A competitive audit selection mechanism with incomplete information

Miloš Fišar *[1], Ondřej Krčál[1], Jiří Špalek[1], Rostislav Staněk[1], and James Tremewan[2]

[1] *Masaryk University, Brno, Czech Republic.*
[2] *HSE University, Moscow, Russia.*

December 3, 2025

Supplementary Materials

*Corresponding author. Email: milos.fisar@econ.muni.cz.

A Additional Results

Table A1: Average declared income - All rounds

	(1)	(2)	(3)	(4)	(5)	(6)
Signal	10.21**	10.65**	10.13**	-15.43***	-16.04***	-16.19***
	(3.604)	(3.495)	(3.294)	(1.648)	(1.556)	(1.617)
Random	-8.420+	-8.616*	-9.960*	-10.46***	-11.83***	-12.40***
	(4.297)	(4.175)	(4.040)	(2.632)	(2.615)	(2.671)
Income		0.690***	0.690***	0.582***	0.581***	0.578***
		(0.0247)	(0.0247)	(0.0271)	(0.0272)	(0.0278)
Signal#Income				0.261***	0.262***	0.267***
				(0.0337)	(0.0337)	(0.0345)
Random#Income				0.0187	0.0181	0.0235
				(0.0502)	(0.0500)	(0.0509)
Female			1.493		1.787	2.144
			(2.273)		(2.242)	(2.255)
Age			-0.765+		-0.792+	-0.736+
			(0.419)		(0.409)	(0.417)
Economics student			-0.128		-0.196	-0.0451
			(2.023)		(2.027)	(2.027)
Audit in *t-1*						-2.245*
						(0.970)
Constant	61.50***	-7.764**	9.293	3.117*	20.71*	20.30*
	(2.673)	(2.916)	(9.823)	(1.377)	(9.759)	(9.998)
Observations	5982	5982	5982	5982	5982	5783
R^2	0.021	0.647	0.648	0.667	0.669	0.671

Note: OLS regressions with robust standard errors clustered at the group level. Observations, where income $I = 0$, are excluded. $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A2: Expected total revenue - All rounds

	(1)	(2)	(3)	(4)	(5)	(6)
Signal	4.699***	5.034***	5.228***	-10.32***	-10.17***	-10.25***
	(0.926)	(0.599)	(0.569)	(0.598)	(0.610)	(0.624)
Random	-1.002	-1.153	-1.263	-7.580***	-7.720***	-7.904***
	(1.328)	(0.847)	(0.817)	(0.557)	(0.578)	(0.597)
Income		0.532***	0.532***	0.456***	0.456***	0.455***
		(0.0115)	(0.0115)	(0.00796)	(0.00795)	(0.00820)
Signal#Income				0.154***	0.154***	0.155***
				(0.00943)	(0.00943)	(0.00969)
Random#Income				0.0641***	0.0641***	0.0660***
				(0.0116)	(0.0116)	(0.0119)
Female			1.413**		1.559***	1.570***
			(0.430)		(0.419)	(0.422)
Age			0.0256		0.0143	0.00361
			(0.0781)		(0.0762)	(0.0767)
Economics student			-0.0609		-0.110	-0.111
			(0.454)		(0.447)	(0.443)
Audit in *t-1*						0.0153
						(0.244)
Constant	51.86***	-1.508	-2.787	6.111***	5.064**	5.389**
	(0.719)	(1.144)	(2.119)	(0.331)	(1.836)	(1.862)
Observations	5982	5982	5982	5982	5982	5783
R^2	0.006	0.938	0.938	0.953	0.953	0.953

Note: OLS regressions with robust standard errors clustered at the group level. Observations, where income $I = 0$, are excluded. $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A3: Inequality (Gini coefficient of net income) - All rounds

	(1)	(2)	(3)
Signal	-0.0620***	-0.0657***	-0.0689***
	(0.00945)	(0.00850)	(0.00910)
Random	0.00639	0.00165	0.00428
	(0.0120)	(0.0126)	(0.0127)
Income		0.590***	0.589***
		(0.0278)	(0.0276)
Female			-0.0226
			(0.0192)
Age			-0.000181
			(0.00338)
Economics student			0.0168
			(0.0215)
Constant	0.354***	0.189***	0.196*
	(0.00593)	(0.00980)	(0.0837)
Observations	1200	1200	1200
R^2	0.080	0.381	0.383

Note: OLS regressions with robust standard errors clustered at the group level. Observations, where income $I = 0$, are excluded. $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

B Theoretical framework

B.1 Model description

This section describes the theoretical model that underlies the experiment. Taxpayer i receives an income I_i drawn form a distribution $F(I)$ with support $[0, \overline{I}]$. The taxpayer chooses the declared income $R_i \in \langle 0, I_i \rangle$. Declared income is taxed at rate τ, so the taxpayer pays a tax τR_i. The taxpayer i is audited with a probability $\pi_i(R_i, R_{-i})$. The formula for π_i depends on the audit selection mechanism used by the tax authority. If the taxpayer is chosen for audit, she pays a fine which depends on the undeclared income, $\phi(I_i - R_i)$ where ϕ is the fine rate and $\phi > \tau$.

We examine three different audit selection mechanisms (ASM) that correspond to our three treatments: *random*, *no-signal*, and *signal*. Under the *random* ASM, the audit probability is the same for all taxpayers regardless of their declared income, $\pi_i = p$. When the *no-signal* ASM is applied, the audit probability depends on both the income declared by the taxpayer and the income declared by other taxpayers. In particular, the form of our audit selection rule resembles the generalized audit selection rule proposed by **?** closely. It has the following form:

$$\pi_i = p - \delta \left(R_i - \frac{\sum R_{-i}}{N-1} \right),$$

where N is the number of taxpayers in a group. This mechanism assigns higher audit probability to the taxpayers with relatively lower declared income. Parameter p defines the basic audit probability, and parameter δ defines the sensitivity of the audit selection rule to the declared income. The random audit selection mechanism is obviously a special case of the audit selection mechanism for which $\delta = 0$.

The *signal* ASM has the same functional form, but it uses undeclared income instead of declared income:

$$\pi_i = p + \delta \left(Z_i - \frac{\sum Z_{-i}}{N-1} \right),$$

where $Z_i = I_i - R_i$ is the undeclared income. The formula for this audit selection mechanism can be rewritten as

$$\pi_i = p - \delta \left(R_i - \frac{\sum R_{-i}}{N-1} \right) + \delta \left(I_i - \frac{\sum I_{-i}}{N-1} \right).$$

We can see that the difference between the *signal* and *no-signal* ASMs is the last term in the previous equation. This term shows that the *signal* ASM accounts also for the differences in actual incomes.

The important aspect of the *no-signal* ASM is that the audit probability depends only on the difference between the individual's declared income and the average declared income of other taxpayers. Therefore, thisASM does not require information concerning each taxpayer's income or the income distribution in their peer group. On the other hand, it loses a substantial amount of information since the differences in the declared income may be caused not only by different compliance but also by differences in actual incomes. While the *signal* ASM can differentiate between these effects, the *no-signal* ASM cannot.

B.2 Equilibrium

Risk-neutral taxpayer i chooses the declared income R_i in order to maximize the expected wealth:

$$W = (1 - \pi(R_i, R_{-i}))(I_i - \tau R_i) + \pi(R_i, R_{-i})((1 - \tau)R_i + (1 - \phi)(I_i - R_i)).$$

Under the *random* ASM, the first derivative of the expected wealth can be written as $p\phi - \tau$. Suppose that the tax rate τ is higher than $p\phi$. If this is the case, then the optimal choice of the risk-neutral taxpayer is to evade paying taxes on all of her income.

Now, we derive the solution for the *no-signal* ASM. Suppose that the other player's strategy is $R(I_i)$, where R is a non-decreasing function. The first order condition is given as follows:

$$\left(p - \delta R_i + \delta \int R(I_i)dF(I)\right)\phi - \tau + \delta\phi(I_i - R_i) = 0$$

We focus on the symmetric equilibria of the model. The solution of the first order condition is a linear function $R = a + bI$. Because the declared income is constrained to the interval $[0, I_i]$, the interior solution need not be feasible for taxpayers with sufficiently low income. The constrained optimum is obtained by projecting the unconstrained solution onto the feasible interval. Thus, if $a + bI > I$ the constrained solution is to declare the actual income $R_i = I_i$; and if $a + bI < 0$ the constrained solution is to evade all income $R_i = 0$. Hence, depending on the value of exogeneous parameters τ, p, δ and ϕ, the model yields two possible equilibrium reporting strategies[1].

The first equilibrium occurs when the tax rate is low compared to the audit probability and the fine. In this situation, the low-income taxpayers declare their whole income, as they face a high probability of being audited. The taxpayers with higher income face lower probabilities of being audited, and they optimally react by evading being taxed on some

[1]These strategies do not form multiple equilibria. Only one of these strategies forms an equilibrium depending on model parameters.

amount of their income. Formally, the equilibrium strategy has the following form:

$$R(I_i) = \begin{cases} I_i & \text{if } I_i < \hat{I} \\ a + bI_i & \text{if } I_i \geq \hat{I}, \end{cases} \tag{1}$$

where $a > 0$. The equilibrium values of the parameters a, b and $\hat{I}$ are given by the solutions of the following three conditions (the last condition is derived from the equation $a + b\hat{I} = \hat{I}$), respectively:

$$a = \frac{p\phi - \tau}{\delta\phi(1 + F(\hat{I}))} + \frac{F(\hat{I})}{1 + F(\hat{I})} E(I|I < \hat{I}) + \frac{1 - F(\hat{I})}{2(1 + F(\hat{I}))} E(I|I > \hat{I})$$

$$b = \frac{1}{2}$$

$$F(\hat{I})E(I|I < \hat{I}) + \frac{(1 - F(\hat{I}))}{2} E(I|I > \hat{I}) - \frac{1 + F(\hat{I})}{2} \hat{I} = \frac{\tau - p\phi}{\delta\phi}$$

The second equilibrium is applicable when the tax rate is relatively high compared to the audit probability and fine. The taxpayers with lower income evade paying taxes on all their income, as they risk losing a small amount of money if they are audited. On the other hand, taxpayers with high income have a lot of money at stake, so they will optimally declare some income. In formal terms, the equilibrium strategy has the following form:

$$R(I_i) = \begin{cases} 0 & \text{if } I_i < \hat{I} \\ a + bI_i & \text{if } I_i \geq \hat{I} \end{cases} \tag{2}$$

The equilibrium values of the parameters a, b and $\hat{I}$ are given by the solution of the following three conditions (the last condition is derived from the equation $a + b\hat{I} = 0$), respectively:

$$a = \frac{p\phi - \tau}{\delta\phi(1 + F(\hat{I}))} + \frac{1 - F(\hat{I})}{2(1 + F(\hat{I}))} E(I|I > \hat{I})$$

$$b = \frac{1}{2}$$

$$\frac{(1 - F(\hat{I}))}{2} E(I|I > \hat{I}) - \frac{1 + F(\hat{I})}{2} \hat{I} = \frac{\tau - p\phi}{\delta\phi}$$

Both equilibria under the *no-signal* ASM result in higher tax compliance compared to using random audits. This result forms part of the Hypothesis 1.

Now, consider the case of the *signal* ASM. The best-response function is defined implicitly by the following first-order-condition in which $Z(I)$ denotes the equilibrium strategy of the other players:

$$\tau - \left(p + \delta Z_i - \delta \int Z(I) dF(I)\right) \phi - \delta \phi Z_i = 0$$

The best-response yields that taxpayers' strategy is to evade taxes on some fixed amount c which means that the declared income is $R_i^* = I_i - c$. Because the declared income is constrained to the interval $[0, I_i]$, this solution is not feasible for taxpayers with low income. The constrained optimum is obtained by projecting the unconstrained solution onto the feasible interval: thus if $I_i < c$ the constrained optimum is $Z = I_i$.

$$Z(I_i) = \begin{cases} I_i & \text{if } I_i < c \\ c & \text{if } I_i \geq c \end{cases} \tag{3}$$

In symmetric Nash equilibrium, the amount of evaded taxes c is determined by the following equation, where $F(\dot{})$ is the cumulative distribution function of actual income:

$$\tau - \left(p + \delta Z_i - \delta \left(\int_0^c I dF(I) + (1 - F(c))c\right)\right) \phi - \delta \phi Z_i = 0.$$

We can observe one stark contrast between the equilibrium strategies under *no-signal* and *no-signal* ASMs. The slope of the equilibrium strategy is lower for the *no-signal* ASM. Therefore, higher-income taxpayers evade more taxes when estimates of individuals' incomes are missing.

C Instructions

Here we provide the instructions for all treatments (translated, original in Czech). Instructions change according to the treatment. The changes between treatments are indicated in *italics*.

Introduction

In this experiment, we study your decision-making as individuals as well as groups. Your earnings are based on your decisions. Therefore, we recommend you to read the following instructions carefully. You will be paid at the end of the experiment in cash and in private.

You will make your decisions without communication with the other participants in the experiment. If you have any question, please raise your hand, and an experimenter will come to answer it in private.

Please do not communicate with other participants during the experiment, do not use your mobile phone or other devices than the computer you are seated at and pay attention to the experiment. In case of disobedience, you will be excluded from the experiment without any reward.

The course of the experiment

(signal and no-signal) The experiment will take place in groups of five people: you and four other participants. Players in your group are in this room, but we don't tell you who belongs to your group. The group is randomized at the beginning of the experiment and does not change throughout the experiment.

The experiment consists of 30 identical rounds. At the beginning of each round, you receive an amount from the range of 0 to 200 CZK from us. Each 1 CZK may be selected with the same probability. You can imagine the amount selection as a draw from a hat that contains 201 balls with numbers from 0 to 200.

In the next step, you will be asked to report the amount. This declared amount may be the same or lower than the amount you received. E.g. if you received 102 CZK, you can declare any integer number between 0 and 102 CZK. We will always deduct 60% from the declared amount and you will keep 40% of the declared amount. The amount of money that you received from us at the beginning of the round and did not report is called the undeclared amount. The fate of the undeclared amount depends on chance. With some probability an adverse event occurs and you lose the entire undeclared amount. Otherwise, you keep the undeclared amount.

(random) The basic probability that an adverse event will occur and you lose the unreported amount is 40%. Whether an adverse event occurs will be drawn in each round again.

(signal) The basic probability that an adverse event occurs and the undeclared amount is lost is 40%. Besides, for each 10 CZK by which your ***undeclared amount*** *is lower than the average* ***undeclared amounts*** *of the four other players in your group, the probability of an adverse event decreases by 4 percentage points. On the other hand, each 10 CZK above the average* ***undeclared amounts*** *of the four other players in your group increases the probability of an adverse event by 4 percentage points.*

For example, if your undeclared amount is 81 CZK below the average ***undeclared amounts*** *of four other players in your group, the probability of an adverse event is* $81 \times 0.4 = 32.4$ *percentage points lower than the basic probability, that means* $40 - 32.4 = 7.6\%$*. On the other hand, if your* ***undeclared amount*** *is 35 CZK more than the average undeclared amounts of the four other players in your group, the probability will be* $40 + (35 \times 0.4) = 54\%$*. Whether the adverse event occurs is drawn in each round.*

(no-signal) The basic probability that an adverse event occurs and the undeclared amount is lost is 40%. Besides, for each 10 CZK by which your ***declared amount*** *is lower than the average* ***declared amounts*** *of the four other players in your group, the probability of an adverse event increases by 4 percentage points. On the other hand, each 10 CZK above the average* ***declared amounts*** *of the four other players in your group decreases the probability of an adverse event by 4 percentage points.*

For example, if your declared amount is 81 CZK above the average ***declared amounts*** *of the four other players in your group, the probability of an adverse event is* $81 \times 0.4 = 32.4$ *percentage points lower than the basic probability, that means* $40–32.4 = 7.6\%$*. On the other hand, if your* ***declared amount*** *35 CZK lower than the average declared amounts of the four other players in your group, the probability will be* $40 + (35 \times 0.4) = 54\%$*. Whether the adverse event occurs is drawn in each round.*

Payoffs

There are two possible scenarios in each round of this experiment:

1. An adverse event occurs. Your payoff for that round is 0.4 × declared amount.
2. No adverse event occurs. Your payoff will be 0.4 × declared amount + undeclared

amount.

(random) At the end of each round, you receive information about whether an adverse event occurred, and what your payoff was in the round. At the end of the experiment, you will receive your payoff in CZK from five randomly selected rounds of the experiment.

(signal and no-signal) At the end of each round, you receive information about whether an adverse event occurred, and what your payoff was in the round. We will also inform you of the average declared amount of other players and of the probability of the adverse event. At the end of the experiment, you will receive your payoff in CZK from five randomly selected rounds of the experiment.