Approaches to Economic and Mathematical Taxation Models in the Conditions of the Shadow Economy

E.L. Fesina and N.N. Kozlova
Institute of Management, Economics and Finance,
Kazan Federal University, 420008 Kazan, Russia

Abstract: In the market economy the taxation system is one of factors of its stabilization and development. Positive development of this process contains such social phenomenon as tax avoidance. The means hidden from the taxation quite often go to the "hadow" turnover, begin to work for criminal structures, generating the corruption phenomena. Complexity and discrepancy of the taxation system created conditions for different kind of schemes of tax avoidance and numerous address privileges and unreliable forms of calculations led to the individual taxation. In the conditions of violation of tax discipline and concealment of a money by entities the importance is allocated for analytical control procedures for its tax discipline and identifications of the facts of tax avoidance.

Key words: Tax avoidance, "gray" schemes, shadow economy, tax offenses, binary and regression models, social

INTRODUCTION

One of the most effective tools of social and economic policy of Russia are taxes. Taxes allow to form more than 80% of its income of territorial budgets. At the same time, tax policy of the state is insufficiently effective that it requires application of effective measures for the taxation system reforming. Transition to market economy in Russia was followed by deep social and economic crisis which revealed essential shortcomings of the taxation system. The high tax burden undermines opportunities even of simple reproduction at fair taxpayers and promotes forming of large-scale shadow sector in the Russian economy.

The sociological survey conducted among heads of the entities in the sphere of small business in the Republic of Tatarstan in October of 2016 years allowed to reveal the factors complicating development of an entrepreneurship and to estimate actions of the entrepreneurs hiding the income from the taxation (Grigoreva and Fesina, 2014).

Results of sociological questioning of entrepreneurs about the factors constraining development of small business in the Republic of Tatarstan are presented in Fig. 1.

The results of sociological questioning of entrepreneurs about estimates of their actions connected with further development of small business are provided in Fig. 2. Analytical data of law enforcement agencies of
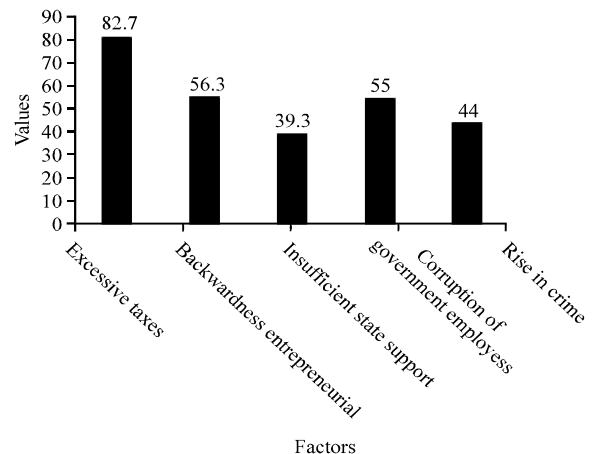

Fig. 1: The factor loadings constraining development of small business in the Republic of Tatarstan (as of 10.01.2016)

Russia demonstrate that over the last 10 years about 50-60% of taxpayer's total number (legal entities and natural persons) didn't perform or not fully performed tax payments in the budget (Grigoreva and Fesina, 2014). From 1st January, 2011 to 1st July, 2016 Russia law enforcement agencies investigated about 100 thousand criminal cases of tax offences and to institute more than 25 thousand persons to criminal proceedings. Only in 2016 the number of the criminal cases was more than 40

Corresponding Author: E.L. Fesina, Institute of Management, Economics and Finance, Kazan Federal University, 420008 Kazan, Russia
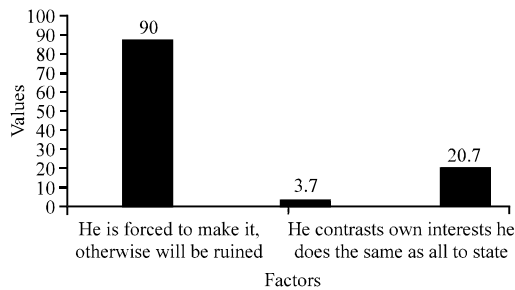
Fig. 2: Estimates of entrepreneur's actions for development of small business in the Republic of Tatarstan (as of 10.01.2016)

thousand which were examined by law enforcement agencies of Russia on fight against tax offenses (Vorontsova and Fesina, 2016). Thus, the change in a legal awareness of taxpayers didn't occur and also it is necessary to recognize that the level of preventive measures on the prevention of tax offenses and crimes is unsatisfactory.

MATERIALS AND MEHTODS

Key question for creation of the strong statistical base for the choice of the taxpayers who are evading paying taxes is availability of good statistical model. The statistical model serves as peculiar "window" through which it is possible to consider the available data on each entity and to solve whether this entity is a promising candidate on conducting exit tax audit or not. In a research models of a binary response (binary models) and regression models were used. In spite of the fact that tax departments of the different countries usually prefer to use one type of models, more preferable option is joint use of models. Models of a binary response allow to determine the level of probability that check of this entity will lead to additional tax charges and regression models about potentially possible amount of additional accruals if check will be productive. For acceptance of reasoned decision, it is desirable to have both types of information (Torgler, 2009).

Theory: Among binary models the discriminant "probit" and "logit" gained the greatest popularity (Tutz and Berger, 2017). In the tax sphere they are used in cases when the quantitative data are analyzed. Such data allow to make a choice between two alternative options to carry out tax audit or not. The modelled variable at the same time always is binary and its expected value shows probability that it will accept unit value. Models are differed by distribution function of dependent variable

using (logit-logistic, probit normal). Both types of models allow to estimate probability that the taxpayer with such characteristics will be "productive" if it will be subjected to documentary check. In the "logit" model the probability of implementation of additional tax charges on results of payer check is described by logistic function:

$$\exp\left(B^{T}X\right)/\left[1+\exp\left(B^{T}X\right)\right]$$

Where:
$X = (X_1, X_2, \ldots, X_n)$ vector of explanatory variables
$B = (B_1, B_2, \ldots, B_n)$ vector of fixed coefficients
T = Transposition sign

As explanatory variables can be fictitious variables accepting value 1 if the payer has some attribute (for example on the tax declaration which arrived from it some nonzero amount of the taxes which are due from it was declared) and 0 otherwise can appear. On the other hand as explaining variables can be also used variables reflecting the amount specified in the report or a ratio of different values. If the explanatory variable is multiplied by the positive coefficient it means that value of this variable is more, than it is a high probability of the fact that check of the enterprise will lead to additional charges. The negative sign of coefficient by which the explanatory variable is multiplied makes an opposite sense. Values of coefficients aren't known in advance, its turn out as a result of statistical assessment of the given specification model by data about the enterprises about which there are both standard data and results of tax audits. For the creation of coefficients assessment of such model it is necessary to have values of the modelled binary variable reflecting charge of surtaxes and also values of the explanatory variables including in model.

The model shall show the level of probability that the entity tax audit will lead to additional tax charges. Having passed set of the submitted declarations through model, it is possible to constitute the list of payers who can be considered as potential candidates on conducting tax audit. As practice, application of any models of a binary response ("logit", "probit" or the discriminant analysis) shows very similar results (Bartelsmana and Beetsm, 2003).

The lack of binary response models is that its model probability of a positive result of documentary check. It means that by results of tax audit additional charges will be made but such model can't predict its size. Its allow to identify more or less precisely the payers trying to evade paying taxes but don't allow to distinguish cases of small and large tax avoidance. Usually tax authorities interest particularly such cases.

Regression models are as a rule, used to explain on what factors value of additional tax charges at payers to whom such charges were made depends. In such models as dependent variable usually appears the variable reflecting effectiveness of documentary check and as explanatory variables as well as in models of a binary response, appear different characteristics of the taxpayer. After determination of unknown coefficients of model, the regression equation can be applied to all set of tax declarations and on each of its can to predict the amount of additional tax charges in case of conducting documentary check. Regression models allow to predict value of additional charges in case of productive check but don't contain information about the probability that check will be productive. Therefore, regression models shall be reinforced by models of a binary response. In tax authorities in practice of the analysis the multiple regression is used which represent regression of a productive sign with two and large number of factors, i.e., follow model:

$$Y = f(X_1, X_2, ..., X_n)$$

Where:
Y = Depend variable
X = Independent (explanatory) variable

In this case, the model of a binary response will give an assessment to probability that check of the entity will be productive and the regression model will allow to determine amount of additional charges when conducting productive check. It will allow to give an assessment of the expected amount of charges. Such information would be useful not only to acceptance of reasoned decision about what entities should be subjected to tax audit but also for the purposes of forecasting of tax revenues.

RESULTS AND DISCUSSION

In case of a choice of taxpayers the regression model and model of a binary response were used. For binary model it was decided to use the logit specification. Assessment of models of such specification was carried out by the maximum plausible method. The form of regression model is selected by logarithmically linear. The assumption is made that the natural logarithm of additional tax charges is the linear function from explanatory variables. The hypothesis was made about that accidental perturbation of regression model is distributed normally with average equal to zero and fixed dispersion. Such specification implies that value of additional tax charges in case of these values of explanatory variables is distributed logarithmically normal

(Berdiev and Saunoris, 2016). The choice of such specification of model is caused by the fact that in practice distribution of additional tax charges usually differs in a high deflection that is small additional charges are appropriated to many taxpayers from among checked. At the same time the small number of taxpayers receives high additional charges (Bayer *et al.*, 2015).

Each of two models also included the absolute term. The coefficient of determination of regression model made 10%. Thus, it turned out that the model is capable to explain only 10% of a cumulative variation of additional charges. If to specify smaller dispersion for random disturbance in the regression equation, then the coefficient of determination would be higher.

The data about tax avoidance were generated for all entities including in to selection. It was made to illustrate the efficiency of choice methodology of the entities for conducting exit tax audits. To make a task more realistic, the assumption is made that results of tax charges are known on performed only for some number of the entities which are selected in a random order. Thus, the choice model of taxpayers was under construction only by data about the entities at which allegedly in the past were performed. For creation of the specification of the "logit" model and regression model three procedures of explanatory variables choice were used: direct selection, the reverse selection and step by step selection. Thus, about three versions of specifications of each model were received. Procedures for specifications choice to the "logit" model were applied to all subselection. All three procedures showed identical result that is the same explanatory variables were selected. These variables were included in the "true" specification according to which basic data were generated. The exception was made by one variable which was included into the "true" specification but didn't enter statistically selected. Procedures of explanatory variables choice were applied only to those enterprises which last check showed the positive results. Application of all three procedures led to an identical set of variables. Only two variables weren't included in the specification. Most likely, those variables were unimportant for a prediction of additional tax charges.

The evaluated "logit" model and regression model with the selected explanatory variables then were used for prediction the amount of potential additional accruals for each of the enterprises which remained unchecked. These expected values are defined as the product of probability of nonzero charge in case of tax audit on the expected value of additional accrual provided that it will be the positive. Value of probability was defined from the "logit" model and the expected value of additional accruals from

regression model. As data which testified tax avoidance, were generated for all entities, the estimation of the chosen statistical model overall performance became possible. Comparison of the result received on model with the greatest possible additional accruals for this purpose is carried out (if check was continuous and when checking the entities which are selected in a random order). The average amount of additional accruals in the amount of 7670 monetary units is received as a result for a case if the entities for check were selected in a random order.

When using the statistical procedure of selection all entities are ranged in increasing order. Thus, the specific amount of additional accruals counting on one check decreased in process of increase in number of checks. Therefore, directed statistical selection of the entities significantly exceeded the random choice on the productivity. It is especially noticeable if few funds are allocated for conducting tax audits and check the small number of the entities.

On the other hand, productivity of procedures of the directed statistical choice was significantly lower, than the greatest possible result. For example, if to foreknow what entities from the available selection hid the maximum amount of profit and didn't pay a tax and would only subject to check these companies, then the average amount of additional accruals counting on one check would constitute 292896 monetary units what is about 4 times higher than average productivity of checks of the first checked entities. The reason of such big discrepancy is that for realness of model parameters of true statistical models were chosen, so that, at the expense of observed characteristics of the entities it would be possible to explain only a small part of a variation of the modelled variable (10%).

As data on what charges were actually carried out by results of tax audits are unknown these data are constructed as a simulated data. For its generation preset values of coefficients of the logit model and regression model. For modeling were selected such values of indexes which were more or less approved with really observable. To give specifications of model more sensible nature for generation of results incomplete specifications of both models were used.

Dispersion of random disturbance in regression model shall be chosen rather big for the reflecting objectively existing rule that the considerable share of additional accruals variation will never to be explained on the basis of the available data (Ehrlich and Saito, 2010). Then by random-number generator it is necessary to construct a number of the simulated amounts which are allegedly in addition added by results of check, so, that

Table 1: The consolidated report of the entities on the income tax

Variables	Values
The number of the entities	584.0
Is added taxes in the conditions of the current	6.5
It is added taxes in the the alternative conditions of legislation,monetary units	6.1
Discrepancy, monetary units	0.4

these amounts were approved both with assumptions of both models, and with specific values of coefficients of these models which were chosen (Vorontsova and Fesina, 2016). As a result, the average size of additional charges on one tax audit constituted 7.3 monetary units whereas the average size of self-charges (charges which the entities specified in the calculations) -50 monetary units. At the same time the positive size of additional charges was appropriated only 52% of the entities that is 48% of total number of the entities were law-abiding taxpayers. For the entities violators the average amount of additional charges constituted 14.1 monetary units. Median value of additional charges on these entities are 460 monetary units that is significantly lower than average value that testifies to high asymmetry of its distribution (Dell'Anno, 2016).

Let's illustrate possibilities of application of statistical model in case of tax legislation change on a dummy example. Let's suppose that in case of the current legislation the total tax liabilities of the entities which entered selection constitute 6.5 million rubles. Let's try to answer a question of what will be amount of tax revenues in case of increase in a rate of the income tax and depreciation charge rate for 1%. For the solution of this task as a basis for forming of the alternative legislation the operating version of the legislation must be used to which the user makes necessary changes (Korndorfer *et al.*, 2014). These changes are set in the dialogue mode through system of the menu then "the tax calculator" is started and the program issues results of calculations which are provided in Table 1.

Aggregative data which show only total change of tax revenues as a result of the offered modification of the tax legislation are provided in Table 1. By results of calculations it is visible that reducing tax revenues is expected only on 0.4 million rubles that is less than for one percent from the taxes accrued in the conditions of the current legislation. At first sight it seems insignificant change, howeverm it is necessary to find out from what entities which included in to selection the tax burden will change. Also, it is necessary to find out in what degree change of the tax legislation will affect spheres of economic activity of these companies. Generally, it is possible to choose any indicator which will carry out a role of a classification sign and to set its boundary values (Mazhar and Meon, 2017). It can be number of personnel,

Table 2: Results of practical application of statistical model

The number of declarations under inspections, coverage of payers have been carried out by tax audits (%)	The greatest possible average amount of additional accruals counting on one check, monetary units	The actualaverage amount of additional accruals (counting on one check) received with use of methods of selection, monetary units	The expected average amount of additional accruals counting on one check (statistical forecast), monetary units
50 declarations (2.0% of selection)	263548	94856	89450
100 declarations (4.0% of selection)	125432	65417	62856
300 declarations (7.0% of selection)	85967	49873	39634
500 declarations (10.0% of selection)	56390	38129	45651

income or cost of fixed assets. Results of calculations are presented in Table 2. These results are made from the assumption that the highest rank has been appropriated to all considered declarations at which its documentary check will undergo. In practice the accuracy of forecasts will depend on stability of random disturbance of regression and the amount of selection on which the model is estimated (Goel and Nelson, 2016). The measure of exact values of additional accruals can be received rather easily and always to construct a confidence interval of the necessary size in which true value of additional accruals will be.

CONCLUSION

In practice it is necessary to deal with huge dispersion in law-abidingness degree among taxpayers with very similar characteristics. Therefore, it is impossible to hope that with 100% probability it will be possible to find the entities which are evading paying taxes. In a research efficiency of application of statistical models which allow to analyze even a small part of a variation of additional accruals but which can be explained by means of observed characteristics of taxpayers is shown. Even, if this part of additional accruals is rather small, the potential benefit from use of this information can be rather powerful. Results from statistical models can be used for forecasting of additional accruals in case of different degrees of a scope of payers tax audits still before these inspections are carried actually out. Such information is demanded not only when planning exit tax audits but also in case of acceptance of reasoned decisions in the field of tax policy when statistical models are irreplaceable sources of its obtaining.

REFERENCES

Bartelsmana, E.J. and R.M.W.J. Beetsm, 2003. Why pay more? Corporate tax avoidance through transfer pricing in OECD countries. J. Publ. Econ., 87: 2225-2252.

Bayer, R.C., H. Oberhofer and H. Winner, 2015. The occurrence of tax amnesties: Theory and evidence. J. Publ. Econ., 125: 70-82.

Berdiev, A.N. and J.W. Saunoris, 2016. Financial development and the shadow economy: A panel VAR analysis. Econ. Modell., 57: 197-207.

Dell'Anno, R., 2016. Analyzing the determinants of the shadow economy with a separate approach: An application of the relationship between inequality and the shadow economy. World Dev., 84: 342-356.

Ehrlich, I. and T. Saito, 2010. Taxing guns vs. taxing crime: An application of the market for offenses model. J. Policy Model., 32: 670-689.

Goel, R.K. and M.A. Nelson, 2016. Shining a light on the shadows: Identifying robust determinants of the shadow economy. Econ. Modell., 58: 351-364.

Grigoreva, E.A. and E.L. Fesina, 2014. Economic security as a condition of institutional support of economic modernization. World Appl. Sci. J., 31: 940-948.

Korndorfer, M., I. Krumpal and S.C. Schmukle, 2014. Measuring and explaining tax evasion: Improving self-reports using the crosswise model. J. Econ. Psychol., 45: 18-32.

Mazhar, U. and P.G. Meon, 2017. Taxing the unobservable: The impact of the shadow economy on inflation and taxation. World Dev., 90: 89-103.

Torgler, B., 2009. Friedrich Schneider: The impact of tax morale and institutional quality on the shadow economy. J. Econ. Psychol., 30: 228-245.

Tutz, G. and M. Berger, 2017. Separating location and dispersion in ordinal regression models. Econ. Stat., 2: 131-148.

Vorontsova, V.L. and E.L. Fesina, 2016. Methodological approaches to the construction of an econometric model of hidden profits at the micro level. J. Econ. Educ. Res., 17: 83-89.