0.000		
0	0.622 **	*		0.098		
R^2	_			0.179		

Notes: p < 0.10, p < 0.05, p < 0.01.

5. Conclusion

The contributions of this paper are as follows. First, we confirm that tax revenue can be improved by deciding the scheme for choosing the audited target, controlling tax and penalty rate, and determining the auditing resources. Although this suggestion is based on theoretical study, we show that this actually works in a laboratory setting.

Second, among the auditing rules considered, the cut-off rule generates the highest total tax revenue in the laboratory setting, which is consistent with our

theoretical prediction. However, as suggested by previous experimental study on the behavioral aspects of tax reporting decisions, the social norm of tax observance, tax awareness, moral cost of tax evasion, other-regarding preferences, tendency to overestimate a small probability, and asymmetry between loss and gain (Alm, 2012) may affect the performance of the rules. Therefore, each rule shows better performance in terms of total revenue in the lab compared to the theoretical prediction based on the rational risk-neutral decision maker. Nonetheless, the ranking of the performances of the schemes in the lab setting is consistent with that based on the theoretical predictions.

Third, the human condition wherein a subject in the role of the auditor chooses the target for the inspection after seeing the reported incomes works well beyond our expectation. In fact, the human condition with the option of irrational excessive audits achieves the highest total tax revenue among all the conditions (including the cut-off rule). This result means that there can be another route to enhance tax compliance, that is, relying on the ambiguous human choice of the target. While building a better scheme for the audit is common in the US and European countries, audit selection based on the intuition of the inspector is emphasized in Japan. The tax auditors of the National Tax Agency Japan identify the taxpayers based on the National Tax Total Management System database. However, the final decision depends on the tax auditors' discretion.¹⁴ The findings of our study imply that both these methods are useful for increasing tax compliance. However, they work for distinct reasons: one depends on the optimal design based on the rational decision-making of the taxpayers; the other depends on the ambiguous selection of the audit target.

Although our current study finds that the human condition works well (beyond what was expected), the reason for this is still uncertain. Since the purpose of this study is to compare the five audit schemes, our experimental design has an obvious limit when it comes to identifying the reason for this result. Ambiguity aversion may provide one promising explanation (Eckel and Grossman, 2008); identifying the reason for this finding could be an interesting topic for future research.

Ambiguity is regarded as an important characteristic in tax law (Long and Swingen, 1987). Given the experimental finding that people exhibit aversion to

¹³ A former national tax auditor said the following about the Japanese tax audit: "The standards to select a respondent virtually don't exist. A respondent is selected based on instincts of an experienced tax auditor" (AERA, 2014).

¹⁴ See OECD (2006).

ambiguity (Eckel and Grossman, 2008), emerging theoretical research offers strategic usage of ambiguity to fight against fraud by ambiguity-averse players in the context of tax evasion (Yoon et al., 2011) and insurance fraud (Lang and Wambach, 2013). Moreover, we observe that the human auditor conditions succeed beyond the theoretical prediction and even exceeds the theoretically optimal Cut-off rule in terms of total revenue. Hence, a fruitful line of future research would be to explore richer non-committed auditor settings in theory as well as experiments in line with the real audit process. Possible extensions include: the auditor receives a signal correlated with the true income, in addition to the reported income, based on the documents or an interview; the true income may not be revealed even after an audit, which will decrease the effective penalty rate; there are repeated interactions where auditors can pick up the inspection targets, based on the reported income history. The present study connects committed (rule-based, principal-agent) and non-committed (human-based, incentivized) auditing schemes, which have been considered separately in prior studies. Our findings of the success of the audit schemes involving human conditions imply the importance of designing institutions that incorporate the psychological aspects of human beings.

Acknowledgments

Y. Kamijo is grateful for the financial support received from JSPS Grant-in Aid for Exploratory Research and Grant-in-Aid for Scientific Research (B).

Appendices

Appendix A. Theoretical analysis of the lowest income reporter audited (LIRA) rule

Let $N = \{1, 2, ..., n\}$, with $n \ge 2$ as the set of taxpayers (individuals or firms) that should report their income to a tax authority. For $i \in N$, true income is denoted by $Y_i \in [Y_\ell, Y_h]$, where Y_ℓ , and Y_h are the lower and upper bounds of income, respectively. In our experiment, they are equal to 0 and 1000, respectively. Each *i* with income Y_i reports $r_i \in [0, Y_i]$ to the tax authority.

In an income reporting game (IRG), taxpayers report their incomes simultaneously. Let $(r_1, r_2, ..., r_n) \in [0, 1000]^n$ be the profile of the reported incomes. A tax authority observes the profile and inspects the individual with the lowest reported income. If there is a tie, a random selection is made from among the tied members.

We assume that the true income of each individual is a random variable. Thus, we model IRG with a strategic inspection as a normal-form game with incomplete information (Harsanyi, 1967). We assume that the true income Y_i of an individual is identically and independently distributed according to a continuous distribution function F on [0,1000]. Let f be a density function of F. Because the IRG with strategic auditing is a normal-form game with incomplete information, the strategy of player i is a function that associates his/her realized true income Y_i with reporting income r_i . Let γ_i be the strategy of player i.

We adopt the symmetric Bayesian Nash equilibrium (BNE) $(\gamma, \gamma, ..., \gamma)$, where every player uses the same strategy γ as a solution criterion to evaluate strategic auditing. We assume the following differentiability condition.

Assumption 1. A Bayesian equilibrium strategy γ is a continuous, differentiable, and increasing function with $\gamma(0) = 0$.

We explore the conditions that should be satisfied by γ . Suppose n-1 individuals, with the exception of player *i* with income Y(type Y player), follow strategy γ . The expected payoff of the type Y player reporting $r \leq Y$ is:

$$U(r,Y) = Y - tr - \left(1 - F(\gamma^{-1}(r))\right)^{n-1} qt(Y - r).$$
(1)

Note that $(1 - F(\gamma^{-1}(r)))^{n-1}$ is the probability of *r* being the lowest reported income among *n* reported incomes. This is a continuous function in the domain [0, Y] when γ is a continuous function.

By differentiating U(r,Y) in r, we obtain the following:

$$\frac{\partial U}{\partial r} = -t - (n-1) \left(1 - F(\gamma^{-1}(r)) \right)^{n-2} \left(-f(\gamma^{-1}(r)) \right) \frac{qt(Y-r)}{\gamma'(\gamma^{-1}(r))} + \left(1 - F(\gamma^{-1}(r)) \right)^{n-1} qt \quad (2)$$

For $(\gamma, \gamma, ..., \gamma)$ to constitute a BNE, this must be a local maximum at $r = \gamma(Y)$. Thus, the following first-order condition should be satisfied:

$$\frac{\partial U}{\partial r} (\gamma(Y), Y) \begin{cases} \geq 0 & \text{if } \gamma(Y) = Y \\ = 0 & \text{if } 0 < \gamma(Y) < Y \\ \leq 0 & \text{if } \gamma(Y) = 0 \end{cases}$$
$$\Leftrightarrow \frac{\left(\frac{1}{q} - \left(1 - F(Y)\right)^{n-1}\right)}{\left(n-1\right)\left(1 - F(Y)\right)^{n-2} f(Y)} \gamma'(Y) \begin{cases} \leq Y - \gamma(Y) & \text{if } \gamma(Y) = Y \\ = Y - \gamma(Y) & \text{if } 0 < \gamma(Y) < Y \\ \geq Y - \gamma(Y) & \text{if } \gamma(Y) = 0 \end{cases}$$

Let Y^* be defined as follows:

$$Y^{*} = F^{-1} \left(1 - \left(\frac{1}{q} \right)^{1/(n-1)} \right)$$
(3)

For $Y < Y^*$, $\frac{1}{q} - (1 - F(Y))^{n-1} < 0$. Because $\gamma' > 0$ from Assumption A1 and

 $Y - \gamma(Y) \ge 0$, $Y = \gamma(Y)$ must hold for $Y < Y^*$. Therefore, a type *Y* taxpayer for $Y \le Y^*$ sincerely reports his/her income.

Next, consider Y that satisfies $Y > Y^*$. The differential equation can be reduced to

$$\gamma'(Y) + A(Y)\gamma(Y) = A(Y)Y$$

where

$$A(Y) = \frac{(n-1)(1-F(Y))^{n-2}f(Y)}{\left(\frac{1}{q} - (1-F(Y))^{n-1}\right)}.$$

and A(Y) > 0 for $Y > Y^*$. A general solution of the above differential equation is

$$\gamma(Y) = e^{-\int A(Y)dY} \left(\int A(Y) Y e^{\int A(Y)dY} dY + C \right).$$

with an initial condition $A(Y) = Y^*$. By using partial integration,

$$\begin{split} \gamma(Y) &= e^{-\int A(Y)dY} \left(Y e^{\int A(Y)dY} - \int e^{\int A(Y)dY} dY + C \right) \\ &= Y - e^{-\int A(Y)dY} \left(\int e^{\int A(Y)dY} dY - C \right). \end{split}$$

Let $a(Y) = \int A(Y)dY$, that is, an indefinite integral of A(Y). Considering the initial condition,

$$\gamma(Y) = Y - \int_{Y^*}^{Y} e^{a(z)} dz / e^{a(Y)} = Y - \int_{Y^*}^{Y} e^{a(z) - a(Y)} dz \text{ for } Y > Y^*.$$

Therefore, we have a candidate for an equilibrium strategy as follows:

$$\gamma(Y) = \begin{cases} Y & \text{for } Y \leq Y^* \\ Y - \int_{Y^*}^{Y} e^{a(z) - a(Y)} dz & \text{for } Y > Y^* \end{cases}$$
(4)

The next theorem states that γ constitutes a BNE.

Proposition 1. Let γ be defined in (4). Strategy profile $(\gamma, \gamma, ..., \gamma)$ is a BNE.

Proof. The payoff of type Y reporting r is given by (1) and is reduced to

$$U(r,Y) = (1-t)Y + t(Y-r)\left(1 - q\left(1 - F(\gamma^{-1}(r))\right)^{n-1}\right)$$
(5)

We consider the following two cases separately: (i) $Y < Y^*$ and (ii) $Y \ge Y^*$.

Case (i) $Y < Y^*$. Because $r \le Y < Y^*$ and $\gamma(r) = r$, the payoff described by (5) is re-written as follows:

$$(1-t)Y + t(Y-r)(1-q(1-F(r))^{n-1})$$
 (6)

Because $r \leq Y < Y^*$ and Y^* satisfies (3), $1 - q(1 - F(r))^{n-1}$ is negative. Therefore, the taxpayer payoff is maximized at r=Y.

Case (ii) $Y \ge Y^*$. When $r \le Y^*$, the payoff is given by (6) and is maximized at $r=Y^*$ in the domain $[0,Y^*]$. Next, suppose $r > Y^*$. The first derivative of U(r,Y) given by (2) is rewritten as follows:

$$\begin{aligned} \frac{\partial U}{\partial r} &= -t + (n-1) \left(1 - F(\gamma^{-1}(r)) \right)^{n-2} f(\gamma^{-1}(r)) \frac{qt(\gamma^{-1}(r) - r)}{\gamma'(\gamma^{-1}(r))} \\ &+ \left(1 - F(\gamma^{-1}(r)) \right)^{n-1} qt + (n-1) \left(1 - F(\gamma^{-1}(r)) \right)^{n-2} f(\gamma^{-1}(r)) \frac{qt(Y - \gamma^{-1}(r))}{\gamma'(\gamma^{-1}(r))} \end{aligned}$$

Because $Y^* < \gamma^{-1}(r) < r$ and from (3), γ must satisfy the following:

$$\frac{\left(\frac{1}{q}-\left(1-F\left(Y\right)\right)^{n-1}\right)}{\left(n-1\right)\left(1-F\left(Y\right)\right)^{n-2}f\left(Y\right)}=\frac{Y-\gamma\left(Y\right)}{\gamma'\left(Y\right)}.$$

Using this, the first derivative is reduced to

$$\begin{split} &\frac{\partial U}{\partial r} = -t + (n-1) \Big(1 - F \Big(\gamma^{-1}(r) \Big) \Big)^{n-2} f \Big(\gamma^{-1}(r) \Big) t \Bigg(\frac{\Big(1 - q \Big(1 - F \Big(\gamma^{-1}(r) \Big) \Big)^{n-1} \Big)}{(n-1) \Big(1 - F \Big(\gamma^{-1}(r) \Big) \Big)^{n-2} f \Big(\gamma^{-1}(r) \Big)} \\ &+ \Big(1 - F \Big(\gamma^{-1}(r) \Big) \Big)^{n-1} qt + (n-1) \Big(1 - F \Big(\gamma^{-1}(r) \Big) \Big)^{n-2} f \Big(\gamma^{-1}(r) \Big) \frac{qt \Big(Y - \gamma^{-1}(r) \Big)}{\gamma' \Big(\gamma^{-1}(r) \Big)} \\ &= -t + t \Big(1 - q \Big(1 - F \Big(\gamma^{-1}(r) \Big) \Big)^{n-1} \Big) + \Big(1 - F \Big(\gamma^{-1}(r) \Big) \Big)^{n-1} qt \\ &+ (n-1) \Big(1 - F \Big(\gamma^{-1}(r) \Big) \Big)^{n-2} f \Big(\gamma^{-1}(r) \Big) \frac{qt \Big(Y - \gamma^{-1}(r) \Big)}{\gamma' \big(\gamma^{-1}(r) \big)} \\ &= (n-1) \Big(1 - F \Big(\gamma^{-1}(r) \Big) \Big)^{n-2} f \Big(\gamma^{-1}(r) \Big) \frac{qt \Big(Y - \gamma^{-1}(r) \Big)}{\gamma' \big(\gamma^{-1}(r) \big)}. \end{split}$$

This is positive for $r \in [Y^*, \gamma(Y))$, negative for $r \in (\gamma(Y), Y]$, and zero if $r = \gamma(Y)$. Thus, U is maximized at $r = \gamma(Y)$.

Therefore, the proof ends.

The following intuition can be gained from the preceding discussion. Because the lowest reporter is audited, the risk of punishment when cheating is high for low-income taxpayers. This implies that sincere reporting is more likely to occur among low-income taxpayers. Assuming that every taxpayer with income less than Y honestly reports his/her true income, the payoff for a taxpayer with income Y when (s)he reports r is given by (6). Therefore, as long as $1-q(1-F(r))^{n-1}$ is negative, the preferred action is to honestly report. The critical value of reporting income honestly is obtained when $1-q(1-F(r))^{n-1} = 0$ i.e., $Y = Y^*$. For a taxpayer whose income exceeds Y^* , honest reporting is never a preferred action. The extent of tax evasion is captured by $\int_{Y^*}^{Y} e^{a(z)} dz / e^{a(Y)}$. The slope of γ in the domain $[Y^*, 1000]$ is

$$\begin{split} \gamma'(Y) &= 1 - \frac{1}{\left(e^{a(Y)}\right)^2} \left(\left(e^{a(Y)}\right)^2 - e^{a(Y)} A(Y) \int_{Y^*}^Y e^{a(z)} dz \right) \\ &= \frac{A(Y)}{e^{a(Y)}} \int_{Y^*}^Y e^{a(z)} dz > 0. \end{split}$$

Thus, the reported income itself is an increasing function, and Assumption A1 is fulfilled. Figure 1 is obtained by applying the formula in (4) to our experimental setting with numerical calculation of the integral.

Appendix B. Theoretical analysis of Human condition

We derive the perfect Bayesian equilibrium (PBE) where taxpayers do not randomize their reporting behaviors. Let $\gamma(Y)$ be the reported income of a type Y taxpayer in the PBE. In the PBE, the auditor must hold a belief that is consistent with the reporting behavior $\gamma(\cdot)$ and choose one of the four taxpayers in order to maximize the expected tax revenue (including penalty).

Our experimental setting assumes the following environment.

Assumption 2. Scarcity condition: n > q.

The following proposition holds true.

Proposition 2. The following profile of the strategies of taxpayers and the auditor and the belief of the auditor constitutes a PBE.

For any Y, γ(Y) = 0,
 For any reported income profile (R₁,...,R_n), the auditor chooses one taxpayer with probability 1 / n,
 For reported income profile (0,0,...,0), the auditor believes that every taxpayer perfectly cheats, and
 For reported income profile (R₁,...,R_n) except for (0,0,...,0), the auditor

Proof. We first check the rationality of the taxpayers' reporting behavior. Since the auditor randomly chooses the target, and the scarcity condition holds, the optimal behavior of a taxpayer is to cheat fully. Next, we check the rationality of the auditor's belief. Since every taxpayer fully cheats, the auditor's belief is consistent with the taxpayer's behavior. Finally, we check the rationality of the auditor's behavior. According to the auditor's belief, the expected payoff of the audit is the same across the taxpayers. Thus, random audit is the optimal strategy for the auditor.

believes that every taxpayer sincerely reports his/her true income.

Thus, there exists a pooling equilibrium where every taxpayer cheats perfectly. This need not be the unique PBE. However, we do not try to explore the PBE further because this is quite difficult compared to theoretical studies such as Reinganum and Wilde (1986); when there are multiple taxpayers, the auditor must choose one of them for any tax reported profile. This implies that the auditor does not choose the probability of audit for each reported income, and the auditor face a fixed budget constraint. This further implies that the auditor's problem is the maximization of total tax revenue subject to this budget constraint.

Nonetheless, we posit that the theoretical prediction for human condition is sufficient because the theorem derives at least one type of PBE, and we already know that all the PBEs of the human condition are inferior to the cut-off rule in terms of tax revenue. Cut-off is the tax revenue maximizing audit rule, and the human auditor does not credibly implement the cut-off as we explained in Section 2.3.

Appendix C. Nonparametric tests using four classifications of income

		Income classification				
Scheme	All data	$10 \le Y \le 250$	$260 \le Y \le 500$	$510 \le Y \le 750$	$760 \le Y \le 1000$	
Random	21.8	28.8	28.6*	14.0**	17.9	
Cut-off	31.7	39.6	34.5	37.4	15.5	
LIRA	18.3*	48.6	12.9^{***}	9.3***	5.2^{***}	
Human_1	22	40.5	23.2	13.3**	11.1**	
Human_2	23.9	43.9	24.7	13.9**	13.6	

Table A1. Frequency of truthful reporting according to income classification

Notes: a) Bold indicates the highest ranked scheme in the column. b) * p < 0.10, ** p < 0.05, *** p < 0.01, relative to the highest ranked scheme in the column.

Table A2. Reporting percentages according to income classification

		Inc	ome classifica	tion	
Scheme	All data	$10 \le Y \le 250$	$260 \le Y \le 500$	$510 \le Y \le 750$	$760 \le Y \le 1000$
Random	51.0***	48.3***	49.3***	57.2	48.0*
Cut-off	59.3	53.4***	56.3	63.1	64.2
LIRA	66.3	80.5	69.9	61.7	54.7*
Human_1	58.4	62.8**	60.7	56.3	53.3*
_Human_2	62.6	69.2**	62.3	60.1	59.0

Notes: a) Bold indicates the highest ranked scheme in the column. b) * p < 0.10, ** p < 0.05, *** p < 0.01, relative to the highest ranked scheme in the column.

References

- Allingham, M.G., Sandmo, A., 1972. Income tax evasion: A theoretical analysis. J. Public Econ. 1, 323-338.
- Alm, J. 1988. Uncertain tax policies, individual behavior, and welfare. Am. Econ. Rev. 78:237-245.
- Alm, J., 2012. Measuring, explaining, and controlling tax evasion: Lessons from theory, experiments, and field studies. Int. Tax. Public Finance 19, 54-77.
- Alm, J., McKee, M., 2004. Tax compliance as a coordination game. J. Econ. Behav. Org. 54, 297-312.
- Andreoni, J., Erard, B., Feinstein, J., 1998. Tax compliance. J. Econ. Lit. 36, 818-860.
- Baldry, J.C., 1986. Tax evasion is not a gamble: A report on two experiments. Econ. Letters 22, 333-335.
- Andreoni, J., Miller, J.H., 1993. Rational cooperation in the finitely repeated prisoner's dilemma: Experimental evidence. Econ. J. 103 (418), 570-585.
- Braithwaite, V. 2003. Dancing with tax authorities: Motivational postures and non-compliant actions. In V. Braithwaite (ed.), Taxing democracy. Understanding tax avoidance and tax evasion. Aldershot: Ashgate.
- Clark, J., Friesen, L., Muller, A., 2004. The good, the bad, and the regulator: An experimental test of two conditional audit schemes. Econ. Inq. 42, 69-87.
- Collins, J.H., Plumlee, R.D., 1991. The taxpayer's labor and reporting decision: The effect of audit schemes. Account. Rev. 66, 559-576.
- Coricelli, G., Joffily, M., Montmarquette, C., Villeval, M., 2010. Cheating, emotions, and rationality: An experiment on tax evasion. Exp. Econ. 13, 226-247.
- Eckel, C.C., Grossman, P.J., 2008. Chapter 113 Men, women and risk aversion: Experimental evidence, in: Charles, R.P., Vernon, L.S. (Eds.), Handbook of Experimental Economics Results. Elsevier, pp. 1061-1073.
- Fischbacher, U., 2007. z-Tree: Zurich toolbox for ready-made economic experiments. Exp. Econ. 10, 171-178.
- Friesen, L., 2003. Targeting enforcement to improve compliance with environmental regulations. J. Environ. Econ. Manage. 46, 72-85.
- Gërxhani, K., 2004. Tax evasion in transition: Outcome of an institutional clash? Testing Feige's conjecture in Albania. Eur. Econ. Rev. 48, 729-745.
- Gërxhani, K., Schram, A., 2006. Tax evasion and income source: A comparative experimental study. J. Econ. Psychol. 27, 402-422.
- Greenberg, J., 1984. Avoiding tax avoidance: A (repeated) game-theoretic approach. J.

Econ. Theory 32, 1-13.

- Harrington, W., 1988. Enforcement leverage when penalties are restricted. J. Public Econ. 37, 29-53.
- Harsanyi, J.C., 1967. Games with incomplete information played by "Bayesian" players, I–III Part I. The Basic Model. Manage. Sci. 14, 159-182.
- AERA, 2014. Is it impossible to predict the targets of tax investigation selected by KSK system? The Asahi Shinbun Company. Retrieved from http://dot.asahi.com/aera/2014020600045.html
- Kastlunger, B., Kirchler, E., Mittone, L., Pitters, J., 2009. Sequences of audits, tax compliance, and taxpaying strategies. J. Econ. Psychol. 30, 405-418.
- Kirchler, E. 2007. The Economic Psychology of Tax Behavior Cambridge University Press, NY.
- Lang, M., Wambach, A., 2013. The fog of fraud: Mitigating fraud by strategic ambiguity. Games Econ. Behav. 81, 255-275.
- Lefebvre, M., Pestieau, P., Riedl, A., Villeval, M., 2014. Tax evasion and social information: An experiment in Belgium, France, and the Netherlands. Int. Tax. Public Finance, 1-25.
- Long, S.B. and Swingen, J.A., 1987. An approach to the measurement of tax law complexity. J. Am. Tax. Assoc. 8, 22–36.
- OECD, 2007. Tax Administration in OECD and Selected Non-OECD Countries: Comparative Information Series (2006). Retrieved from http://www.oecd.org/ctp/administration/CIS-2006.pdf
- Reinganum, J.F., Wilde, L.L., 1986. Equilibrium verification and reporting policies in a model of tax compliance. Int. Econ. Rev. 27, 739-760.
- Sánchez, I., Sobel, J., 1993. Hierarchical design and enforcement of income tax policies. J. Public Econ. 50, 345-369.
- Slemrod, J., 2007. Cheating ourselves: The economics of tax evasion. J. Econ. Perspect. 21, 25-48.
- Webber, C., Wildavsky, A.B., 1986. A history of taxation and expenditure in the Western world. Simon and Schuster.
- Yitzhaki, S., 1974. Income tax evasion: A theoretical analysis. J. Public Econ. 3, 201-202.
- Yoon, S.-S., Yoo, S.-W., Kim, J., 2011. Ambiguity, audit errors, and tax compliance. Asia-Pac. J. Account. Econ. 18, 11-26.