Contributions to Game Theory and Management, XV, 155–177

Tax Authority and Taxpayers: How Does Mutual Collecting of Information Affect the Effectiveness of Tax Control

Suriya Sh. Kumacheva[1] **and Galina A. Tomilina**[2]

[1] *Faculty of Applied Mathematics and Control Processes, St. Petersburg State University, 7/9, Universitetskaya nab., St. Petersburg, 199034, Russia*
E-mail: s.kumacheva@spbu.ru

[2] *EPAM Systems, 12, Pevchesky per., St. Petersburg, 197101, Russia*
E-mail: g.tomilina@yandex.ru

Abstract Lots of works have been devoted to the modelling of tax control. In the classic game-theoretic formulation, the subject of study is the interaction between tax authority and taxpayers. However, along with such attitude to the problem, some works allow us to study auxiliary problems related to tax collection and tax control. Two of these issues are worth mentioning. The first is the dissemination of information about future tax audits among the taxable population. The second is the gathering and analysis of information on the tendency of taxpayers to evasion. It was shown that both processes can be used to adjust the tax authority's strategy and optimize audits in order to improve tax collection. However, the mutual collecting of information on tax audits by the taxpayers and the tax authority decreases the total tax revenue. The study of formulated problem is accompanied by the modelling of both processes, their simulation, and comparative analysis of the results.

Keywords: tax control, tax evasions, risk propensity, structured network, Bayesian approach, mutual collecting of information.

1. Introduction

In the modern world, information is a key factor influencing decision-making in the fields of economy, business, politics and other processes important for the development of society. The lack and incompleteness of information negatively affect the choice of a strategy of rational behavior, and sometimes force the parties to make decisions that are completely opposite to the optimal ones in their consequences. That is why in various management processes it is necessary to take into account the search, collecting and analysis of information as a part of a solution to a larger problem.

Since taxation is one of the most important areas of the state economy, regulation and social policy, many studies, such as (Antunes et al., 2006; Antocia et al., 2014; Bloomquist, 2006; Chander and Wilde, 1998; Lamantia and Pezzino, 2021; Macho-Stadler and Perez-Castrillo, 2002; Sanchez and Sobel, 1993; Vasin and Morozov, 2005), are devoted to the research of the tax system and its basic functions (Samuelson and Nordhaus, 2007): fiscal, social, regulatory and control.

https://doi.org/10.21638/11701/spbu31.2022.12

The mentioned classic works (Chander and Wilde, 1998; Sanchez and Sobel, 1993; Vasin and Morozov, 2005) apply different approaches to mathematical modelling of the studied problem: they can be represented by the mathematical decision theory and its separate branches that are contract theory and mathematical game theory. Most of these models are based on the hierarchical one-shot static games according to the "principal-to-agent" scheme, in which agents are assumed to be risk-neutral. But in (Antunes et al., 2006; Antocia et al., 2014; Bloomquist, 2006; Lamantia and Pezzino, 2021; Kumacheva and Gubar, 2015) the heterogeneity of the taxable population is emphasized.

The classic approach associated with the decision theory suggests the following forms of solutions (which are the modifications and extensions of similar formalizations): "agent's optimal reporting rule" (Chander and Wilde, 1998), "optimal contract" (Sanchez and Sobel, 1993; Chander and Wilde, 1998), "threshold rule" (Vasin and Morozov, 2005) etc. But it would be more realistic to take into account that the tax authority has a significantly limited budget. This makes achieving optimal values for audit probabilities an extremely rare event. Therefore, the tax authority needs to find the way to optimize tax collections. One of such possible ways, the using an additional information about taxpayers, was mentioned in (Macho-Stadler and Perez-Castrillo, 2002) and studied further in (Bure and Kumacheva, 2005) and (Kumacheva, 2012). The issue of the exchange and dissemination of information on tax audits between taxpayers was also investigated in (Gubar et al., 2016; Gubar et al., 2017; Kumacheva et al., 2018) and (Gubar et al., 2019). The necessity to analyze and collect information is also caused by the fact that most game-theoretic models are assumed to be games with complete information, which is hardly realizable in the real world.

The current work offers a new approach to the investigation of information as a key factor in tax control models. On the one hand, the tax authority (principal) needs to collect and analyze information about the heterogeneous taxable population. On the other hand, taxpayers (agents) also gather and disseminate information about future tax audits. Thus, the interaction of the tax authority and taxpayers, as before, obeys to a hierarchical scheme. But now it is proposed to be considered taking into account the mutual collecting of information by the both parties of the process.

This paper is structured as follows. Section 2 is devoted to the brief description of the static model that lays in the basis of our research and discussion of the most important problems of this model. In Section 3 we propose the model of interaction between tax authority and taxpayers based on the repeated hierarchical "principal-to-agent" game. Section 4 describes the structure of the taxable population in assumption of its heterogeneity. In Section 5 we consider the process of auditing from the standpoints of principal and agents and formulate several assumptions which are useful to study the principle of mutual collecting of the information. Section 6 represents the numerical experiments and parametric analysis based on the mentioned assumptions. And, finally, in Section 7 the conclusions we made from our modelling and simulating are formulated.

2. Static Model

Further research is based on a static model similar to the game-theoretic model of tax control (Bure and Kumacheva, 2010). Following (Chander and Wilde, 1998;

Vasin and Morozov, 2005; Lamantia and Pezzino, 2021) and (Kumacheva et al., 2019), here and below we suppose that there are only two possible levels of income that can be earned. Without loss of generality, a larger number of income gradations can be considered, as it was done, for example, in (Bure and Kumacheva, 2005) or (Kumacheva, 2012). But the basic premises and way of reasoning will remain the same. Therefore, we are restricting the number of gradations by two, formally classifying taxpayers into rich and poor.

2.1. The Model of Tax Control with Two Levels of Income of Taxpayers

Considering the model represented in (Bure and Kumacheva, 2010) as a pattern of interaction between principal and risk-neutral agents, let's assume that there is a population of n taxpayers. Each of them can be characterized by her/his true income ξ and can declare income η, where $\eta \leq \xi$. Incomes ξ and η can take two values, L or H, where $0 < L < H$. Therefore, there are two types of taxpayers: with high and low levels of income. The cardinalities of the corresponding sets are n_H and n_L: $n_L + n_H = n$.

Another important characteristic of auditing process is probability P_L of audit of those agents, who declared $\eta = L$. If the agent is a tax evader $(\eta(H) = L))$ and his evasion was revealed, she/he should pay $(\theta + \pi)(\xi - \eta)$, where constants θ and π are tax and penalty rates correspondingly.

Summarizing and simplifying the results obtained in (Bure and Kumacheva, 2010), we formulate that risk-neutral taxpayers decide whether to evade or not, depending on whether the actual probability P_L of audit is greater or less the value

$$P^* = \frac{\theta}{\theta + \pi},$$

called the optimal value of auditing probability.

2.2. Problems of the Model

Analysing the static model described above and summarizing all the weaknesses of such models, mentioned in Section 1, we can formulate the most important problems underlying the game-theoretic approach to tax control modelling.

The **Problem 1** is related to the static structure of the considered model. Such formulation can be used to study only one tax period without answering the question "what happens next?". How the results of interaction between the tax authority and taxpayers can be taken into account and used to describe it during the future periods? But it is more significant to investigate this process in dynamics taking into account the results of each tax period. The idea to consider the taxable population as an evolutionary system was proposed previously in such works as (Antocia et al., 2014; Antunes et al., 2006) and studied further in (Kumacheva and Gubar, 2015; Gubar et al., 2017; Gubar et al., 2019) and (Kumacheva and Tomilina, 2021).

Problem 2 is in the assumption about risk-neutrality of taxpayers according to the fact that the tax authority knows nothing on the population structure. But in a lot of studies the society is assumed to be heterogeneous. For example, (Antocia et al., 2014; Antunes et al., 2006; Bloomquist, 2006) and (Lamantia and Pezzino, 2021) consider population of agents with various propensities to risk. Such ideas were investigated in (Kumacheva and Gubar, 2015; Kumacheva et al., 2018;

Kumacheva et al., 2019) and (Gubar et al., 2019). In reality there is no information about each agent's risk propensity. But the idea to analyse and use additional information about taxpayers were discussed in many works, such as (Macho-Stadler and Perez-Castrillo, 2002), (Bure and Kumacheva, 2005; Kumacheva, 2012) and (Kumacheva and Tomilina, 2021).

The **Problem 3** is in the fact that the game-theoretic formulation of the problem (Chander and Wilde, 1998; Sanchez and Sobel, 1993; Vasin and Morozov, 2005; Bure and Kumacheva, 2010) used to research cases with complete information. That is, taxpayers are assumed to choose their strategy as the best response on the strategy of the tax authority. But in practice, similar information can be only indirect. Accordingly, agents do not know the actual value of probability of audit and can only estimate it. Some assumptions regarding the methods of such an assessment have been made in (Gubar et al., 2016; Kumacheva et al., 2019; Gubar et al., 2020; Kumacheva et al., 2020) and (Kumacheva and Tomilina, 2021).

The **Problem 4** can be formulated as the limited budget of the tax authority, which makes reaching the optimal (threshold) value extremely rare event. Therefore, the tax authority needs to find the way to optimize tax collections. This problem was discussed previously in (Kumacheva, 2012; Kumacheva et al., 2018; Kumacheva et al., 2019).

The indicated problems lead to the logical conclusion that information is a key factor in modelling of tax control. Previously, the issue of disseminating information about possible tax audits was explored as a way to stimulate agents to pay taxes honestly. Various concepts have been proposed to simulate the process of information spreading. For example, an application of epidemic processes, based on SIR and SIS models was investigated in (Gubar et al., 2016) and (Gubar et al., 2017). Another attitude, based on evolutionary processes on the network, was considered in (Gubar et al., 2017; Kumacheva et al., 2018) and (Kumacheva et al., 2019). The ideas of opinion dynamics based on Markov process were studied in (Gubar et al., 2019). Finally, the idea of Bayesian updating of the information about taxpayers' propensities to evasion was studied in (Bure and Kumacheva, 2010; Kumacheva and Gubar, 2015) and (Kumacheva and Tomilina, 2021). But now we should generalize these ideas into the model of interaction between the tax authority and taxpayers, in which both contreparties choose their strategies basing on the received information.

3. The Interaction Model: Repeated Game

The static model described in Section 2 allows one to study the one-stage process, i.e. one occurrence of auditing. Thus, the **problem 1**, described in the previous section, is formed. An extension of this model allows us to consider a multistage interaction between the tax authority (principal) and the risk-neutral taxpayer (agent). Thus, we can represent an audit process as a repeated game. At each stage the taxpayer has two strategies, that are, to evade taxation $\eta(H) = L$ or to be honest $\eta(H) = H$ (let's denote them **e** and **h** correspondingly).

The tax authority can choose, to audit (**a**) or to skip inspection (**s**). The Figure 1 shows the scheme of such multistage interaction. The circles in the figure represent the vertices in which the choice of strategy is made by taxpayers, the squares represent the vertices corresponding to the moves of the tax authority.

The previously studied static models do not answer the question of how the situation at each stage affects the choice of a rational agent's strategy at subsequent

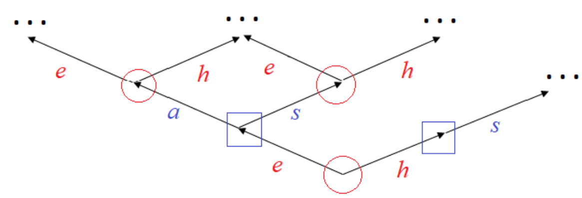

Fig. 1. Multistage game between tax authority and risk-neutral taxpayer

stages of the game. In fact, there can be two scenarios for a taxpayer: the behaviour of a bad pupil or of a clever pupil. Bad pupil will be reasoning like this: "If I had been audited and detected as an evader, I was fined and punished enough, therefore, the next time I won't be audited once again and I can evade". Clever pupil will have another way of reasoning: "If I had been audited and detected as an evader, then the next time I will be audited once again with high probability". We assume that the second scenario seems to be more realistic in accordance with the ideas of bounded rationality, and further we will consider it as a basis.

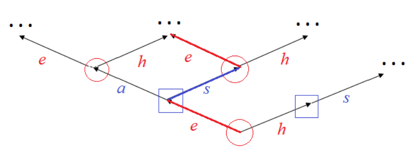

Fig. 2. Multistage game: the case of unrevealed evasion

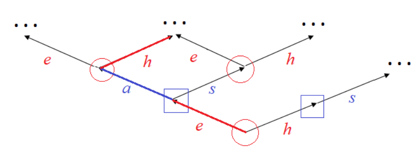

Fig. 3. Multistage game: the case of revealed evasion

The Figure 2 illustrates the scheme of interaction for the case when the evader was not audited and, therefore, continues to evade. Another case, when the tax evasion is revealed, is represented in the Figure 3. Extending the described scheme to subsequent stages of interaction, it is possible to study the following tax periods, occurrences of evasion and audit.

4. The Structure of Population

Now let's study the **problem 2**, indicated in Subsection 2.2: the assumption of risk-neutrality of the entire taxable population significantly removes the studied model from the real processes of tax control. To avoid this disadvantage, we need to differ agents by their various risk propensities.

As it was assumed in (Kumacheva and Gubar, 2015; Kumacheva et al., 2018; Kumacheva et al., 2019), let's suppose that real taxable population consists of three subgroups: risk-averse (ν_a), risk-neutral (ν_n) and risk-loving (ν_l), where $\nu_a + \nu_n + \nu_l = 1$. These subgroups can be characterized by the various agents' behaviour profiles in the similar external conditions and, therefore, by the different response to the same information.

Let γ_L be the portion of those who declared their income as L and γ_H — of those who declared H. The total population consists of these two portions: $\gamma_L + \gamma_H = 1$.

5. The Auditing Process

In the current study the tax authority and taxpayers are supposed to interact to each other according to the repeated game represented in Section 3. Therefore, the auditing is sophisticated process which can be described from two different standpoints.

Above all, let's investigate this process analyzing the tax authority's behaviour. The fact that the principal knows nothing about distribution of the risk-propensities of the population should be taken into account. This can be mentioned as the deepening and expanding the **problem 2** discussed previously.

By the other hand, the **problem 3** also appears: the taxpayers do not have any information on the actual value of probability P_L. This indicates the need to solve the problem of incomplete information in the game being studied.

5.1. Auditing from the Principal's Standpoint

The tax authority has two main purposes providing tax auditing. The first one is obviously the **tax collection** (Samuelson and Nordhaus, 2007): the principal's purpose is to collect taxes and penalties from the taxable population and, therefore, to increase the **net tax revenue**

$$R = n\,[\gamma_H\,\theta\,H + \gamma_L(\theta\,L - P_L\,c)] + n_H\,\gamma_L\,P_L(\theta + \pi)(H - L), \tag{1}$$

where c is the unit cost of audit.

But in real state economy the part of public funds, allocated for tax control, is strongly limited (**problem 4** formulated in Subsection 2.2). Let's denote this amount as B – the budget of the tax authority.

It means that the tax authority's goal is to maximize the net tax revenue (1)

$$\max_{P_l \le P^*} R(P_L)$$

with restriction

$$P_L \le \frac{B}{n\gamma_L\,c}, \tag{2}$$

where $\gamma_L > 0$. The assumption that this inequality is fulfilled seems to be natural, because within the studied model, there always exist taxpayers who declared $\eta = L$.

The second tax authority's purpose is the **population monitoring** (Macho-Stadler and Perez-Castrillo, 2002; Bure and Kumacheva, 2005; Kumacheva, 2012). It means that the tax auditing is used to update information about audited taxpayers and, thus, about their disposition to evasion. Finally, it helps to refresh the information about the distribution of risk propensities among the population. Further we suppose that the structure of this process can be represented by the *Bayesian updating*, which takes into consideration the prior and posterior probabilities based on the principal's knowledge about each taxpayer ex-ante and ex-post correspondingly.

Bayesian Updating. Let's focus on the second purpose, population monitoring. To formalize the process of information collecting based on the observation of data on taxable population, we assume the Bayesian approach (De Groot, 1970; Zellner, 1997) as the mathematical base of this process. To study tax auditing from this position, we need to take into consideration the term *Tax Story* which represents an accumulated knowledge about each agent's propensity to evade. To construct the k-th taxpayer's tax story, we can consider the method previously proposed in (Bure and Kumacheva, 2010) and investigated further in (Kumacheva and Tomilina, 2021).

Within the framework of the current study, the taxpayers are assumed to have one of three types in accordance with their risk propensity. At the same time, we suppose that the agents' risk statuses are associated with their disposition to evade taxation. In reality, there can be much more gradations of this tendency to evasion than only three mentioned classes of risk propensity. Accordingly, the number of scenarios for the auditing of taxpayers with various risk statuses is significantly increasing.

Let W_k be the random variable, which characterizes the k-th taxpayer's risk propensity (or, in the discussed problem, disposition to evade). It is assumed to be beta-distributed with parameters α_k, β_k. As this agent's *tax story* we will determine a characteristic of the taxpayer's behaviour in the previous periods, which is a result of observation (inspection), presented as a Bernoulli-distributed random variable X_j (where j is the number of current tax period):

$$X_j = \begin{cases} 1, & \text{evasion} \\ 0, & \text{no evasion.} \end{cases} \tag{3}$$

After each tax period j the information about audited taxpayers is updated, thus, posterior disposition to evade is recalculated taking into account the agent's tax story.

Next, we can turn to the theorem from (De Groot, 1970), which is quoted verbatim in the appendix 8.2. That is, it can be concluded that the family of beta distributions is conjugate to the family of Bernoulli distribution — see the appendix 8.1.

Now the procedure of Bayesian Updating for the taxpayer k can be generalized in the following **algorithm**:

– Let's consider initial state (period $j = 1$): suppose that this taxpayer was not audited in the previous periods, therefore, there is no information of her/his risk propensity.
– As far as prior information is absent, let's assume that W_k is a uniform distributed (beta distributed with parameters $\alpha_{0_k} = 1$ and $\beta_{0_k} = 1$).
– The tax authority audits a population with some fixed probability P_L.
– The tax story is formed as a result of observation (inspection), presented as a Bernoulli-distributed random variable X_j, (here $j = 1$).
– Using the tax story (3) and applying the mentioned theorem about conjugate distributions (De Groot, 1970), tax authority can compute the new values of parameters α_{j_k}, β_{j_k} of a posteriori distribution of W_k, which can be considered as a priori distribution for the next tax period.
– Starting the next iteration ($j = 2 \ldots$), related to the next tax period.

It should be noted that if there was no audit, there is no renewing of the tax story. In this case the initial distribution of W_k should be considered as a prior distribution for the following period. In all other cases, fixed values (for example, median or some other quantile) of the posterior distribution obtained in this period can be considered as a priori value of disposition to evade in the next period. If this value exceeds some predetermined threshold value, considered as a key point for making an evasion decision (for example, 0.5), this can be considered as a signal to audit k-th agent. The greater the resulting value exceeds the specified threshold, the higher the priority of the inspection. And vice versa: the lower the value compared to the threshold, the less the need to audit this taxpayer.

As the first audit's result (for the period $j = 1$) two various curves can describe the propensity to evasion: beta-distribution with parameters $\alpha_{1_k} = 2$ and $\beta_{1_k} = 1$ for evaders, and with parameters $\alpha_{1_k} = 1$ and $\beta_{1_k} = 2$ for non-evaders . This case is illustrated in the Figure 4: the upper curve refers to the considered risk propensity of those taxpayers, who, according to the results of the first audit, have revealed evasion, the lower one – to those who was detected as non-evader.

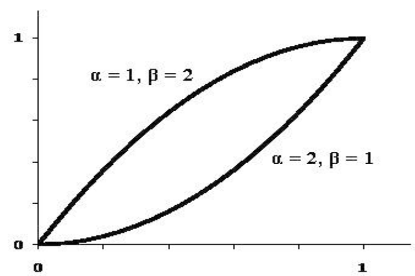

Fig. 4. The distribution of W_k after the first audit

After the second audit ($j = 2$) we obtain three possible groups of agents: the evasion propensity of twice-evaded is represented by the distribution with parameters $\alpha_{2_k} = 3$ and $\beta_{2_k} = 1$, for those, who alternate (sequencing does not matter), it has parameters $\alpha_{2_k} = 2$ and $\beta_{2_k} = 2$, and for those, who were detected as non-evaded twice, the distribution has parameters $\alpha_{2_k} = 1$ and $\beta_{2_k} = 3$ – see the Figure 5.

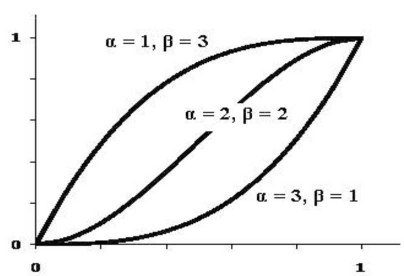

Fig. 5. The distribution of W_k after the second audit

After the third audit ($j = 3$) (see the Figure 6) the number of risk statuses is increasing: for those, who evaded 3 times, it can be represented by the distribution with parameters $\alpha_{3_k} = 4$ and $\beta_{3_k} = 1$, for those, who evaded twice, – with the parameters $\alpha_{3_k} = 3$ and $\beta_{3_k} = 2$, for those, who did not evade twice, – with parameters $\alpha_{3_k} = 2$ and $\beta_{3_k} = 3$, and, finally, for non-evaders during all 3 periods – with $\alpha_{3_k} = 1$ and $\beta_{3_k} = 4$.

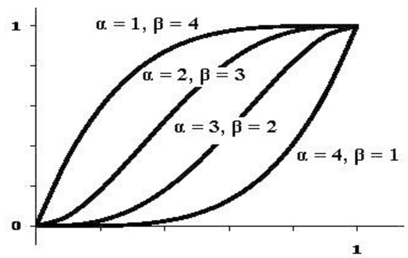

Fig. 6. The distribution of W_k after the third audit

Following periods and posterior risk propensities can be modeled in a similar way.

Pressing or Information Collecting? Thus, tax auditing allows to collect information about the tendency to evade among the studied population. But in real life the budget B of the tax authority is significantly limited. Therefore, the principal faces a dilemma: what is more important – to encourage honest payments (to choose the strategy of **pressing** as a pattern of behaviour) or to gather the data on the taxpayers evasions (to choose the pattern of **information collecting**)?

In the first case, when the tax authority decides to choose **pressing**, the budget should be distributed so as to audit all the evaders first. In the case the tax authority prefers to **collect information** the resource should be used to audit the largest number of agents unaudited previously.

Priorities. To describe the priorities of auditing for the different behavioural patterns chosen by the tax authority, we can analyze sequent inspections and the information, received as the result of monitoring during them. Initially, we restrict our study to three audits, but their number can be increased using similar reasoning.

For the case of **pressing**, we obtain the following distribution of priorities. After the first audit an evader gets the highest priority (1) for the future inspection, an honest taxpayer is marked with the lowest priority (3), and those who were not audited get priority 2. After the second audit the taxpayers who were detected as twice-evaders get priority 1, those who alternate get priority 2, the agents determined as twice-honest get priority 4, those who were not audited get priority 3. After the third audit we obtain three-times-evaders (priority 1), twice-evaders (priority 2), twice-honest (priority 3), three-times-honest (priority 5) and those who were not audited (priority 4). And so on. At each step the tax authority first seeks to re-audit those who were identified as evaders.

If the **information collecting** is considered, we obtain the following distribution of priorities. After the first audit an evader gets the priority 2, honest – priority

3, not audited – priority 1. After the second audit the taxpayers who were identified as twice-evaders get the priority 2, those who alternate get priority 3, twice-honest – priority 4, not audited – priority 1. And after the third audit three-times-evader gets the priority 2, twice-evader gets the priority 3, twice-honest – the priority 4, three times honest – the priority 5, not audited – the priority 1.

5.2. Auditing from the Agents' Standpoint

As it was discussed earlier, there are three subgroups in the taxable population depending on the risk-statuses they possess. Risk-averse taxpayers prefer to pay taxes honestly. Risk-loving agents are tax evaders. And only risk-neutral players make their choice to pay or to evade depending on the external conditions. Previously we considered the repeated game in which risk-neutral taxpayer observe her/his own story of the audits occurred before. But it is natural to assume that every agent tries to get an information about auditing story of other persons using social connections to communicate with other agents.

Previously, in (Gubar et al., 2016; Gubar et al., 2017; Gubar et al., 2019; Kumacheva et al., 2018; Kumacheva et al., 2019; Kumacheva et al., 2020) and (Gubar et al., 2020) the approach, which assumes that the information about future tax audits is spread in a population of taxpayers, was investigated. In the current study, we exclude the factor of deliberate dissemination of information, but at the same time we take into account that agents communicate with each other, sharing the information. Therefore, we come to the idea of the mutual information collecting both by the tax authority and the taxpayers in the studied population.

Let's consider a risk-neutral taxpayer k. Suppose that she/he estimates the probability $\widehat{p}_j^k$ of her/his inspection for the next period after j-th audit according to the collected information including taxpayer's own auditing story.

Let the variable Y_j be an indicator of tax audit (where j is the number of current tax period):

$$Y_j = \begin{cases} 1, & \text{audited and punished} \\ 0, & \text{non-audited.} \end{cases} \tag{4}$$

Then after the j-th period the estimator $\widehat{p}_j^k$ of the probability of audit for risk-neutral agent k is

$$\widehat{p}_j^k = Y_j \cdot P^* + (1 - Y_j)\frac{l_a^k}{l^k}, \tag{5}$$

where l^k is the number of the k-th taxpayer's social connections (friends, relatives, colleges etc.), l_a^k is the number of the audited agents among k-th taxpayer's social connections in the current period. Comparing the obtained estimator $\widehat{p}_j^k$ with known optimal value P^* of the probability of audit, the taxpayer makes her/his decision to evade or not. By the other words, if

$$\widehat{p}_j^k \geq P^*,$$

the risk-neutral agent will not evade during the next tax period. In practice, the last inequality means that $\frac{l_a^k}{l^k} \geq P^*$. Thus, the taxpayer needs only to compare the values $\frac{l_a^k}{l^k}$ and P^* to make her/his decision.

5.3. The Network Representation of the Mutual Collecting of Information

Continuing the research carried out in (Nekovee et al., 2007) and (Riehl and Cao, 2015), we suppose that the population of taxpayers can be considered as unoriented network $G = (N, L)$, where $N = \{1, \ldots, n\}$ is a set of economic agents (taxpayers), $L \subset N \times N$ is an edge set (where each edge is a social connection between two taxpayer). Previously, a similar approach was applied in (Gubar et al., 2017; Gubar et al., 2019; Kumacheva et al., 2019; Kumacheva et al., 2020) and (Kumacheva and Tomilina, 2021).

If we consider a network representation of the taxable population, then l^k can be represented as the number of connections of the taxpayer k with other nodes (agents). The interaction of this taxpayer with other agents leads to the information exchange between them. In the current study the simplest case of interaction, when agents are honest with each other, is considered. It is possible to investigate more general case, when the information distortion is taken into account, thus, the information confidence coefficient should be less than one. But here we will start from the fact that the k-th taxpayer receives an information from the agents connected with him, whether they have been audited or not, and she/he can trust this information. The information spreading in the network of agent is supposed to be similar to the algorithm based on the ideas of the "natural" dissemination of rumors, viruses, ideas and information (see Nekovee et al., 2007). According to this ideology, the information received by the taxpayer is compared with the total number of the agent's social connections (edges). It means that the taxpayer k knows the portion of her/his audited neighbours and, thus, can estimate the probability $\widehat{p}^k$ of the future audit according to the formula (5) and compare it with a known value P^*.

To investigate our model, the social connection of each taxpayer can be represented as networks of various modifications. In the current study we consider random networks (strongly or weakly connected) and grids. Taking these networks as possible representation of the taxable population, we conduct a scenario analysis both of the interaction of taxpayers, their exchange of the information on future tax inspections and posterior assessment of the possibility of audits, and the process of monitoring on the taxable population and analyzing data received by the tax authority.

6. Numerical Simulation

To answer the question, how mutual information collecting affects the effectiveness of tax control, we need to study the following issues:

- How does Bayesian updating impact on the effectiveness of tax control?
- Does the gathering of information by the taxpayers change this process?

To investigate these issues, the numerical simulation was conducted. In each experiment we initiated a network with fixed initial distribution of evaders and honest taxpayers and investigate separate and simultaneous collecting of the information by both parties during several iterations.

The simulation algorithm can be briefly described as follows. First, the network is initialized. Then, the **process A**, consists of inspection and information collecting by the tax authority is realized. After this, the **process B**, an information dissemination over the network, is conducted. After these processes in the network are finished, we simulate updating authority's strategies, based on the computing of estimates of the taxpayer's propensity to evade taxation, and updating taxpayers' strategies, based on the estimates of the future values of the probabilities of audit. With these new strategies the new iteration (new inspection) is started. The stopping point of the process is at the state when all agents (excepting risk-loving) do not evade or after 10 iterations.

To illustrate our results, we represent examples on the networks with $N = n = 49$ nodes (the size of total population), but the experiments with the smaller networks ($N = n = 25$) were also conducted. We considered various relations of the portions of evaders and honest taxpayers at initial time moment, therefore, various initial distributions of risk statuses in the population:

– 17% of risk-averse, 65% of risk-neutral and 18% of risk-loving (in accordance with the results presented in psychological study devoted to the risk-propensity (Niazashvili, 2007));
– 25% of risk-averse, 65% of risk-neutral and 10% of risk-loving (based on the assumption about Normal distribution of risk statuses (Kendall and Stuart, 1966));
– 10% of risk-averse, 65% of risk-neutral and 25% of risk-loving (based on the similar reasoning).

As the simulation parameters in the experiments, we use the empirical data based on the distribution of the income among the population of Russian Federation in 2020 (The web-site of the Russian Federation State Statistics Service, 2020) (see Table 1).

Table 1. Two modeled groups and average income

group	income interval (rubles/month)	average income (rubles)	share of population (%)
L	less 25000	$L = 12500$	51
H	more 25000	$H = 50000$	49

Our numerical simulation is conducted for the following values of parameters: tax rate is $\theta = 13\%$ and penalty rate is $\pi = 13\%$, thus, optimal value of the probability of audit is $P^* = 0.5$ (according to the results of (Bure and Kumacheva, 2010)). At the same time, we suppose that the actual value of probability of audit is less than P^*: in the simulation the values $P_L = 0.4$ or $P_L = 0.3$ were used. Let's also assume that only risk-averse agents are non-evaders, therefore, the share of non-evaders in the initial time moment can be $\nu_a = 17\%$ (or 25%, or 10%). Unit cost of one audit is $c = 7455$ rubles (corresponds to the level of minimum wage in St. Petersburg (The web-site of the Russian Federation State Statistics Service, 2020)).

In our experiments the process represented on the network should be visualised using different graphical notations:

– the risk statuses of taxpayers are represented as the various forms of the nodes: risk-averse agents are represented as triangles, risk-neutral agents are squares and risk-loving are circles;

– the agents' real income is graded by the size of the nodes: agents with low level L of income are the small nodes, agents with high level H are large nodes.

In each experiment we obtain four networks which represent different views (different knowledge) on the taxable population.

1. **Network 1** gives graphic representation of the information about the taxpayers' declared income (the population profile from the tax authority's standpoint): in the figures the taxpayers' declared low level L of income are denoted by red nodes, the taxpayers' declared high level H of income are denoted by green nodes.
2. **Network 2** visualizes the information about compliance of the declared and true income: if the agent's declared income value corresponds to the real one, then the color of node is yellow, if the values do not correspond to each other (the declared value is less than the true) the node is blue.
3. **Network 3** shows the result of the current audit: if there is no evasion, the agent is represented by the yellow node, if there is a revealed evasion, she/he is represented by the blue node, if the taxpayer is not audited, the node has no color (or, we say, grey color).
4. **Network 4** represents cumulative result of all previous inspections: if the taxpayer is not audited, she/he is the node which has no color (grey), in the opposite case, the number of revealed evasions corresponds to the gradient of colors from red to green.

To illustrate the proposed approach to the modelling of tax control and visualizing the experimental side of the research, let's study several examples of such imitation modelling and analyse the results of the conducted simulation.

6.1. Example 1

First, let's study the case of the network with the structure of grid. The initial distribution considered in the current simulation is 17% of risk-averse, 65% of risk-neutral and 18% of risk-loving (see the Niazashvili, 2007). For this example we fix the actual value of the probability of audit $P_L = 0.3$.

At the initial moment of time there are two graph representations of the taxable population: in the Figure 7 you can see the network 1 (how it is perceived by the tax authority) and in the Figure 8 there is the network 2, which is the illustration of the compliance between the declared and true levels of income of the taxable population. The value of the net tax revenue 1 is $R = 74309.58$ (rubles) at this state of the system.

Let's study the dynamics of the system during the following tax periods. First of all, we consider the case of direct auditing for five sequent tax periods. That is, we can not say that the tax authority does not use any additional information about population to choose the strategy during these periods at all: the principal simply prefers to punish the revealed evaders first and audit others second. After the fifth inspection we obtain new state of the system, represented in Figure 9, and cumulative results after 5 first inspections, represented in Figure 10.

We obtain the dynamics of the revenue represented in Figure 11: at the start of the process the net tax revenue falls, but after the first inspection it begins to grow until the third period, and tends to decrease only after it.

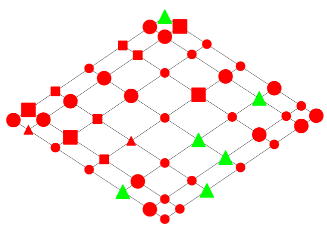

Fig. 7. Network 1. Initial state of the population: declared income

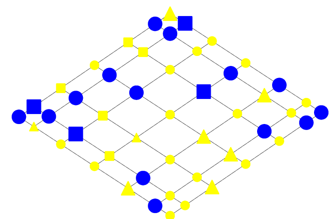

Fig. 8. Network 2. Initial state of the population: compliance of the declared income with true value

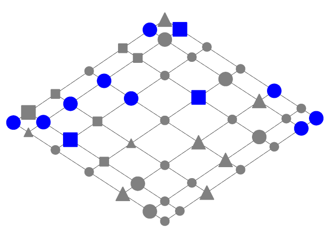

Fig. 9. Network 3. Results of the 5th audit

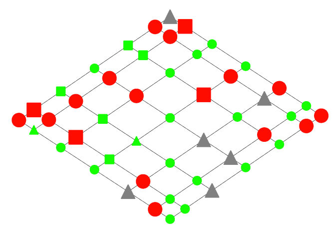

Fig. 10. Network 4. Cumulative results of 5 sequent audits

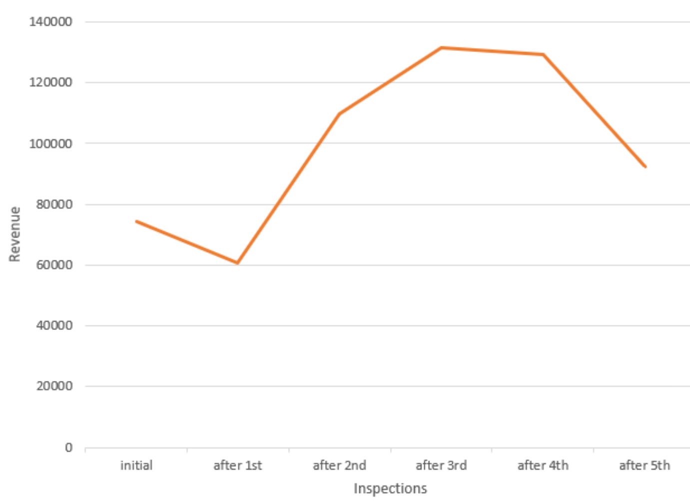

Fig. 11. The revenue dynamics for the grid: case of the direct auditing

The next step is to compare the results of direct auditing and of the case when the tax authority collects the information about the taxable population. In the Figure 12 we can see that in the case of information collecting the revenue also grows until the fifth inspection, but slower than in the first case. The light broken line in the presented plot illustrates a direct auditing, while the dark line represents the case when only the tax authority collects information.

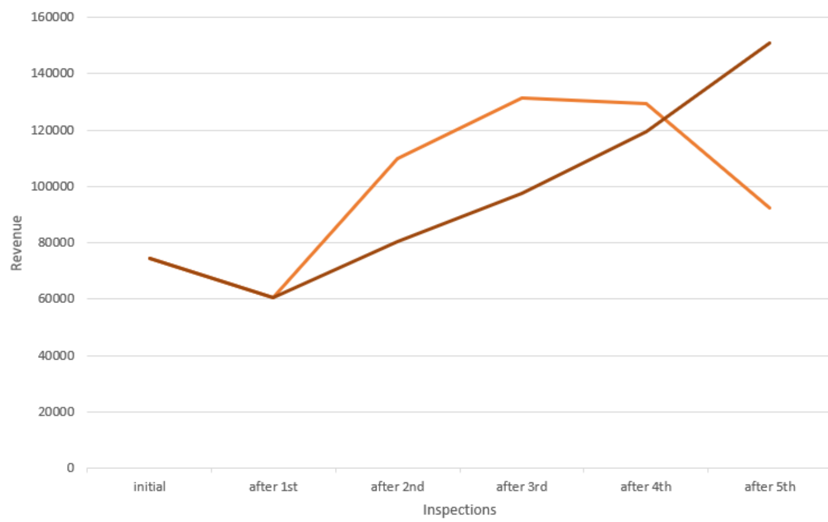

Fig. 12. The revenue dynamics for the grid: comparison of the direct auditing and information collecting

Now we consider dynamics for the both cases during the following periods and compare the values of the net tax revenue (1) with the cases when the agents begin to gather information about possible inspections in the cases of two studied strategies of the tax authority – pressing and information collecting. Figure 13 illustrates that the agents' information collecting decreases the level of revenue independently on the strategy chosen by the principal. As before, the light orange line represents the direct auditing, the dark line represents the case where only the tax authority collects information, the yellow bar graph represents the case of mutual collecting of information, and the green bars illustrate the case when taxpayers collect information against the strategy of pressing.

The Figures 14 and 15 shows the state of the studied grid for the case of direct auditing.

In the considered example the net tax revenue (1) has the following average values for different models of interaction after 10 inspections:

- the case of direct auditing: $R = 127301.33$ (rubles);
- the case when only tax authority collects information: $R = 128630.88$ (rubles);
- the case when tax authority prefers pressing and taxpayers collect information: $R = 117630.42$ (rubles);
- the case of mutual collecting of information: $R = 119663.59$ (rubles).

So, we can see that the case when only tax authority collects information is the most efficient in order to increase net tax revenue, i.e. to realize the fiscal function of tax system.

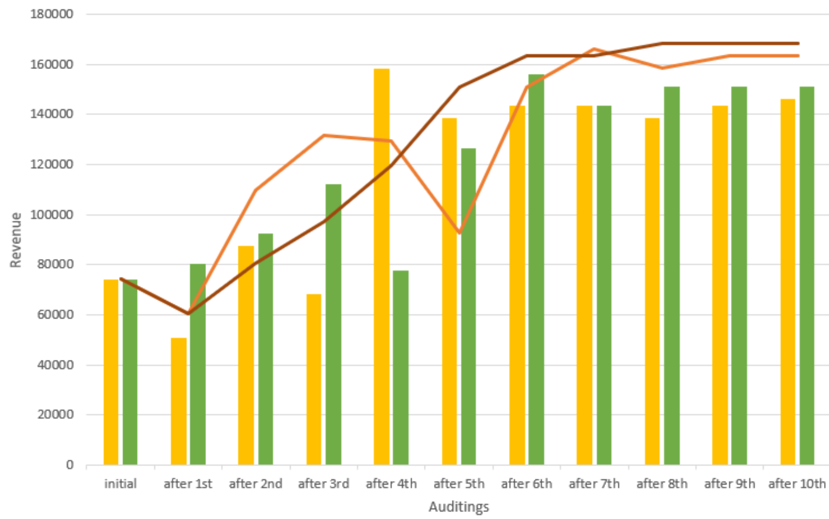

Fig. 13. The revenue dynamics for the grid: different types of interaction

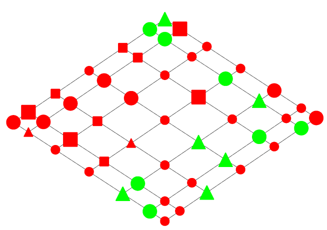

Fig. 14. After the 10$^{\text{th}}$ audit: declared income

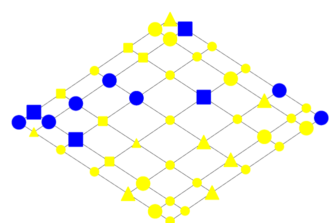

Fig. 15. After the 10$^{\text{th}}$ audit: compliance of the declared income with true value

6.2. Example 2

Now let's study the population modeled as a random weakly connected network. The initial distribution of risk propensities is still supposed to be 17% – 65% – 18% of risk-averse, risk-neutral and risk-loving correspondingly.

Initial state of the population is represented in the Figures 16 and 17.

If the tax authority prefers to collect and analyse the information about taxable population, after the 10$^{\text{th}}$ inspection the modeled network will get the state represented in Figures 18 and 19.

In the framework of our approach, it is interesting to answer the question: does the increasing of the number of audited taxpayers influence on the net tax revenue positively? This question can be formulated in another form: can the probability P_L be interpreted as a univocal tool to improve the fiscal function of tax control? To illustrate the ambiguity of the answer to the question, we compare the dynamics obtained for two different cases.

The first case, when the value of probability $P_L = 0.3$ is illustrated in the Figure 20. As in the previous example, studying various dynamics, we consider the direct

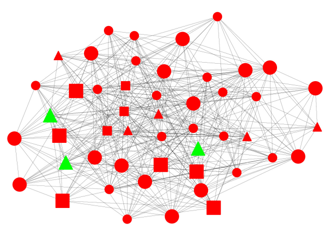

Fig. 16. Initial state of the population: declared income

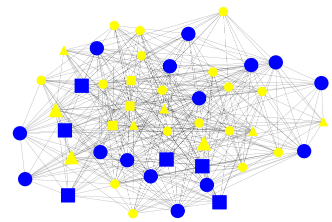

Fig. 17. Initial state of the population: compliance of the declared income with true value

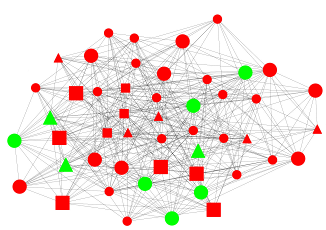

Fig. 18. After the 10^{th} audit: declared income

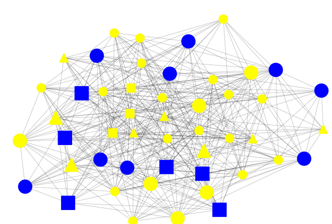

Fig. 19. After the 10^{th} audit: compliance of the declared income with true value

auditing, i.e. pressing without information dissemination (on the plot this dynamics is represented by the orange line), the strategy of pressing when the taxpayers also gather the information (green histogram), and the mutual collecting of information by both parties (yellow histogram).

The second case, when probability $P_L = 0.4$, is illustrated in the Figure 21. On the given plot we can see a fast growth in tax revenue, then a dramatic fall after the second inspection, then a rapid rise again in the fourth period, and fluctuations around the same values in subsequent periods.

Comparing the studied cases, we can find noticeable differences in several periods but quite a similar trend in considered dynamics. The latter illustrates the non-obviousness of the impact of an increase in the number of audits on a growth of the profit of the tax authority.

Now we can focus only on the case of mutual collecting of information and compare the dynamics for two fixed values of the probability P_L. The result of such comparison is shown in the Figure 22: the blue broken line illustrates the changes in revenue (1), when the value of probability is $P_L = 0.3$, and the orange line represents the case of $P_L = 0.4$. As you can see, the blue line lyes over the orange line: the values of the tax revenue are not less for the case of lower probability.

From this comparative analysis we can conclude that the bigger portion of audited taxpayers does not matter the bigger profitability and, therefore, the more efficiency.

Fig. 20. Net tax revenue: case of $P_L = 0.3$

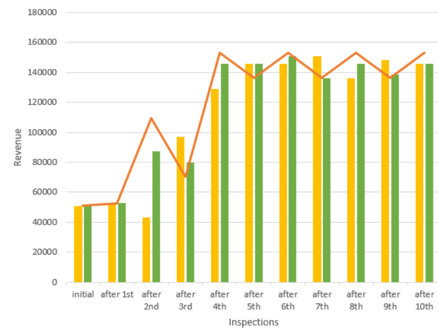

Fig. 21. Net tax revenue: case of $P_L = 0.4$

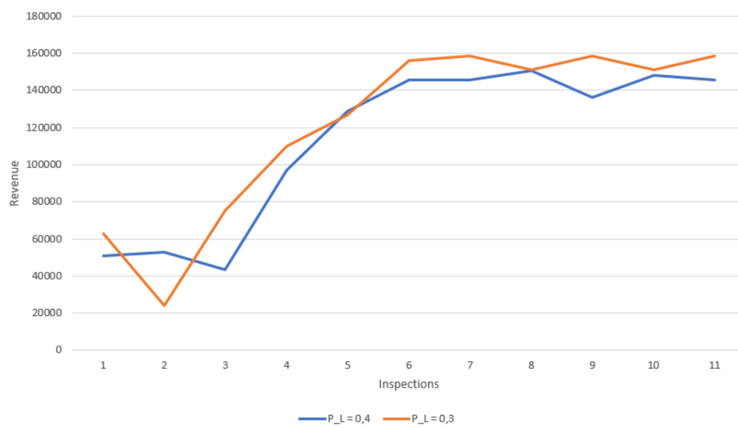

Fig. 22. Mutual collecting. Comparison of the revenues

Now let's compare the efficiency of the fiscal function for different patterns of the principal's and agents' behaviour for the fixed value of the probability $P_L = 0.3$. Figure 23 shows the dynamics of the total revenue (1) for the direct auditing, the case when only tax authority collects information about population, and the corresponding cases (pressing and information collecting), taking into consideration the possibility of mutual gathering of the information about future inspections by the taxpayers. Color representation in the plot remains the same.

The plot represented in the Figure 23 shows that the direct auditing is more profitable during three first inspections. But since the 4^{th} inspection, when the received information grows and let the tax authority chose rational approach to the auditing, the strategy of the collecting of information gives better results. But both dynamics converge at the stage of later inspections. The agents' collecting of information has a negative effect on the fiscal function of the system: the revenue of the tax authority is decreasing in comparison both with the direct auditing and with the strategy of collecting of information. But this fact relates to the long-term prospect. If we talk about short-term period (inspections from 3^{th} to 6^{th}), we can

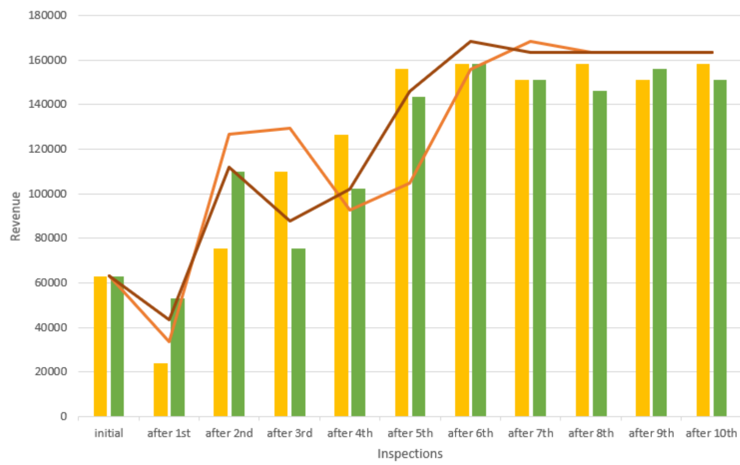

Fig. 23. Comparison of the revenue dynamics for all cases

see that such natural dissemination of the information about future audit is more profitable than its absence.

6.3. Results of Numerical Simulation

The examples considered above, together with many other previously performed experiments, allow us to formulate the following conclusions.

First, it was shown that the gathering of information about the structure of the population brings a positive effect for the total tax revenue and increases the total amount of honest taxpayers in long term. It means that, applying the strategy of information collecting, the tax authority can improve the implementation of not only the fiscal, but also the social (distributive) function of the tax system. But if this strategy is considered as the basic one in the short term, its effectiveness is not so evident. For a small number of sequent inspections, direct auditing (pressing without collecting information) yields higher tax revenue and a higher number of non-evaders among the total number of taxpayers.

If the collecting of information is mutual, oppositely, the efficiency of the fiscal system increases only in the short term, and then decreases (becomes lower than direct audit) in long term. In other words, the natural dissemination of information about future audits among the taxable population, resulting from the gathering of information by agents, reduces the effectiveness of the tax authority's knowledge about taxpayers' propensity to risk.

Second, let's try to reveal, which type of the network gives better results for the case, when both parties collect the information about each other relatively to the process of auditing. Comparison of the revenue dynamics for grid and random graphs (weakly and strongly connected) in the case of mutual information collecting is illustrated by the Figure 24. As anyone can see, the fall in the revenue changes for an opposite tendency after initial inspections independently on the type of the network.

We also obtained that the type of network affects the efficiency of information use either by tax authority or by taxpayers. In the case of mutual collecting of

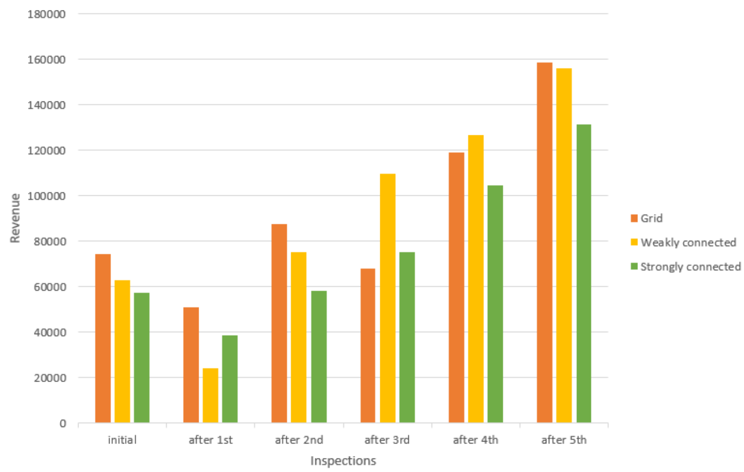

Fig. 24. Mutual information collecting for different networks

information, the most efficient network in order to increase the average tax revenue is random weakly connected. But if we take the increasing of the portion of evaders/non-evaders as a criteria we will not get a steady trend.

Third, the experiments showed that an increase in the initial share of risk averse agents has no stable influence on the dynamics of revenue in the network. But it was found that the initial distributions with the shares "25-65-10" and "17-65-18" give rather close results.

Finally, it was found that the convergence of the process (achieving system stability) is possible only for some types of networks with a small number of nodes. The conducted series of simulation revealed the case of convergence for random weakly connected graph with the strategy of mutual collecting with the number of nodes $n = 25$. For networks with a bigger number of nodes ($n = 50$) the stabilization was not reached.

7. Conclusions

The current work is devoted to the problem of tax control studied in the framework of game-theoretical approach and simulation modelling. According to the classic hierarchical scheme (Vasin and Morozov, 2005), principal is the tax authority and agents are taxpayers in the considered model. The taxable population is represented as a network of various configurations, and possible scenarios of the auditing process are investigated.

Along with the fiscal function, the tax auditing is considered as the process of monitoring of taxable population and gathering information about taxpayers. After each inspection the tax story of audited agent is updated and her/his propensity to evade is re-estimated. Thus, such collecting of the information about risk statuses of taxpayers is assumed to be similar to the Bayesian procedure (Zellner, 1997) and modelled due to the properties of the conjugate families (De Groot, 1970).

The taxpayers are also supposed to be able to get the information about the conducted inspections and, therefore, to estimate the probability of future audit.

The "natural" way of information spreading is modelled in the current research similar to the works (Nekovee et al., 2007) and (Riehl and Cao, 2015).

So, we simulate the mutual information collecting and can compare this case with others, when only one side gathers the facts about another one.

To investigate the proposed approach, a series of simulations were run. The simulation results allow to analyse the impact of various parameters of the audit conducted on the net tax revenue. We can generalize the results of our study, focusing on the following key moments.

First, the current research shows that the collecting of information about the structure of the population brings a positive effect for the revenue of the tax authority (1) and increases the total number of honest taxpayers. It means that the proposed strategy helps to improve the fiscal and distributive functions of the tax system. Moreover, the bigger number of inspections we assume, the more complete information about audited taxpayers is supposed to be received, the more efficiency of this strategy is reached as a result.

Second, the performed simulation shows that the "natural" way of the information dissemination makes the previously mentioned positive effect weaker in comparison with the case when only principal gather the facts about evasions in the population. In other words, the mutual collecting of information obviously decreases the tax revenue (1) and the number of non-evaders in the network of taxpayers.

Third, the dependence between the net tax revenue and the initial distribution of risk propensities is not stable for each type of interaction between the tax authority and taxpayers: various types of networks affect the revenue of the tax authority in different ways. The experiments also showed that an increase in the initial share of risk avoidant has no stable influence on the dynamics of the tax revenue. It was also found that the process convergence (achieving system stability) is possible only for some types of networks with a small number of nodes ($n = 25$). For networks with a bigger number of nodes ($n = 50$) the stabilization was not attained.

Thus, this research is devoted to the idea of the mutual collecting of the information in the application to the game-theoretic modelling of the tax control. The intuitive hypotheses about possible scenarios were supported by the conducted numerical simulation and following parametric analysis.

8. Appendix

8.1. First Appendix

Let's define Beta distribution (see Kendall and Stuart, 1966):

Definition 1. The random quantity X is **beta distributed** with parameters α and β ($\alpha > 0,\ \beta > 0$), if X is distributed absolutely continuously with the density

$$f(x|\alpha,\beta) = \begin{cases} \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}, \text{when } 0 < x < 1, \\ 0, \text{in other cases,} \end{cases}$$

where $B(\alpha,\beta) = \int\limits_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx$ is the beta-function.

8.2. Second Appendix

Let's formulate the theorem about conjugate distribution families (see De Groot, 1970).

Theorem 1. *Let $X_1, X_2, \ldots, X_n$ be a sample from the Bernoulli distribution with unknown value of parameter W. Suppose, that prior distribution of W is the beta distribution with parameters α and β ($\alpha > 0$, $\beta > 0$). Then the posterior distribution W when $X_j = x_j$ ($j = \overline{1,n}$) is the beta distribution with parameters $\alpha + y$ and $\beta + n - y$, where $y = \sum\limits_{j=1}^{n} x_j$.*

References

Antocia, A., Paolo Russua, P. and Zarrib, L. (2014). *Tax Evasion in a Behaviorally Heterogeneous Society: An Evolutionary Analysis.* Economic Modelling, **42(C)**, 106–115.

Antunes, L., Balsa, J., Urbano, P., Moniz, L. and Roseta-Palma, C. (2006). *Tax Compliance in a Simulated Heterogeneous Multi-agent Society.* Lecture Notes in Computer Science, **3891**, 147–161.

Bloomquist, K. M. (2006). *A comparison of agent-based models of income tax evasion.* Social Science Computer Review, **24(4)**, 411–425.

Bure, V. and Kumacheva, S. (2005). *A model of audit with using of statistical information about taxpayers' income.* Vestnik of Saint Petersburg University. Series 10. Applied Mathematics. Computer Science. Control Processes, **1–2**, 140–145 (in Russian).

Bure, V. and Kumacheva, S. (2010). *A game theory model of tax auditing using statistical information about taxpayers.* Vestnik of Saint Petersburg University. Series 10. Applied Mathematics. Computer Science. Control Processes, **4**, 16–24 (in Russian).

Chander, P. and Wilde, L. L. (1998). *A General Characterization of Optimal Income Tax Enforcement.* Rev. of Econ. Studies, **65**, 165–183.

De Groot, M. H. (1970). *Optimal Statistical Decisions.* MsGraw-Hill Company.

Gubar, E., Kumacheva, S., Zhitkova, E. and Porokhnyavaya, O. (2016). *Impact of Propagation Information in the Model of Tax Audit.* Recent advances in game theory and applications, "Static& Dynamic Game Theory: Foundations & Applications", pp. 91–110, Birkhauser, Springer.

Gubar, E., Kumacheva, S., Zhitkova, E., Kurnosykh, Z. and Skovorodina, T. (2017). *Modelling of Information Spreading in the Population of Taxpayers: Evolutionary Approach.* Contributions to Game Theory and Management, **10**, 100–128.

Gubar, E., Kumacheva, S., Zhitkova, E. and Tomilina, G. (2019). *Modelling of the impact of information on tax audits on the risk statuses and evasions of individuals.* Vestnik of Saint Petersburg University. Applied Mathematics. Computer Science. Control Processes, **15(2)**, 245–258 (in Russian).

Gubar, E., Kumacheva, S., Zhitkova, E., Tomilina, G. and Sanchez Carrera, E. J. (2020). *Games and network structures on corruption, income inequality, and tax control.* Contributions to Game Theory and Management, **13**, 152–172.

Kendall, M. G. and Stuart, A. (1966). *Distribution Theory.* Moscow: Nauka Publ. (in Russian).

Kumacheva, S. Sh. (2012). *Tax Auditing Using Statistical Information about Taxpayers.* Contributions to Game Theory and Management, **5**, 156–167.

Kumacheva, S. Sh. and Gubar, E. A. (2015). *Evolutionary model of tax auditing.* Contributions to Game Theory and Management, **8**, 164–175.

Kumacheva, S., Gubar, E., Zhitkova, E. and Tomilina, G. (2018). *Evolution of Risk-Statuses in One Model of Tax Control.* Frontiers of Dynamic Games, pp. 121–138, Birkhauser, Cham.

Kumacheva, S., Gubar, E., Zhitkova, E. and Tomilina, G. (2019). *Modeling the behaviour of Economic Agents as a Response to Information on Tax Audits*. In: Agarwal N., Sakalauskas L., Weber G.-W. (eds.). Modeling and Simulation of Social-Behavioral Phenomena in Creative Societies. First International EURO Mini Conference (MSBC 2019) Proceedings, pp. 96–111. Springer International Publishing.

Kumacheva, S., Gubar, E., Zhitkova, E. and Tomilina, G. (2020). *Analysis of Economic Behaviour in Evolutionary Model of Tax Control Under Information Diffusion*. Frontiers of Dynamic Games. Static & Dynamic Game Theory: Foundations & Applications, pp. 121–140, Birkhauser, Verlag AG.

Kumacheva, S. and Tomilina, G. (2021). *Information Collecting and Dissemination in the Network of Taxpayers: Bayesian Approach*. Contributions to Game Theory and Management, **14**, 236–256.

Lamantia, F. and Pezzino, M. (2021). *Social norms and evolutionary tax compliance*. The Manchester School, **89(4)**, 385–405.

Macho-Stadler, I. and Perez-Castrillo, J. D. (2002). *Auditing with signals*. Economica, **02**, 1–20.

Nekovee, A. M., Moreno, Y., Bianconi, G. and Marsili, M. (2007). *Theory of rumor spreading in complex social networks*. Physica A, **374**, 457–470.

Niazashvili, A. (2007). *Individual differences in risk propensity in different social situations of personal development*. Moscow: Moscow University for the Humanities.

Riehl, J. R. and Cao, M. (2015). *Control of Stochastic Evolutionary Games on Networks*. 5th IFAC Workshop on Distributed Estimation and Control in Networked Systems, Philadelphia, PA, United States, 2015, 458–462.

Samuelson, P. and Nordhaus, W. (2007). *Economics*. Moscow: Williams Publ. (in Russian).

Sanchez, I. and Sobel, J. (1993).*Hierarchical design and enforcement of income tax policies*. Journal of Public Economics, **50**, 345–369.

The web-site of the Russian Federation State Statistics Service *http://www.gks.ru/* (accessed 14.08.2022).

The web-site of the Russian Federal Tax Service. *https://www.nalog.ru/* (accessed 14.08.2022).

Vasin, A. and Morozov, V. (2005). *The Game Theory and Models of Mathematical Economics*. Moscow: MAKSpress Publ. (in Russian).

Zellner, A. (1997). *Bayesian Analysis in Econometrics and Statistics: The Zellner View and Papers*. Edward Elgar.