TAX EVASION: THE DISCOURSE AMONG GOVERNMENT, BUSINESS AND SCIENCE COMMUNITY BASED ON BIBLIOMETRIC ANALYSIS

ABSTRACT. This article analyzes the publications relating to the problem of tax evasion. This topic is attractive not only for the academic community, but also for public at whole. The article explores to what extent the scientific publications on tax evasion correspond to practical issues discussed among the stakeholders. We used the electronic database of e-Library as a source of scientific publications on the subject. The principal stakeholders directly dependent on the taxation are the taxpayers and public authorities. We used the electronic database of publications «Kommersant» publishing house and the «Rossiyskaya Gazeta» to reflect issues discussed among the stakeholders. We selected for analyze 301 publications for the period of 2013–2015. The study was conducted by comparing the publication activity by types and period of publications. In the first stage of the study we have done the qualitative content analysis by identification the common themes discussed in hole sample of publications. Then, a quantitative analysis was conducted by comparing the distribution of publications on a particular topic from each source. We used bibliometric analysis method for the quantitative and bibliographic mapping method to visualize the results of research. Calculations were performed using the software QDA Miner v.5.0 module WordStat v.7.1.7. As a result, studies have concluded that the most popular topics of interest for which no changes are: changes in legislation, legislation and increased enforcement. Using the results of the conducted study, we can identify the main similarities and differences between the monitored sources. We can see the special attention to the: Legislation changes, Law enforcement, Entrepreneurship. Marked reduction of interest can be noted regarding to the following topics: International aspects of taxation, Shadow economy, Ownership, property, investment. The growth of interest can be noted in relation to the following topics: Directorship, Article of the Tax Code, Short-lived companies, Arrears and fines. The study revealed a certain disparity between the topics discussed among academic community and stakeholders. The topics discussed in the majority of scientific texts (shadow economy, corruption, the firm one-day, social security contributions), a much rarer can be found in the publication of «Kommersant» and «Rossiyskaya Gazeta» which focuses mainly on matters of legislation. Analysis of the relationships in the texts according to the source and year of publication showed that research topics converge with issues considered by the public authorities. The business community more involved in discussion the legal issues, because the government notion works upon the impression about tax evasion of the business community and academia. Thus, bibliometric text analysis techniques can be used for research, preparation of literature reviews and thematic information retrieval.

KEYWORDS: tax evasion, bibliometrics, bibliometric analysis, bibliometric mapping, content analysis, scientific publications, electronic publications database.

FUNDING. This work was financially supported by the State Task № 26.1348.2014/К to fulfill research work as part of Project 1348 «Influence of shadow economy on the quality of life in Russia and Ukraine: a comparative analysis» (State registration number in FGASI CITaS: 114091140015).
УКЛОНЕНИЕ ОТ УПЛАТЫ НАЛОГОВ:
БИБЛИОМЕТРИЧЕСКИЙ АНАЛИЗ ТОЧЕК ЗРЕНИЯ ВЛАСТИ,
БИЗНЕСА И НАУКИ

АННОТАЦИЯ. Статья посвящена анализу публикаций, касающихся проблемы уклонения от уплаты налогов. Эта тема привлекает пристальное внимание не только научного сообщества. В статье исследуется соответствие проблематики научных работ по уклонению от уплаты налогов практическим вопросам, обсуждаемым заинтересованными лицами. В качестве источника научных публикаций по данной тематике использовалась электронная база e-Library. В круг заинтересованных лиц, напрямую зависящих от правил налогообложения, входят бизнес-сообщество и государственные органы. Для них источниками информации по исследуемой теме являются электронная база публикаций издательского дома «Коммерсантъ» и «Российская газета». Для анализа отобрана 301 публикация за 2013–2015 гг. Изучение соответствия проблематики проводилось путем сравнения публикационной активности в разрезе видов публикаций. На первом этапе исследования был выполнен качественный контент-анализ посредством выявления общих тем, обсуждаемых в публикациях. Затем проводился количественный анализ через сравнение распределения публикаций по конкретной теме из каждого источника. Для количественного анализа и визуализации результатов использовались методы библиометрического анализа и картирования. Расчеты производились с помощью программного продукта QDA Miner v.5.0 модуль WordStat v.7.1.7. В результате исследования были сделаны выводы, что самыми популярными темами, интерес к которым не меняется, являются: изменение законодательства, законотворчество и усиление принуждения. Темы, к которым за рассматриваемый период снизился интерес, касаются международных аспектов налогообложения, теневой деятельности, собственности и инвестиций. Отмечено выраженное возрастание интереса общества к фирмам-однодневкам, руководству компаний, а также к вопросам штрафов и пеней. Исследование позволило выявить определенное несоответствие тем, обсуждаемых бизнесом и властью, по сравнению с темами научных публикаций. Распространенные в научных публикациях темы (теневая экономика, коррупция, фирмы-однодневки, взносы на социальное страхование), гораздо реже встречаются на ресурсах издательского дома «Коммерсантъ» и «Российской газеты», сосредотачивающих основное внимание на вопросах законотворчества и обсуждения изменений в законодательстве. Анализ взаимосвязей в текстах в соответствии с источниками и годом публикации показал, что темы научных исследований сближаются с проблемами, рассматриваемыми властью, а бизнес-сообщество в большей степени вовлекается обсуждение правовой проблематики, т. е. точка зрения власти во многом определяет обсуждение темы уклонения от налогов бизнес-сообществом, и научными кругами. Таким образом, библиометрические методы анализа текстов могут применяться для проведения научных исследований, составления обзоров литературы и тематического поиска информации.

КЛЮЧЕВЫЕ СЛОВА. Уклонение от уплаты налогов, библиометрия, библиометрический анализ, библиографическое картирование, контент-анализ, научные публикации, электронная база публикаций.

ФИНАНСИРОВАНИЕ. Статья подготовлена при финансовой поддержке государственного задания № 26.1348.2014/К на выполнение работ в сфере научной деятельности в рамках базовой части проекта № 1348 «Влияние теневого сектора экономики на качество жизни населения в России и Украине: сравнительный анализ» (номер госрегистрации в ФГАНУ ЦИТИС 114091140015).
Introduction

Setting research questions. The problem of tax evasion is extremely urgent, because it is one of the largest by the breadth and complexity of manifestation. Tax evasion has a destructive impact not only on the scope of taxation, but in general—on all spheres of society, economy and state. At the same time, considerable scopes of tax evasion are explained by the tolerant attitude of society to such offenses. So we decided to find out just how important is the topic of tax evasion from the state, business and science point of view.

Which aspects of tax evasion are the most popular at the moment? How different from each other are the topics in scientific and journalistic literature on issues related to tax evasion? What are the issues and how often they are discussed on the pages of various publications? What topics which are actively discussed by the press have not yet found their scientific understanding? Elucidation of these issues is also necessary to select the scientific priorities and forecasting of future research.

To answer these questions, we decided to analyze several sources of information which reflect different points of view on relevant aspects of tax evasion. Scientific publications contained in the library system «eLibrary» reflect the views of the scientific community, publications in periodicals, articles in «Rossiyskaya Gazeta» which in our opinion expresses government point of view on the problem of tax evasion and the publications of the «Kommersant» publishing house, which can be considered as reflecting the views of business community.

The methodological basis of this research consisted in the special researching techniques of documentary information flows: for the qualitative research we used the content analysis method and for the quantitative research—we used bibliometric analysis method and bibliographic mapping method as an important objective of the research was to visualize its results. Visualization of analysis results in the form of «maps» will present the problem of tax evasion in a variety of problems related to this phenomenon by using a combination of figures and graphs.

Literature review. Special researching methods of documentary information flows bibliometrics, informetrics, scientometrics, webmetrics, cybermetrics, documentrics, mediametrics are actively used for selecting research priorities and forecasting the future research. The essence of these methods is in counting, combining, interpretation, comparison of certain elements of information flow, acting on each other in the process of developing scientific field.

As the first works in this field foreign researchers often points to the work of F. Cole and N. Ealesin (1917) [1], in which they carried out a statistical analysis of literature on comparative anatomy. O. Voverene [2], followed by science historian S. V. Altshuller [3] says that the name of the Russian scholar P. Walden is unfair forgotten, and in 1911 he was the first who used the method of citation analysis to study the contribution of scientists in individual countries in the development of chemistry [4], and thus, should be recognized as the founder of this method, the most commonly used at present time in evaluating the scientific performance. The first attempts of quantitative study of Russian literature flows were made by A. Storch and F. Adelung [5]. They analyzed their own index, which reflects the literature of 1801–1806 gg. according to various parameters. In 1895, the term «Bibliographic Statistics» [6] was introduced for similar works.

In 1969 V. V. Nalimov and Z. M. Mulchenko introduced the term «scientometrics» [7], which later became widespread. It relates to the field of science, science studies, structure, dynamics of scientific activity, the interaction between science and other spheres of material and spiritual life of society, which means, it explores processes and phenomena of scientific activities in the broader context. In 1979, in articles of German authors L. Blackert and S. Siegel [8] and O. Nacke new term «informetrics» was firstly introduced, and it was originally defined as a set of mathematical methods for the study about objects of informational science, the description and analysis of their properties, as well as the
laws in order to optimize these objects when making decisions.

Quantitative research of documentation flow was initially called «statistical bibliography». Then the term «bibliometrics» was introduced, which, according to E. Garfield [9], was first mentioned in the works of P. Otte (1934) and A. Pritchard (1969) [10]. A. Pritchard in his famous article in 1969 defined the bibliometrics as «the application of mathematical and statistical methods in relation to books and other communication media».

The subject of content analysis, or analysis of the content is the content of text arrays and communicative correspondence products. In the national research in tradition content analysis is defined as the quantitative analysis of texts and text arrays for subsequent meaningful interpretation of identified numerical laws1. The term content analysis was first used in the late XIX — early XX centuries by American journalists Matthew B., A. Tenney, D. Spiid, D. Uipkins [6; 11].

In modern Russian literature, devoted to the study of texts on economics, bibliometric studies are rare. We may noted the concept of system-innovation bibliometric analysis and mapping of economic literature (SIBAMEL). The Department of Economics of the National Research University — Novosibirsk State University and the Institute of Economics and Industrial Engineering of the Siberian Branch of the Russian Academy of Sciences created and developed this concept SIBAMEL [12]. The authors elaborated special techniques and software for the SIBAMEL [13].

The research was based primarily on the analysis of non Russian, and foreign economic literature. The first studies were carried out on the basis of the electronic bibliography EconLit. In the future, the authors began to use data from journal articles JEL issues; Social Science Research Network electronic library (www.ssrn.com); electronic bibliography EconLit. From the Russian sources the flows of new books in all sciences, which went on State Public Scientific Technical Library were analyzed. Among the subject areas to which the authors paid particular attention to we can highlight: the financial economy [14]; regional economy [15]; health economics; agricultural economics [16]; environmental economics; energy economy; government regulation [17]; entrepreneurship; accounting; marketing; innovation and others.

A detailed analysis of data about the use of bibliometric methods in the Russian scientific literature cited in the study of I. Yu. Popova, let him made a conclusion that bibliometrics in the overwhelming majority of Russian studies is used for comparative evaluation of research performance of countries, organizations, publications, and sometimes individual scientific fields and very few studies in which the focus is on the identification of new areas of research, especially at the intersections of different subject areas [18].

Tax evasion is precisely located at the intersection of domains, at the crossroads of economic, legal and even political trials.

**Methodology**

**Bibliometric mapping.** Pritchard (1969) [10] explained bibliometrics as «the application of mathematical and statistical methods to books and other media of communication». Later, Hawkins (2001) [19] defined bibliometrics as «the quantitative analysis of the bibliographic features of a body of literature». Within a bibliometric analysis, mainly books, monographs, reports, theses, and papers in serials and periodicals are analyzed; however papers which are published in journals seem to be the most suitable studies for bibliometric research (Glänzel (2003) [20]). For analyzing research literature production (to identify patterns in the literature) bibliometric analysis uses quantitative methods (De Bel-lis (2009) [21]). Moreover, Garfield (2006) [22] is convinced that with bibliometric analysis, we can also examine the history and structure of a field, the flow of information into a field, the growth of the literature, the patterns of collaboration amongst scientists, the impact of journals, and the long-term citation impact of a work.

Bibliometric mapping is used with the purpose to visually represent scien-

---

tific publications based on bibliographic data. With bibliometric mapping we can produce different bibliometric maps which provide an overview of the structure of the scientific publications in a specific research field. One of the most popular ways to use bibliometric mapping is to identify specific research areas within a selected science field, with the purpose of getting a view of the size of the field and relevant subfields, and how they relate to each other (van Eck (2011) [23]). In this way, we can understand the broader aspects of the particular research field (Börner, et al. (2012) [24], (2012) [25]). Visualization of Similarities (VOS) is the new mapping technique, which has been used to create bibliometric maps in various studies (van Eck and Waltman (2007) [26]; van Eck et al. (2010) [27]; Waaijer et al. (2011) [28]). The VOS mapping technique has been implemented in the computer program called VOS viewer (Leiden University, Netherlands) (van Eck and Waltman (2013) [29]). The VOS viewer software has visualization capabilities, therefore bibliometric maps can be displayed in various different ways and consequently emphasize on different aspects of a map. Additionally, VOS viewer allows using different colors to indicate clusters of objects. Moreover, the VOS viewer software also merges terms that may be closely related to term clusters denoted by the same cluster color (van Eck (2011) [23]). According to van Eck, the proximity of the terms can be interpreted as an indication of their relatedness. VOS viewer Version 1.6.0 additionally enables the creation of maps in which terms are colored according to the year of the term’s appearance in the scientific literature.

**Data set and its analysis.** Issues, which attract attention not only of academic community, but also of public at whole, are topics in the field of social sciences. One of such debated topics is taxes, and issues related to tax payment, including tax evasion. In connection with this, the research question of this paper is the following: to what extent the problems of academic articles on this subject correspond to the practical issues which are discussed by stakeholders?

Among the stakeholders we include a circle of subjects who directly depend on the established rules of taxation: taxpayers and public authorities. This article studies issues of correspondence by means of comparison, using the data on publication activities by types of publications, both at the level of qualitative and quantitative analysis. Qualitative analysis here is the identification of common topics discussed in the academic and non-academic publications. Quantitative analysis here is a comparison between the percentage of academic and non-academic publications, which discuss a particular topic, at the total number of publications of this type.

In this research we carry out a comparison of the discourse by the business community, government agencies, and academia. For this purpose, we identify content and frequency of the main topics discussed in the texts from three sources: PH Kommersant (business discourse), Rossiyskaya Gazeta (the discourse of authorities), and eLibrary (discourse of the scientific community). Period of study is 2013–2015.

As an object of the research we selected all the texts in the considered period which contained keywords «evasion, payment, taxes» (including the possibility of changing keywords according to the rules of Russian language). Text characteristics are presented in table 1.

For the analysis of texts we used the program QDA Miner v.5.0 with module

<table>
<thead>
<tr>
<th>Source</th>
<th>2013</th>
<th>2014</th>
<th>2015</th>
<th>Total, all years</th>
<th>2013</th>
<th>2014</th>
<th>2015</th>
<th>Total, all years</th>
</tr>
</thead>
<tbody>
<tr>
<td>eLibrary</td>
<td>13</td>
<td>16</td>
<td>19</td>
<td>48</td>
<td>11</td>
<td>15</td>
<td>24</td>
<td>16</td>
</tr>
<tr>
<td>Kommersant</td>
<td>50</td>
<td>40</td>
<td>30</td>
<td>120</td>
<td>44</td>
<td>37</td>
<td>38</td>
<td>40</td>
</tr>
<tr>
<td>Rossiyskaya Gazeta (RG)</td>
<td>51</td>
<td>52</td>
<td>30</td>
<td>133</td>
<td>45</td>
<td>48</td>
<td>38</td>
<td>44</td>
</tr>
<tr>
<td>Total all sources</td>
<td>114</td>
<td>108</td>
<td>79</td>
<td>301</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 1
In this study we use an approach that combines elements of quantitative and qualitative analysis of texts, partly we follow the methodology described by A. Oleynik [30].

In order to bring the text to a uniform manner, at the preliminary stage of our investigation the initial information has been processed, thus the analysis used only texts, headers, and headers of graphics from the text; i.e. such attributes as date of publication, the author of the publication, graphic objects (tables, figures and diagrams) have been removed.

In the first phase of the research we compiled the dictionary of categories, which included words, which occurred in 10 % and more texts. Words that are most frequently used, but have the least significance (for example, they, do, is) were exclude from the content analysis. Also keywords for search («evasion*», «*payment*», «tax*») were excluded from the content analysis (where the * replaces any character, and designation «---» refers for word forms). The dictionary includes 133 categories. Characteristics of the texts and the dictionary are given in table 2.

The second stage was implemented by extraction of the text’s topics. Extraction of topics was carried out by applying the method of factor analysis. Module WordStat allows using following units for the extraction of topics: document (i.e., all the texts used in the analysis); paragraph; sentence.

For the analysis it can be taken as variables: words or set phrases, or categories defined by researcher.

Factors (latent variables) are recoverable topics which are identified basing on the values of the variables (frequency of variables usage) in the unit of analysis.

Module WordStat can extract topics in two possible ways:
- without usage of the dictionary categories (thereby words are extracted in such forms as they are used in the texts; as well set phrases can be extracted);
- with usage of dictionary categories (which are defined by researcher).

According to the rules of implementing factor analysis, the obligatory condition is the following quantitative restriction: the number of observations should not be less doubled number of variables. In our research the number of analyzed texts is equal to 301; that’s why the number of variables (number of words / categories) should be less than 150. Thus the problem of extracting topics is linked with possibility of lemmatization for the language used in the texts. Such option is absent for the Russian language in this version of WordStat, so the ability of inflexion is achieved by specifying the categories manually. Due to these limitations, topics extracted with the option «without usage of the dictionary categories» are used in our content analysis not for directly quantitative analysis of texts, but as a reference that is taken into account when filling dictionary categories.

Owing to the fact that we identify the dominant topics of publications in all texts selected for analysis, so we extract topics at the document level (not at the level of smaller segments of the text, such as paragraphs or sentences).
WordStat offers the following options: to extract 3 topics, or 8, or 13, or 18, or 23, or 28 (we did not consider here options with more numbers of topics). Trying to get the maximum number of topics which contain meaningful idea, we chose option «23 topics». Thus, 23 topics have been extracted using the document as the unit of analysis, the total percentage of variance explained — 48.6 %. The last of the topics is expressed by a single word — category named «AUTHORIT*». Criterion for determining the number of factors is Kaiser criteria (to drop all components with eigenvalues under 1.0).

Thus, the result of extraction (the listed topics) reflects joint occurrence of words in all analyzed documents, where «joint occurrence» means within one document. Studying the keyword list is of interest in terms of understanding issues which are being most frequently discussed in the context of tax evasion in the Russian sources.

The next step was carried out as a qualitative analysis of the extracted keywords to identify the meaning of each extracted topic. It was found that extracted topics 2 «CRIM*; FELON*» and 7 «SLED-KOM; LTD» are linked by a joint meaning, «Law enforcement». Similarly, topics 14 «DECLAR*; VAT» and 15 «COST*» are connected by the joint keyword «VAT». Topics 21 «PARTNER*; COMMER*» and 22 «BUSINESS*» are also linked by a joint meaning «Business management, entrepreneurial activity». Topic 23 «AUTHORIT*» was excluded from further analysis because it was formulated in one word, that is, the context of its use was not specified. These subjects were grouped according to joint meaning, so the 19 topics are analyzed further.

The final coding were carried out in the following way: the code was been assigned to the minimal segment of text. To do this, it has been made extraction of stable statements from entire list of keywords.

In order to assess the agreement of coding used the kappa coefficient of Cohen, which is a statistical measure of agreement between coders (in this case — the agreement between the different approaches to coding).

Cohen’s kappa measures the agreement between two raters both classify N items into C mutually exclusive categories.

The equation for $k$ is:

$$\kappa = \frac{p_o - p_e}{1 - p_e},$$

where $p_o$ is the relative observed agreement between raters; and $p_e$ is the hypothetical probability of chance agreement, which uses the observed data to calculate the probabilities what each observer randomly says each category.

If the raters are in complete agreement then $k = 1$.

If there is no agreement between the raters other than one, that would be expected by chance (as given by $p_e$), $k \leq 0$.

There isn’t clear-cut agreement on what constitutes high or low levels of agreement based on Cohen’s kappa, common, but not always so useful set of criteria is: less than 0 % no agreement, 0–20 % poor, 20–40 % fair, 40–60 % moderate, 60–80 % good, 80 % or higher very good.

According to the information received, $p_o = 75 \%, p_e = 44 \%$, thus, $k = 55 \%$ (moderate level of agreement).

**Results**

**Qualitative analysis.** Using the results of the conducted content analysis, we can identify the main similarities and differences between the monitored sources.

The first source RG is an official publication — state documents come into force after publishing in it. At the same time, it also contains news, reports and interviews of government officials, as well as competent comments on documents. Speaking of tax evasion, RG reveals itself, trying to highlight the diversity of the problem from all possible angles. Anti-offshores economics and making of various laws contributing to this are very relevant topics for the RG. Very important topic, which is often illuminated in the RG is the punishment for tax evasion.

The second source is the Publishing House Kommersant. By category of
news, reports and articles, it is similar to the RG, but has one major difference in the form of a more in-depth view in the business sphere on taxation, as well as international tax evasion. Also, this source provides materials on measures to counter tax evasion in the business. For Kommersant journalists, the Russian and the international tax control is a very topical theme. Especially when it comes to «break the bank» sums.

The third source which was analyzed is the electronic library (eLibrary). Materials from this source mostly have scientific nature, abundant number of scientific terms, charts and economic indicators. Scientists publish their works in this source mainly emphasize its focus on such topics as tax control, tax system, tax competition and the shadow economy. For example, one of the authors writes about unfair tax competition as a factor contributing to tax evasion. He analyzes the impact of interstate tax competition on financial flows in the world and identifies main negative effects of tax competition. He specifies that a high level of tax burden leads to the shadow economy, and encourages taxpayers to move their business in the most favorable from a tax point of view jurisdiction. He lights problem of poorly developed international cooperation in the fight against tax crimes and offenses.

Content analysis revealed a number of publications which were most often considered by the media and by scientific community.

Most often, in connection with the tax evasion, the topic of changes in tax legislation is considered. In this context, the most common suggestions are to strengthen the responsibility for the illegal tax optimization. For this purpose, for example, the Tax and Criminal Codes changes are proposed to clarify the concept of unjustified tax benefit and to toughen the penalties for its receipt, including the confiscation of property.

Also the company’s proposals to harmonize their tax optimization schemes with the Federal Tax Service (FTS) are described. It is assumed that each year before the December 1, organization must assert processes with agency for the preparation of fiscal benefits for the next year.

This issue also includes a discussion of draft laws and legislative activities related to tax evasion. For example, a discussion about anti-offshores package of laws that have been prepared and submitted to the State Duma by the Federation Council members, headed by the Chamber Chairman Valentina Matvienko. The starting point in the creation of anti-offshore measures was a report from the Accounting Chamber, which reported that the customs office fixed the foreign trade operations with low-tax zones in the amount of $ 329 billion. Using offshore companies more than 40 % of foreign trade turnover of the country was carried, over the past two decades from 800 billion to $ 1 trillion were derived from Russia to offshores. This became the reason for the development of measures which would reduce the outflow of capital from the country and prevent the erosion of the tax base.

Another frequently discussed topic is the work of law enforcement and judicial authorities in preventing tax evasion. Media clearly describes the conduct of search and seizure of documents in the framework of tax evasion. The interest of the press to the operational activities of law enforcement agencies in the fight against tax evasion is quite understandable. Such events always look very impressive. For example, searches in a criminal proceeding on the fact of tax evasion, amounting to more than 380 million rubles in National Reserve Bank. In order to identify persons involved in the evasion of taxes, main office of the bank in Moscow and buildings of affiliated companies were searched, all documentation of the bank’s activities was withdrawn. A separate section of this issue is devoted to the action of the Investigative Committee of Russia. For example, the description of the criminal proceeding against the CEO of the Galich distillery «KOSTROMAHLEBPROM» on suspicion of tax offenses amounting to
over 56 million rubles. Press reported at the same time that all assets were derived virtually from the enterprise, goods are not in stock, so the Investigative Committee of Russia initiated two criminal proceedings under articles «evasion of taxes and duties with the organization on a large scale» and «concealment of money» to the director of the company, by means of which the collection of taxes should be done.

However, when the brightest stages consisting of searches and arrests ends, interest in the investigation of such proceedings in the press is lost due to the extremely low percentage of criminal liability for tax evasion. Information about actual punishment of specific individuals is never highlighted neither by the media nor by the law enforcement agencies.

An interesting section of this topic is the discussion on the actions of experts and expertise in relation to tax proceedings. The source of information for law enforcement agencies may be represented by all banks that are engaged in cashing funds, audit companies which accumulate data on violations and customer documents. Often the source of information on fraud are owners or employees of the company. In the situation of a corporate conflict they go to law enforcement agencies, and talk about crimes in the organization. Also, companies often respond to a request of law enforcement agencies for documents, and not always even analyze the validity of the requirements to provide documents.

The third popular topic is the topic about international taxation, offshore, and optimization of tax payments. A striking example is the article about the Russian offshore companies in Seychelles. After the escape of Russian business from the British Virgin Islands, as well as Belize, Seychelles has become one of the most popular jurisdictions for the registration of so-called non-transparent shareholders. These are companies that are «opaque» shareholders of other companies in jurisdictions such as Cyprus. It is a practice, when, instead of individuals shareholders indicate Seychelles companies.

In addition, Seychelles is one of the «favorite» jurisdictions of private consultants and financiers providing consulting services outside the territory of Russia. These companies are often used for property registration. For example, a mansion in the center of Moscow may be registered as an offshore company. If the Seychelles really begin the exchange of tax information with Russia in 2017, the domestic interested individuals will either have to go back to Russia, or to search for other «convenient» jurisdiction, which Russian Finance Ministry does not consider as an offshore.

We can note the issue of the shadow economy and fighting shadow activity as a fairly common problem. For media there persist questions like: why we do not like to pay by checks, and salary is put into envelopes. The materials devoted to this topic have scientific research with statistics and research methodology. For example, the publication of the report about sociological monitors of RANEPA (Russian Presidential Academy of National Economy and Public Administration). The results of the survey of 1,600 respondents conducted in 20 regions of the Russian Federation show that currently every ninth respondent (11,9 %) is occupied only in the shadow labor market. State authorities do not know about his work and earnings. In general, the proportion of the working population, which uses shadow employment schemes — 41 %. That’s about 30 million economically active citizens of Russia.

The theme of evasion of social (insurance) premiums is not less important. In the framework of the topic the huge debts to non-budgetary funds are being discussed. Insurance premiums after the abolition of single social tax in 2010 were excluded from the scope of criminal legal regulation. For debts to non-budgetary funds the employer would face only administrative fines, depending on the amount of arrears. In 2012–2014, Russian insurers hadn’t paid to the pension fund more than 100 billion rubles. Nearly two billion rubles — the amount of arrears for the same period through the Social Insurance Fund.
The stricter legislation will not be effective while the current rate of insurance premiums exist, experts believe. While the size of the standard payment to FSS, FIU and HIF would be about 30 per cent of the wage bill, entrepreneurs do not give up the «gray» wages. Even employers who pay the «white» salary and are faced with economic difficulties, can not always find the money for payroll taxes.

Banks actions are also one of the most frequently discussed topics. Articles on banking topics discuss banks’ participation in the activities of offshore companies and ephemeral firms and issues about bank employees who sell information about their customers.

The topic of tax incentives is also very popular. In this regard, we consider two opposing points of view. Business point of view is led to expanding the list of benefits and increasing their size. And the point of view of the State was announced recently, by the Head of the Ministry of Finance Anton Siluanov. According to the Ministry of Finance, it is necessary not only to talk about granting new benefits, but also to abolish the existing privileges, which are not used for stimulating economy and are ineffective. For this an interagency working group of the Accounts Chamber of the Ministry of Finance, the Federal Tax Service, Federal Customs Service, Ministry of Economic Development, the RSPP and «SUPPORT of Russia» is formed for the analysis of the effectiveness of tax breaks. Efficiency benefits will be evaluated through the analysis of economic indicators based on the methodology, which is now developed by the Ministry of Economic Development. Benefits of having a social orientation (such as benefits for the Heroes of Russia or organizations that employ disabled), removed behind the scope of this analysis.

For all other benefits so-called «benefits passport» would be created. It allows you to assess their economic feasibility. Federal and regional benefits would be analyzed.

Practice has shown that tax incentives are also used for tax evasion. Firstly, the overwhelming majority of organizations which have tax incentives use them to hide profits from commercial transactions, which often do not comply with their regulations and subject to preferences, type of activity. Secondly, officials of preferential organizations use the funds and property of these organizations for their own gain. Thirdly, organizations with benefits are used by other companies in the schemes of minimizing tax evasion.

Another often discussed topic is the topic of struggle with ephemeral firms. Publications on the given issue reflects the tendency to tighten the fight with any means of tax evasion. Therefore, many entrepreneurs now have to rethink their way of doing business, to close the companies, in which is impossible to remedy the situation and to move to a new, more secure way of doing business. There is a definite relation between the theme of ephemeral firms and the problem of VAT declaration. For example, fraudulent VAT refund schemes are carried out, usually with the help of ephemeral firms. However, the system of analysis of declarations introduced in 2015 in conjunction with data books of purchases and sales of books, as well as a comparison of the data of buyers and sellers significantly limits the potential for abusing.

**Quantity analysis**

_Coding Frequencies._ Statistics of topics mentioned in the text is given in the table 3. It should be noted that the number of words relating to each code is depends on designated coding rules (is it search for a particular word or phrase, is it search for a particular word or phrase, is encoded segment sentence or paragraph).

We can see the special attention to the topics discussed in more than ¾ of texts (following list is descending by decreasing of coding frequency), see. table 3:

- 19 Entrepreneurship, business management, profit.
- 01 Legislation changes.
- 02 Law enforcement.
- 16 Registration, employment and income.
- 04 Directorship.
- 08 Banking information.
Table 3

Frequencies of the topics

<table>
<thead>
<tr>
<th>Code</th>
<th>Given for the topic</th>
<th>Quantity of codes</th>
<th>Share of codes in total quantity, %</th>
<th>Quantity of Cases, were this code occurs</th>
<th>Share of cases, were this code occurs in total quantity, %</th>
<th>Quantity of coded words</th>
<th>Share of coded words in the topic in total quantity of words, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>Legislation changes</td>
<td>3047</td>
<td>15,1</td>
<td>274</td>
<td>91,0</td>
<td>155000</td>
<td>54,2</td>
</tr>
<tr>
<td>02</td>
<td>Law enforcement</td>
<td>2805</td>
<td>13,9</td>
<td>268</td>
<td>89,0</td>
<td>131799</td>
<td>46,1</td>
</tr>
<tr>
<td>03</td>
<td>International aspects of taxation</td>
<td>1808</td>
<td>9,0</td>
<td>204</td>
<td>67,8</td>
<td>88069</td>
<td>30,8</td>
</tr>
<tr>
<td>04</td>
<td>Directorship</td>
<td>1439</td>
<td>7,1</td>
<td>245</td>
<td>81,4</td>
<td>78980</td>
<td>27,6</td>
</tr>
<tr>
<td>05</td>
<td>Shadow economy</td>
<td>475</td>
<td>2,4</td>
<td>141</td>
<td>46,8</td>
<td>21843</td>
<td>7,6</td>
</tr>
<tr>
<td>06</td>
<td>Lawmaking</td>
<td>416</td>
<td>2,1</td>
<td>124</td>
<td>41,2</td>
<td>18896</td>
<td>6,6</td>
</tr>
<tr>
<td>07</td>
<td>Social security contributions</td>
<td>774</td>
<td>3,8</td>
<td>129</td>
<td>42,9</td>
<td>40889</td>
<td>14,5</td>
</tr>
<tr>
<td>08</td>
<td>Banking information</td>
<td>1624</td>
<td>8,0</td>
<td>245</td>
<td>81,4</td>
<td>82982</td>
<td>29,0</td>
</tr>
<tr>
<td>09</td>
<td>Monitoring</td>
<td>728</td>
<td>3,6</td>
<td>156</td>
<td>51,8</td>
<td>33687</td>
<td>11,8</td>
</tr>
<tr>
<td>10</td>
<td>Tax allowance</td>
<td>260</td>
<td>1,3</td>
<td>82</td>
<td>27,2</td>
<td>12698</td>
<td>4,4</td>
</tr>
<tr>
<td>11</td>
<td>Article of the Tax Code</td>
<td>507</td>
<td>2,5</td>
<td>102</td>
<td>33,9</td>
<td>16835</td>
<td>5,9</td>
</tr>
<tr>
<td>12</td>
<td>Short-lived companies</td>
<td>100</td>
<td>0,5</td>
<td>51</td>
<td>16,9</td>
<td>5010</td>
<td>1,8</td>
</tr>
<tr>
<td>13</td>
<td>Value-added tax</td>
<td>424</td>
<td>2,1</td>
<td>113</td>
<td>37,5</td>
<td>19678</td>
<td>6,9</td>
</tr>
<tr>
<td>14</td>
<td>Expert examination</td>
<td>241</td>
<td>1,2</td>
<td>124</td>
<td>41,2</td>
<td>10021</td>
<td>3,5</td>
</tr>
<tr>
<td>15</td>
<td>Ownership, property, assets, capital, investment</td>
<td>814</td>
<td>4,0</td>
<td>183</td>
<td>60,8</td>
<td>42810</td>
<td>15,0</td>
</tr>
<tr>
<td>16</td>
<td>Registration, employment and income</td>
<td>1630</td>
<td>8,1</td>
<td>258</td>
<td>85,7</td>
<td>84825</td>
<td>29,7</td>
</tr>
<tr>
<td>17</td>
<td>Corruption</td>
<td>84</td>
<td>0,4</td>
<td>45</td>
<td>15,0</td>
<td>3859</td>
<td>1,4</td>
</tr>
<tr>
<td>18</td>
<td>Arrears and fines</td>
<td>315</td>
<td>1,6</td>
<td>120</td>
<td>39,9</td>
<td>15859</td>
<td>5,5</td>
</tr>
<tr>
<td>19</td>
<td>Entrepreneurship, business management, profit</td>
<td>2693</td>
<td>13,3</td>
<td>283</td>
<td>94,0</td>
<td>130476</td>
<td>45,7</td>
</tr>
</tbody>
</table>

Next, we can note some trends: the dynamics of frequencies for certain topics. Dynamics of topic frequencies (for certain topics) is presented at the Fig. 1–2. One can mark out a range of topics, which became less interesting for the years analyzed, or on the contrary, other became more interesting for community. Topics which frequency has no clear trend are the following: Legislation changes, Law enforcement, Lawmaking, Social security contributions, Banking information, Monitoring, Tax allowance, Value-added tax, Registration, Employment and Income, Corruption, Entrepreneurship, Business management, Profit.

Marked reduction of interest can be noted regarding to the following topics (Fig. 1): International aspects of taxation, Shadow economy, Expert examination, Ownership, property, assets, capital, investment.

The interest to the topic «03 International aspects of taxation» decreased after 2013 largely due to the new political course of the Russian Federation taken after the reunification of the Crimea with Russia. The topic «15 Ownership, prop-
erty, assets, capital, investment» has also been linked with international interaction in the field of taxation.

The growth of interest can be noted in relation to the following topics (Fig. 2): Directorship, Article of the Tax Code, Short-lived companies, Arrears and fines.

A notable increase on frequency of the codes «11 Article of the Tax Code», «12 Short-lived companies», «18 Arrears and fines» is associated with the tightening of law enforcement practices in this field.

**Codes Co-occurrences.** In order to analyze codes’ co-occurrences, we use information about the proximity or the co-occurrence of codes within documents to investigate probable relationships among them and similarities among cases. This information allows implementing cluster analysis.

We implemented cauterization of the codes by usage of the data on similarity matrix. The result is presented in the form of a dendrogram (here encoded segments can overlap each other) (Fig. 3). The vertical axis is made up of the codes, and the horizontal axis represents the clusters formed at each step of the clustering procedure.

All proximity values calculated on all included items (codes) are represented by the concept map (2D) (Fig. 4). These maps allow detecting meaningful dimensions that may explain observed similarities between items. Each sphere depicts a code; the distances between pairs of spheres indicate how likely these codes tend to appear together. Size of the sphere represents number of times this code has been used (count). On the figure can be seen that the topics having joint issues for discussion, it is 01, «Legislation changes» and 02 «Law enforcement»; 04 «Directorship» and 19 «Entrepreneurship, business management, profit»; 03 «International aspects of taxation» and 08 «Banking information».
Coding by Variables. Analysis according to variables allows investigating the relationship between codes assigned to cases defined by values of a variable. We analyze each variable (SOURSE, YEAR) and for combination of these variables such measure as Code Occurrence — the number of cases where this code appears at least one time. The most representative results are shown in the following table 4.

![Figure 4. 2D Concept Map](image)

<table>
<thead>
<tr>
<th>Code</th>
<th>Given for the topic</th>
<th>All years</th>
<th>All sources</th>
<th>RG</th>
<th>Kommer- sant</th>
<th>ELIB</th>
</tr>
</thead>
</table>

Note: Minimum value (0%) | Average value (50%) | Maximum value (100%)
Codes “06 Lawmaking” and “14 Expert examination” have notably higher frequency at the source “Kommersant PH” and “Rossiyskaya Gazeta”, and we believe that it is due to the target audience specifics of these sources.

At the same time topics “05 Shadow economy”, “07 Social security contributions”, “12 Short-lived companies”, “16 Corruption” get more coverage in academic articles as compared with other resources; this can not be associated solely with the language features of academic texts. We believe that the reasons are the following. Compared to academic publishers, source RG has less freedom of publications concerning these topics. Simultaneously, source Kommersant needs of lessening the negative attitude of readers regarding these topics, what is achieved by “reticence” of data on these unabiding phenomena.

Heatmap represents relations between code frequencies and variables. Codes occurring frequently have bright color, whereas if the code is not used often, its cell has a dark color (Fig. 5).

For heatmap plot clustering analysis are performed on the cross tabulation tables.

The similarity index (computed for two codes and used for clustering) measures the similarity of their distribution by variables (e.g., “01 Legislation changes” and “02 Law enforcement” for Elib).

Analogous, two variables will be considered near to each other if the distributions of codes in those two groups are similar.

The correspondence analysis graphically represents the relation between assigned codes and subgroups of an independent variable (it is based on usage of the frequency data). The results are presented by a three-dimensional map (Fig. 6).

Our correspondence analysis is implemented by dividing the sample in compliance with combined variables (each of them are combination of two initial vari-
ables: SOURCE and YEAR), here are denoted as «cohort»\(^2\).

Here data are depicted according to the following information:

1. Similarity of the code distribution among certain cohort to the total distribution of all codes within cohorts.
2. Similarity of the code distribution between cohorts.
3. Similarity of the codes.

Relationship among codes is following:

1. The more similar the distribution of a code among cohorts is to the total distribution of all codes within cohorts, the closer it will be to the origin. Codes which are depicted far from point of origin have singular distributions (such as code «12 Short-lived companies» in our analysis).
2. If two codes have similar distributions among cohorts, their points in the correspondence analysis plot will be close together (such as codes «01 Legislation changes» and «02 Law enforcement»).

Relationship among cohorts:

1. The more distribution of a code for a cohort is, compared with the distribution of these codes for the entire sample; the farther this cohort will be from the point of origin (such as RG in comparison with Elib).
2. If two cohorts have similar distribution of a code, they will be depicted near each other (such as RG&2014 and RG&2015).

\(^2\) Cohort — a generational group as defined in demographics, statistics, or market research (http://www.thefreedictionary.com/cohort)
Relationship between codes and cohorts:
1. The distance between codes and cohorts has no meaning for interpretation.
2. The angle between this code point and a cohort’s point from the origin is meaningful:
   - an acute angle indicates that the two characteristics are correlated (such as «04 Directorship» and «16 Registration, employment and income»);
   - an obtuse angle indicates that the two characteristics are negatively correlated (such as «17 Corruption» and «18 Arrears and fines»).

Analysing the correspondence map, it can be noticed that source RG moves around the origin of coordinates, source Elib moves towards the center, and source Kommersant leaves out to the periphery during considered years, and we can give the following interpretation for this: academic topics become closer to issues considered by the authorities; and the language of business community is considering more legal issues («13 Value-added tax», «18 Arrears and fines», «11 Articles of the Tax Code»).

Results and discussion
That result of correspondence analysis has research significance. As a result, studies have concluded that the most popular topics of interest for which no changes are: changes in legislation, legislation and increased enforcement. Using the results of the conducted study, we can identify the main similarities and differences between the monitored sources. We can see the special attention to the: Legislation changes, Law enforcement, Entrepreneurship. Marked reduction of interest can be noted regarding to the following topics: International aspects of taxation, Shadow economy, Ownership, property, investment. The growth of interest can be noted in relation to the following topics: Directorship, Article of the Tax Code, Short-lived companies, Arrears and fines. The study revealed a certain disparity between the topics discussed among academic community and stakeholders. At the beginning of our study we represented characteristics of our sample; its general structure is as follows: 44 % of cases are the texts of source RG, 40 % — «Kommersant», and 16 % — Elibrary. We would assumed that the discourse of the source RG and «Kommersant» will have approximately equal influence on the results of the correspondence analysis, and both of these sources will be approximately equidistant from the origin of coordinates. However, the actual results show that the origin of coordinates is surrounded by graphical representations of a source RG for different years, i.e. discourse of all analyzed sources largely coincides with the source RG discourse for 2014 and 2015. At the same time, the discussion topics at the source «Kommersant» is not typical for all texts. Thus the correspondence analysis indicates that rhetoric of authority is largely the dominant in description of tax evasion, both for business community and for academic circles.

References
4. Val’den P. I. The development of chemistry in Russia. Dnevnik Vtorogo mendeleevskogo s”ezda po obshchei i prikladnoi khimii i fizike. Saint Petersburg, 21–28 dekabrya 1911 g. [The Diary of


Authors

Ekaterina N. Nevzorova — PhD in Economics, Associate Professor, Baikal State University, Irkutsk, Russia (11 Lenin St., 664003, Irkutsk, Russia); e-mail: nevzorova_kat@mail.ru.

Samo Bobek — PhD in Economics, Professor, Head of Department of E-Business, Faculty of Economics and Business University of Maribor, Maribor, Slovenia, (Razlagova 14, 2000 Maribor, Slovenia); e-mail: samo.bobek@um.si.

Anna P. Kireenko — Doctor of Economics, Professor, Head of the Department of Taxation and Customs, Baikal State University, Irkutsk, Russia (11 Lenin St., 664003, Irkutsk, Russia); e-mail: KireenkoAP@bgu.ru.

Roman A. Sklyarov — PhD Student, Department of Taxation and Customs, Baikal State University, Irkutsk, Russia (11 Lenin St., 664003, Irkutsk, Russia); e-mail: free9sonjas@gmail.com.

Информация об авторах

Невзорова Екатерина Николаевна — кандидат экономических наук, доцент, Байкальский государственный университет, г. Иркутск, Россия (664003, г. Иркутск, ул. Ленина, 11); nevzorova_kat@mail.ru.

Бобек Само — доктор экономических наук, профессор, заведующий кафедрой электронного бизнеса, факультет экономики и бизнеса университета Марибор, г. Марибор, Словения (2000, Slovénie, g. Maribor, ul. Razlagova, 14); e-mail: samo.bobek@um.si.

Киренко Анна Павловна — доктор экономических наук, профессор, заведующая кафедрой, Байкальский государственный университет, г. Иркутск, Россия (664003, г. Иркутск, ул. Ленина, 11); e-mail: kireenko-ap@isea.ru.

Скляров Роман Андреевич — аспирант, кафедра налогов и таможенного дела, Байкальский государственный университет, г. Иркутск, Россия (664003, г. Иркутск, ул. Ленина, 11); e-mail: free9sonjas@gmail.com.