

Information Collecting and Dissemination in the Network of Taxpayers: Bayesian Approach

Suriya Sh. Kumacheva¹ and Galina A. Tomilina²

¹ *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 The current research is based on the assumption that the result of tax inspections is not only collection of taxes and fines. The information about audited taxpayers is also collected and helps to renew *a priori* knowledge of each agent's evasion propensity and obtain new *a posteriori* estimate of this propensity. In the beginning of the following tax period the fiscal authority can correct auditing strategy using updated information on every taxpayer. Each inspection is considered as a repeated game, in which the choice of agents to audit is associated with their revealed tendency to evade. Taxpayers also renew the information on the number of inspected neighbors using their social connections, represented by networks of various configurations, and estimate the probability of auditing before the next tax period. Thus, the application of the Bayesian approach to the process of collecting and disseminating information in the network of taxpayers allows to optimize the audit scheme, reducing unnecessary expenses of tax authority and eventually increasing net tax revenue. To illustrate the application of the approach described above to the indicated problem, numerical simulation and scenario analysis were carried out.

Keywords: tax control, tax evasions, risk propensity, structured network, a posteriori information, Bayesian approach.

1. Introduction

Taxation is one of the most important areas of the state economy, regulation and social policy. In accordance with the basic definitions of the tax system (Samuelson and Nordhaus, 2007), it has the following basic functions: fiscal, social, regulatory and control. It is worth to remember that tax control is one of the most important tools for the administration of the tax system. Therefore, it is clear that many works, such as (Antunes et al., 2006; Antocia et al., 2014; Bloomquist, 2006; Chander and Wilde, 1998; Macho-Stadler and Perez-Castrillo, 2002; Sanchez and Sobel, 1993; Vasin and Morozov, 2005), are devoted to the research in this area of the economy.

The mentioned models of tax control apply different approaches to mathematical modelling of the studied problem. A lot of them can be represented by the mathematical decision theory and its separate branches – contract theory and mathematical game theory, and also by various directions of probabilistic modelling.

Most of the works which develop these approaches have a set of common features. Most of them are based on the models with hierarchical structure similar to the

<https://doi.org/10.21638/11701/spbu31.2021.18>

“principal-to-agent” scheme. In general case agents are assumed to be risk-neutral. In the mentioned studies cases the problems have 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.

One of modifications of the solutions from the described class was proposed in (Bure and Kumacheva, 2010). In this model risk-neutral taxpayers refuse their evasion from high level of income H to low level of income L if the probability P_L of audit of the agents, who declared low level of income, is not less than the optimal value of auditing probability P^* . But the application of this result has a number of disadvantages and limitations, both in practice and in theoretical assumption underlying the model. Among them there are the following problems, common with the previously listed studies.

The **Problem 1** can be formulated as the limited budget of the tax authority, which makes reaching the optimal values of auditing probabilities to be 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 and Gubar, 2015; Kumacheva et al., 2018; Kumacheva et al., 2019).

The **Problem 2** 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) is most often considered as models 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, such information can be only indirect. Accordingly, agents do not know the probability of audit and can only estimate it.

Problem 3 is in that the tax authority knows nothing on the population structure. There is no information about each agent’s risk propensity. This is an obstacle to the choice of optimal auditing strategy, corresponding to the real distribution of risk statuses of the population. Previously the ideas about using additional information about taxpayers were discussed in (Macho-Stadler and Perez-Castrillo, 2002; Bure and Kumacheva, 2005; Kumacheva, 2012).

The listed 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., 2015; Gubar et al., 2016) and (Gubar et al., 2017b). Another attitude, which is the modelling based on evolutionary processes on the network, was considered in (Gubar et al., 2017a; Gubar et al., 2017b) and (Kumacheva et al., 2018). The ideas of opinion dynamics based on Markov process were studied in (Gubar et al., 2019).

The current work proposes a new approach based on a combination of the processes of disseminating information about inspections among the taxable population and collecting information on the risk propensity of economic agents by the tax authority. Such information makes it possible to re-estimate the probabilities of audit of certain taxpayers using the Bayesian procedure (De Groot, 1974; Zellner, 1997).

This paper is structured as follows. Section 2 is devoted to the brief description of the static model which lays in the basis of our research. In section 3 we consider the process of auditing from the standpoints of principal and agents and formulate several assumptions useful to make the studied model and the simulation built on it closer to reality. Section 4 describes the numerical experiments and parameter analysis based on the mentioned assumptions. And, finally, in section 5 we formulate the conclusions we made from our modelling and simulating, and prospects for the future study of the proposed model.

2. Static Model of Tax Control with Two Levels of Income of Taxpayers

The model formulated further is based on the static model of tax control (Bure and Kumacheva, 2010). But in the current study we will restrict our consideration by two levels of income received by the taxpayers as it was done in (Chander and Wilde, 1998; Vasin and Morozov, 2005). Without restricting generality, it can be expanded to more number of income levels. However, the interaction pattern will be the same as described here for the two income levels.

Following (Bure and Kumacheva, 2010), let's consider a population of n taxpayers. Each of them has her/his true income ξ and can declare income η , where $\eta \leq \xi$. Both 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 cardinality of the corresponding sets are n_H and n_L : $n_L + n_H = n$.

In the further work on the model to approximate it to reality, it is possible to consider a larger number of income gradations, as was done earlier, for example, in (Bure and Kumacheva, 2005) or (Kumacheva, 2012). But the basic premises and way of reasoning will remain the same. Therefore, we will adhere to the principle previously applied in (Macho-Stadler and Perez-Castrillo, 2002; Vasin and Morozov, 2005) and (Kumacheva et al., 2019) and consider only two levels of income of the population, formally classifying taxpayers into rich and poor.

Taxpayers who declare a high level of their income cannot be classified as evaders. Thus, the tax authority does not need to audit such agents and, therefore, should focus on those who declared the low level $\eta = L$. Let P_L be the probability of audit of such agents. In the case of revealed tax evasion the agent should pay $(\theta + \pi)(\xi - \eta)$, where the constants θ and π are tax and penalty rates correspondingly. According to the results obtained in (Bure and Kumacheva, 2010), the optimal value of auditing probability is

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

As it was assumed in (Kumacheva and Gubar, 2015; Kumacheva et al., 2018; Kumacheva et al., 2019), let's suppose that the real structure of 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 on 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: $\gamma_L + \gamma_H = 1$.

3. The Auditing Process

In the current study we assume that tax authority and taxpayers interact according to the classical hierarchical scheme “principal-to-agent”. Therefore the tax auditing is sophisticated process which can be described from two different standpoints: the tax authority’s and the taxpayers’ ones correspondingly.

3.1. Auditing from the Principal’s Standpoint

The tax authority has three various goals 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:

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

where c is the unit cost of auditing.

The second one is the **population monitoring** (Macho-Stadler and Perez-Castrillo, 2002; Bure and Kumacheva, 2005; Kumacheva, 2012), which purpose is the updating information about audited taxpayers and, thus, about their evasion propensities. Finally, it helps to refresh the information about the distribution of risk statuses 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.

The third goal is the **information dissemination** (Gubar et al., 2016; Gubar et al., 2017a; Gubar et al., 2019; Kumacheva et al., 2019; Kumacheva et al., 2020). Being strongly bounded by the state budget, the tax authority can consider the information spreading as a tool to stimulate tax payers for honest payments. As such information the message about high possibility of the future tax auditing among the population can be considered, despite its real values which can be much smaller.

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, 1974; 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 propose the method previously investigated in (Bure and Kumacheva, 2010).

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, evasion propensity). 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):

$$X_j = \begin{cases} 1, & \text{no evasion} \\ 0, & \text{evasion.} \end{cases} \quad (2)$$

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, 1974), which is quoted verbatim in the appendix 6.1. That is, it can be concluded that the family of beta distributions is conjugate to the family of Bernoulli distribution — see the appendix 6.2.

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 .
- Using the tax story (2) and applying the mentioned theorem about conjugate distributions (De Groot, 1974), tax authority can compute the new values of parameters $\alpha_{j_k}, \beta_{j_k}$ of a posteriori distribution of W_k , which can be considered as a priori distribution for the next tax period.
- Starting the next iteration ($j = 2 \dots$), 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 the probability of tax evasion in the next period.

As the first audit's result (for the period $j = 1$) two various curves can describe the propensity to evasion: beta-distribution with parameters $\alpha_{1_k} = 1$ and $\beta_{1_k} = 2$ for evaders, and with parameters $\alpha_{1_k} = 2$ and $\beta_{1_k} = 1$ for non-evaders. This case is illustrated in the figure 1: 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 not an evader.

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

After the third audit ($j = 3$) (see the figure 3) the number of risk statuses is increasing: for those who evaded 3 times it can be represented by the distribution with parameters $\alpha_{3_k} = 1$ and $\beta_{3_k} = 4$, for those who evaded twice – with the parameters

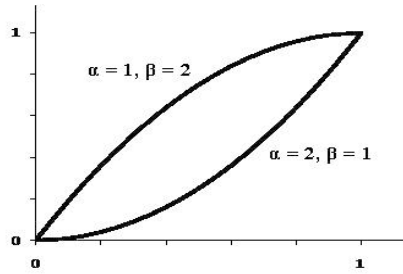


Fig. 1. The Results of the First Audit

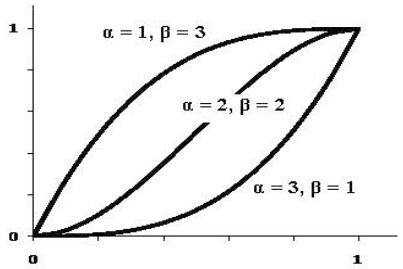


Fig. 2. The Results of the Second Audit

$\alpha_{3_k} = 2$ and $\beta_{3_k} = 3$, for those who did not evade twice – with parameters $\alpha_{3_k} = 3$ and $\beta_{3_k} = 2$, and, finally, for non-evaders during all 3 periods – with $\alpha_{3_k} = 4$ and $\beta_{3_k} = 1$.

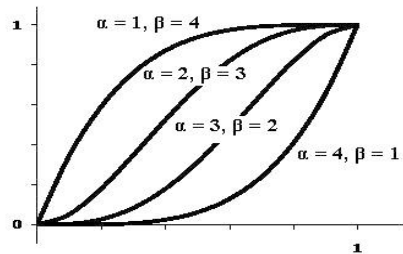


Fig. 3. The Results of the Third Audit

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 of the population to evade. But in real life the budget 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**) or to gather the data on the taxpayers evasions (to choose **the strategy 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 distributed due to audit the largest number of agents unaudited previously.

3.2. Auditing from the Agents' Standpoint

Since in formulation of the model it was initially assumed that agents can have one of three risk statuses, the choice of their strategies depends on which group they belong to. If a taxpayer is risk-averse, she/he prefers to pay the taxes honestly, in accordance with her/his true level of income. Risk-loving agents are tax evaders. And only risk-neutral taxpayers will make their choice, according to their bounded rationality – to pay or to evade. They make their decision comparing known value P^* with their own estimation of actual probability P_L . Therefore they also need to collect information about possible auditing. We suppose that agents can gather such information from their social contacts, following to the previous ideas developed in (Gubar et al., 2015; Gubar et al., 2017b; Gubar et al., 2019; Kumacheva et al., 2018; Kumacheva et al., 2019; Kumacheva et al., 2020; Gubar et al., 2020).

3.3. Information Dissemination: the Network Approach

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 indirect network $G = (N, L)$, where $N = \{1, \dots, n\}$ 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., 2017a; Gubar et al., 2019; Kumacheva et al., 2019; Kumacheva et al., 2020).

Let the k -th taxpayer has l_k connections 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 an opposite case, when the degree of information distortion is taken into account, thus, the information confidence coefficient should be not equal to one. But here we will start from the fact that the k -th taxpayer receives information from the agents connected with him, whether they have been audited or not, and she/he believes 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): 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 can calculate the portion of her/his audited neighbours and, thus, estimate the probability \hat{P}_k of the future audit and compare it with a known value P^* .

In our modelling 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. Modeling networks of various modifications, 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 and analyzing data on the taxable population received by the tax authority.

4. Numerical Simulations

The main questions to be answered by the numerical experiment are: do the dissemination of information among taxpayers and the collecting of information about taxpayers (their propensity to evade) by the tax authority affect the increase in tax collection?

To investigate these problems we conduct a numerical simulation, in which we initiate a network with some initial distribution of evaders and honest taxpayers and investigate the collecting and dissemination of information over the network during several iterations.

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

- 17% of risk-avoiding, 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-avoiding, 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-avoiding, 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 2019 (The web-site of the Russian Federation State Statistics Service, 2019) (see Tab. 1).

Table 1. Two modeled groups and average income

group	income (rub./month)	interval	average income (rub.)	share of population (%)
<i>L</i>	less 25000		$L = 12500$	51
<i>H</i>	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 let's suppose that the actual value of probability of audit is $P_L = 0.4$. 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 auditing is $c = 7455$ rub. (corresponds to the level of minimum wage in St. Petersburg (The web-site of the Russian Federation State Statistics Service, 2019)).

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 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 true value is less then the real value) the node is blue.
3. **Network 3:** shows the result of auditing: 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, grey color).
4. **Network 4:** represents cumulative result of all previous audits: 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.

4.1. Example

To illustrate the proposed approach to the modelling of tax control and visualizing the experimental side of the research, we will study one of the examples obtained as a result of the conducted simulation in detail. We will discuss the case of the network with structure of strongly connected network. 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). The tax authority chooses the strategy of pressing.

The first iteration. In the initial moment of time there are two graph representations of the taxable population: in the figure 4 you can see the network 1 (how it is seen by the tax authority) and in the figure 5 there is the network 2 – the illustration of the compliance between the declared and true levels of income of the taxable population.

In the initial moment the total tax revenue can be obtained as the expected value of the tax authority's income which can be computed according to the equation (1). For the considered example its value is $R = 38836.47$ (rub.).

After the first audit the information about taxpayers' income and their propensities to evasion is received. Thus, we obtain two new networks: the network 3 (the figure 6), which represents the result of the first audit, and the network 4 (the figure 7), which is the cumulative result of previous inspections. We can see that, in fact, the network 3 and the network 4 coincide in this case, because we are taking into

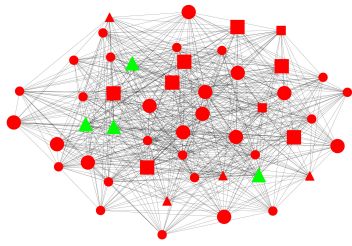


Fig. 4. Network 1. Initial State of the Population: Declared Income

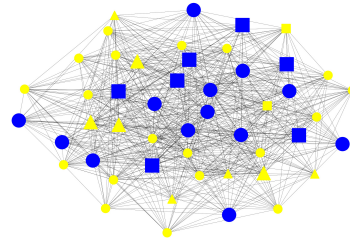


Fig. 5. Network 2. Initial State of the Population: Compliance of the Declared Income with Real

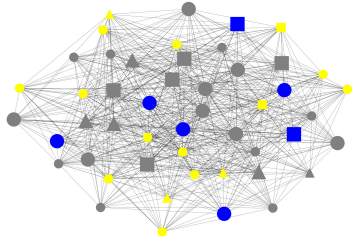


Fig. 6. Network 3. The Result of the First Audit

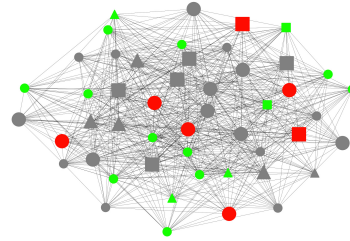


Fig. 7. Network 4. Cumulative Result of the Previous Audits

account only one auditing and the information obtained from it only. The net tax revenue has decreased – after the first audit its value is $R = -4090.0$ (rub.). The inspection is unprofitable at the current moment, because in the studied example the tax authority acts according to the strategy of pressing. It leads to the small number of the audited evaders and, therefore, small value of collected taxes and penalties.

The second iteration. At the beginning of the second period, we see a slight increase in the number of those who declared high income, in other words, those who interrupt tax evasion. The changes are reflected in the network 1 and network 2, represented in the figures 8 and 9.

In accordance with the chosen strategy of pressing, during the second audit the tax authority tries to audit those who have already evaded first of all. We obtain the network 3 (figure 10), corresponding to the results of the audit carried out, and the network 4 (figure 11), which is a visualization of the information about the taxable population accumulated over 2 passed inspections.

After the second audit the net tax revenue has increased: $R = 81650.00$ (rub.). It has happened due to the fact that the information collected about the population contributes to a more accurate forecast of tax evasion and, thus, the choice of a more rational strategy of inspections.

The third iteration. Before the third audit (figures 12 and 13), we can observe the state of the network, similar to the one that was in the previous period.

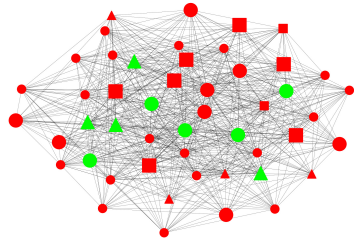


Fig. 8. Network 1. The State of the Population after the First Audit: Declared Income

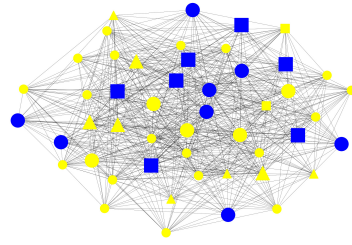


Fig. 9. Network 2. The State of the Population after the First Audit: Compliance of the Declared Income with Real

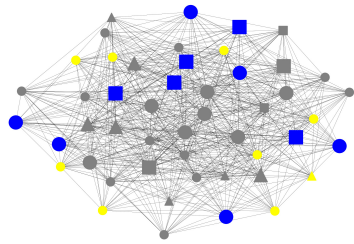


Fig. 10. Network 3. The Result of the Second Audit

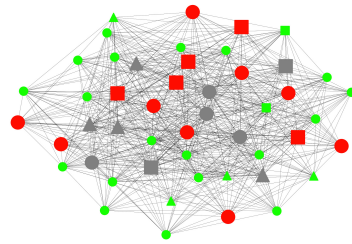


Fig. 11. Network 4. Cumulative Result of the Previous Audits

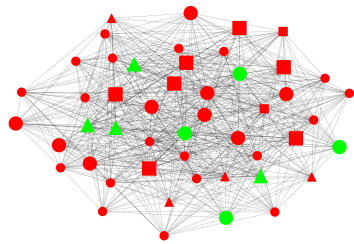


Fig. 12. Network 1. The State of the Population after the Second Audit: Declared Income

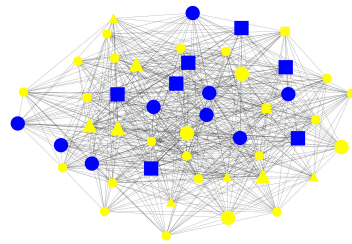


Fig. 13. Network 2. The State of the Population after the Second Audit: Compliance of the Declared Income with Real

The results of the third audit are represented as network 3 and network 4 for the current iteration represented in the figures 14 and 15 correspondingly.

Comparing the results of this inspection with the previous one we can see that the number of non-evaders does not change, but the tax authority's data on the taxable population has changed, which allows to increase the net tax revenue through taxes and penalties collected from evaders.

The 4-th iteration. Consider the fourth period. The population state after the third audit is represented in figures 16 and 17. The net tax revenue is $R = 125525.0$ (rub.).

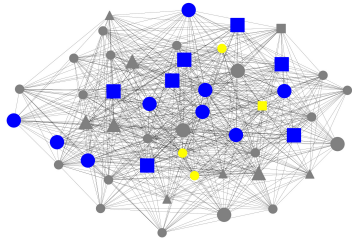


Fig. 14. Network 3. The Result of the Third Audit

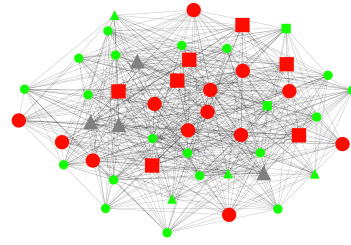


Fig. 15. Network 4. Cumulative Result of the Previous Audits

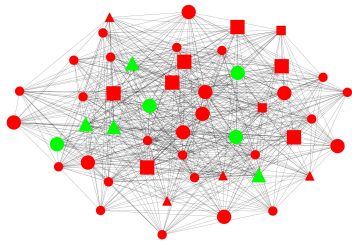


Fig. 16. Network 1. The State of the Population after the Third Audit: Declared Income

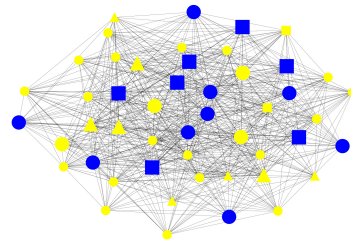


Fig. 17. Network 2. The State of the Population after the Third Audit: Compliance of the Declared Income with Real

The results of the fourth inspection are represented in the figures 18 and 19. The trend revealed after the previous inspection remains constant. The number of honest taxpayers is not increasing, but the tax authority’s income is still growing.

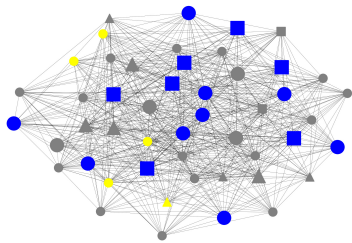


Fig. 18. Network 3. The Result of the Fourth Audit

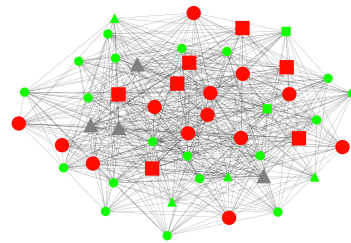


Fig. 19. Example 4. Cumulative Result of Previous Audits

4.2. Examples of Comparative Analysis

The analysis similar to that presented in the considered example was carried out for other modifications of the network and other initial distributions of risk statuses of agents in the population. Further several results of the conducted analysis are demonstrated.

Comparison of two strategies. Let's compare the profitability of the strategies of *pressing* and *information collecting*. For this purpose we consider the network of the same configuration and with the same initial distribution of risk statuses on it. We will study strongly connected network and following initial state of the population: 17% of risk-avoiding, 65% of risk-neutral and 18% of risk-loving agents. For the given initial conditions the strategy of pressing was studied in the previous example. For the strategy of information collecting the following results were obtained:

- Initial state: $R = 38836.47$ (rub.)
- After the first audit: $R = 34910.0$ (rub.)
- After the second audit: $R = 59855.0$ (rub.)
- After the third audit: $R = 116060.0$ (rub.)

The dynamic of the net tax revenue (1) for both types of the strategies is represented in the figure 20. In the diagram the blue bars represent the tax authority's revenue after the respective inspections using the strategy of pressing, the red ones refer to the strategy of information collecting. Obviously, after a sharp drop in its value

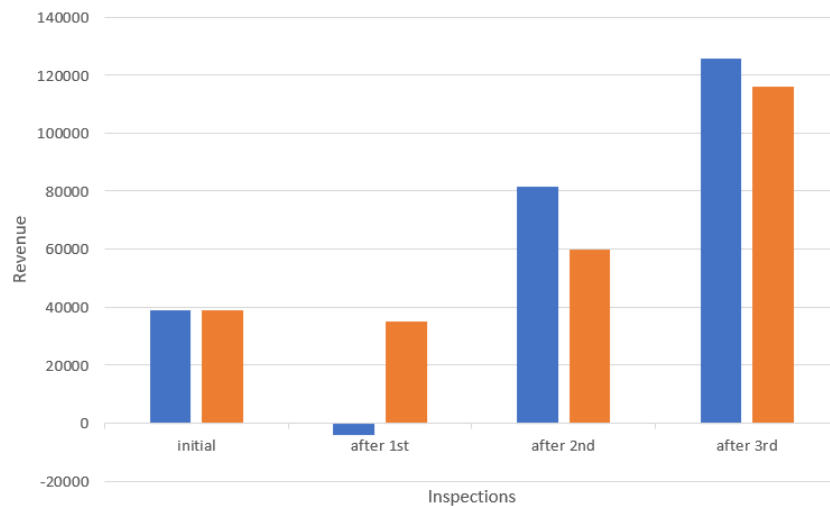


Fig. 20. The Revenue Dynamics for Different Strategies of Auditing

after the first audit, net tax revenue has a steady upward trend, more evident for the strategy of pressing, than for the strategy of information collecting.

When studying the dynamics over a long-term period (figure 21), for example, 10 inspections, which is equal to 10 years, it is easy to see that net tax revenue stabilizes at approximately the same values, regardless of the chosen strategy. However, in the short term (less than 5 years), the strategy of pressing is obviously more profitable. The similar results were obtained for other network configurations (grid and weakly connected network) used for performed simulations.

Comparison of the network configurations. Next, let's investigate the effect of the type of the considered network on the tax collection. As one can see from the previous example, a large number of inspections smooths out the differences

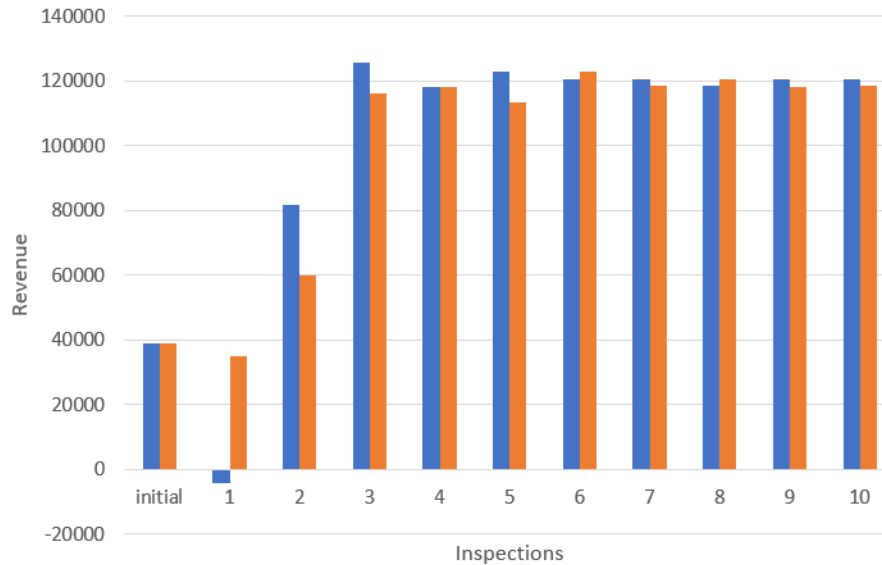


Fig. 21. The 10-years Revenue Dynamics

between strategies, therefore, we focus on the first iterations of the method under consideration. Let's estimate the cumulative system revenue for the first three checks for the same initial distribution of risk statuses (17% of risk-avoiding, 65% of risk-neutral and 18% of risk-loving) and different network configurations (strongly and weakly connected random graphs and grid).

First, let's consider the results of simulations obtained for the strategy of pressing. Comparison of all three network configurations by their profitability is illustrated in the figure 22: the green polyline demonstrates the revenue dynamic modeled on the weakly connected graph, yellow polyline shows dynamic for the strongly connected network, and red color corresponds to grid.

It can be seen from the graphs that in the long term period, the largest model revenue is possible on a strongly connected network, while in the short term, the simulation gives the most effective results for a weakly connected network. The grid is proving to be the most tax-ineffective model.

The differences in the dynamics of net tax revenue obtained on the networks of all three configurations for the information collecting strategy are less obvious (see figure 23).

But still, as earlier, the most efficient in terms of collecting taxes is the strongly connected network (yellow polyline), the most ineffective is the grid (red polyline).

Comparison of the Initial Distributions of Risk Statuses. For a change, let's consider the results of simulating a strategy of information collecting for a network type "grid" and obtain the revenue dynamics for different initial distribution of risk statuses among the taxable population.

Figures 24 – 29 show the initial states of the network, represented as a grid, for the corresponding initial distributions of the risk propensity among the agents of the population.

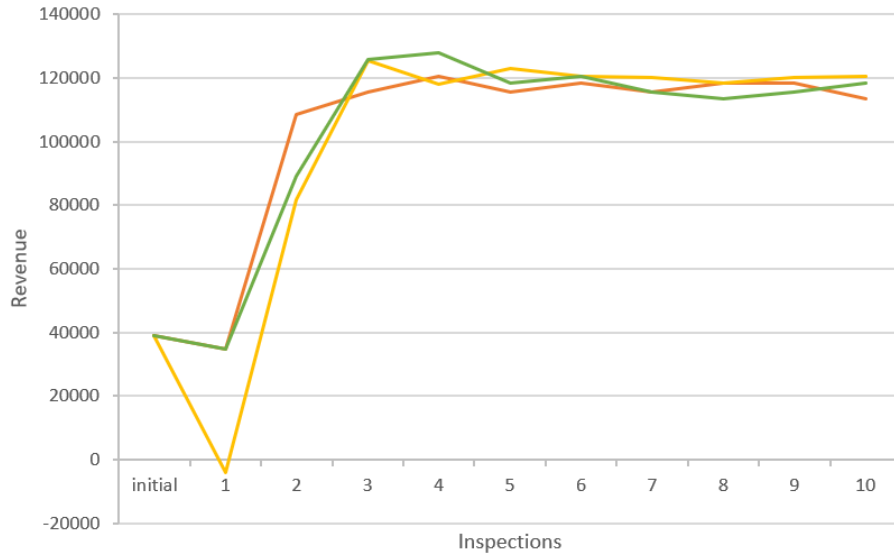


Fig. 22. The Revenue Dynamics for Different Network Configurations: the Strategy of Pressing

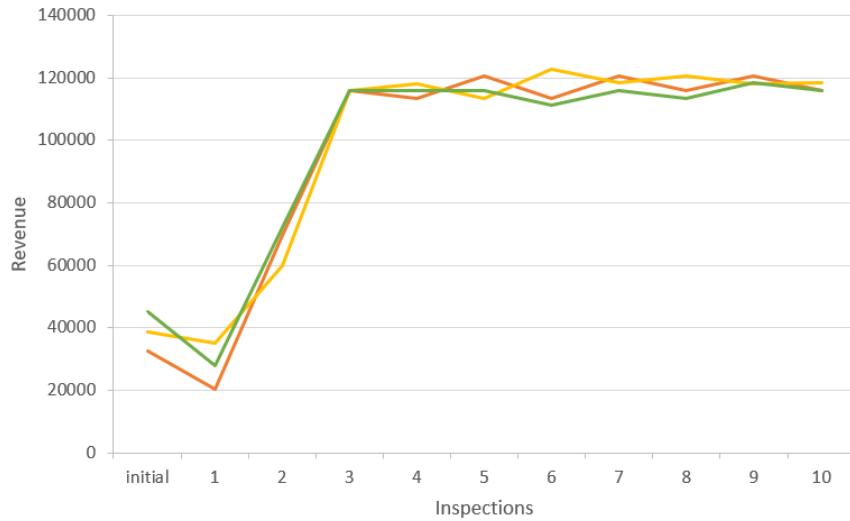


Fig. 23. The Revenue Dynamics for Different Network Configurations: the Strategy of Information Collecting

Figure 24 illustrates network of type 1 and figure 25 illustrates network of type 2 for the case when the portions at the initial time moment are following: 25% of risk-avoiding, 65% of risk-neutral and 10% of risk-loving (according to the assumption

that the portion of risk-averse can be represented as a quantile of the given distribution).

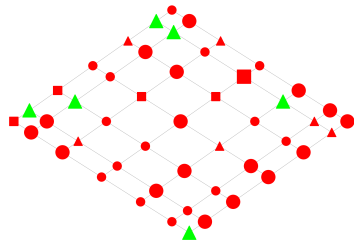


Fig. 24. Network 1. Initial State of the Population: Declared Income

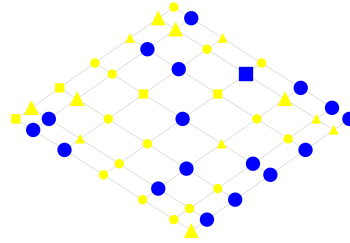


Fig. 25. Network 2. Initial State of the Population: Compliance of the Declared Income with Real

Another initial distribution – 17% of risk-avoiding, 65% of risk-neutral and 18% of risk-loving – corresponds to the case of simulation related to the psychological study mentioned before (Niazashvili, 2007). This case of modelling is illustrated in the figures 26 and 27.

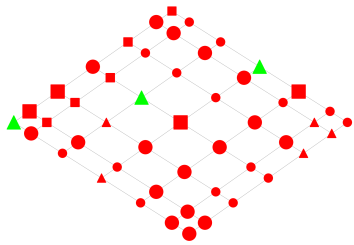


Fig. 26. Network 1. Initial State of the Population: Declared Income

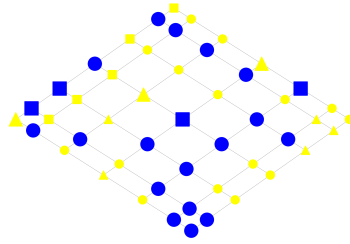


Fig. 27. Network 2. Initial State of the Population: Compliance of the Declared Income with Real

The third case of initial distribution is an opposite to the first one: 10% of risk-avoiding, 65% of risk-neutral and 25% of risk-loving. It is represented in the figures 28 and 29.

Since in the previous examples we saw that the most significant leaps in revenue dynamics occur during the first inspections, we will restrict our consideration by the results of simulations of first five inspections. The figure 30 illustrates comparison of the results obtained on the grid for different initial distributions of risk propensity. In the diagram the green bars represent the tax authority’s income after the respective inspections for the case of 10% of risk-avoiding, 65% of risk-neutral and 25% of risk-loving, the blue bars – 17% of risk-avoiding, 65% of risk-neutral and 18% of risk-loving, and the yellow bars – 25% of risk-avoiding, 65% of risk-neutral and 10% of risk-loving.

It can be seen from the diagram 30 that, starting from the third audit, for all subsequent inspections it is not significant which initial distribution of risk statuses

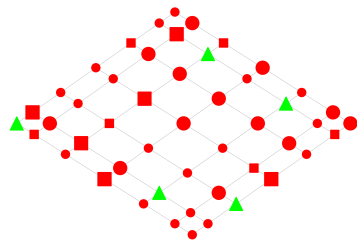


Fig. 28. Network 1. Initial State of the Population: Declared Income

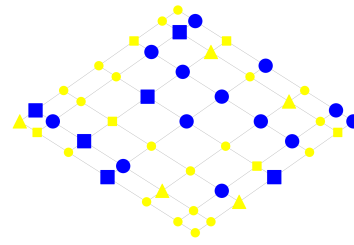


Fig. 29. Network 2. Initial State of the Population: Compliance of the Declared Income with Real

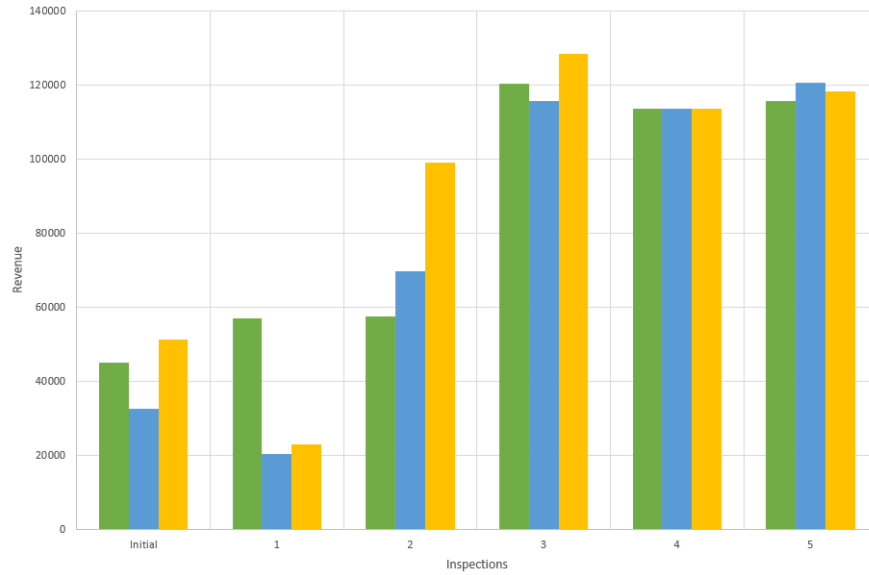


Fig. 30. The Revenue Dynamics for Different Initial Distributions on the Grid and Strategy of Information Collecting

was assumed in the simulation. The impact of different shares of risk averse on the revenue of tax authority is smoothed out. Considering the first iterations, it is worth noting that the higher the proportion of risk averse, the greater the drop in the net tax revenue after the first audit and the higher the jump after the second. After the second inspection, the largest share of risk averse agents (25%) brings the maximum revenue, the same result is obtained after the third one, but there is no longer a direct relationship here. The smallest proportion of risk-averse in the population (10%) leads to a slower but more stable growth in tax revenue.

Simulations carried out for networks of other configurations and using the strategy of pressing give the similar conclusions.

4.3. Numerical Simulation: Results

A series of the experiments carried out before gives the following conclusions.

First, it was found that the looping process occurs early enough – no later than the in the fifth iteration, each of which corresponds to a simulation of a new tax audit. From the standpoint of assessing the revenue of the tax authorities, it means that its values will not change significantly during subsequent inspections, only fluctuating around the values reached in the first iterations. Obviously, the number of evaders in the network will also not change significantly, although actual evasions can be redistributed among agents.

Second, the results of experiment show that the strategy of pressing is more effective (in terms of increasing net tax revenue) than information collecting strategy. The chosen auditing strategy and network configuration also affect the rate of change in the population structure (system stabilization). Moreover, the type of network that best simulates the social connections of the studied population also has a significant impact on the efficiency of the process in terms of tax collection. The current work considered simulations on three types of networks. If they are arranged in descending order of efficiency, the following sequence will be obtained: strongly connected network, weakly connected network, grid.

Third, it was found that an increase in the initial share of risk averse has no stable influence on the dynamics of the tax authority's revenue in long-term period (for a big number of iterations) and gives big jumps in short-term period (during first several iterations).

5. Conclusions and Prospects

Such important problem as the tax evasions was studied in the framework of game-theoretical modelling based on the idea of Bayesian updating. The tax auditing is considered not only as the tool to collect taxes, but also as the process of monitoring of population and gathering information about taxpayers. According to our assumption, another task the tax control is designed for is the dissemination of information about future inspections in order to stimulate taxpayers to pay taxes honestly. This multipurpose process is modeled as a repeated game which is based on a static model of interaction between the tax authority and taxpayers according to the “principal-agent” hierarchical scheme and the Bayesian procedure for re-evaluating the probabilities of tax evasion by the principal, on the one hand, and the probabilities of tax audits by agents, on the other hand.

To investigate the proposed approach, a series of simulations were run. The simulation results allow to study 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 fiscal system and increases the total number of honest taxpayers.

Second, the performed simulation shows that the “natural” way of the information dissemination is not so effective as it was expected: it provides a very slow spreading of information among taxpayers. In addition, the circulating of information only among the nearest neighbors leads to a fairly fast “looping” of the process. In order to bring the simulation results closer to real processes, in the future it is planned to replace this method of information dissemination by a Markov random process in the network. Moreover, to make the studied model closer to real life, we

should take into account the information confidence coefficient, which is not equal to one.

Third, the dependence between the net tax revenue and the initial distribution of risk statuses is not stable for both types of the strategies of tax auditing. On the other hand, we obtained that the dependence between the final distribution of the evasions and the network configuration is very high. Moreover, different types of networks affect the revenue of the tax authority in different ways. From this fact another prospect for future research follows: to obtain a statistical distribution of final states of the network running the simulation for more number of times. Moreover, the idea to run a series of simulations for the network which contains a large number of nodes, seems to be a very attractive. We assume that it will help to make the current research closer to reality and, probably, increase the rate of information dissemination.

6. Appendix

6.1. First Appendix

Let's formulate the theorem about conjugate distribution families (see De Groot, 1974).

Theorem 1. *Let X_1, X_2, \dots, 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 $y = \sum_{j=1}^n x_j$.*

6.2. Second 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 $B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx$ is the beta-function.

References

- Antocia, A. and Paolo Russua, P. and Zarrib, L. (2014). *Tax Evasion in a Behaviorally Heterogeneous Society: An Evolutionary Analysis*. Economic Modelling, **42**, 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., Kumacheva, S. (2005). *A model of audit with using of statistical information about taxpayers' income*. Vestnik SPbGU, series 10, **1–2**, 140–145 (in Russian).

- Bure, V., Kumacheva, S. (2010). *A game theory model of tax auditing using statistical information about taxpayers*. Vestnik SPbGU, series 10, 4, 16–24 (in Russian).
- Chander, P., Wilde, L.L. (1998). *A General Characterization of Optimal Income Tax Enforcement*. Rev. of Econ. Studies, **65**, 165–183.
- De Groot, M.H. (1974). *Optimal Statistical Decisions*. Moscow: Mir (in Russian).
- Gubar, E. A., Kumacheva, S. Sh., Zhitkova, E. M., Porokhnyavaya, O. Yu. (2015). *Propagation of information over the network of taxpayers in the model of tax auditing*. International Conference on Stability and Control Processes in Memory of V.I. Zubov, (SCP 2015). Proceedings, IEEE Conference Publications. INSPEC Accession Number: 15637330, pp. 244–247.
- Gubar, E., Kumacheva, S., Zhitkova E., 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. (2017). *Evolutionary Behavior of Taxpayers in the Model of Information Dissemination*. Constructive Nonsmooth Analysis and Related Topics (Dedicated to the Memory of V.F. Demyanov), (CNSA 2017). Proceedings, IEEE Conference Publications, 1–4.
- Gubar, E., Kumacheva, S., Zhitkova, E., Kurnosykh, Z., 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., & Tomilina, G. (2019). *Modelling of the impact of information on tax audits on the risk statuses and evasions of individuals*. Vestnik SPbGU, **15(2)**, 245–258 (in Russian).
- Gubar, E., Kumacheva, S., Zhitkova, E., Tomilina, G., & 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 (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., Gubar, E. A. (2015). *Evolutionary model of tax auditing*. Contributions to Game Theory and Management, **8**, 164–175.
- Kumacheva, S., Gubar, E., Zhitkova, E., 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., 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., & 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.
- Macho-Stadler, I. and Perez-Castrillo, J. D. (2002). *Auditing with signals*. *Economica*, **02**, 1–20.
- Nekovee, A. M., Moreno, Y., Bianconi, G. and M. Marsili. (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., 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, 458–462.

- Samuelson, P. and Nordhaus, W. (2007). *Economics*. Moscow: Williams (in Russian).
- Sanchez, I. and Sobel, J. (1993). *Hierarchical design and enforcement of income tax policies*. *Journal of Public Economics*, **50**, 345–369.
- Vasin, A., Morozov, V. (2005). *The Game Theory and Models of Mathematical Economics*. Moscow: MAKSpres (in Russian).
- Zellner, A. (1997). *Bayesian Analysis in Econometrics and Statistics: The Zellner View and Papers*. Edward Elgar, ISBN 978-1-85898-220-5.
- The web-site of the Russian Federation State Statistics Service. <http://www.gks.ru/>
- The web-site of the Russian Federal Tax Service. <https://www.nalog.ru/>