THE APPLICATION OF NON-LIFE INSURANCE RATEMAKING METHODS IN MODELLING TAX ABUSE

ŚLĄSKI PRZEGLĄD STATYSTYCZNY
Nr 17(23)

ISSN 1644-6739
e-ISSN 2449-9765

Marek Leśkiewicz
Wrocław University of Economics, Poland
e-mail: marek.leskiewicz@ue.wroc.pl
ORCID: 0000-0002-1666-8174

DOI: 10.15611/sps.2019.17.03
JEL Classification: C58

Summary: In the most general terms, tax evasion is a breach of tax law. Such a practice is characterised by the economic effect of tax burden elimination or reduction. This paper is aimed at adapting the fixed effects models that are applied in the non-life insurance ratemaking process to modelling tax abuses. The common characteristics of tax risk assessment and non-life insurance ratemaking include for both cases: the number of a given type of events is counted in the examined time period, the distribution of the amount levied as a tax liability for a single event is examined, with the distribution characterised by strong right-skewed asymmetry, and finally the expected amount levied as a tax liability on a single insured/entity is estimated based on these results. In the case of insurance ratemaking, the obtained expected value serves directly the purpose of estimating the insurance premium. In the area of tax risk, the obtained expected values serve the purpose of estimating the scale of abuses arising from the application of a given behaviour pattern.

Keywords: tax abuse, counting variable, compound Poisson distribution, ZIP model.

1. Introduction

The Ordinance of the Minister of Development and Finance on directions and development of the National Revenue Administration (NRA) activity for 2017-2020[1] formulates the following vision:

The National Revenue Administration is a modern organisation that is effective in enforcing taxes, enjoys public trust, supports honest taxpayers and entrepreneurs, combats tax and customs fraud, and protects markets and society.

The document also states and demonstrates the directions and development of NRA activity in a chart and emphasises the necessity to:

- improve the effectiveness of combating financial crimes and abuses,
- develop information technologies,

[1] Official Journal of the Minister of Development and Finance of 2 August 2017, item 144: https://www.mf.gov.pl/c/document_library/get_file?uuid=5d385dc4-69c3-4945-9284-000c6498a59a&groupId=764034.

ŚLĄSKI PRZEGLĄD STATYSTYCZNY
Nr 17(23)

- increase the effectiveness of task performance by NRA employees and officers.

The obvious purpose of the activity of the revenue apparatus is effective tax collection, and an improved collection rate will enable reduction (or at least avoidance of increase) of tax rates. Hence the need to understand all patterns of behaviour that breaches the tax law and to detect its abuses more effectively.

On the one hand, the NRA puts a strong emphasis on minimising unreasonable oppressiveness of the revenue control apparatus as much as possible. On the other hand, entities that are delinquent in terms of tax ought to be detected in a confident and prompt manner. Therefore, one of the NRA's main tasks is to enhance the effectiveness of selecting the entities to be subject to audit. A well-targeted audit will also translate directly into reducing the costs of the activities taken by the revenue apparatus. Automated tax risk assessment is supposed to help select proper entities for audit, inspections and other administrative operations, e.g. in the case of decisions on accelerated VAT refunds, to an increasing extent. The applied risk assessment should rely on statistical and econometric methods (alternatively, on machine learning methods) developed by mathematicians and data analysts cooperating with experts in key disciplines of tax knowledge. Such activities are conditioned on the development of the IT infrastructure in line with the latest trends in the IT industry, in particular the centralisation of database resources and central data processing.

Automated risk assessment of an entity is not a computer-based indication of the entities to be audited. It is supposed to significantly support NRA employees in performing their operations by, among others:

- separating law-abiding entities, which will permit the shift of their efforts to the entities with increased and high tax risk,
- distinguishing, based on desired criteria, high tax risk entities, the number of which corresponds to the office's auditing capacity,
- providing any digitally available information which facilitates verification work immediately and conveniently such that decisions, selections and other operations could be accelerated and optimised.

In low risk situations, the taxpayer may be requested to perform simple operations only, e.g. adjust tax returns, as a result of which the proper status, as per the tax law, is achieved.

The main objective of this study is to present the possibilities to adapt selected basic non-life insurance ratemaking methods to modelling the magnitude of tax abuses. In particular:

1) abuse frequency,
2) value of a single abuse,
3) total value of abuses

ŚLĄSKI PRZEGLĄD STATYSTYCZNY

Nr 17(23)

of an economic entity when it follows one of the risk behaviour patterns. It is assumed that a catalogue of such patterns is defined and that the vector of the characteristics permitting the construction of econometric models for the enumerated values is given for each entity. The great importance of the total value of abuses is due to the fact that it describes in a fully financial manner the effect of tax abuse committed by the taxpayer once or repeatedly. After summing up the estimated amounts for all taxpayers, we obtain an estimate of the financial effect of using a particular tax fraud pattern in the country, which is useful in estimating the tax gap.

There is no literature that deals with estimating tax abuses. This paper fills this gap. In particular, an original idea based on the adaptation of insurance ratemaking techniques in modelling tax abuse is presented in this article. Unfortunately, econometric criteria, the techniques in detecting tax fraud used in the NRA and detailed empirical examples cannot be disclosed in the content of this article because of the Tax Secrecy law which the author is obliged to respect.

2. Tax audit

The primary objective of the analytical support of tax audit is the effective selection of the economic entities that breach (in one or a variety of ways) the tax law from among the business community. The potential behaviour of the entity in breach of the tax law is to the largest possible (not always) extent described and catalogued.

Some of the selected entities are subjected to tax audit, which, even if costly, time-consuming and highly inconvenient to the taxpayer, is the fundamental and crucial task of the revenue apparatus. A sufficiently comprehensive description of a given economic entity cannot be obtained until the auditing procedure has been instituted. The information available earlier is the one that the taxpayer is obliged to submit directly or indirectly to the revenue apparatus initially, periodically (on a monthly, quarterly, annual basis) and in the case of exceptional events (e.g. change of address, company composition, entry regarding the Polish Classification of Activities in the database of the Central Registration and Information on Business or the National Court Register). In most cases, the information is processed and delivered digitally .

Since at present the greatest tax abuses concern VAT, the discussion will focus on this type of tax.

The basic and most significant pre-audit digital data include registration data, VAT returns, and Standard Audit Files – Tax (JPK-VAT). Advanced analytics on the data obtained from the Clearance House ICT System

(System Teleinformatyczny Izby Rozliczeniowej, STIR[2]), which cover mutual exchange of information between the banking sector and the revenue administration, is currently under implementation.

VAT returns provide summary (monthly or quarterly) information about purchases and sales. JPK-VAT data concern aggregated values from individual invoices. In order to obtain complete information about individual items in an invoice, the revenue clerk still has to visit the economic entity, or ask it to deliver such invoices to the revenue office, or require the information to be submitted in the form of JPK in a more detailed scope than is defined in JPK-VAT. From among all the JPK types (templates) only JPK-VAT is obligatory (for every active VAT payer).

Tax audit is aimed at checking whether the audited entity fulfils its obligations arising from tax laws. Therefore, compliance of the facts with the law is examined. If the audit proves:

- positive – non-compliance is found – the audited entity will be obliged to take measures to remedy the non-compliance or its consequences, if any, and sanctions may be imposed on the audited entity,
- negative – no non-compliance is found – the audited entity will not suffer any negative consequences.[3]

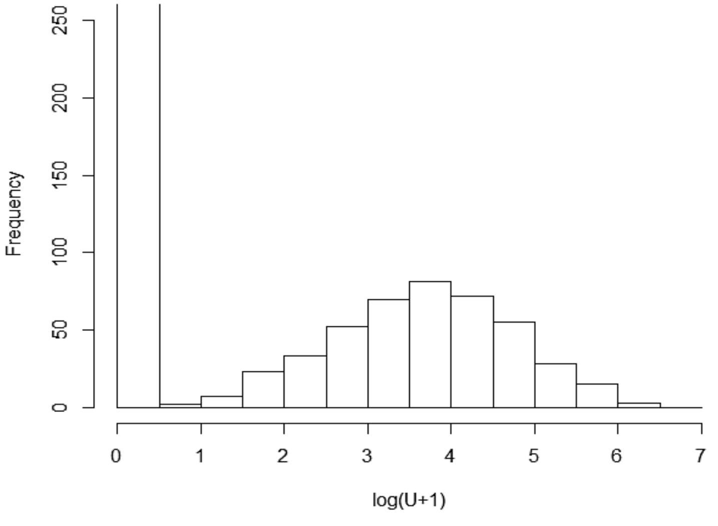

Fig. 1. Distribution of the amount of unlawful reductions of the tax payable levied as a tax liability to entities in breach of the tax law following one of the behaviour patterns. Variable *U* and frequency have been rescaled and the height of the first bar has been trimmed to hide their actual levels

Source: the author's computations based on the results of the Polish NRA's fiscal controls.

---

[2] http://www.itexcellence.pl/aktualnosci/system-teleinformatyczny-izby-rozliczeniowej.

[3] https://www.zakiewicz-adwokaci.pl/prawo/kontrola-podatkowa.html.

ŚLĄSKI PRZEGLĄD STATYSTYCZNY

Nr 17(23)

The 'abuse amount' is a tax liability levied in every positive audit. Such a tax liability is called an unlawful reduction of the tax payable. It is the most important (albeit not the only) financial parameter indicating the scale of abuse and reasonability of the audit's commencement.

Figure 1 presents a histogram for non-negative unlawful reductions of the tax payable levied as a lax liability on entities in breach of tax law following one of the behaviour patterns. Most of the controls are negative or with no financial result, which is expressed in the histogram in the highest number in discrete zero. The positive values of control results are characterized by extreme right-sided asymmetry. The distribution of positive results has a normal or near-log normal distribution. Often the population consists of several subpopulations, where each distribution is mixed discrete-continuous, discrete zero with continuous lognormal distribution.

3. Risk assessment

The amount of an unlawful reduction of the tax payable is a random variable. For the sake of simplification, let us assume that the completed audits are divided into two groups. One is positive audits with a positive unlawful reduction of the tax payable, and the other is the remaining audits, i.e. either negative audits or positive audits with a non-negative unlawful reduction of the tax payable. Let random variable U adopt a value equal to the obtained unlawful reduction of the tax payable in the case of a positive audit with a positive reduction of the tax payable (group one) and the zero value in other cases (group two). Despite the simplification, variable U reflects the level of the abuse diagnosed. The examinations carried out to date have indicated that the distribution of positive unlawful reductions of the tax payable, like most continuous numerical characteristics present in tax risk assessment, is lognormal or quasi-lognormal.

Let us assume that the list of undesirable behaviour patterns designated as $S_1, S_2, \dots, S_m$ is given and that we wish to estimate whether the *i-th* VAT payer ($i \in \{1,2\dots,v\}$), in a time period *T*, follows the *j-th* pattern, $j \in \{1,2\dots,m\}$, and to characterise the *i-th* taxpayer with the use of the vector $[N_i^{S_1}, N_i^{S_2}, \dots, N_i^{S_m}, Y_i^{S_1}, Y_i^{S_2}, \dots, Y_i^{S_m}]$, where:

1. $N_i^{S_j}$ represents the number of breaches of the tax law in pattern S_j by *i-th* taxpayer in time period *T*.
2. $Y_i^{S_j}$ represents the amount of the unlawful reductions of the tax payable obtained as a result of an audit regarding a breach of the tax law in pattern S_j by *i-th* taxpayer in time period *T*.

The expected values of these two variables serve the purpose of estimating the scale of abuse arising from the application of a given (from the list) behaviour pattern. This will permit the auditing activities of the revenue apparatus to be better targeted and adapted to the scale of breaches in terms of cost. Obviously, the cost of the mentioned auditing activities should not be at the level of the abuse amounts, it ought to constitute a small fraction of it, e.g. 1 percent, and, at a further stage of the analysis it should allow for the fact that the cost of the activities taken by the audit apparatus is characterised by decreasing monotony – the relative cost of obtaining high post-audit tax liabilities is smaller than for low ones.

$N_i^{S_j}$, $Y_i^{S_j}$ can be estimated only based on a given entity's observable characteristics $(X_1, X_2, \ldots, X_m)$, yet only those values which are determined (and available) prior to the commencement of auditing activities. Additionally, we assume that the values of these characteristics are available in digital form, with an acceptably low delay.

When used for risk assessment, characteristics X_k $(k \in \{1,2\ldots,m\})$ fulfil a dual role:

1. They are supposed to be statistically significant econometric criteria in at least one of the estimation models for $N_i^{S_1}, N_i^{S_2}, \ldots, N_i^{S_m}, Y_i^{S_1}, Y_i^{S_2}, \ldots, Y_i^{S_m}$ levels.
2. They should be useful information in the process of approving the decision on commencing the audit or later, in the course of the activities taken.

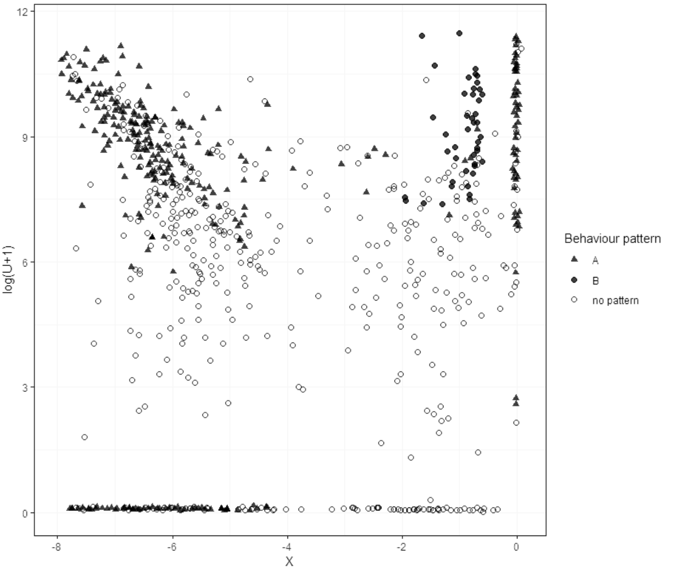

Fig. 2. The amount of the unlawful reductions of tax payable *U* depending on the selected variable *X*. Two behaviour patterns (*A* and *B*) were identified. *Y* axis was rescaled to hide the real level of variable *U*

Source: author's computations based on the results of the Polish NRA's fiscal controls.

It needs to be emphasised that in the majority of cases, the characteristics used for modelling are dichotomous, e.g. "the entity pursues its activity in a virtual office" or "incomplete contact data have been provided".

Figure 2 illustrated the amount of the unlawful reductions of the tax payable depending on the selected variable *X* (the meaning of the variable cannot be disclosed). Two behaviour patterns were identified in the analysed data. This highlights the possibility of constructing variables that allow for building models that effectively identify the behaviour according to the analysed behaviour patterns.

We can construct a vector of signal characteristics $(X_1, X_2, \ldots, X_m)$, which is expanded along with the entity's life, econometric models permitting the estimation of both the number and value of each entity's abuse in a specified time period.

4. Non-life insurance ratemaking

On conclusion of the insurance contract, certain observable factors which describe the insuring party, the object of the insurance, and spatial variables (regarding the geographical region) are known. The observable factors divide the policy portfolio into risk groups, e.g. drivers from the lowest age group are assigned a clearly higher risk level.

The risk level of a given policy depends on unobservable factors such as the driver's skills or driving style, the condition of local roads or the actual condition of the car. Save for the still rare cases where the car has telemetric devices installed, the assessment must be carried out based on the vector of characteristics $(X_1, X_2, \ldots, X_m)$ the value of which is known on signing the contract (normally for a one-year term).

There are usually many policies in non-life insurance, in particular motor insurance portfolios. Their number is designated as *n*. The crucial objective of the examination in this area is modelling the total value of the claims generated by the *i*-th policy

$$L_i = Y_{i1} + \cdots + Y_{iN_i},$$

where $i = 1, \ldots, n$, N_i means the number of claims for the *i*-th policy in the portfolio, while Y_{ik} means the value of the *k*-th claim for $k = 1, \ldots, N_i$. The number of claims (a typical example of the so-called counting variable) is a discrete random variable which assumes a value from the set of non-negative integers. The value of a single claim Y_i is a non-negative random variable which is assumed to be right-skewed.

If:

1) $N_i \sim Pois(\lambda_i), i = 1, \ldots, n$, where λ_i is the expected value of the number of claims,

2) $Y_{i1}, \ldots, Y_{iN_i}$ are independent and have the same distributions and are independent of N_i ,

then the random variable L_i has the so-called compound Poisson distribution. The pure premium π_i is the expected value of the random variable L_i.

If L_i has the compound Poisson distribution, then [Wolny-Dominiak 2014]

$$\pi_i = E(L_i) = E(Y_i)E(N_i)$$

and hence, when estimating the expected value of a single claim and the expected value of the number of claims separately, the estimation of the pure premium $\hat{\pi}_i$ is obtained as a product of estimations for $E(Y_i)$ and $E(N_i)$.

Moreover, in order to calculate the security loading, the following needs to be determined [Wolny-Dominiak 2014]

$$Var(L_i) = E^2(Y_i)Var(N_i) + E(N_i)Var(Y_i), i = 1, \ldots, n$$

and hence, when estimating the expected value of a single claim and the expected value of the number of claims separately, the pure premium $\hat{\pi}_i$ is estimated as a product of estimations for $E(Y_i)$ and $E(N_i)$. The current practice of insurance companies in *a priori* ratemaking is the application of two regression models:

1) the claim severity model for $E(Y_i)$ estimation,
2) the claim frequency model for $E(N_i)$ estimation.

They use the information about the number of claims and value of claims for individual policies to this end [Dimakos, Di Rattalma 2002; Wolny-Dominiak 2011; Antonio,Valdez 2012].

The models where risk factors are determined separately for the number of claims and for the value of risk allow the estimation of the pure risk premium. The alternative solution provides for the use of only one regression model where the dependent variable is the total value of claims for a single policy [Wolny-Dominiak 2014].

One of the classical counting models applied in insurance ratemaking is GLM-POIS, which is based on two components: $N_i \sim Pois(\lambda_i)$ and $\lambda_i = E_i \exp(x_{(N)i}^T \beta_{(N)})$, where $\beta_{(N)=}(\beta_{(N)0,\ldots,}\beta_{(N)k})^T$ is a vector of the risk factors influencing the number of claims, whereas $x_{(N)i}$ is a vector of observations for the i-th policy. Parameter E_i is risk exposure, that is a technical adjustment in the case of a non-standard (usually other than a one year) contract term.

The Poisson distribution conditioned on the values of independent variables through λ_i dependent on certain dependent variables characterising the insured and the object of the insurance, e.g. gender, age, engine capacity, is applied for modelling the number of claims. We get

$$\lambda_i = \mu_i = E_i \exp(x^T_{(N)i}\beta_{(N)}) .$$

It should be noted that the expected number of claims is positive for every combination of independent variables. When applying this model, it is assumed that the expected value is equal to variance, that is it does not allow for the frequent situation of the over-presence of zero values (there was no claim for the majority of the policies), which automatically translates into over-dispersion (that is clearly higher than the average one). In the ZIP (Zero-Inflated Poisson) model, independent variables N_i assume zero values with probability ω_i or values from the Poisson distribution with probability 1-ω_i, that is for $i = 1, \dots, n$ [Wolny-Dominiak 2013]. Then

$$P(Y_i = 0) = \omega_i + (1 - \omega_i)e^{-\lambda_i} ,$$

$$P(Y_i = y) = (1 - \omega_i)\frac{{\lambda_i}^y e^{-\lambda_i}}{y!} \text{ if y} > 0 .$$

Such an operation requires the introduction of the model for ω_i in the form

$$ln\left(\frac{\omega_i}{1 - \omega_i}\right) = \gamma_0 + \sum_{j=1}^{t} \gamma_{ji} Z_{ji} ,$$

where $\gamma_{(N)}=(\gamma_{(N)0,\dots,}\gamma_{(N)k})^T$ is a vector of risk factors contributing to the over-presence of zero values, and $(Z_1, Z_2, \dots, Z_t)$ is a vector of the characteristics used in the model.
In the ZIP model:

$$E(Y_i) = \lambda_i(1 - \omega_i) ,$$

$$Var(Y_i) = (1 - \omega_i)(\lambda_i - \omega_i\lambda_i^2) .$$

Over-dispersion, if any, which often occurs for counting variables, results in the overestimation of the statistics testing the significance of the parameters, yet it does not influence the consistency of the estimators [Wolny-Dominiak 2013].

A classical claim value model employed in insurance is GLM-gamma in the form $Y_i \sim Gamma(\mu_i, \varphi_{(Y)})$, $\mu_i = \exp(x^T_{(Y)i}\beta_{(Y)})$, where $\beta_{(Y)}=(\beta_{(Y)0,\dots,}\beta_{(Y)k})^T$. A vector of the parameters has the form $(\beta_{(Y),}\varphi_{(Y)})^T$, where $\varphi_{(Y)}$ is an unknown dispersion parameter of the gamma distribution and needs to be estimated [Wolny-Dominiak 2014].

5. Tax risk assessment and insurance premium

The classical exercise concerning non-life insurance ratemaking is formulated as follows:

Let us consider a *portfolio of n non-life insurance policies*. One random variable with a specific distribution, designated as L_i, where $i = 1, \dots, n$, corresponds to one *policy*. Let us designate the *number of claims for the i-th policy in the portfolio* as N_i, and the corresponding *value of a single claim* as Y_{ik}, where $k = 1, \dots, N_i$. Then, the variable L_i has the form: $L_i = Y_{i1} + \cdots + Y_{iN_i}$ and specifies the *total value of claims generated by the i-th policy* [Wolny-Dominiak 2014].

The expected value of the total value of claim L_i is estimated based on the observable risk factors which fulfil the role of independent variables for regression modelling.

The tax auditing apparatus also has an observable set of factors characterising an economic entity (incompletely) in terms of risk level and the main objective is to estimate the total value of abuses of L_i. Therefore, the classical exercise concerning non-life insurance ratemaking for the purpose of tax audit would be as follows:

Let us consider *n economic entities in the area of VAT*. The abuse amount, i.e. a random variable with a specified distribution designated as L_i, where $i = 1, \dots, n$, is levied as a tax liability on each law-breaching taxpayer. Let us designate the *number of abuses in pattern A of the i-th taxpayer* as N_i and the corresponding *value of the k-th abuse* as Y_{ik}, where $k = 1, \dots, N_i$. Then, the variable L_i has the form: $L_i = Y_{i1} + \cdots + Y_{iN_i}$ and specifies the *total value of the abuses generated in pattern A by the i-th taxpayer*.

The common characteristics of tax risk assessment and non-life insurance ratemaking include for both cases:

- the number of a given type of events is counted in the examined time period,
- the distribution of the amount levied as a tax liability for a single event is examined, with the distribution characterised by strong right-skewed asymmetry,
- the expected amount levied as a tax liability on a single insured/entity is estimated based on these results.

In the case of insurance-rate-making, the obtained expected value π_i (called a pure premium or a net premium) serves directly the purpose of estimating the insurance premium (the net premium increased by administration and acquisition costs and security loading, various surcharges and adjustments, is called the gross premium).

As regards modelling the number of abuses, the ZIP model presented above could be adapted from the theory of non-life insurance. It is necessary to take into account the excessive number of zeros in the area of audit. The number of irrelevant audits or ones with non-positive tax liabilities should be reduced, yet they will never be completely eliminated. When counting the number of claims, apart from the Poisson distribution, the application of the binomial or the negative binomial distribution could also be considered [Szekli 2016].

What is common to both described areas of analytical works is (on the ratio scale) the distribution of cash amounts, often right-skewed claim values in the case of insurance, and always an extremely asymmetrical right-skewed post-audit tax liabilities in the tax area. Apart from the mentioned gamma distribution, the Tweedie subfamily (with variance in the form $Var(Y_i) = \mu^p$), inverse normal, normal, and lognormal distributions are applicable in rate-making in the area of insurance modelling [Wolny-Dominiak 2014]. In the area of tax audit, analytical works should be concentrated, at least at the first stage of the examination, on the lognormal distribution, that is the GLM model with the dependent variable in the form $\ln(Y_i)$ and $\mu_i = \exp(x_{(Y)i}^T \beta_{(Y)})$, where $\beta_{(Y)}=(\beta_{(Y)0,\dots,}\beta_{(Y)k})^T$ is a vector of factors influencing the tax risk level, while $x_{(N)i}$ is a vector of observations for the i-th taxpayer.

Table1. Information regarding the common terms for non-life insurance ratemaking and tax abuse analytics

Term	*Non-life insurance*	*Tax abuse*
N_i	number of claims	number of abuses
Y_{ik}	value of a single claim	value of a single abuse
distribution Y_{ik}	exponential family of distributions	lognormal distribution
estimation method	one and two-model	two-model
modelling purpose	determination of the net premium	adaptation of auditing expenditure to the level of abuse

Source: author's own elaboration.

Table 1 summarises the information regarding the common terms for non-life insurance ratemaking and tax abuse analytics. For N_i and Y_{ik}, the differences concern only the nomenclature. The type of distribution in the case of tax analysis can be narrowed down only to lognormal distribution. The method of estimation can be limited to the one-model.

6. Conclusion

The aim of this paper was to signal the possibilities of adapting the classical models used in insurance ratemaking to tax risk assessment. In the case of modelling the number of tax abuses for individual tax behaviour patterns, the described ZIP model could be applied directly. As regards modelling the value of the abuse, the post-audit data imply that analytical works could be narrowed down to the lognormal distribution of the abuse value.

Bibliography

Antonio K., Valdez E.A., 2012, *Statistical concepts of a priori and a posteriori risk classification in insurance*, AStA Advances on Statistical Analysis, 96(2), Journal of the German Statistical Society.
Dimakos X.K., Di Rattalma A.F., 2002, *Bayesian premium ratemaking with latent structure*, Scandinavian Actuarial Journal, 3.
Lambert D., 1992, *Zero-inflated poisson regression, with an application to defects in manufacturing*, Technometrics, vol. 34, no. 1.
Szekli R., 2016, *Matematyka ubezpieczeń majątkowych i osobowych*, Wydawnictwo uniwersytetu Ekonomicznego we Wrocławiu, Wrocław.
Wolny-Dominiak A., 2011, *Analiza porównawcza modeli mieszanych szacowania stóp taryf w ubezpieczeniach majątkowych z wykorzystaniem kroswalidacji*, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, nr 207.
Wolny-Dominiak A., 2013, *Modelowanie liczby szkód w ubezpieczeniach komunikacyjnych w przypadku występowania dużej liczby zer, z wykorzystaniem procedury kroswalidacji*, Studia Ekonomiczne, 124.
Wolny-Dominiak A., 2014, *Taryfikacja w ubezpieczeniach majątkowych z wykorzystaniem modeli mieszanych*, Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach, Katowice.
Zarządzenie Ministra Rozwoju i Finansów w sprawie kierunków działania i rozwoju Krajowej Administracji Skarbowej na lata 2017-2020, Dz. Urz. Min. Roz. i Fin. z dnia 2 sierpnia 2017 r., poz. 144.

ZASTOSOWANIE METOD TARYFIKACJI UBEZPIECZEŃ MAJĄTKOWYCH W MODELOWANIU NADUŻYĆ PODATKOWYCH

Streszczenie: Unikanie opodatkowania jest zachowaniem sprzecznym z prawem podatkowym. Tego typu działanie charakteryzuje się skutkiem ekonomicznym w postaci wyeliminowania lub redukcji obciążenia podatkowego. Celem artykułu jest adaptacja modeli z efektami stałymi stosowanych w procesie taryfikacji ubezpieczeń majątkowych do modelowania częstości nadużyć podatkowych, wartości pojedynczego nadużycia i łącznej wartości nadużyć. Cechami łączącymi ocenę ryzyka podatkowego i taryfikację ubezpieczeń majątkowych jest to, że w obu przypadkach zliczamy w badanym przedziale czasowym liczbę zdarzeń określonego typu, badamy rozkład kwoty przypisanych pojedynczemu zdarzeniu, przy czym rozkład charakteryzuje się mocną asymetrią prawostronną, na podstawie tych wyników szacujemy kwotę oczekiwaną przypisaną do pojedynczego ubezpieczonego/podmiotu. W przypadku taryfikacji ubezpieczeniowej otrzymana wartość oczekiwana służy bezpośrednio do oszacowania składki ubezpieczeniowej. W obszarze ryzyka podatkowego uzyskane wartości oczekiwane służą do oszacowania skali nadużyć wynikających ze stosowania poszczególnego schematu zachowań.

Słowa kluczowe: nadużycie podatkowe, zmienna zliczająca, złożony rozkład Poissona, model ZIP.