Saint Petersburg State University

Manuscript

Moldobaev Temirlan Shaiyrbekovich

TYPICAL FINANCIAL MODELS IN THE ASSESSMENT OF INSOLVENT OF ORGANISATIONS

Scientific specialty

5.2.4. Finance

Research is submitted

for the degree of candidate of science in economics

Translation from Russian

Scientific supervisor:

Doctor of science in Economics, Professor

Kovalev V.V.

Saint Petersburg

2025

CONTENTS

INTRODUCTION

Relevance of the research topic. The problem of financial insolvency is a subject of interest for the scientific community, which is confirmed by the rapid growth in the number of domestic and foreign studies devoted to this topic. In particular, the number of Russian-language articles in specialised journals on financial insolvency increased by 131.2% in 2010-2014 compared to 2005-2009 (growth from 1525 to 3526). In turn, the number of publications for 2015-2019 exceeds 2010-2014 by 129.8% (8,201 vs. 3,526 respectively), and by 16.1% in 2020-2024 compared to the previous five-year interval (9,405 vs. 8,201 respectively). Meanwhile, since 2015, there has been a rapid increase in domestic research articles on the development of new and review of existing financial insolvency assessment tools. Thus, for the period 2005-2024, the total number of such publications totalled 1,288, of which more than 75 per cent fall within the period 2015-2024[1]. In foreign scientific practice, a different trend is observed: for 2005-2024, a total of 537400 papers were published with an even distribution by year, including 75400 studies on financial insolvency assessment tools[2]. The above statistics shows a consistently high level of interest of the foreign scientific community to the problem of financial insolvency throughout the time intervals under consideration, while the domestic scientific community is characterised by a rapid increase in the relevance of the topic since 2015.

The evaluation of financial insolvency constitutes a pivotal concern for the professional community, encompassing the administration of entities susceptible to associating with partners encountering financial challenges. Such cooperation may result in a number of adverse consequences, including, but not limited to: the suspension of production processes due to delayed delivery of raw materials and supplies; increased logistics and storage costs; increased expenses on searching for

[1] Calculated by the author based on data from the scientific electronic library eLibrary.ru. URL: https://www.elibrary.ru/

[2] Calculated by the author based on data from the Google Scholar search engine for scientific publications. URL: https://scholar.google.ru/

alternative suppliers; and non-repayment of advances. To overcome these challenges, it is important for organisations to pursue a preventive strategy in their interaction with both current and potential counterparties by monitoring their activities for signs of financial insolvency. In the event of a delayed response to the deterioration of the counterparty's financial condition, the organisation's losses may increase significantly. These losses include, in addition to the basic amount of compensation for obligations until the bankruptcy case is initiated, additional costs associated with the involvement of relevant specialists to represent the organisation in the arbitration process, and alternative income that the organisation misses as a result of the freezing of liquid assets.

One of the most popular tools of financial insolvency assessment presented in modern domestic and classical foreign scientific works are bankruptcy forecasting models (henceforth referred to as BFMs). BFMs are a class of models developed with the use of classification algorithms, which are based on the use of pre-labelled data on a given attribute to train the model (in the case of financial insolvency, this attribute is the initiation of bankruptcy proceedings). The rapid development of machine learning technologies and the availability of large volumes of financial data on organisations has led to a current trend in domestic economic science related to the development of BFMs.

Despite the extensive availability of many BFMs presented in both domestic and foreign scientific publications, not all of them have retained high predictive accuracy in current realities. In order to identify models that have retained a high predictive ability, quality metrics are used to assess the forecast accuracy using specially prepared labelled data that were not previously used by the authors of these models. In the present study, selected domestic and foreign BFMs are considered, for which a quality metric has been calculated to assess the ability of models to correctly identify financially insolvent organisations. The results of the study, which included the assessment of the quality of BFMs using accounting indicators of financially insolvent organisations, confirmed a number of provisions stated in the existing scientific works about the short-term nature of the application of these models, as well as about the low level of predictive ability of some domestic BFMs.

The present analysis aims to identify the peculiarities of coefficient analysis (as a method of analyzing financial statements) which allow variability in calculation of predictors of tested BFMs; the use of limited sets of key predictors that do not cover all indicators of financial statements; the training of models on a limited unbalanced sample; and the high level of fluctuation of the values of financial indicators of organisations in the context of types of economic activities (hereinafter referred to as TEA) for different periods. These problems have a negative impact on the predictive ability of the existing insolvency assessment models problems and are the focus of this research.

Notwithstanding the high importance and widespread utilisation of coefficient analysis, which is predicated on the utilisation of accounting indicators as an information base, this method has a number of drawbacks associated with the variability of financial ratio calculation. This complicates the comparison of the values of one and the same indicator across different organisations and leads to different financial analysts, using previously developed BFMs, calculating one and the same predictor of the model differently and consequently arriving at different final predictions.

Following a comprehensive analysis of both foreign and domestic BFMs, it has been determined that the authors of the models employ a restricted set of financial indicators. The selection of these indicators is achieved through one or a combination of the following methodologies: the formulation of a comprehensive list of financial ratios and the execution of statistical tests to identify the most substantial factors; or the utilisation of an expert method, founded on the professional discernment of researchers, to ascertain significant indicators. It is evident that both of these approaches result in BFMs including a limited set of predictors that do not take into account all financial reporting indicators. This, in turn, can potentially signal the risk of financial insolvency of the organisation. It is also important to note the difference in the composition of factors between the BFMs considered in this study, which indicates a disagreement in understanding which aspects (criteria) of financial condition can be used as key indicators assessing financial insolvency: liquidity and solvency; financial stability; business activity; profitability.

Furthermore, a distinctive feature of domestic BFMs is their training on a restricted sample of organisations established with the utilisation of reference and analytical systems (henceforth referred to as RAS), which consolidate data from multiple open sources. Concurrently, the number of exported records from RAS, as stipulated by the terms of the user agreement, is often considerably lower than the total number of records in primary sources integrated with RAS.

The utilisation of limited sampling engenders a model that fails to account for all major patterns and relationships in the data, thereby diminishing its capacity to effectively apply its conclusions and predictions to new, previously unobserved data. To address this limitation, this study proposes the utilisation of a general population derived from primary sources, namely open government data processed using specialised software tools.

The mean values of financial ratios employed as predictors of BFMs demonstrate significant variation between TEAs, owing to discrepancies in business models, capital intensity and operating cycles. Consequently, a model developed for one TEA may not be applicable or less accurate for another TEA. Furthermore, the time series of financial indicators exhibit varying degrees of stationarity, contingent on the TEA. In TEAs with highly cyclical or seasonal business models, there is often considerable volatility in financial ratios, which poses challenges in the construction of reliable BFMs. In light of the aforementioned features of the use of financial ratios as predictors of sectoral BFMs, it is imperative to acknowledge that, on the one hand, enhancing the accuracy of forecasting financial insolvency necessitates adapting models to specific TEAs; on the other hand, not all TEAs can be utilised to create stable and accurate models, as the statistical properties of indicators included in the independent variables of BFMs may evolve over time.

In order to surmount the issues delineated above, this study posits the development of clusters in the context of TEA through the utilisation of an array of financial statements of over 2 million Russian organisations, inclusive of financially insolvent organisations. Each cluster is defined as a group of firms that are analogous in the composition of assets and liabilities, and is characterised by a prototype, designated as a centroid, for which the coordinates are defined in the form of average specific values of the sections of the statement of financial position (balance sheet).

Further, we will call a separate centroid of a cluster a typical (average) financial model. The use of this term is due to the following reasons: the centroid characterizes the average or typical values for the cluster, which facilitates the understanding and description of the main characteristics of the group; at the same time, the coordinates of the centroid are presented in the form of average specific values of the sections of the statement of financial position. Consequently, the aforementioned statement of financial position is regarded as the ideal financial model for the organisation, as it furnishes a structured and systematised perspective on its financial position, is connected to other forms of accounting reports through separate indicators, and serves as a foundation for financial analysis.

The identification of several typical financial models for each TEA is facilitated by the clustering of organisations, which results in the formation of groups of objects with similar financial attributes, for which the corresponding centroids are established.

The distribution of financially insolvent organisations into clusters will facilitate the identification of a typical financial model characteristic of such organisations, or the confirmation of its absence within a particular TEA.

The typical financial models of insolvent organisations identified in this study are proposed to be used as models for assessing financial insolvency.

Utilising the average balance sheet as a financial model enables the expansion of the scope beyond the limitations of coefficient analysis, which operates with individual elements of financial statements, thereby hindering a comprehensive overview of the organisation's financial position. Specifically, the assessment of a given organisation's financial insolvency is reduced to the determination of the degree to which its balance sheet conforms to the financial model of insolvent organisations, with consideration for the TEA, without the preliminary allocation of a restricted set of criteria in the form of financial ratios.

The degree of development of the research topic is considered. Domestic and foreign studies consider the main theoretical and practical aspects of assessing the financial insolvency of an organisation using economic and mathematical methods.

Notable contributions in this area include works by foreign researchers, including W. Beaver, E. Altman, R. Tuffler, G. Tishaw, R. Lis, G. Springate, J. Olson, M. Odom, R. Shadr, L. Salchenberger, E. Chinar, and N. Lash have examined in detail mathematical methods used to develop BFMs presented as: a system of financial ratios with specified interval values; models based on multiple discriminant analysis (MDA-models); logistic regressions to solve the problem of binary classification (Logit-models); neural networks. The features of domestic BFMs are presented in the works of I.P. Boyko, A.V. Kazakov, A.V. Kolyshkin, E.A. Fedorova, L.E. Khrustova, D.V. Chekrizov, and F.Y. Fedorov. Russian researchers have adapted foreign MDA-models and Logit-models, taking into account industry specifics and using open government data in the form of an array of financial statements of Russian organisations. The works of B.B. Demeshev, A.S. Tikhonova, O.V. Kolokolova, P.E. Razumov, A.D. Batrasova, T.V. Konovalova, and P.I. Komarov consider the application of more advanced machine learning methods, including algorithms based on the construction of decision trees and rigid and fuzzy cluster analysis algorithms, for bankruptcy forecasting.

The works of N.V. Generalova and N.A. Sokolova consider the variability of interpretation of accounting information at all stages of the accounting process. These stages begin with the moment of registration of the facts of economic activity of the organisation and end with the analysis of its financial statements. The described concept elucidates the impact of distortions in the calculations of financial ratios on the quality of BFMs. Additionally, the contributions of A.M. Patrov, N.G.

Akulova, D.I. Ryakhovsky and O.A. Lvova are noteworthy, as they elucidate the limitations and disadvantages of the coefficient analysis that underpin financial insolvency assessment models. Val. V. Kovalev, Vit. V. Kovalev, and M.L. Pyatov posit that the statement of financial position, formed in the system of dual accounting, is among the most informative models, allowing for the estimation of productive and financial capacities, as well as the expediency of using the economic potential of the organisation. This concept formed the foundation for the development of a model to assess the financial insolvency of organisations. The model utilises the statement of financial position indicators as input data, rather than the ratios derived from them.

A review of successful cases of applying machine learning algorithms to solve applied problems in finance has also been conducted by O. Sezer, A. Ozbayoglu, S. Selvin, R. Vinayakumar, E. Gopalakrishnan, V. Menon, K. Soman, A. M. Karminsky, R. N. Burekhin, W. Liu, H. Fan, M. Xia, W. Chen, H. Zhang, M. Mehlawat, L. Jia, I. Fisher, M. Garnsey, M. Hughes, P. Coulombe, M. Leroux, D. Stevanovic, and S. Surprenan. The researchers in their respective papers propose new models for financial time series forecasting, credit scoring, investment portfolio optimisation, textual data analysis in financial analysis and auditing, and macroeconomic forecasting.

The aim and objectives of the research. The aim of the present dissertation research is to create models for assessing the financial insolvency of commercial organisations using clustering algorithms. In order to achieve the aforementioned objective, it is necessary to complete the following objectives:

1. To investigate scientific publications in the field of financial insolvency assessment, based on the evolution of mathematical tools;
2. To identify and classify the distinguishing features of modern methods of assessing the financial insolvency of Russian organisations;
3. To evaluate the predictive ability of some domestic and foreign bankruptcy forecasting models on new data;

4. To entail the systematisation of data sources and the development of algorithms for their interfacing, with the aim of creating a unified information base containing data on the financial and economic activities of Russian organisations;

5. To argue the necessity of creating models for assessing financial insolvency in the context of types of economic activity;

6. To develop typical financial models for different types of economic activity and to identify those that correspond to the greatest extent to financially insolvent organisations.

7. To study the prospects of using clustering methods beyond the financial insolvency assessment.

The object of the study. The object of study is Russian commercial organisations that provide annual statistical accounting reports to the Federal Tax Service of the Russian Federation (hereinafter referred to as the FTS).

The subject of the present study is the methods of assessing the financial insolvency of organisations.

The information base of the dissertation research is open data sources presented in the form of the register of financial statements and other registers describing certain aspects of financial and economic activities of Russian organisations and published on the official websites of government agencies in different formats, the processing of which is implemented using the Python programming language. The classification and interfacing of data from diverse sources was facilitated by the utilisation of All-Russian classifiers and conjugation keys, which have been endorsed by government agencies and are accessible on their official websites.

The research methodology employed in the present study is rooted in scientific principles. The present study is based on the application of a wide range of methodological tools. In particular, general scientific methods such as abstraction, formalisation, deduction, induction, analysis, synthesis, comparison, proof, mathematical modelling including regression and cluster analyses have been applied. In addition to the above-mentioned methods, the research also used qualitative analysis methods, including content analysis and thematic analysis. The graphical presentation of information was accorded particular attention.

Compliance with the passport of scientific speciality. The area of research corresponds to paras: 15. 'Corporate finance. Financial strategy of corporations. Financial management'; 17. 'The system of financial control in corporations: content, forms, methods and tools of implementation' of passport of speciality 5.2.4. Finance.

The scientific novelty of the dissertation research lies in the development and substantiation of a method for utilising clustering algorithms in the processing of large data sets, with the aim of assessing the financial insolvency of commercial organisations.

The main scientific results, which contain components of scientific novelty obtained during the research and put forward for defence, consist in the following:

1. The author's characterisation of the limitations of the known bankruptcy forecasting models as a tool in the assessment of financial insolvency is given [Kovalev, Moldobaev, 2021; Bakunova, Koltsova, Moldobaev, 2019; Ilysheva, Savostina, Moldobaev, 2018].

2. A critique of coefficient analysis of reporting from the position of its use in the assessment of financial insolvency of organisations has been proposed [Kovalev, Moldobaev, 2021; Bakunova, Prisyazhny et al., 2019; Medvedev et al., 2019].

3. The advantages of multi-criteria assessments in diagnosing the financial insolvency of organisations using cluster analysis are characterised and the requirements to the data sets necessary for the formation of such assessments are justified [Kovalev et al., 2022].

4. The possibility of forming typical financial models of organisations taking into account industry specifics on the basis of using algorithms for clustering large data sets has been proved [Kovalev, Moldobaev, 2021].

5. The text also demonstrates the feasibility of evaluating the financial insolvency of a commercial entity by correlating its financial statements with the average model of financially insolvent organisations, with adjustments for the industry-specific characteristics of its operations [Kovalev, Moldobaev, 2021].

6. The directions of using typical financial models outside the tasks of assessing the financial insolvency of organisations have been determined [Moldobaev, 2022].

The provisions submitted for defense:

1. Popular foreign and domestic models of bankruptcy forecasting, as tools in assessing the financial insolvency of commercial organisations, have limited applicability in modern realities due to a number of factors that can be conditionally divided into economic and information-statistical. The need to develop new methods for assessing the financial insolvency of organisations is due to the need to improve the accuracy of diagnostics of this condition, taking into account industry specifics.

2. Ratio analysis of financial statements has significant limitations when applied to assess the financial insolvency of commercial organsations. This approach reflects only a fragmented picture of the financial position of the organisation, which does not allow for a complete picture of its current and potential financial solvency. In addition, the variability of the calculation of the same financial ratios used as predictors of bankruptcy forecasting models leads to statistically incomparable and potentially unreliable results.

3. Cluster analysis, as a basis for multi-criteria assessment of financial insolvency of organisations, has a number of advantages compared to traditional linear and nonlinear models built on the basis of machine learning algorithms to solve the classification problem. The main advantages of cluster analysis are its scalability in terms of expanding the list of analyzed indicators, as well as ease of supporting the updating of the formed clusters and their associated prototypes - centroids.

4. By applying cluster analysis to large arrays of financial statements, it has become possible to establish the similarity of the structure of the property complex and sources of financing of commercial organisations related to various sectors of the economy. In addition, this technology is the basis for creating typical financial models that can serve as a reliable tool for assessing the financial insolvency of

organisations, and also opens up new opportunities for analytical research beyond financial insolvency.

5. The financial insolvency of a commercial organisation can be assessed by comparing its reporting with a typical (average) financial model, typical of most insolvent companies in a specific industry. Based on the results of cluster analysis, the introduced index of the frequency of occurrence of insolvent organisations in clusters, as well as the analysis of the level of displacement of cluster centoids for different periods, a system of typical financial models in the context of economic sectors has been created, among which stable over time and typical financial models typical of most financially insolvent companies have been identified.

6. The directions of practical application of standard financial models beyond the assessment of financial insolvency of commercial organisations are established. These models are considered as tools in substantiating management decisions aimed at financial recovery of organisations by implementing merger and acquisition strategies or introducing innovative technologies taking into account industry specifics.

The theoretical significance of the dissertation research lies in the development of principles for constructing typical (average) financial models in the context of foreign economic activity, the formation of which was previously impossible due to technological limitations and the lack of big data presented in the form of financial statements of Russian organisations in order to use them for further assessment of financial insolvency. The developed set of models makes it possible to overcome the limitations of the coefficient analysis, to include financial statement indicators rather than derived ratios, to take into account all commercial organisations registered in the Russian Federation, and to identify those TEAs for which the financial insolvency assessment methodology proposed in this study is applicable. A salient benefit of the proposed methodology is its scalability, which encompasses not only components of the statement of financial position, but also other indicators of financial statements, along with the utilisation of typical financial models to assess other dimensions of the financial condition of the organisation. This enhances the scientific and methodological apparatus in the domain of corporate finance.

The practical significance of the dissertation research consists in the application of the developed typical financial models for: early detection of signs of financial insolvency in order to strengthen the financial stability and risk management of the organisation; more accurate assessment of debtors' creditworthiness, which allows financial institutions and lenders to optimise credit policy and reduce the risk of non-repayment; development of more effective policies and regulations aimed at stabilising certain industries and preventing crises. The results of the thesis can also be used in scientific and educational work when studying the courses 'Financial Management' and 'Solving Business Problems Using Python Programming Language'.

The validity and reliability of the scientific provisions, results and conclusions of the dissertation are ensured by the correct use of methods of logical and mathematical analysis as applied to open government data presented in the form of a register of financial statements of Russian organisations and other registers and published on the official websites and services of the Federal State Statistics Service of the Russian Federation (hereinafter referred to as Rosstat) and the FTS. Concurrently, the testing of extant BFMs and the development of the author's methodology for assessing financial insolvency are grounded in the utilisation of population data, the processing of which was facilitated by the employment of specialised tools for working with big data. This approach took into account the established technical controls, which enabled the exclusion of low-quality data from the study.

Approbation of the research results. The primary provisions and conclusions of the thesis were presented and received a favourable evaluation at a number of international and all-Russian scientific and practical conferences. These included 'Import substitution: resources, opportunities, challenges for entrepreneurship' (St. Petersburg, 2022), 'Economic security in the conditions of transformation' (Tyumen, 2020), 'Regional dimension of digital transformation'

(Moscow, 2019), and 'Entrepreneurship and reforms in Russia' (St. Petersburg, 2019).

Publications. With regard to the thesis, seven scientific papers have been published, totalling 5.2 of the author's sheet. (author's contribution – 2.2 of the author's sheet) have been published, including 3 articles with a total volume of 2.6 of the author's sheet. (author's contribution – 0.8 of the author's sheet) in the publications indexed in the international citation databases Scopus and WoS, and 4 articles with a total volume of 2.6 of the author's sheet. (author's contribution – 1.4 of the author's sheet.) in peer-reviewed publications recommended by VAK under the auspices of the Ministry of Education and Science of Russia:

1. Kovalev V. V., Moldobaev T. Sh. V., Moldobaev T. Sh. Testing of foreign and domestic models of bankruptcy forecasting at Russian enterprises // Development of territories. – 2021. – №. 3 (25). – P. 10-19.

2. Moldobaev T. Sh. Influence of innovations on the efficiency of enterprises: sectoral and regional analysis // Economics and Entrepreneurship – 2022. – № 11 (148). – P. 494-498.

3. Kovalev V. V. V., Moldobaev T. Sh., Molitvin M. N., Suyazov V. V. Analysis of the effectiveness of programmes to support Russian universities (2010-2020) // Vestnik of St. Petersburg University. Economics. – 2022. – T. 38. – №. 2. – P. 208-234.

4. Bakunova T. V., Prisyazhny A. V., Detkov A. A., Moldobaev T. Sh., Taubaev A. A. A. Spectral-ballistic method for analysing financial indicators of commercial enterprises // AIP Conference Proceedings. – AIP Publishing, 2019. – Vol. 2116. – №. 1.

5. Medvedev M. A. A., Detkov A. A., Moldobaev T. Sh. Analysing the competitive environment in the sectoral markets using information systems //AIP Conference Proceedings. – AIP Publishing, 2019. – Vol. 2172. – №. 1.

6. Bakunova T. V., Koltsova T. A., Moldobaev T. Sh. Peculiarities of accounting for operating and financial leases under IFRS // Accounting and Statistics. – 2019. – №. 1 (53). – P. 10-17.

7. Ilysheva N. N., Savostina O. V., Moldobaev T. Sh. Peculiarities of accounting for contractor agreements in construction organisations under the requirement of IFRS // Discussion. – 2018. – №. 3 (88). – P. 110-119.

The structure and content of the work are as follows. The thesis is comprised of 167 pages, including an introduction, three chapters, the general conclusions of the dissertation work, a list of literature comprising 142 names, 12 tables and 17 figures.

CHAPTER 1. PREREQUISITES FOR THE USE OF CLUSTER ANALYSIS FOR FINANCIAL INSOLVENCY ASSESSMENT

1.1 Bankruptcy forecasting models as tools in assessing financial insolvency of organisations in the historical retrospective

There are variable approaches to the interpretation of the concept of insolvency (bankruptcy), including a legal approach according to which an organisation is considered insolvent if the arbitration court recognizes the inability of this organisation to fulfill its obligations. In the context of this study, the priority objective of which is to develop a method for assessing financial insolvency, an organisation is considered financially insolvent if it meets all of the following criteria:

– bankruptcy proceedings have been initiated against these entities (and not necessarily subsequently declared bankrupt);

– their total assets and sales revenue for the year before the bankruptcy proceedings are greater than zero, confirming that they are operationally and financially active;

– correctly financial statements used as inputs for the assessment of financial insolvency.

The monitoring of an organisation's activities to assess the risk of financial insolvency, both in relation to itself and its current and potential counterparties, is an essential condition for business continuity. For instance, a significant challenge in dealing with financially insolvent organisations is the inability to obtain reimbursement of liabilities from them, along with the additional costs associated with engaging specialised professionals to obtain at least part of the reimbursement and mitigate losses. According to the data of the Unified Federal Register of Legally Significant Information on the Facts of Activity of Legal Entities, Individual Entrepreneurs and Other Subjects of Economic Activity (hereinafter - Fedresource), the average duration of bankruptcy proceedings has increased, which is due to the use by unscrupulous debtors of complex schemes to withdraw assets, which require a lot of time to challenge transactions and recover property. Furthermore, in 2019, the proportion of satisfied claims in bankruptcy cases was 4.7%, while the proportion of cases in which creditors received nothing at the conclusion of bankruptcy proceedings was 68%[3]. The above statistics demonstrates the pertinence of implementing various measures aimed at forecasting the probability of initiating bankruptcy proceedings against counterparties, with a view to reducing potential losses.

The scientific literature on the subject of financial insolvency forecasting can be approached from the perspective of historical shifts in researchers' perspectives on the issue of bankruptcy, as well as the evolution of mathematical approaches employed for its resolution. Consequently, a number of domestic researchers have distinguished three stages in the evolution of financial insolvency forecasting methods [Kolyshkin et al., 2014].

The initial stage. During the first half of the 20th century, financial indicators calculated on the basis of accounting statements began to be widely used to assess the financial condition of organisations. A notable study from this era is that of W. Chudson, published in 1945. [Chudson, 1945]. In this seminal article, Chudson established the existence of a similar financial structure in organisations of the same industry and the same scale of activity. The findings of this analysis were subsequently employed by researchers in the development of BFMs. This period is distinguished by the use of individual indicators for the prediction of financial insolvency.

The second stage of the process was marked by the publication of an article by W. Biver in 1966. This article proposed a system of indicators for the purpose of forecasting financial insolvency, with the indicators themselves being drawn from the liquidity, profitability and financial stability of the organisation in question.

[3] Bankruptcies of companies – Federal Resource statistics for 2019. URL: https://fedresurs.ru/news/7b3c8884-b159-4ee7-b5fb-7770d9d941da

W. Beaver selected six financial ratios from the initial 30 indicators for the subsequent analysis: return on assets ratio; ratio of operating cash flow to total liabilities; financial leverage ratio; ratio of assets coverage by own working capital; current liquidity ratio; ratio of quick-liquid assets (representing the sum of cash and cash equivalents, short-term receivables and short-term financial investments) to average daily operating expenses. For each coefficient, W. Beaver has defined intervals that allow for the classification of organisations into three groups: those deemed to be financially stable, those with a high risk of bankruptcy within five years, and those with a high risk of bankruptcy within one year [Beaver, 1966]. The main disadvantages of the approach under consideration are the lack of an integral indicator that unambiguously assesses the risk of bankruptcy of an enterprise and the complexity of interpreting the results, according to which an organisation may be financially sound according to one set of indicators but bankrupt according to another.

The contemporary domestic practice employs a system of ratios to assess financial insolvency. In academic studies, this approach is termed 'normative', entailing the comparison of a given indicator's value with its normative value or the assessment of its position within a specified interval. The majority of domestic studies are devoted to the specification of statutory norms using special statistical methods and taking into account industry specifics. The universality of statutory norms, which do not take into account industry specifics, as well as the low level of predictive power, which according to the results of testing is set by researchers in the range from 50% to 60%, are the key prerequisites for specifying the normative values of financial indicators. Domestic researchers have also confirmed a significant deviation of the specified normative values from the universal values established by legislation for individual industries, and they have achieved an increase in the classification ability of the system of updated parameters up to 70-80% [Fedorova et al., 2015; Fedorova et al., 2017].

The authors employ a range of statistical methodologies to refine the normative values of financial ratios. For instance, the works by E.A. Fedorova, M.A.

Chukhlantseva, D.V. Chekrizov, and Y.V. Timofeev employ classification trees and the Gini coefficient [Fedorova et al., 2015; Fedorova et al., 2017]. Conversely, D.P. Zharsky, M.S. Kurazhenkov and A.A. Akifiev employ a different methodology. They define quartiles of an ascending-ordered series with the values of a single financial ratio within a particular economic sector. They consider that the organisation falls into the area of financial insolvency if the value of the indicator is below the boundary of the first quartile [Zharsky et al., 2023].

The paper by E.A. Fedorova, S.O. Musienko and F.Y. Fedorov considers the use of machine learning algorithms that are considered to be more advanced ('random forest', 'boosting', 'bagging') in order to take into account the non-linear dependence of bankruptcy on financial and non-financial indicators. The algorithms under discussion are based on the construction of a set of decision trees. Each decision tree is designed to categorise an organisation as either a potential bankrupt or a financially healthy entity. The forecasts are then averaged to obtain a summary model score. A distinctive feature of the work under review, in comparison with other domestic studies, is the use of advanced algorithms to refine the normative values of the main legislative and most popular Western indicators used as single-criteria BFMs for small and medium-sized enterprises. The employment of the 'random forest' algorithm is instrumental in attaining the maximum overall predictive ability of the variables. Furthermore, the scientific community has successfully substantiated the importance of segmenting all small and medium-sized enterprises into discrete groups for the purpose of bankruptcy forecasting [Fedorova et al., 2018].

In the domestic context of implementing this approach, the following limitations can be identified:

1. The models employed are single-criteria, with a decision on the financial insolvency of the organisation being made on the basis of one coefficient reflecting one of the aspects of its activity. This approach is limited in its capacity to address other areas of the organisation's activity, which may signal a high probability of bankruptcy [Dyagel et al., 2008].

2. The variability of the normative values of financial ratios, defined by researchers due to different initial data sets and statistical approaches to calculating the threshold values, is also a salient issue. For instance, within the body of work produced by the same authors' team, there is often found to be a divergence of estimates when a single financial ratio is calculated for a specific economic sector across different periods.

The third stage of the process is as follows. Utilising discriminant analysis, E. Altman has developed a set of regression BFMs corresponding to industry, country and organisational and legal characteristics of organisations. The initial model was presented in one of the author's works in 1968 and was intended for public manufacturing organisations. In developing this model, Altman analysed the activities of 66 organisations for the period from 1946 to 1965, 33 of which went bankrupt, while the rest continued their activities [Altman, 1968]. The scientist's approach enables the calculation of an integral indicator (Z-score), also termed the composite bankruptcy ratio, the value of which for a specific organisation is then compared with pre-calculated intervals corresponding to varying degrees of financial stability, including financial insolvency.

The initial model developed by E. Altman incorporated five financial ratios, one of which involved the calculation of the ratio of market capitalisation to total liabilities. Subsequently, E. Altman refined his initial model by substituting market capitalisation with equity, thereby expanding the scope of application of the new model to encompass non-public organisations, as all predictors were calculated on the basis of accounting statements [Altman, 1983]. Subsequent to this, Altman proceeded to develop models for manufacturing and non-manufacturing organisations, as well as for developed and emerging markets [Lvova, 2015]. The implementation of E. Altman's ideas by other foreign economists included the works of R. Taffler and G. Tisshaw [Taffler, Tisshaw, 1977], R. Lis [Edmister, 1972] and G. Springate [Springate, 1978]. In their research, Altman's followers either utilised the same set of financial indicators, but with recalculated weighting coefficients that took into account the national and industry-specific characteristics of the analysed organisations' activities, or they created new models that included other financial indicators. Along with discriminant analysis, logit models estimating the probability of bankruptcy on the interval from 0 to 1 have been developed since 1980 [Ohlson, 1980].

The foreign BFMs described above formed the basis for many domestic scientific studies that assess the applicability of these models in Russian economic conditions and describe the development of author's BFMs. The creation of the author's BFMs can be approached in three distinct ways: firstly, the specification of normative values of the composite bankruptcy coefficient of the original foreign BFMs; secondly, the preservation of the composition of predictors of the original foreign BFMs with the specification of the weighting coefficients corresponding to the predictors; and thirdly, the formation of an author's set of predictors, including some indicators from existing models, as well as other financial and non-financial indicators.

The article by A.V. Kazakov and A.V. Kolyshkin presents a review of 35 scientific publications on domestic BFMs. The researchers observe a paucity of work in this area, attributing this to the fact that the test sample for assessing the predictive power of the model was used in only 15 papers. The formed samples were unrepresentative due to their small size (most of the models were built on a sample of less than 1000 organisations). In some publications, the authors did not reflect the formulaic representation of BFMs at all. The authors of the article also provide a justification for the high popularity of balance sheet BFMs in Russian science, which is based on accounting indicators, in comparison to market models. The underdeveloped financial market is identified as a significant constraint on the implementation of market models in Russia [Kazakov, Kolyshkin, 2018].

In this study, the results of Russian scientists' testing of individual BFMs will be considered. In their study (henceforth referred to as the "first paper"), E.A. Fedorova, L.E. Khrustova and D.V. Chekrizov evaluated the forecast accuracy of 10 classical models, including nine foreign and one domestic, within the context of eight economic sectors. The mean forecast accuracy, incorporating the accurate categorisation of both bankrupt and financially stable organisations, for all industries, was 73% (for 9 foreign BFMs - 72.8%, for 1 domestic BFM - 74.4%). In a separate study, E.A. Fedorova, S.E. Dovzhenko and F.Y. Fedorov (henceforth referred to as the second paper) evaluated the forecasting capability of four foreign and four domestic BFMs within the context of ten economic sectors. The mean accuracy was determined to be 66% (71.2% for foreign BFMs and 57.1% for domestic BFMs). In contrast, a lower level of forecast accuracy was obtained for 35 BFMs under consideration, including classic and most popular foreign models, as well as domestic models with the highest level of citation (Kazakov & Kolyshkin, 2023). Therefore, having established the model adequacy threshold (forecast accuracy) at 70%, the scientists ensured that all models tested for the five economic sectors demonstrated forecast accuracy below the adequacy threshold for both bankrupt and financially healthy organisations. Concurrently, the researchers contend that from an economic perspective, errors of the first kind, wherein the BFM erroneously classifies a bankrupt entity, are significantly more perilous than errors of the second kind, wherein the BFM erroneously categorises a financially robust entity as bankrupt, as in the latter scenario, the consequence is merely the loss of economic profit. This observation is further substantiated by the finding that the model adequacy threshold for financially insolvent organisations was only observed in eight out of 35 models in the manufacturing industry, one out of 35 models in agriculture, and two out of 35 models in the service sector. In all the aforementioned works, the researchers arrived at the same conclusion: namely, that the quality of the tested BFMs was low. This finding served as the basis for the subsequent development of new author's BFMs. The first and second papers utilise the same approaches to model creation, namely the specification of normative values of the composite bankruptcy coefficient of existing BFMs, with consideration for industry specifics; the formation of an author's set of financial indicators calculated on the basis of accounting data and the creation of one BFMs with the specification of normative values of the composite bankruptcy coefficient for different economic sectors. The specification of sectoral normative values of the composite bankruptcy ratio for the existing BFMs was performed using classification trees and the Gini index, which allowed increasing the average forecast accuracy for all industries from 73% to 77% and from 66% to 72.6% in the first and second papers, respectively. The authors of these studies also developed BFMs presented as logistic regressions. Both financial indicators employed in the tested BFMs and indicators not covered by the study were considered as predictors of the models. Concurrently, the composition of independent variables and their corresponding weighting coefficients remained constant across diverse economic sectors; industry specificity manifested through the specification of the final normative values of the aggregate bankruptcy coefficient. The regression models obtained demonstrated a high level of prediction accuracy in the test sample, with values of 82% and 78.9% achieved for the first and second papers, respectively.

In the third paper, the author's BFMs are also presented in the form of a logistic regression model. The distinguishing characteristic of this model pertains to the utilisation of diverse financial indicator sets as predictors, which are selected based on the economic sector. The researchers achieved an average forecast accuracy for all industries in the test sample of 71.9%, which exceeds the threshold of model adequacy established by the authors of the study.

In addition to the linear models described above, non-linear models have also gained significant relevance. Consequently, B.B. Demeshev and A.S. Tikhonova have developed a series of BFMs utilising the methodologies of logistic regression (a linear method) and 'random forest' (a non-linear method) for small and medium-sized Russian non-public organisations operating within the wholesale and retail trade industry. Four groups of indicators were considered: financial ratios of the model of E. Altman and D. Sabato for small and medium-sized organisations [Altman, Sabato, 2007] in combination with one non-financial indicator (age) and a set of non-financial indicators (age, legal form, size, region); financial ratios, which are most often mentioned in bankruptcy studies, in combination with one non-financial indicator (age) and a set of non-financial indicators (age, legal form, size, region). The predictive accuracy of each model was assessed using the area of the

ROC curve, the maximum value of which was obtained for the author's BFMs. These were constructed using the random forest algorithm and including explanatory variables in the form of financial ratios, the most frequently mentioned in the studies, and a set of non-financial indicators. The non-linear dependence of the probability of bankruptcy on financial and non-financial indicators indicates the low suitability of linear models for predicting the financial insolvency of Russian small and medium-sized organisations [Demeshev, Tikhonova, 2014].

It is evident from a number of Russian scientific publications that cluster analysis is also being applied within the framework of financial insolvency assessment. For instance, O.V. Kolokolova presents an algorithm for evaluating the likelihood of bankruptcy, utilising cluster analysis as a foundation. This approach operates under the premise that organisations within a shared cluster, exhibiting comparable characteristics, are categorised within the same risk group, which encompasses the potential for bankruptcy. The study incorporates both quantitative and qualitative indicators as characteristics of an organisation. However, it should be noted that the numerical representation of qualitative indicators imposes certain limitations in terms of the selection of an appropriate clustering method and the rule for calculating distances between the objects of study. In light of this circumstance, O.V. Kolokolova arrived at the conclusion that the utilisation of fuzzy cluster analysis is imperative, predicated on the premise that an organisation can be ascribed to multiple clusters with disparate probability. Moreover, the Manhattan distance is deemed more appropriate for the calculation of distances between the objects of study. Additionally, the paper observes that organisations falling into the same cluster and concurrently identified as financially insolvent possess a distinct life span (the period from the organisation's establishment to its recognition as financially insolvent). Consequently, the ultimate evaluation of the likelihood of bankruptcy for the borrowing organisation should encompass, in addition to its cluster affiliation, the intensity of bankruptcy, which is conceptualised by the author as the probability of bankruptcy during a specific time interval [Kolokolova, 2007].

Another feature of domestic research related to the use of cluster analysis to assess financial insolvency is the utilisation of a limited set of financial ratios and a small sample size of organisations from one industry. P.E. Razumov's work employs a hierarchical method of clustering 30 agricultural organisations in Krasnoyarsk Krai by two financial ratios: own working capital ratio and current liquidity ratio. This resulted in the identification of five distinct clusters, including those exhibiting a high risk of bankruptcy. The distribution of organisations into groups of financial stability on the basis of cluster analysis coincided with the results obtained in the course of testing other foreign BFMs (models of E. Altman, R. Tuffler and R. Lis) [Razumova, 2015].

A.D. Batrasova, T.V. Konovalova, and P.I. Komarov regard clustering as a methodology for conducting research into financial stability. Utilising the k-means clustering algorithm, the scientists allocated 65 Russian IT organisations with analogous values of six financial ratios, calculated on the basis of accounting data. The researchers obtained six clusters and identified the cluster with the largest number of organisations. They then proposed the use of the coordinates of the centroid of this cluster as normative values corresponding to financially stable organisations [Batrasova et al., 2022].

Despite the high accuracy of BFMs forecasts developed using more advanced machine learning methods based on the construction of decision trees and taking into account non-linear dependence between different variables, their main disadvantage is the lack of possibility of their use by third-party stakeholders to solve scientific and applied problems due to the difficulty of visualising deep decision trees[4] and their combination when constructing, for example, a random forest. Moreover, an analysis of domestic research in the field of bankruptcy forecasting based on cluster analysis has revealed that this area is underdeveloped, as evidenced by the sectoral nature of works, the small number of organisations under study, and the limited set of financial ratios.

[4] The depth of a decision tree is the maximum number of edges from the root to any leaf node.

In light of the aforementioned limitations of individual statistical methodologies employed in the construction of BFMs, the subsequent sections of this study will assess the applicability of several classical Western and contemporary domestic balance sheet BFMs, which have been developed on the basis of discriminant analysis, in addition to models in the form of logistic regressions. The selection of Western models was based on their wide recognition in scientific and professional circles, as well as their use as a theoretical basis in domestic studies in the development of the author's BFMs. The selection of domestic BFMs was based on the following criteria: high level of citation of scientific publications; selection of explanatory variables based on the analysis of their frequency of use in other popular foreign and domestic BFMs; applicability to several economic sectors; large sample size.

1.2 Peculiarities of modern approaches to the assessment of financial insolvency of organisations

Among the many studies in the field of bankruptcy forecasting, two approaches to classifying organisations as bankrupt (insolvent) can be distinguished: legal and economic.

The legal approach is based on the legal interpretation of the concepts of bankruptcy and debtor. Thus, in accordance with Federal Law No. 127-FZ 'On Insolvency (Bankruptcy)', bankruptcy (insolvency) is defined as 'the debtor's inability, recognised by an arbitration court, to fully satisfy creditors' claims on monetary obligations, on payment of severance payments and (or) on payment of wages to persons employed or working under a labour contract, and (or) to fulfil the obligation to pay obligatory payments'. The term ‘debtor’ is defined as ‘a citizen, including an individual entrepreneur, or a legal entity that has been unable to meet the requirements of creditors on monetary obligations, on payment of severance payments and (or) on payment of wages to persons employed or working under a labour contract’[5].

Consequently, from a legal perspective, the debtor, specifically the legal entity under scrutiny in this study, is deemed bankrupt only subsequent to the relevant decision of the arbitration court. However, it is important to note that the debtor organisation may encounter financial difficulties prior to the recognition of bankruptcy, but subsequent to the initiation of legal proceedings, for instance, at one of the following stages: supervision, financial recovery or external management. It is noteworthy that the duration of consideration by the arbitration court of a filed bankruptcy application is up to seven months from the date of filing the application, and the duration of rehabilitation procedures (financial recovery and external supervision) is up to two years[6].

---

[5] Federal Law of October 26, 2002 N 127-FZ "On Insolvency (Bankruptcy)". URL: https://www.consultant.ru/document/cons_doc_LAW_39331/

[6] How bankruptcy and recognition of a debtor as bankrupt proceed: features of the procedure. URL: https://pravobez.ru/news/kak-prohodit-bankrotstvo-i-priznanie-dolzhnika-bankrotom-osobennosti-procedury.html

Following each stage of the aforementioned process, the arbitration court may elect to declare the debtor bankrupt and initiate bankruptcy proceedings. Statistics demonstrate that, on average, these proceedings exceed two years[7]. Irrespective of the stage of bankruptcy proceedings at which the debtor organisation is at, the creditor's involvement in such a protracted process invariably leads to negative economic consequences. To elaborate, creditors encounter a decline in the liquidity level of accounts receivable, an elevated probability of recognising doubtful debts as uncollectible, and a heightened risk of disruption to business continuity due to delays in the supply of inventories and other resources necessary for business operations. Conversely, creditors must allocate additional material and temporal resources to engage subject matter experts to maximise recovery from the debtor counterparty, as well as to engage in various activities within the bankruptcy proceedings, which the organisation could otherwise devote to business development. Consequently, the initiation of bankruptcy proceedings against a debtor counterparty constitutes a significant event that engenders the risk of financial losses in the event of interaction with this partner.

The economic approach to the definition of bankruptcy is predicated on the notion that a debtor organisation is incapable of meeting its obligations as a consequence of a decline in liquidity and solvency, in addition to a deterioration in financial stability over several periods. Such a negative change in the financial position of the organisation can be caused by both internal (ineffective marketing strategy, suboptimal use of resources in production, problems in financing, reduced sales, etc.) and external (inflation, changes in legislative regulation of activities, scientific and technological progress, etc.) factors that lead to bankruptcy. Consequently, it is imperative to preliminarily identify a subset of debtors experiencing financial difficulties prior to the initiation of bankruptcy proceedings.

In order to identify such a subset of organisations, it is proposed to analyse changes in the dynamics of financial ratios selected due to their use in legislative

[7] Review: Cunning and stagnation have prolonged bankruptcies. URL: https://fedresurs.ru/news/ccfc7b66-7065-4eef-9bc2-a2426e0e895d?attempt=1

documents and their frequent occurrence in domestic and foreign scientific studies on BFMs. In order to take into account the impact of various aspects of the organisation's financial and economic activities on its financial condition, including susceptibility to bankruptcy, the following ratios are identified: current liquidity, solvency on current liabilities and financial stability.

The FTS employs the aforementioned indicators to record and analyse the financial stability and solvency of commercial organisations[8]. Additionally, bankruptcy trustees utilise these indicators during the course of conducting financial analysis in bankruptcy proceedings[9]. A negative trend in the current liquidity and financial stability ratios, coupled with a positive trend in the current liabilities solvency ratio, is indicative of a deterioration in the organisation's liquidity and solvency, as well as financial instability, over the ensuing years. This may, in turn, result in the organisation's bankruptcy.

Taking into account the legal and economic features of bankruptcy, in order to assess the predictive ability of individual foreign and domestic BFMs, a sample of debtor organisations was formed that simultaneously meet the following criteria:

1. Bankruptcy proceedings have been initiated against an organisation. The reporting period utilised for the purpose of bankruptcy forecasting is the year preceding the initiation of bankruptcy proceedings.

During the three years preceding the year of initiation of bankruptcy proceedings, according to the financial statements of the organisation, there was a negative trend in the current liquidity and financial stability ratios and a positive trend in the solvency on current liabilities.

The organisations selected in accordance with the aforementioned criteria will be subjected to further consideration as ***financially insolvent***. The subsequent objective is to assess the predictive capability of extant BFMs within the framework

---

[8] Order of the Ministry of Economic Development of the Russian Federation dated 21.04.2006 N 104 (as amended on 13.12.2011) "On approval of the Methodology for the Federal Tax Service to conduct accounting and analysis of the financial condition and solvency of strategic enterprises and organisations". URL: https://www.consultant.ru/document/cons_doc_LAW_61032/

[9] Resolution of the Government of the Russian Federation of 25.06.2003 N 367 "On approval of the Rules for conducting financial analysis by an arbitration manager". URL: https://www.consultant.ru/document/cons_doc_LAW_42901

of this study, with the aim of identifying debtor organisations subject to bankruptcy proceedings and encountering financial difficulties during the three-year period prior to the initiation of bankruptcy proceedings.

It is also important to identify and systematise the features of modern domestic BFMs for their subsequent use as prerequisites and limitations when creating the author's methodology for assessing the financial insolvency of organisations. The subsequent analysis will focus on two groups of features of the tested models: *economic* and *information-statistical* [Kovalev, Moldobaev, 2021; Bakunova, Koltsova, Moldobaev, 2019; Ilysheva, Savostina, Moldobaev, 2018].

The Economic Features of BFMs

Following a comprehensive analysis of the accounting reporting data of Russian organisations, it was evident that certain entities exhibited deficiencies in the quality of reflected data, characterised by the presence of illogical and contradictory values of indicators, in addition to a substantial number of omissions with regard to specific reporting items [Boiko et al., 2017]. As N.A. Sokolova and N.V. Generalova have noted, accounting information undergoes several transformations at all stages of the life cycle of the accounting process, starting from the reflection of business operations by an accountant and ending with the analysis of financial statements by an interested party. Consequently, accounting standards regulated by law and generally accepted methods of financial statement analysis allow for variability, which enables an organisation to adapt accounting procedures and analytical approaches to the specifics of its activities, while ensuring compliance with regulatory requirements. Concurrently, this variability engenders risks of abuse, thereby providing organisations with the opportunity to manipulate data in order to present a more optimistic representation of financial position, which may not correspond to the actual state of affairs [Generalova, Sokolova, 2012; Generalova, Sokolova, 2013]. The latter problem has a negative impact on the predictive accuracy of balance sheet BFMs that use accounting statements as a source of data for calculating explanatory variables [Generalova & Sokolova, 2012; Generalova & Sokolova, 2013].

In order to comprehend the intricacies of the accounting process, it is imperative to consider the various stages at which it is subject to variability. The following stages are of particular note, as they are subject to diverse accounting procedures and analytical approaches:

1. The initial stage is the reflection of business operations in accounting.

In accordance with the prevailing legislation in the domain of accounting regulation, organisations possess the prerogative to opt for one of several accounting methodologies for their business operations. This enables accountants to manipulate the values of accounting data, thereby enhancing the relevance and appeal of reports for users. It is important to note that certain transactions may have a significant impact on the assessment of the financial position, and these include depreciation accrual using the linear or non-linear method, write-off of inventories using the FIFO or weighted average cost method, inclusion of management expenses in the cost of production or writing them off to expenses of the period, and selection of the appropriate allocation base.

2. Compilation of financial statements.

The standards that govern the preparation and presentation of financial statements permit variability in the grouping and degree of aggregation of items in the financial statements. This variability may consequently affect the accuracy of financial analyses.

3. Analysing financial statement

In order to predict bankruptcy, researchers identify groups of relative financial indicators reflecting liquidity, solvency, financial stability and profitability of the organisation [Boyko et al., 2017]. The final result of the calculation of financial indicators is dependent upon the manner in which the content of the compared sections of the accounting statements is completed. This includes the inclusion of deferred income and estimated liabilities in equity or liabilities, the exclusion of doubtful receivables, long-term receivables, illiquid inventories from current assets, or the maintenance of this section in its original state.; non-comparability of valuations of balance sheet items used to calculate financial indicators: depending on the chosen accounting policy, non-monetary items in assets may be valued at historical prices that do not correspond to the fair value of assets as of the reporting date [Sokolova, 2011; Petrov, 2020; Akulova, Ryakhovsky, 2014; Lvova, 2021].

Information and statistical features of BFMs.

The advent of a plethora of information services, which function as aggregators of data pertaining to the activities of Russian organisations, has had a profound impact on the evolution of contemporary BFMs. Researchers now have the opportunity to import extensive arrays of accounting data and utilise them to evaluate the accuracy of previously created BFMs, as well as to generate new author's BFMs by processing large data sets and leveraging the accumulated scientific and practical experience of previous researchers.

The advent of automated data processing can be largely attributed to the accounting reporting form developed by Rosstat. This form comprises statistical registers that are utilised for the determination of TEA, organisational and legal form, form of ownership, and other pertinent metrics. For instance, domestic scientists employ the All-Russian Classifier of Types of Economic Activities (hereinafter – OKVED) to categorise the studied organisations into economic branches, which are a combination of specific OKVED codes. This practice is not without its drawbacks. Primarily, given the hierarchical structure of the OKVED code, in order to determine whether an individual organisation belongs to a particular industry, it is necessary to take the first digits of the OKVED code, which define a particular class, and compare them with the corresponding codes that form individual industries. However, it should be noted that each class differs from other classes of the same section in terms of its economic essence, as evidenced by the volume and structure of assets, the sources of its financing, the generation of income from operating and other activities, and the peculiarities of costing. Secondly, under the legislation, organisations have the right to engage in several TEAs. For this purpose, they need to specify the main type of activity (hereinafter referred to as MTA) and additional ones during registration. Concurrently, the definition of MTA may be constrained by tax optimisation, encompassing the selection of the most advantageous taxation system, the acquisition of benefits for employee contributions, and the reduction of rates for injury contributions, among other factors. These shortcomings have negative consequences when using OKVED for the distribution of organisations by industries and further construction of BFMs on the basis of an unrepresentative sample.

The accounting reports provided by Rosstat do not include a statistical register, which would make it possible to establish at what stage of bankruptcy proceedings the organisation under study is, except for bankruptcy proceedings, after the end of which the company is already bankrupt and is excluded from the register of legal entities. Consequently, it can be assumed that the sample population used by researchers in the field of financial insolvency does not include organisations that are subject to bankruptcy proceedings and are at one of the stages: observation, financial rehabilitation or external management. During such periods, organisations typically encounter financial challenges and, in numerous instances, are ultimately liquidated as a consequence of being declared bankrupt by the arbitration court. In order to solve the identified problem, it is proposed to test or create BFMs at the stage of data preparation to develop a consolidated database combining the register of accounting data[10] and the register of debtors, starting from the stage of initiation of bankruptcy proceedings[11].

Following a thorough analysis of domestic BFMs, it has been determined that the authors of the models utilise both relative and absolute profitability indicators as regressors. Concurrently, researchers seek to minimise the measure of dispersion by calculating the logarithm function of financial indicators. This, in turn, imposes restrictions on the area of definition of the logarithmic function. The study of financially insolvent organisations revealed 29% of organisations with negative gross profit, 47% with negative profit on sales, 54% with negative profit before tax, and 56% with negative net profit. The analysis of profitability on a set of debtor organisations confirmed the inexpediency of using logarithmic functions that limit

[10] Rosstat. Provision of accounting data upon user requests. URL: https://www.gks.ru/accounting_report
[11] Unified Federal Register of Bankruptcy Information. URL: https://bankrot.fedresurs.ru/DebtorsSearch.aspx

the calculation of the final assessment of financial insolvency forecasting [Kovalev, Moldobaev, 2021; Bakunova, Prisyazhny et al., 2019; Medvedev et al., 2019].

In the creation of BFMs through the utilisation of the discriminant analysis method and logit models, scientists undertake the division of the statistical population into two distinct subsets: the subset of bankrupt organisations and the subset of financially healthy organisations. Accordingly, the proportion of corporate bankruptcies in the Russian Federation in 2018 was 0.3% of the total number of registered organisations. Consequently, the share of organisations continuing entrepreneurial activity is equal to 99.7%[12]. In light of the aforementioned statistics, it can be concluded that when constructing models, researchers have the capacity to formulate a sample of financially robust organisations in a manner that ensures the attainment of a model characterised by elevated values of accuracy indicators for both the training and test samples [Boyko et al., 2017].

Moreover, a clear distinction emerges between classical foreign models and modern domestic models with respect to the volume of statistical sample and the period of observation of financial performance of organisations. For instance, E. Altman developed a five-factor model by studying the activities of 66 organisations in 1946-1965, of which 33 were bankrupt and 33 were financially healthy [Altman, 1968]. Conversely, R. Taffler and G. Tisshaw developed a model using a sample of 46 bankrupt organisations and 46 financially healthy organisations for 1969-1975 [Taffler and Tisshaw, 1968]. [Taffler, Tisshaw, 1977]. In contrast to classical foreign models, modern domestic BFMs are based on the study of large data arrays imported from various information systems. The article by A.V. Kazakov and A.V. Kolyshkin assesses the efficacy of BFMs using a sample size of 31377 (25871 and 5506 representing the number of bankrupt and financially healthy organisations, respectively) for the period 2014-2015. [Kazakov, Kolyshkin, 2018] and the BFMs of E.A. Fedorova, S.E. Dovzhenko, F.Y. Fedorova with a sample size of 8573 (2136 and 6437 – the number of bankrupt and financially healthy organisations

[12] Review: The number of corporate bankruptcies in the Russian Federation in 2018 decreased to a natural level. URL: https://fedresurs.ru/news/5e75d843-000b-4ce1-9925-9e02d89fa0e7?attempt=1

respectively) for 2011-2013 [Fedorova et al. [Fedorova et al., 2016]. It can be hypothesised that differing approaches to the formation of a sample of organisations for the development of BFMs have a bearing on the accuracy of the forecast. Classical foreign models were created on the basis of a study of a small number of organisations with accounting data for a long period of time, which allowed scientists to analyse the activities of each organisation in detail, using additional financial information and consulting with experts who were directly or indirectly related to the analysed enterprises. In the case of the BFMs, the approach used by domestic researchers is limited to solving a mathematical problem without giving economic content to the results obtained. This assertion is substantiated by the existence of contradictory dependencies between the assessment of bankruptcy probability and individual explanatory variables.

It is also important to note that a central problem in the design of BFMs remains the definition of financial insolvency criteria. This issue pertains to both the economic and information-statistical characteristics of BFMs. On the one hand, researchers have access to extensive data sets in the form of accounting statements comprising numerous financial indicators utilised for the calculation of financial ratios. Conversely, the model must be straightforward and accessible for utilisation by other stakeholders, which necessitates the utilisation of a restricted set of financial indicators. The selection of financial indicators must, therefore, satisfy two conditions: firstly, that it assesses the aspects of the organisation's financial and economic activity which have the greatest impact on financial insolvency; secondly, that the indicators used as explanatory variables in the model are statistically significant. It is acknowledged that both conditions permit variability in approaches to meet them, which in practice leads to diversity in the composition of financial ratios used as explanatory variables for different models, despite the common task of assessing the probability of bankruptcy of the organisation.

1.3 Testing foreign and domestic bankruptcy forecasting models

In this study, we conducted a detailed examination of the foreign and domestic balance sheet BFMs, with the objective of identifying the financial indicators that exert the most significant influence on the ultimate value of BFMs. This analysis was undertaken from the perspective of the values of weighting coefficients that mirror the rate of change in the integral assessment when a factor undergoes alteration in a linear model. The ranges of final values of the MPB are also described and their economic interpretation is given.

1. E. Altman's five-factor model for non-public organisations [Altman, 1983]

$$Z = 0{,}717 * A_1 + 0{,}847 * A_2 + 3{,}107 * A_3 + 0{,}42 * A_4 + 0{,}998 * A_5, \quad (1)$$

where A_1 – the ratio of working capital to total assets; A_2 – the ratio of retained earnings to total assets; A_3 – the ratio of EBIT (earnings before interest and taxes) to total assets; A_4 – the ratio of equity to total liabilities; and A_5 – the ratio of revenue to total assets.

Depending on the final value of the function, the organisation can be attributed to one of the classes of financial stability:

- $Z > 2.9$ – area of financial stability;
- $2.9 \geq Z > 1.23$ – the area of uncertainty;
- $Z \leq 1.23$ – the area of bankruptcy.

The consolidated assessment of bankruptcy is predominantly influenced by the coefficients of return on total assets (3.107) and turnover of total assets (0.998), as evidenced by the weighting coefficients of E. Altman's model. Concurrently, the remaining ratios demonstrate a marginal decline in significance, ranging from 0.42 to 0.847.

2. The model of R. Lis [Edmister, 1972]

$$Z = 0{,}063 * L_1 + 0{,}092 * L_2 + 0{,}057 * L_3 + 0{,}001 * L_4, \quad (2)$$

where L_1 – the ratio of working capital to total assets; L_2 – the ratio of EBIT to total assets; L_3 – the ratio of retained earnings to total assets; L_4 – the ratio of equity to total liabilities.

Ranges of values of the function that determine the class of financial stability of the organisation:

- Z > 0,037 – the area of financial stability;
- Z 0,037 – the area of bankruptcy.

The greatest contribution to the final assessment of the R. Lis model is made by the return on total assets ratio with a weight coefficient of 0.092. No less significant are the indicators reflecting the share of working capital in total assets and the ratio of retained earnings to total assets.

3. The model of R. Taffler, G. Tisshaw [Taffler, Tisshaw, 1977]

$$Z = 0{,}53 * T_1 + 0{,}13 * T_2 + 0{,}18 * T_3 + 0{,}16 * T_4, \quad (3)$$

where T_1 – the ratio of profit from sales to short-term liabilities; T_2 – the ratio of current assets to total liabilities; T_3 – the ratio of short-term liabilities to total assets; T_4 – the ratio of revenue to total assets.

Ranges of values of the function that determine the class of financial stability of the organisation:

- Z > 0,3 – area of financial stability;
- 0,3 ≥Z > 0,2 – the area of uncertainty;
- Z≤0,2 – the area of bankruptcy.

The final value of R. Tuffler and G. Tishaw's model depends to a greater extent on the profitability of short-term liabilities. The other three predictors make approximately the same contribution with weight coefficients in the range from 0.13 to 0.18.

4. The G. Springate model [Springate, 1978]

$$Z = 1{,}03 * S_1 + 3{,}07 * S_2 + 0{,}66 * S_3 + 0{,}4 * S_4, \quad (4)$$

where S_1 – the ratio of working capital to total assets; S_2 – the ratio of EBIT to total assets; S_3 – the ratio of profit before tax to short-term liabilities; S_4 – the ratio of revenue to total assets.

Ranges of values of the function determining the class of financial stability of the organisation:

– Z > 0,862 – financial sustainability area;

– Z≤0,862 – bankruptcy area.

In G. Springate's model, the rate of change of the dependent variable is mostly influenced by the return on total assets with a weighting factor of 3.07. For the other factors the weighting coefficients vary from 0.4 to 1.03.

5. The model of A.V. Kazakov, A.V. Kolyshkin [Kazakov, Kolyshkin, 2018]

$$P = 1/(1 + e^{-Y}), \quad (5)$$

where Y is the linear classifier function, which is defined for 4 industries:

Linear classifier in agriculture:

$$Y = -6{,}903 - 16{,}416 * K_1 - 0{,}43 * K_2 - 0{,}326 * K_3 + 0{,}335 * K_4 \quad (6)$$

Linear classifier in construction:

$$Y = -17{,}603 - 0{,}038 * K_5 + 0{,}007 * K_6 + 0{,}961 * K_4 \quad (7)$$

Linear classifier in commerce:

$$Y = -22{,}329 - 1{,}825 * K_1 - 0{,}181 * K_7 - 0{,}039 * K_8 + 1{,}395 * K_4 \quad (8)$$

Linear classifier in the service industry:

$$Y = -16{,}478 - 3{,}493 * K_1 - 0{,}215 * K_9 + 0{,}888 * K_4 \quad (9)$$

where K_1 – the ratio of cash to short-term liabilities; K_2 – the ratio of net profit to total liabilities; K_3 – the ratio of equity to total liabilities; K_4 – the natural logarithm of total liabilities; K_5 – the ratio of equity to current assets; K_6 – the ratio of total liabilities to revenue; K_7 – the natural logarithm of total assets less receivables; K_8 – the natural logarithm of revenue; K_9 – the ratio of profit before tax to total assets.

The following ranges of values of the function are of significance in determining the class of financial stability of the organisation:

– P > 0,5 – bankruptcy area;

– P≤0,5 – financial sustainability area.

A comparison of the modules of weighting coefficients of the integrated model of A.V. Kazakov and A.V. Kolyshkin reveals that the significance of the coefficients differs for the presented industries. Specifically, in the agricultural and service sectors, the integral value of the model is found to be more strongly associated with the absolute liquidity ratio. In the construction sector, financial insolvency is found to be more closely linked to the magnitude of total liabilities. In the trade sector, the absolute liquidity ratio and the natural logarithm of total liabilities contribute more significantly to the model's performance.

6. The model of E.A. Fedorova, S.E. Dovzhenko, F.Y. Fedorova [Fedorova et al., 2016]

$$P = 1/(1 + e^{-Z}), \tag{10}$$

where Z is the logistic regression linear classifier formula, which is as follows:

$$Z = -2,2 - 0,29 * D_1 + 1,45 * D_2 - 0,42 * D_3 - 8,24 * D_4 - 0,9 * D_5 + 1,01 * D_6 + 0,94 * D_7 - 0,06 * D_8 - 0,58 * D_9 + 0,00002 * D_{10}, \tag{11}$$

where D_1 – the ratio of revenue to average current assets for the period; D_2 – the ratio of short-term liabilities to total liabilities; D_3 – the ratio of net working capital to total assets; D_4 – return on assets; D_5 – autonomy ratio; D_6 – receivables to total assets ratio; D_7 – working capital to total assets ratio; D_8 – natural logarithm of tangible assets; D_9 – natural logarithm of EBIT to interest payable ratio; D_{10} – inverse absolute liquidity ratio.

The threshold values of Z, which determine the class of financial stability of organisations, were presented for 10 industries (see Table 1). In the event that the calculated Z indicator for the analysed enterprise is higher than the threshold value, it can be deduced that the enterprise is experiencing an unsatisfactory financial condition and is likely to become bankrupt.

Table 1 – Threshold values of the linear classifier of the logistic regression model by E.A. Fedorova, S.E. Dovzhenko, F.Y. Fedorova [Fedorova et al., 2016]

Economic sector	**Threshold value Z**
Information and communications	-0,895
Science and Technology	-1,118
Real estate	-1,531
Manufacturing industry	-1,001
Agriculture	-1,476
Construction	-1,158
Wholesale	-1,280
Finance and insurance	-1,010
Electric power industry	-0,816
Transport	-0,870

Following a thorough analysis of the weighting factors employed within the model developed by E.A. Fedorova, S.E. Dovzhenko and F.Y. Fedorov, it is evident that the rate of change in the function is predominantly influenced by the return on assets ratio and the proportion of short-term liabilities in total liabilities. Concurrently, the aforementioned factors prove to be the most significant for all ten industries studied by the authors, as the BFM is applicable to all industries, taking into account individual threshold industry values of the function, based on which the financial solvency class of the organisation is determined.

7. The model of E.A. Fedorova, L.E. Khrustov, D.V. Chekrizova [Fedorova et al., 2018]

$$P = 1/(1 + e^{-Z}), \tag{12}$$

where Z is the logistic regression linear classifier formula, which is as follows:

$$Z = -3{,}04 + 0{,}91 * F_1 + 2{,}41 * F_2 - 0{,}12 * F_3 - 0{,}25 * F_4 + 0{,}14 * F_5 - 0{,}19 * F_6, \tag{13}$$

where F_1 is the ratio of current assets to total assets; F_2 = 1 if total assets are less than total liabilities, otherwise F_2 = 0; F_3 is the ratio of revenue to total liabilities; F_4 is the profitability of production; F_5 is the natural logarithm of tangible assets; F_6 is the natural logarithm of the ratio of EBIT to interest payable.

The authors of the model calculated normative values for each industry that delineate the areas of financial stability of organisations. The results obtained are presented in Table 2. In instances where the observed value of the target function surpasses a predetermined threshold, it is deemed that the organisation is susceptible to a high degree of risk of bankruptcy.

Table 2 – Threshold values of the model of E.A. Fedorova, L.E. Khrustov, D.V. Chekrizova [Fedorova et al., 2018]

Economic sector	Threshold value Z
Manufacturing	-0,64630
Agriculture	-0,47827
Real estate	-0,43117
Construction	-0,53444
Transport	-0,82901
Hotels and Catering	-0,89084
Science	-0,53033
Trade	-0,55229

In the model proposed by E.A. Fedorova, L.E. Khrustov and D.V. Chekrizov, it is demonstrated that the following financial indicators, characterised by their maximum modulo weight coefficients, exert a more substantial influence on the ultimate value of the function: the value of net assets, which is articulated in the BFMs as a binary predictor; and the proportion of current assets in total assets.

Following a thorough analysis of both foreign and domestic BFMs, the subsequent phase entailed the evaluation of their quality. This process involved the assessment of the accuracy of the classification of financially insolvent organisations utilising open government data. The following information sources were used for the research:

1. The Open Data provided by Rosstat is presented in the form of an array of data relating to the accounting statements of organisations for the years 2016 to 2018.[13]

2. The Unified Federal Register of Information on Bankruptcy is a compendium of data concerning organisations subject to bankruptcy proceedings in 2019.[14]

The initial phase of data processing encompasses the integration of data from the aforementioned registers, thereby establishing a database that is subsequently utilised for the purpose of conducting further testing of the BFMs. The sample size

---

[13] Rosstat. Provision of accounting data upon user requests. URL: https://www.gks.ru/accounting_report

[14] Unified Federal Register of Bankruptcy Information – portal for disclosure of information on bankruptcy procedures. URL: https://bankrot.fedresurs.ru/bankrupts

was 11,021 organisations subject to bankruptcy proceedings. The subsequent stage entailed the processing of data pertaining to the identification of organisations exhibiting signs of deteriorating financial condition. This was achieved through a meticulous analysis of the dynamics of three financial ratios over a period of three years, prior to the initiation of bankruptcy proceedings. The indicators selected for this purpose include the current liquidity ratio, the current liabilities solvency ratio and the financial stability ratio [Kovalev V.V., Kovalev Vit. V., 2019]. If at least one of the indicators exhibited negative dynamics, it was assumed that the organisation experienced financial difficulties, which, in turn, led it to the extreme in the form of initiating bankruptcy proceedings. Consequently, these organisations were not excluded from the database initially formed, and were instead utilised for further research. Conversely, the remaining organisations that did not satisfy the adopted heuristic rule were excluded from the register. Consequently, following the process of data cleansing, the sample size was reduced to 5,881 entities.

The data were collected, processed and analysed using the Python programming language. The processing of the data, which included arrays in the form of accounting statements, the register of financially insolvent organisations and liquidity, solvency and financial stability indicators calculated for them, as well as predictions of tested BFMs, was achieved by utilising the non-relational database MongoDB. Queries were implemented using the pymongo library. Utilising the aforementioned tools, a programme code was developed to assess the quality of each BFM on a substantial data set.

In order to assess the predictive ability of the models, the author's metric was utilised, which is based on the calculation of the relative recall (completeness) indicator [Powers, 2020]. This indicator reflects the share of financially insolvent organisations that are correctly classified by the tested BFMs in the total number of such organisations from the database formed and described above. Modification of the initial recall indicator also implies taking into account organisations with financial statements for which at least one predictor of BFMs cannot be calculated. The following formula is employed to calculate the author's metric for assessing the predictive ability (henceforth referred to as the 'quality metric', QM) of the tested models:

$$QM = \frac{TP}{TP + FN + NA} \times 100\%, \tag{14}$$

where TP – the number of financially insolvent organisations correctly classified by the model; FN – the number of financially insolvent organisations incorrectly classified by the model; NA – the number of financially insolvent organisations for which it is impossible to calculate at least one predictor of the model according to the accounting data.

The proposed quality metric has been demonstrated to offer several advantages. Primarily, it has the capacity to evaluate the precision of the categorisation of financially insolvent entities. Additionally, it serves as a reflection of the calibre of the model in terms of the relevance of financially insolvent entities' financial statements for the calculation of the indicators utilised within the BFMs. Consequently, if a particular model is found to be suitable specifically for debtor organisations, i.e. all predictors can be calculated using the accounting data for all financially insolvent organisations, the aggregate quality metric will increase. Conversely, if it is not possible to calculate at least one indicator of the model due to mathematical constraints (e.g. zero financial statement line item used as the denominator of the relative indicator; negative retained earnings used as the argument of the logarithmic function), the summary quality metric will decrease.

It is also important to note that within the framework of the present study, the testing of the predictive ability of BFMs for the correct classification of financially healthy organisations was not carried out. It is imperative to acknowledge that forecast errors of financially insolvent organisations when they are financially healthy are less hazardous than forecast errors of financially healthy organisations when they are financially insolvent. In the context of the present study, the author's quality metric does not encompass the scenario where a financially healthy organisation is erroneously predicted to become financially insolvent (i.e. 'reinsurance' and lost profit [Kazakov, Kolyshkin, 2018]).

The BFMs testing was conducted on two sets of data.

The first set comprised raw data concerning debtor organisations subject to bankruptcy proceedings, without separating these into a separate group for organisations for which there was a deterioration in financial condition.

Secondly, the second set of data comprised processed data with financially insolvent organisations in respect of which bankruptcy proceedings have been initiated and for which deterioration of financial condition was observed over the last 3 years by negative dynamics of at least one of the financial ratios: current liquidity, financial stability, solvency on current liabilities.

As a result of the application of the BFMs, all debtor organisations are divided into 4 classes: the financial stability area; the uncertainty area; the bankruptcy area; and organisations with accounting reports (AR) deemed unsuitable for BFMs calculation. Concurrently, the assessment correlating exclusively with the bankruptcy area (TP) was deemed to be correct. All other classes were designated as erroneous assessments (FN or NA).

The findings of the conducted tests for both the initial and processed data are presented in Tables 3 and 4, respectively. It is evident from the data that the maximum prediction accuracy is observed for three foreign models (E. Altman, R. Lis and G. Springate). The classification quality metric (QM) for these models ranges from 71% to 84% for raw data and from 84% to 94% for processed data. Furthermore, when a separate database was created by separating financially insolvent organisations from all debtors, the accuracy of the forecast increased by more than 10% for each of the three listed foreign BFMs, and the share of organisations with accounting statements unsuitable for analysis decreased. With regard to contemporary domestic models, the findings indicate a low degree of forecast accuracy, exhibiting variations across all models within the range of 7-26% for initial data and 2-9% for processed data. The limited forecasting value of domestic models can be attributed to the utilisation of logarithmic functions from profitability indicators as explanatory variables, which are unable to be calculated for negative values. It is noteworthy that the proportion of organisations with negative net profit among the total number of debtor organisations was 56%. Additionally, the final group of models under scrutiny incorporates additional relative indicators, where the denominator is zero, rendering the calculations mathematically infeasible. Conversely, the application of domestic models to a sample of financially insolvent organisations resulted in a substantial decline in forecast accuracy. In contrast, foreign models exhibited an enhanced capacity for accurate prediction.

Table 3 – Testing of models on initial data [Compiled by the author]

Authors of BFMs	**Number of organisations, units**				**Accuracy (QM)**
	Financial stability area (FN)	**Uncertainty area (FN)**	**Bankruptcy area (TP)**	**Organisations with (AR) deemed unsuitable for BFMs calculation (NA)**	
Model of E. Altman	1 121	1 154	7 831	915	71%
Model of R. Lis	896	0	9 210	915	84%
Model of R. Taffler, G. Tisshaw	6 999	1 540	1 525	957	14%
Model of G. Springate	1 443	0	8 621	957	78%
Model of A.V. Kazakov, A.V. Kolyshkin	2 455	0	1 462	1 697	26%
Model of E.A. Fedorova, S.E. Dovzhenko, F.Y. Fedorova	0	0	1 028	9 227	10%
Model of E.A. Fedorova, L.E. Khrustov, D.V. Chekrizova	184	0	679	9 348	7%

Table 4 – Testing models on processed data [Compiled by the author]

Authors of BFMs	Number of organisations, units				Accuracy (QM)
	Financial stability area (FN)	Uncertainty area (FN)	Bankruptcy area (TP)	Organisations with (AR) deemed unsuitable for BFMs calculation (NA)	
Model of E. Altman	339	522	4 918	102	84%
Model of R. Lis	264	0	5 515	102	94%
Model of R. Taffler, G. Tisshaw	3 946	942	880	113	15%
Model of G. Springate	451	0	5 317	113	90%
Model of A.V. Kazakov, A.V. Kolyshkin	2 154	0	43	627	2%
Model of E.A. Fedorova, S.E. Dovzhenko, F.Y. Fedorova	0	0	520	4 959	9%
Model of E.A. Fedorova, L.E. Khrustov, D.V. Chekrizova	280	0	143	5 030	3%

One of the features of the models under consideration is their application for short-term forecasting. As demonstrated by A.V. Kazakov and A.V. Kolyshkin, classical foreign balance sheet BFMs demonstrate optimal forecasting estimation at a one-year horizon [Kazakov, Kolyshkin, 2018]. However, forecasting financial insolvency for two or more periods prior to the initiation of bankruptcy proceedings substantially reduces the accuracy of the forecast. The hypothesis under investigation is thus proposed to be tested using classical foreign and modern domestic models on the processed data of accounting statements for the years 2016-2018 of organisations that were at one of the stages of bankruptcy proceedings during 2019, but at the same time bankruptcy proceedings have not yet been initiated against them in 2018, i.e. we consider that in 2018 the forecast is made 1 year before the initiation of bankruptcy proceedings, 2017 - 2 years, 2016 - 3 years. The results obtained are presented in Figure 1.

It has been demonstrated that the accuracy of foreign models' forecasts is diminished when the time interval between the year in which the forecast was calculated and the year of initiation of bankruptcy proceedings increases. The observed regularity corroborates the applicability of the studied models for forecasting financial insolvency in the short-term period. It is noteworthy that the proportion of organisations with financial statements unsuitable for analysis was the highest in the earliest period (2016). This may have had a bearing on the inaccuracy of the models under consideration. With regard to the domestic models, the obtained results demonstrate multidirectional dynamics with remarkably low values of forecast accuracy when expanding the forecast horizon. This may signify the random nature of the final forecast estimates in subsequent applications of BFMs, as they do not fully take into account the regularities in the accounting data of debtor organisations.

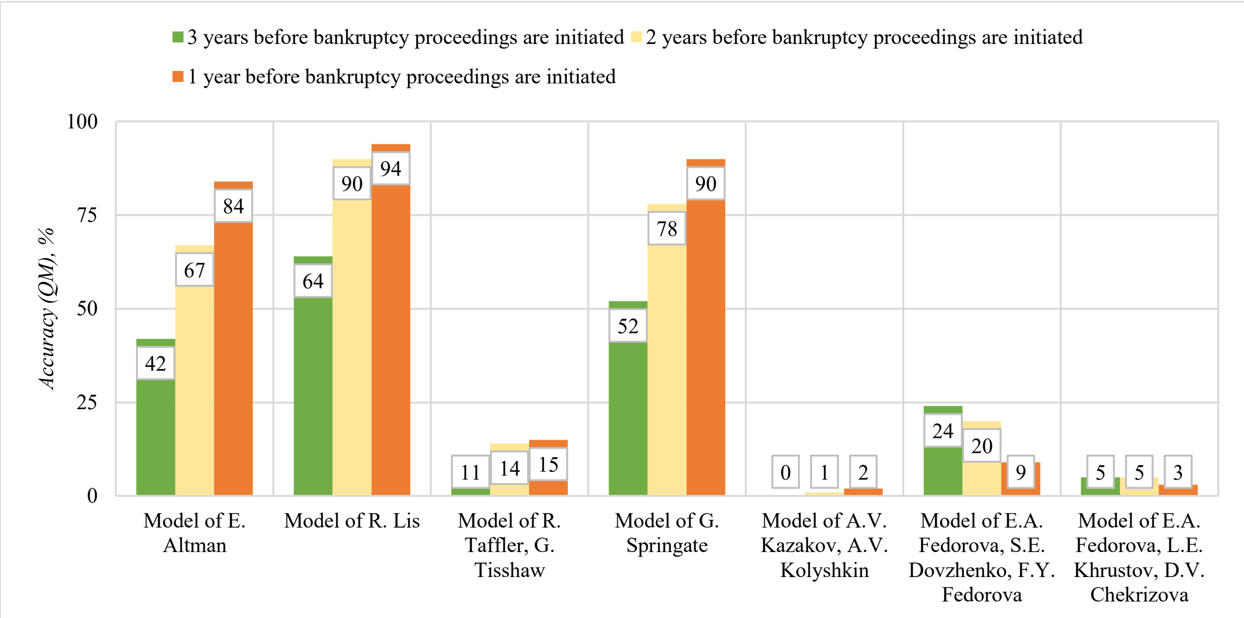


Figure 1 – Change in forecast accuracy of foreign and domestic BFMs as the period before bankruptcy proceedings increases

Chapter conclusions

A comprehensive analysis of extant literature pertaining to BFMs, a popular tool in the assessment of financial insolvency, has been conducted. This analysis has resulted in the identification of three distinct stages of evolution of statistical methods utilised for the resolution of this issue. The initial two stages entail the utilisation of individual financial ratios, alongside the establishment of normative values for these ratios. The limitations of this approach are evident, as single-criteria models fail to address other dimensions of organisational activity; different estimates of normative values for the same indicator result from the variability of the applied statistical methods; and the absence of a composite bankruptcy ratio. The third stage is characterised by the development of multi-criteria models, which are used to calculate the composite bankruptcy ratio. All models within this stage can be divided into two groups: linear (MDA-models, logistic regressions) and non-linear (decision trees, random forest, etc.). In the domestic practice of developing BFMs from the first group, it has been observed that existing foreign BFMs are refined in terms of refining the composite bankruptcy coefficient, or the composition of explanatory variables is preserved with the specification of weighting coefficients, or the combination of the most frequent indicators used as predictors in other BFMs. The second group of models demonstrated high accuracy in bankruptcy forecasting; however, their primary disadvantage lies in the restriction of BFMs data usage by interested parties due to the absence of formulaic representation (in contrast to linear models).

The features of modern domestic BFMs are also identified and classified. The peculiarities are divided into two groups: economic (the influence of accounting procedures and analytical approaches on the registration of facts of economic life of the organisation, formation and analysis of financial statements) and informational and statistical (restriction on importing arrays of accounting statements from RAS; the conditionality of grouping organisations by economic sectors using statistical classifiers; the need to combine data from different sources; mathematical limitations in the use of financial ratios; the observation period). Furthermore, the necessity to ascertain the criteria of financial insolvency is identified as the primary challenge confronting the studied BFMs.

The predictive ability of certain classical foreign and modern domestic balance sheet BFMs has been evaluated through the utilisation of an author-developed quality metric. This metric incorporates a multifaceted assessment, encompassing the accuracy of financial insolvency classification and the proportion of debtor organisations whose accounting statements are deemed unsuitable for calculating the predictors of the examined models.

The classical foreign models demonstrated a high level of forecast accuracy for financially insolvent organisations. Conversely, the performance of modern domestic models, both for initial and processed data, exhibited low accuracy indicators, underscoring the necessity for substantial enhancement.

The experimental findings further substantiated the hypothesis concerning the short-term forecasting horizon of classical foreign BFMs. It was also demonstrated that as the period of bankruptcy forecasting increases, the accuracy of classical foreign BFMs decreases significantly.

CHAPTER 2. METHODOLOGICAL BASIS FOR THE USE OF CLUSTERING ALGORITHMS IN THE FORMATION OF TYPICAL FINANCIAL MODELS IN THE CONTEXT OF ECONOMIC SECTORS

2.1 Formation of a consolidated database for building typical financial models

In order to solve complex problems, including the justification of accounting for the industry specifics of Russian organisations' activities, the construction of typical financial models in the context of industries and their application both for the assessment of financial insolvency and beyond, it is necessary to form a unified database based on the interfacing of large data sets from various sources using identification and classification attributes, as well as interfacing keys. This will be achieved within the framework of the current and next chapters of the thesis. The systematised data sources described below, as well as the developed algorithms for their interfacing, will allow for prompt updating of data to adjust the estimates used in solving research problems.

In the contemporary context, data has emerged as a pivotal asset, providing organisations and society at large with a competitive edge. Moreover, there is an exponential growth of data associated with the intensive development of technology, which has led to the emergence of a separate area of research in the field of big data (Big Data).

Big Data can be defined as structured or unstructured data sets of large volume, which are intended to be processed with the aid of specialised automated tools for further use in statistics, analysis, forecasting and decision-making.

The primary characteristics of big data can be categorised as follows:

– The Volume. This characteristic signifies that data is generated and collected in such a manner that it becomes challenging or impossible to process using conventional tools;

– The Velocity. Big data is characterised by its rapid generation and arrival. Examples of such data include data from social media and financial transactions.

The processing of such data primarily involves the ability to analyse and react to information in real time.

– The Variety. The variety of data types is also a characteristic of big data. Data can be structured, semi-structured, or unstructured, and can be in the form of tables, text, images, audio, or video. The heterogeneity of data formats gives rise to challenges related to integration and analysis;

– The Veracity. Large data sets, generated from a variety of sources, may contain errors, inaccuracies, duplicates or falsified information, which can have a direct impact on the results of analyses and decision-making. Big Data tools facilitate automated procedures to verify data for errors and anomalies, and to rectify data when necessary;

– The Variability. Streaming data is characterised by peaks and downturns, e.g. due to seasons or social events. The more unstable and variable the data stream, the more difficult it is to analyse;

– The Value. The determination of the value of data involves the following: firstly, the assessment of the potential economic benefits that data can bring to the organisation; secondly, the assessment of the ability of data to provide insights; thirdly, the assessment of the relevance of data to a specific task; fourthly, the assessment of whether the benefits derived from the use of data analysis results exceed the costs of data collection, storage and processing; and fifthly, the assessment of the competitive advantage that an organisation can gain by using the results of data analysis [Kapil, et al., 2016; Saeed, Husamaldin, 2021].

The term 'Big Data' was coined by Clifford Lynch, editor of Nature, in 2008. The term was introduced to describe the volumes of data that require new methods and tools for their collection, storage, processing and analysis. Since then, the concept of Big Data has become a pivotal one in the fields of information technology, data analytics and business [Clifford, 2008].

Prior to 2011, the development of Big Data was predominantly focused in the domain of scientific research. However, from 2011 onwards, prominent universities worldwide began to actively incorporate Big Data-related disciplines into engineering and IT courses [Wixom B., et al., 2014]. Since 2011, major IT corporations such as Microsoft, IBM, Oracle, EMC, and subsequently leading technology industry giants like Google, Apple, Facebook[15] and Amazon have also joined the collection and analysis of Big Data [Blasiak, 2014].

In the contemporary business landscape, big data has emerged as a pivotal tool for large organisations across diverse industry sectors, as well as for government entities. It has become a pivotal element in decision-making processes, optimising business operations, and formulating strategies for sustained success.

Big Data tools are also employed when working with individual open government data. In the context of this study, this is considered to be one of the Big Data areas, since standard solutions cannot be used to collect, store and process this kind of data.

Open State Data refers to both structured and unstructured data sets that are published by state bodies, their territorial bodies, local self-government bodies, or organisations subordinate to state bodies or local self-government bodies.

The following events are of particular significance in the context of the development of open government data in Russia[16]:

1. The President of the Russian Federation signed the Decree No. 601 of 7 May 2012 entitled 'On the main directions of improving the system of public administration'. This decree stipulates that by 15 July 2013, Internet access must be provided to open data contained in the information systems of public authorities of the Russian Federation.

2. In 2013, representatives of the G8 countries signed an agreement to develop open data in their respective countries, based on the following principles:

- the provision of data in an open, machine-readable format;
- the assurance of high quality and quantity of data;

---

[15] The organisation is included in the list of public and religious associations and other organisations in respect of which a court decision on liquidation or prohibition of activities has entered into legal force on the grounds provided for by Federal Law No. 114-FZ of 25.07.2002 "On Combating Extremist Activity". URL: https://minjust.gov.ru/ru/documents/7822/

[16] Open Data in the Russian Federation. Open Data Bulletin, July 2015. URL: https://ac.gov.ru/files/publication/a/5572.pdf

– the standardisation of metadata by government;

– the publication of standardised datasets;

– the provision of data to innovative organisations.

In 2013, the Government of the Russian Federation developed and approved the Concept for the placement of information on their activities in the form of open data by state bodies and local authorities (hereinafter - the Open Data Concept), in pursuance of the Decree. The anticipated outcomes of the implementation of the Open Data Concept are as follows:

– ensuring transparency in the work of state bodies and creating an information base for public control;

– creating conditions for the emergence of new services for citizens and organisations, including in the form of applications and services operating on the basis of information in the form of open data.

The main aggregator of datasets is the Open Data Portal of the Russian Federation[17]. Concurrently, the official websites of government agencies are obliged to publish data registers in the designated section for open data, in accordance with the stipulated dates as per the orders issued.

The active development of open data in Russia is a consequence of the rapid progress of the modern innovation economy, which is based on advances in digital technologies. This new digital culture is shaped by citizens, businesses and the state, and ultimately leads to qualitative changes in the principles of interaction between people, the operation of industries and markets. The disclosure of information by government agencies and their integration are directly related to the implementation of the Digital Economy of the Russian Federation programme. However, the development of this programme is hindered by many factors, including the identification of the digital economy with automation, manifested in the implementation of projects aimed at optimising existing processes rather than creating innovative products, and the lack of a digital environment of trust between

[17] Open Data Portal of the Russian Federation. URL: https://data.gov.ru/

the state (Analytical Centre under the Government of the Russian Federation, 2023). Experts also note the high level of fragmentation of state open data, which describe the same object according to the established comparative criterion in different ways [18].

In the course of the research, open data presented on various official websites of public authorities were collected and analysed. The collected data were divided into two groups: those describing the financial and economic activities of Russian commercial organisations, which included accounting data, information on related parties, amounts of taxes paid and others; and those containing data on the movement of inventory, including production, shipment to other regions of Russia, exports and imports. The analysis of these registers revealed that developers of individual databases had independently set requirements to their format, structure and content. This discrepancy led to challenges in integrating data from disparate sources into existing information systems. However, these challenges could be surmounted through the utilisation of programming tools.

In order to establish an information and statistical foundation based on the processing and integration of open state data, and utilised in this study to develop models for assessing the financial insolvency of organisations, it is necessary to:

1. To describe and systematise the codes of state statistics used in automatic processing and analysis of large data sets.

2. To systematise the sources and formats of data presentation.

3. To develop the scheme of data interfacing, taking into account the differences in the codes of state statistics and data presentation formats.

Codes of the state statistics

Codes of government statistics are unique numeric or alphanumeric designations used for the identification, classification and organisation of data in the field of statistics and economy.

[18] Head of the Analytical Center under the Government: SMEV has accumulated a technological lag. URL: https://www.cnews.ru/news/top/2018-11-20_glava_proektnogo_ofisa_tsifrovoj_ekonomiki_raskritikoval

The role of codes of government statistics is pivotal in the domains of data collection, analysis and presentation on various societal and economic facets. The utilisation of these codes facilitates the following:

1. To carry out identification and classification. Codes of state statistics are used for identification and classification of different objects and phenomena, such as enterprises, goods, services, regions, social groups, which allows to standardise data and to ensure their comparability at different levels and in different reports.

2. To perform systematisation and analysis. Data coding facilitates the collection, storage and processing of information. Codes can be used to analyse data, identify trends, determine relationships and draw conclusions based on comparable categories.

3. To reduce errors. Using codes reduces the likelihood of errors in data entry and interpretation, which increases the accuracy of statistical reports and studies.

4. To improve data transparency and accessibility. State statistical codes make data more accessible and understandable to a wide audience, including researchers, government agencies, businesses, and the public.

5. To compare data across states. Many codes are used in several countries at the same time, which facilitates the comparison of statistical data between different countries and ensures international comparability of information [Surinov, 2018; Sorokin, Popova, 2023].

Figure 2 presents the scheme of grouping of codes of state statistics used in the development of models.

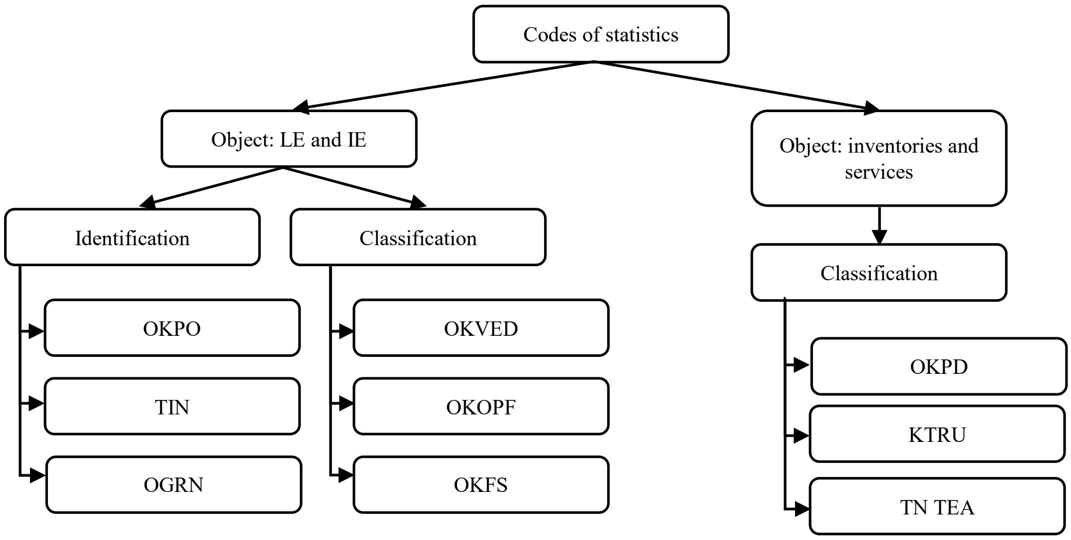


Figure 2 – Codes of state statistics characterising the activity of legal entities and individual entrepreneurs and movement of inventories and services [Compiled by the author]

All codes of state statistics at the first level of hierarchy are divided into two groups, peculiar to: legal entities (hereinafter – LE) and individual entrepreneurs (hereinafter – IE); inventories and services. In its turn, the second level of hierarchy implies division of codes for legal entities and individual entrepreneurs into identification and classification codes. The codes used to reflect the movement of inventories and services relate only to the classification group.

The purpose of statistical identification codes is to ensure unambiguous identification of different entities, including countries, regions and organisations. The utilisation of distinctive codes facilitates the identification and tracing of diverse data elements.

The classification codes of statistics offer the possibility to group data by certain categories, branches, types of goods, services and other attributes, which serves to simplify the structuring of information for the purpose of more convenient analyses.

The following discussion will examine each statistical code presented in figure 2 in turn.

Statistical Identification Codes for Les and IEs

The All-Russian Classifier of Enterprises and Organisations (hereinafter -– OKPO) is a system of coding and classification of organisations and enterprises in Russia, designed to identify and systematise legal entities, individual entrepreneurs, branches, representative offices and other organisational units. The allocation of the OKPO code is the responsibility of the FTS[19].

The peculiarities of OKPO codes assigned to LEs and IEs, from the point of view of their further processing by software tools, are as follows:

– OKPO code has 8 digits for LEs and 10 digits for IEs;

– the first 2 digits of the code define the sphere of activity: natural and labour resources; products of labour and production activity; subjects of national economy; management and documentation;

– the last digit is the control digit, for its determination the method of calculation of the control number established by the standardisation rules is used;

– the remaining digits, except for the last one - individual code generated by the system in the process of registration. It is this combination (excluding the first two digits and the final digit) that must be unique.

The Taxpayer Identification Number (hereinafter referred to as the 'TIN') is a unique numerical identification code assigned to IEs and Les in Russia for the purpose of taxation and identification with various tax and state authorities[20].

The main characteristics of TIN codes include:

– The code consists of 10 digits for LEs and 12 digits for IEs;

– The first 2 digits define the code of the constituent entity of the Russian Federation;

– Digits 3 and 4 reflect the number of the local tax inspectorate;

[19] Order of Rosstat dated 29.03.2017 N 211 "On approval of the Regulation on the All-Russian Classifier of Enterprises and Organisations (OKPO) and classifiers related to it". URL: https://www.consultant.ru/document/cons_doc_LAW_215015/

[20] Order of the Federal Tax Service of Russia dated 29.06.2012 N MMB-7-6/435@ "On approval of the Procedure and conditions for assigning, using, and changing the taxpayer identification number" (Registered in the Ministry of Justice of Russia on 14.08.2012 N 25183).URL: https://www.consultant.ru/document/cons_doc_LAW_134082/

– The remaining digits reflect the number of the taxpayer's tax record in the territorial section of the Unified State Register of Taxpayers.

The Main State Registration Number (hereinafter referred to as the OGRN) is a unique registration number assigned to LEs in Russia upon their creation or registration in accordance with federal legislation[21].

The main characteristics of OGRN codes are as follows:

– The code consists of 13 digits;

– The first digit takes the value 1 or 5, which serves as an indirect confirmation of the code's association with a LE;

– Digits 2 and 3 reflect the last digits of the year of entry;

– Digits 4 and 5 represent the region code;

– Digits 6 to 12 are the number of the entry made in the Unified State Register of Legal Entities (hereinafter – USRLE);

– The thirteenth digit is the control digit. The determination of this digit is made using the method of calculation of the control number established by the standardisation rules.

Classification codes of statistics for LEs and IEs

The All-Russian Classifier of Economic Activities (OKVED) is a system of classification and coding of various types of economic activities carried out by organisations and enterprises in Russia[22].

The main features of OKVED codes include the following:

– The main and auxiliary types of economic activities are indicated for LEs and IEs;

---

[21] Order of the Ministry of Finance of the Russian Federation dated 30.10.2017 N 165n (as amended on 19.12.2022) "On approval of the Procedure for maintaining the Unified State Register of Legal Entities and the Unified State Register of Individual Entrepreneurs, making corrections to information included in the records of the Unified State Register of Legal Entities and the Unified State Register of Individual Entrepreneurs on electronic media that do not correspond to the information contained in the documents on the basis of which such records were made (correction of a technical error), and on recognizing as invalid the order of the Ministry of Finance of the Russian Federation dated 18 February 2015 N 25n" (Registered with the Ministry of Justice of Russia on 16.01.2018 N 49645). URL: https://www.consultant.ru/document/cons_doc_LAW_288080/

[22] "OK 029-2014 (KDES Rev. 2). All-Russian Classifier of Types of Economic Activities" (approved by Order of Rosstandart dated 31.01.2014 N 14-st) (as amended on 30.11.2023). URL: https://www.consultant.ru/document/cons_doc_LAW_163320/

– Codes have a hierarchical structure, according to which all codes are divided by levels, starting with general categories and subdivided into more specific subcategories. Thus, the first two digits reflect a class, the first three digits reflect a subclass, the first four digits reflect a group, the first five digits reflect a subgroup, and the first six digits reflect a species;

– Codes are subject to review and periodic updating in response to changes in the economy and legislation. These changes result in the addition of new codes or the deletion of old codes. Additionally, in Russia, with effect from 01.01.2017, all IEs and LEs transitioned to the new OKVED code system. This transition has introduced further complications in terms of data integration during research, particularly across different periods.

The All-Russian Classifier of Organisational and Legal Forms (hereinafter referred to as the OKOPF) is a classification system developed in Russia for the establishment and recording of various organisational and legal forms of business entities. It also determines the method of property allocation and utilisation by business entities, as well as the resulting legal status and objectives of entrepreneurial activity[23].

OKOPF codes are characterised by a hierarchical structure, which consists of four levels of classification.

The All-Russian Classifier of Forms of Ownership (hereinafter referred to as the OKFS) is a classification system employed in Russia for the definition and categorisation of various forms of ownership. It delineates the relationship between a subject and an object, whereby an object is attributed to a subject that possesses the exclusive right to dispose of, own and use the object[24].

The specific features of OKFS codes include:

– Two-digit codes;

[23] "OK 028-2012. All-Russian classifier of organisational and legal forms" (approved by Order of Rosstandart dated 16.10.2012 N 505-st) (as amended on 14.03.2023) (together with "Explanations to OKOPF positions"). URL: https://www.consultant.ru/document/cons_doc_LAW_139192/

[24] Resolution of the State Standard of Russia dated 30.03.1999 N 97 (as amended on 22.09.2023) "On the adoption and implementation of All-Russian classifiers" (together with "OK 027-99. All-Russian classifier of forms of ownership") (date of introduction 01.01.2000). URL: https://www.consultant.ru/document/cons_doc_LAW_26587/

– Hierarchical structure of the classifier, which implies that individual groups are included in subgroups;

– The first level of the classifier assumes distribution of all objects into the following groups: Russian property, foreign property, joint Russian and foreign property.

Statistical codes for inventories and services

The All-Russian Classifier of Products by Types of Economic Activity (hereinafter referred to as the OKPD) is a classification system employed in Russia for the purpose of categorising a wide range of products and services associated with economic activity[25].

The primary characteristics of OKPD codes are as follows:

– The utilisation of codes is primarily for statistical analysis, public procurement, and document compilation;

– The code possesses a hierarchical structure, with the potential to comprise between two and nine digits;

– The first 6 digits correspond with OKVED codes.

The catalogue of goods, works and services (hereinafter referred to as KTRU) is a classification system used in Russia to categorise various types of goods, works and services[26].

The peculiarities of KTRU codes are as follows:

– they are formed on the basis of OKPD codes;

– KTRU codes detail OKPD codes;

– the purpose of using KTRU codes is exactly the same as OKPD codes.

[25] "OK 034-2014 (KPES 2008). All-Russian classifier of products by types of economic activity (OKPD 2)" (approved by Order of Rosstandart dated 31.01.2014 N 14-st) (as amended on 04.02.2022).URL: https://www.consultant.ru/document/cons_doc_LAW_163703/

[26] Resolution of the Government of the Russian Federation of 08.02.2017 N 145 (as amended on 27.03.2023) "On approval of the Rules for the formation and maintenance in the unified information system in the field of procurement of a catalog of goods, works, services to meet state and municipal needs and the Rules for the use of the catalog of goods, works, services to meet state and municipal needs". URL: https://www.consultant.ru/document/cons_doc_LAW_212534/

Commodity Nomenclature of Foreign Economic Activity (hereinafter - TN TEA) is an international system of classification of goods, which is used to describe and classify goods transported across the borders of different countries[27].

The primary characteristics of TN TEA codes are as follows:

– These codes are developed in accordance with international standards and accepted by most countries worldwide, including Russia;

– The coding of goods is used for unambiguous classification of goods that move across the customs border, as well as to simplify automated processing of information and statistics;

– Each commodity is assigned a TN TEA code, which is used for the purpose of declaration or collection of customs duties;

– the structure of the ten-digit TN TEA code is based on the decimal system and includes the code of the group, commodity item, sub-item and sub-sub-item;

– a commodity group includes the first two digits, a commodity position - four digits, a subposition - six digits, and a subposition - ten digits.

In the subsequent stage, consideration should be given to the data sources utilised in this study and their respective presentation formats.

In a manner analogous to the systematisation of codes of government statistics, all data sources at the first level of the hierarchy can be divided into two groups: those related to LE and IE; and those related to inventories and services. In turn, data sources related to LE and IE can be divided into two subgroups: financial and non-financial data. The categorisation of data sources pertaining to inventories and services is based on the nature of market movement, distinguishing between domestic and foreign activities.

[27] Decision of the Council of the Eurasian Economic Commission of 14.09.2021 N 80 (as amended on 27.12.2023) "On approval of the unified Commodity Nomenclature of Foreign Economic Activity of the Eurasian Economic Union and the Single Customs Tariff of the Eurasian Economic Union, as well as on amending and recognizing as invalid certain decisions of the Council of the Eurasian Economic Commission" (as amended and supplemented, entered into force on 08.02.2024). URL: https://www.consultant.ru/document/cons_doc_LAW_397176/

Figure 3 presents a scheme of data source grouping.

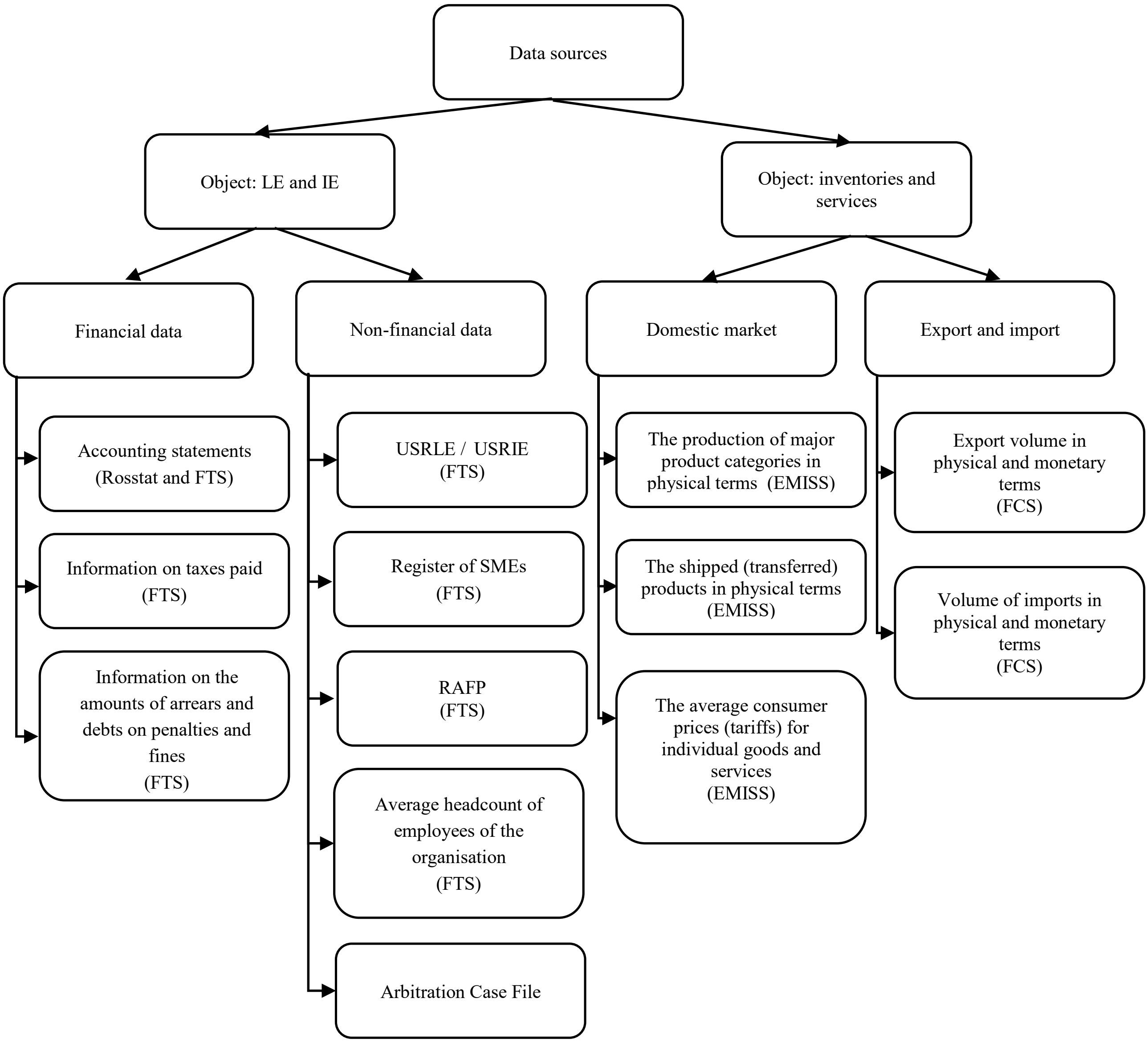


Figure 3 – Data sources describing LE and IE activities and movement of inventory and services [Compiled by author].

The financial data pertaining to LE and IE is presented in the following registers:

1. LE accounting statements, including balance sheet, profit and loss statement, statement of changes in capital, cash flow statement, appendix to the balance sheet and report on the targeted use of funds received. The accounting statements data for the period 2012-2018 is available on the official website of Rosstat, whereas data from 2019 onwards is available on the website of the Federal

Tax Service. The data format employed on the Rosstat website is CSV (Comma-Separated Values)[28], whereas the FTS utilises XML (Extensible Markup Language)[29].

2. The following information is provided on the amounts of taxes and levies (for each tax and levy) paid by an organisation in the calendar year preceding the year in which the above information was posted on the information and telecommunications network 'Internet'. It should be noted that the amounts of taxes (levies) paid in connection with the importation of goods into the customs territory of the Eurasian Economic Union, the amounts of taxes paid by a tax agent, and the amounts of insurance contributions are excluded. This register was first published on 01.10.2019 on the website of the FTS. The register is presented in XML format[30].

3. The set of data encompasses information pertaining to amounts of arrears of arrears and arrears of fines and penalties. The set contains information regarding the amounts of arrears, arrears of penalties, arrears of penalties and fines (for each tax and levy, insurance contributions for which the organisation has arrears and/or arrears of penalties and fines); and the total amount of arrears and arrears of penalties and fines. The initial publication date of the data set is 01.12.2019. The data is presented in XML format[31].

Among the registers relating to LEs and IEs and describing non-financial data are:

1. The Unified State Register of LE (USRLE) and IE (USRIE) is a comprehensive database that serves as a central repository for official records related to LEs and IEs in the Russian Federation. The register contains registration data,

---

[28] Accounting (financial) statements of enterprises and organisations for 2018.URL: https://rosstat.gov.ru/opendata/7708234640-7708234640bdboo2018

[29] A monthly updated complete database of accounting (financial) statements of organisations throughout Russia. URL: https://bo.nalog.ru/

[30] Information on the amounts of taxes and fees paid by the organisation in the calendar year preceding the year of posting the said information on the Internet in accordance with paragraph 1.1 of Article 102 of the Tax Code of the Russian Federation (for each tax and fee), excluding the amounts of taxes (fees) paid in connection with the import of goods into the customs territory of the EAEU, the amounts of taxes paid by the tax agent, and the amounts of insurance premiums. URL: https://file.nalog.ru/opendata/7707329152-paytax/structure-20180110.xsd

[31] Information on the amounts of debt for taxes, fees and insurance premiums, penalties and fines to the budgets of the budgetary system of the Russian Federation. URL: https://data.nalog.ru/opendata/7707329152-debtam/structure-20181201.xsd

including identification and classification codes of state statistics, as well as ancillary information useful for research, including the date of LE registration, date of LE liquidation, information on the director and founders, the amount of authorised capital, etc. Access to the data is granted by the FTS via the application programming interface (henceforth - API). Access to the data is provided by the FTS through the application programming interface (hereinafter - API). The format of data received via API is XML[32].

2. The Register of small and medium-sized enterprises (henceforth referred to as the SME Register) is a comprehensive database that serves as a crucial tool for the administration and management of small and medium-sized enterprises (SMEs) in the Russian Federation. This set comprises registration data of LEs and IEs that have been designated as SMEs and are eligible to apply for supplementary state support. The FTS is the proprietor of the data set, and the initial publication date is 01.08.2016. The data is presented in XML format[33].

3. The State Register of Accredited Branches, Representative Offices of Foreign LEs (RAFP) is a comprehensive compendium of foreign LEs operating within the Russian Federation. The data is presented on the FTS website, employing an XML data format, and the set was first published on 29.03.2018[34].

4. The following data set contains information regarding the average number of employees of an organisation. The FTS is the proprietor of the data set, and the data set is presented in XML format. The date of first publication was 01.08.2019[35].

5. The arbitration case file is a compendium of judicial decisions pertaining to all instances involving LE, IE, federal and local authorities[36].

---

[32] Integration of information from the Unified State Register of Legal Entities and the Unified State Register of Individual Entrepreneurs into the information systems of interested parties. URL: https://www.nalog.gov.ru/rn77/service/egrip2/egrip_vzayim/

[33] Unified register of small and medium-sized businesses. URL: https://file.nalog.ru/opendata/7707329152-rsmp/structure-10062023.xsd

[34] State Register of Accredited Branches and Representative Offices of Foreign Legal Entities (RAFP). URL: https://data.nalog.ru/opendata/7707329152-rafp/structure-16032022.xsd

[35] Information on the average number of employees of the organisation. URL: https://file.nalog.ru/opendata/7707329152-sshr2019/structure-20200408.xsd

[36] Electronic justice. URL: https://kad.arbitr.ru/

6. The present document constitutes a register of intellectual property objects. It comprises distinct data sets, encompassing databases, inventions, utility models, computer programmes, industrial designs, trade secrets (know-how), selection achievements, and integrated circuit topologies. The proprietor of the data set is the Federal Service for Intellectual Property (hereinafter referred to as Rospatent), and the data format is CSV. The date of first publication is 19.10.2017[37].

In order to describe the movement of stocks and services in Russia (domestic market), the following data registers were used, which are presented in XLSX format on one of the services of Rosstat - in the Unified Interdepartmental Information and Statistical System (hereinafter – EMISS), with details by regions of Russia:

1. The production of major product categories in physical terms since 2017 (operational data in accordance with OKPD)[38].

2. The shipped (transferred) products in physical terms from 2017 (operational data in accordance with OKPD)[39].

3. The weekly average consumer prices (tariffs) for individual goods and services are also provided[40].

The official website of the Federal Customs Service of the Russian Federation (FCS) presents data on imports and exports of inventories in the form of CSV files. The data on exports and imports of products is reflected using the TN TEA codes and detailed by months, counterparty countries and administrative-territorial units of the Russian Federation[41].

Pairing data from different sources

In the course of conducting a variety of surveys, there is frequently a requirement to link data from multiple sources. This task is complicated by the fact that, in order to establish links between disparate data, it is necessary to analyse in

[37] Open data of the Federal Service for Intellectual Property. URL: https://rospatent.gov.ru/opendata
[38] Production of main types of products in physical terms since 2017 (current data in accordance with OKPD2). URL: https://www.fedstat.ru/indicator/57783
[39] Products shipped (transferred) in kind since 2017 (current data in accordance with OKPD2).URL: https://www.fedstat.ru/indicator/57786
[40] Weekly average consumer prices (tariffs) for individual goods and services. URL: https://fedstat.ru/indicator/37426
[41] Export and import of the Russian Federation by goods. URL: https://customs.gov.ru/statistic

advance which codes of state statistics should be used as keys for further data merging.

Utilising the codes of state statistics presented in figure 2 and the data sources delineated in figure 3, it is feasible to generate typical scenarios for merging data from disparate groups and subgroups:

1. Data fusion by identification code. Thus, for all sources from the LE and IE group, including financial and non-financial registers, the pairing is performed using the TIN.

2. Data merging by pairing keys. In the context of calculating the volume of the regional sectoral market, defined as the difference between the sum of production and imports and the sum of shipments to other regions of Russia and exports, it is imperative to record the parameters affecting the final result in various classification systems. Specifically, production and shipments are to be recorded in OKPD codes, while exports and imports are to be recorded in TN TEA codes. It is imperative to standardise the parameter values to align with a single system when quantifying market volume. To this end, the OKPD2-TN TEA transition keys, developed by the Ministry of Economic Development of the Russian Federation, are employed[42].

3. By-digit comparison of classifiers. In order to analyse the structure of the regional industry market and determine which IEs and LEs fill it, it is necessary to establish a link between two groups of data sources. The first group consists of data sets on IEs and LEs, while the second group consists of data sets on inventories and services. The solution to this problem is to match the first six digits of the OKVED2 and OKPD2 codes. These codes are found to have a similar structure.

[42] Transition keys between the Commodity Nomenclature of Foreign Economic Activity of the Eurasian Economic Union TN TEA EAEU and the All-Russian Classifier of Products by Types of Economic Activity (OKVED) OK 034-2014 OKPD2. URL: https://economy.gov.ru/material/file/8fe3bac6d1fec0a3b943272e28212592/%D0%A2%D0%9D%D0%92%D0%AD%D0%94_%D0%9E%D0%9A%D0%9F%D0%942_20_07_2023.xlsx

As illustrated schematically in Figure 4, the coupling of situations 2 and 3, as previously described, is demonstrated.

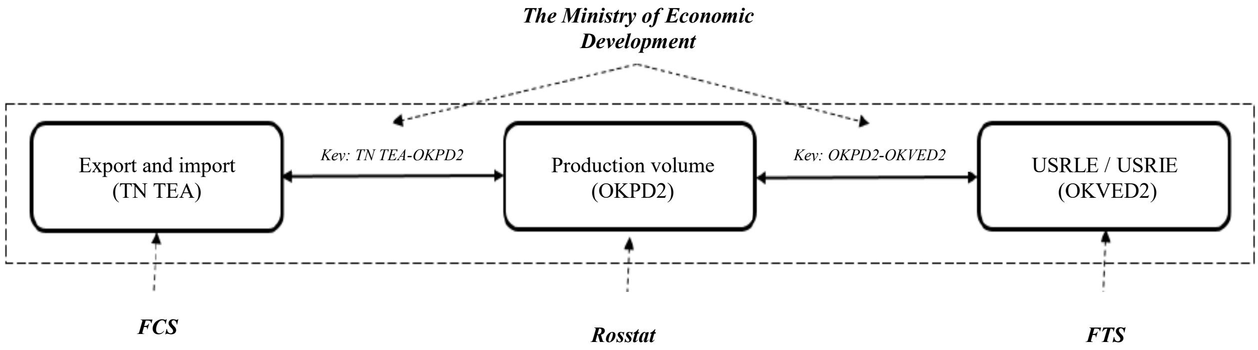


Figure 4 – Scheme of consolidation of databases with different statistical classifiers [Compiled by the author]

The heterogeneity of codes of state statistics, data sources, and data presentation formats, in conjunction with the myriad methods of data interfacing, renders the manual approach to data collection for a relatively small sample a laborious process, and for a large array, practically unfeasible. To address this challenge, the study uses the Python programming language and its integrated libraries for data collection, processing and storage, encompassing:

- csv – the library provides functions for reading and writing CSV files;
- xmltodict – the library allows you to convert XML-format data into a dictionary, which is easy to work with;
- pymongo – the library allows you to connect to the MongoDB non-relational database, write and unload data from it;
- numpy – the library contains one-dimensional data structures for which statistical functions are implemented;
- pandas – the library contains both one-dimensional and two-dimensional data structures, for which many statistical functions are defined.

The following libraries were utilised in the construction of machine learning models and the visualisation of data:

– sklearn – the library contains many classes for creating various machine learning models;

– matplotlib – the library supports data visualisation tools.

Following the establishment of an information base comprising a set of open, disparate financial and non-financial data, as well as the subsequent establishment of the keys for interfacing different registers, it is proposed that, at the subsequent stage, separate slices of the developed consolidated database be utilised for the industry analysis of financial ratios, the construction of typical financial models in the context of TEA, and their further application in the assessment of financial insolvency, and for the justification of decisions on M&A transactions in horizontally integrated companies.

2.2 Industry specifics in assessing the financial insolvency of organisations

A review of contemporary domestic BFMs reveals a tendency among scientists to prioritise the industry-specific characteristics of models, despite a lack of substantiation for this approach [Kazakov, Kolyshkin, 2018]. The sectoral specification of BFMs is presented in two ways: either as a set of models with different sets of explanatory variables, or as a single model for which the normative values of the composite bankruptcy ratio are specified depending on the economic sector.

The purpose of this stage of the research is to test the hypothesis that it is necessary to stratify TEA organisations when developing models. The financial and non-financial performance indicators of Russian organisations were considered as the parameters to be analysed. All financial indicators were calculated according to the accounting statements and grouped into blocks: liquidity, financial stability, business activity, profitability [Kovalev V.V., Kovalev Vit. V., 2019]. The age of the organisation was considered as a non-financial indicator, which was calculated as the difference between the following dates: 31.12.2020 and the date of registration.

The USRLE[43] was used to obtain information on the date of registration and MTA of the organisations under study, and the register of accounting statements of Russian organisations[44] was used to obtain financial indicators (see Fig. 3). The listed databases are available on the official websites of Rosstat and FTS. To match the data from the two sources, the TIN of the organisations was used (see Fig. 2).

The selection of objects was conducted on the basis of the following criteria:

1. Firstly, total assets must equal total liabilities.
2. Revenue for 2020 and total assets as of 31.12.2020 are greater than zero.

---

[43] Integration of information from the Unified State Register of Legal Entities and the Unified State Register of Individual Entrepreneurs into the information systems of interested parties. URL: https://www.nalog.gov.ru/rn77/service/egrip2/egrip_vzayim/

[44] A monthly updated complete database of accounting (financial) statements of organisations throughout Russia. URL: https://bo.nalog.ru/

The initial detection of distortions in the balance sheet is contingent on the fulfilment of the first condition [Pyatov, 2014]. In the context of accounting, the principle of balance (equality) of total assets and total liabilities is of paramount importance. The second condition enables the selection of only active organisations that generate income and possess the resource potential to support their activities.

Following the application of the aforementioned rules, the sample size was reduced to 1,580,510 organisations, which were then grouped by TEA at the subsequent stage. The Python programming language was utilised for the processing of the voluminous data set.

In order to demonstrate the differences in the activities of organisations from different sectors of the economy, six TEAs were randomly selected from all processed TEAs (a total of 89 classes in OKVED). The sectoral features identified by other scientists were systematised for these six TEAs, and the financial and non-financial indicators described above were analysed. The list with OKVED codes and names is presented below:

1. Extraction of oil and natural gas (OKVED 06).
2. Extraction of metal ores (OKVED 07).
3. The manufacture of tobacco products (OKVED 12)
4. The provision of electricity, gas and steam; air conditioning (OKVED 35)
5. The provision of pollution remediation and other services related to waste disposal (OKVED 39).
6. The retail trade, with the exception of trade in motor vehicles and motorbikes, is categorised under OKVED 47.

The oil and natural gas production sector is characterised by numerous peculiarities, including:

1. A strong dependence on the dynamics of energy prices. In the context of declining energy resource costs, enterprises in this sector may encounter significant financial challenges, necessitating a strategic approach to risk management [Sokolov, 2019].

2. A high degree of capital intensity is exhibited, signifying that enterprises possess substantial financial resources, enabling extensive investments in diverse domains, including infrastructure, technical equipment, and research and development. [Vyakina, Garanikova, 2015].

3. The geographical diversity of oil and gas production, coupled with the inherent geopolitical risks, underscores the imperative for active interaction and cooperation with a range of national and international actors within this sector. [Gnilitskaya, 2002].

4. The extraction of oil and gas resources is associated with an increased level of social and environmental responsibility, and with negative consequences for the environment and society as a whole (see Gorbunova, Kanitskaya, 2017; Schwartz et al., 2015; Yao et al., 2011).

5. The global nature of operations, which implies that one organisation operates in several countries, which, first of all, has an impact on logistics management [Leonov, Voronov, 2017; Syrovetsky, 2020; Erokhin, 2021].

6. Digital transformation and technological progress represent pivotal structural components of the strategic framework employed by oil and gas organisations. Enterprises operating within this sector must continuously invest in modern technologies in order to improve production processes, reduce negative environmental impact and ensure long-term competitiveness [Zhukinsky, 2022].

7. Regular cyclicality in oil and gas production is attributable to systematic fluctuations in economic activity, which affect the dynamics of demand and price formation for energy resources [Gil-Alana, Gupta, 2014].

The features of the TEA 'Mining of metal ores' bear a strong resemblance to the extraction of oil and natural gas. For this particular TEA, we can also highlight the following:

1. High level of energy intensity. The process of metal mining requires significant energy costs, which has a significant impact on the financial stability of organisations in this sector [Kaplunov, Yukov, 2016].

2. The technical intricacy of mining operations entails the utilisation of sophisticated technological complexes and advanced equipment, encompassing drilling apparatuses, processing facilities, and other specialised apparatus [Martins, 2019].

The production of tobacco products is characterised by its own distinctive characteristics, which include the following:

1. A high level of state regulation in terms of the development and application of norms on product quality, packaging, and warning about the harm of smoking [Budarin, Perepechkina, 2018].

2. Tobacco products are subject to significant excise duties, which often has a significant impact on the price setting of finished products and the financial results of enterprises in this industry [Salomatin et al., 2021].

3. Dependence on the dynamics of consumer preferences, influenced by individual habits and socio-cultural factors. Enterprises in this sector respond to changes in consumer preferences by adapting production and marketing strategies [Pidyashova et al., 2019; Migunova Y.V., 2020].

The TEA ‘Supply of electricity, gas and steam; air conditioning’ is characterised by the following features:

1. High infrastructure dependence, which implies the presence of developed infrastructure for the production and distribution of electricity, gas and steam [Butakova et al., 2022].

2. Regulation and licensing. This type of activity is often subject to strict government regulation and licensing, and includes safety standards, quality norms and environmental protection requirements [Kologermanskaya, 2020].

3. The seasonality of demand is another salient factor. Demand for electricity, gas and steam can vary according to the season and climatic conditions. For instance, electricity consumption has been observed to increase during winter months due to heating requirements, and in summer months due to air conditioning [Munirov, 2023].

4. Reliability and continuity requirements are also crucial factors in the provision of these essential services. Electricity and gas are critical resources, and their supply must be reliable to ensure the continuous functioning of society and enterprises [Fadeev & Fadeeva, 2020].

5. Technical complexity. The generation, distribution and maintenance of infrastructure for the supply of electricity, gas and steam are technically complex tasks that require specialised knowledge and equipment. This underscores the significant investment required in research and development to ensure the effective management of these resources [Butakova et al., 2022].

The provision of pollution remediation and other waste management services is a specific industry characterised by the following features:

1. A high level of government regulation. Waste management and environmental remediation are frequently subject to stringent state and local regulation, encompassing standards for the treatment, removal and disposal of hazardous waste [Sarkisov et al., 2022].

2. The technical complexity of pollution elimination and hazardous waste disposal processes, which involve the use of sophisticated equipment and specialised technologies and the availability of highly qualified specialists [Saraev, 2018].

3. Continuous training and certification are also essential components of this process. Specialists in this industry must constantly train and update their skills, as requirements and technologies are constantly changing [Svyatokho & Timayev, 2020].

4. A high level of investment in research and development is also required. The development of new methods and technologies for more efficient and environmentally sustainable elimination of waste and pollution is becoming increasingly important [Zabortseva et al., 2017].

5. Significant amount of diverse resources required to fulfil large pollution elimination and waste management projects, including equipment, personnel and finances [Chernov, 2020].

TEA ‘Retail trade, except trade in motor vehicles and motorbikes’ is characterised by the following features:

1. The variability of product assortment. The sale of a wide range of goods is encompassed within the purview of retail trade, ranging from food and clothing to electronics and household goods. The effective management of this assortment is predicated on a comprehensive understanding of both external factors, including demand for a particular product, price category, product life cycle, and internal factors, including enterprise type, location, and strategy [Petrova, 2021].

2. High level of competition. Retailing is often characterised by a high degree of competition, especially in large cities [Munshi et al., 2022].

3. Seasonality of sales. Certain categories of goods may have seasonal demand, so retailers have to take this factor into account when planning and managing inventory [Butters, Sacks, Seo, 2023].

4. Trade organisations may operate in low-margin segments, therefore effective management of costs and turnover of goods is imperative for maintaining or improving the financial condition of organisations in this sector [Kovalenkova, 2023].

5. High level of expenditure on marketing and promotion. Advertising and marketing, as key components of operational activities, are aimed at actively attracting consumers, building and strengthening brands, and stimulating an increase in sales. The integration of modern marketing methodologies, encompassing digital technologies, analytical tools and personalisation strategies, has emerged as a pivotal factor in achieving competitive advantage within the retail sector. [Grishina, 2022].

6. Inventory management. Inventory control is necessary to minimise additional costs associated with the storage of goods and reduce the risk of losses from obsolete goods [Novikova, Shchepina, 2023; Sharokhina, 2023].

7. The development of e-commerce is another key factor in this regard. In the contemporary context, online trade and e-commerce have emerged as pivotal sales channels for trade organisations, as the development of their online presence has

been shown to result in a substantial augmentation in sales [Popenkova, Stukalova, 2022].

The characteristics of individual TEAs described above have an effect on the values of financial and non-financial indicators in one way or another. Consequently, by demonstrating the discrepancy in average sectoral estimates of financial and non-financial indicators, including frequently used financial ratios in BFMs, a substantiated foundation for the advancement of typical financial models by TEA is established.

Let us consider the median age of organisations with non-zero revenues and total assets (Fig. 5). It is evident that activities related to oil and natural gas extraction, the provision of electric power, gas and steam, and the extraction of metal ores exhibit some of the highest age values, with median ages of 10, 8, and 6 years, respectively. The total number of organisations engaged in the first type of activity is 708, for the second it is 13,222, and for the third it is 1,712. The high median age values of organisations in these sectors can be attributed to the presence of substantial barriers to entry for new competitors, which include significant infrastructure investments. The third oldest sector is retail trade, with an average lifespan of seven years. This sector is characterised by constant changes in consumer demand and competition. Furthermore, the sector is characterised by heterogeneity, with retail organisations offering a diverse range of products from multiple, distinct product markets. Consequently, a more detailed analysis of individual organisations is necessary to identify homogeneous age groups within this sector. The TEAs 'Provision of pollution remediation and other waste management services' and 'Manufacture of tobacco products' exhibit the lowest median ages of 4 and 2 years, respectively. A total of 133 organisations are present in the Russian market for the first type of activity, and 235 organisations for the second type of activity. The presence of low age values may be indicative of a high level of competition and dynamism within these sectors.

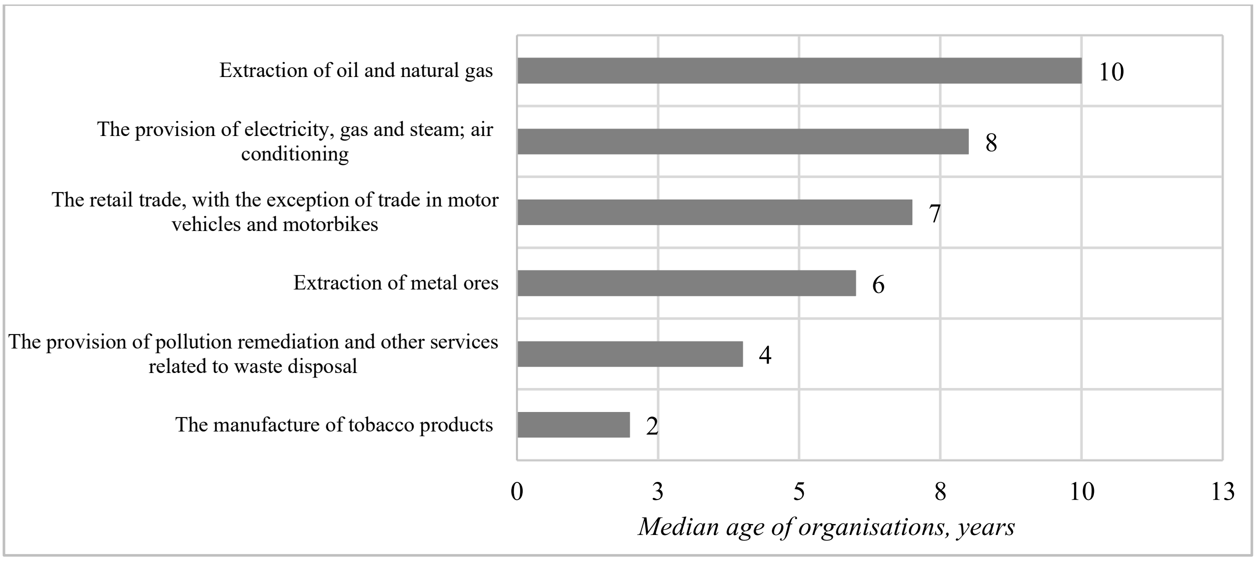


Figure 5 – Median age of TEA organisations as of 31.12.2020 [Compiled by the author]

Figure 6 shows a graph of the distribution of organisations by age in terms of the TEAs under consideration.

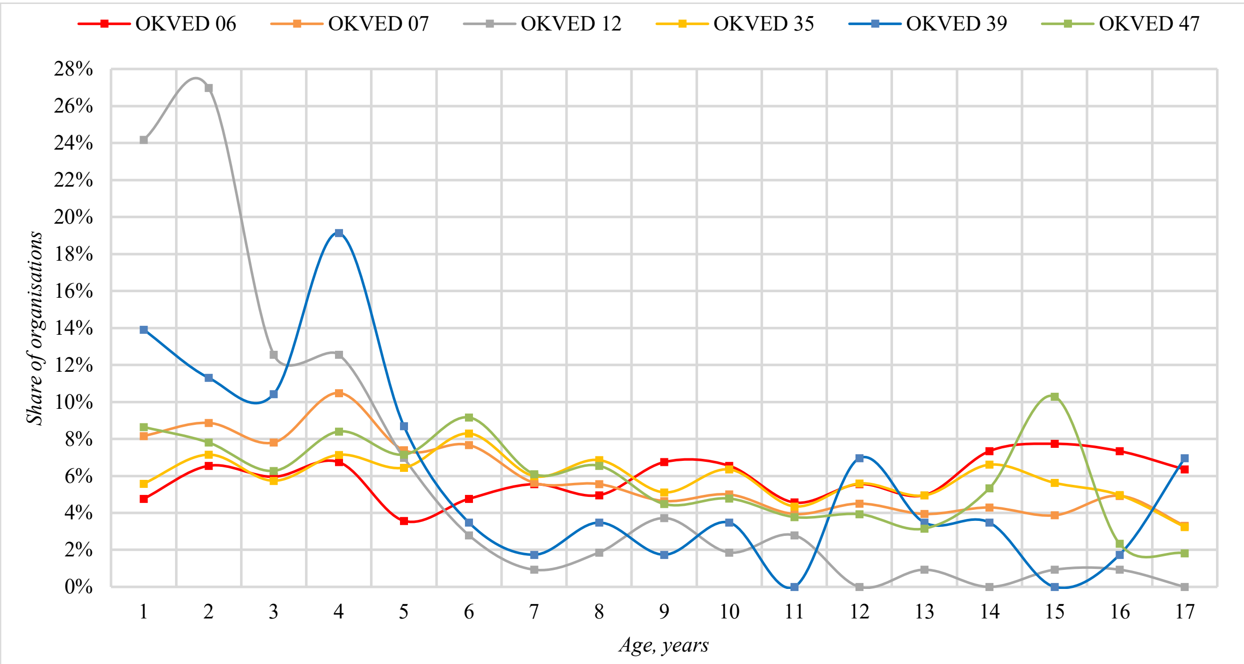


Figure 6 – Distribution of organisations by age by TEA as of 31.12.2020 [Compiled by the author]

In the oil and gas extraction sector (OKVED 06), organisations with a duration of 15 years represent the largest share of all organisations in this sector, accounting for 7.7%. Conversely, organisations with less than six years of existence (i.e. those between one and five years old) account for approximately 27.6% of all organisations in this sector. This finding suggests that new players are actively entering the sector due to the growing interest in oil and gas extraction. Conversely, organisations with a history of over 10 years (exhibiting an age range of 11 to 17 years) account for approximately 43.8%, denoting the existence of a substantial number of well-established entities that possess extensive experience in the industry. The age representation of organisations in this industry is quite diverse, indicating the dynamic and heterogeneous nature of the oil and natural gas extraction business.

The following investigation will consider the age distribution of organisations in the metal ore mining sector (OKVED 07). The predominant share of organisations (10.5%) is 4 years old. The proportion of organisations with an age of less than six years (in the range of one to five years) is 42.8%, while enterprises exceeding 10 years of age (in the range from 11 to 17 years) account for a share equal to 28.7%. This finding suggests the presence of both financially attractive sectors for young organisations and stable, experienced participants.

In the tobacco products manufacturing sector (OKVED 12), organisations with a maximum of two years of existence account for the largest share of 27%. Furthermore, organisations that have been in existence for less than six years account for 83.4% of all organisations in this sector. This may indicate the dominance of relatively young organisations and the dynamic emergence of new players in the sector. Conversely, organisations with a duration of over 10 years occupy a negligible proportion of 5.5%, suggesting a paucity of long-term entrants. This distribution is indicative of a concentration on young enterprises, which may be attributed to the high proportion of start-ups and new entrants in the industry.

The predominant proportion is accounted for by 6-year-old organisations with TEA 'Provision of electricity, gas and steam; air conditioning' (OKVED 35) (8.3%). It is also noteworthy that organisations younger than 6 years (in the interval from 1 to 5 years) account for approximately 32%, indicating the emergence of new market participants. Organisations that have crossed the 10-year mark (in the range of 11-17 years) account for 35.3%, indicating a significant number of participants who have been in the industry for a long period of time. The distribution of organisations by age in this industry is relatively even, which may indicate the absence of dominant players and moderate competition.

The following investigation will consider the distribution of organisations providing pollution remediation and other services related to waste disposal (OKVED 39). The largest share of organisations in this industry is 4 years old, accounting for 19.1%. Conversely, organisations that have been in existence for less than six years account for approximately 63.4%, thereby indicating that new market entrants are actively entering the sector, a phenomenon that may be attributed to the growing demand for pollution remediation and waste management services. A notable proportion of organisations that have been in existence for 11 years and over account for 22.7% of the total, suggesting a substantial presence of long-standing entities in the sector. This trend may be indicative of the presence of well-established entities with extensive experience in the field. The industry's age distribution is characterised by a concentration of young and medium-sized organisations. This phenomenon could be indicative of the dynamic nature of the industry, as well as the presence of a limited number of well-established entities.

Within the context of retail trade (OKVED 47), two distinct groups of organisations can be identified: those that are nascent, with an operational lifespan of between one and five years, and those that are established, with an operational lifespan of between 11 years and beyond. The largest share of organisations in this sector is 15 years old, accounting for 10.3%, which indicates the presence of stable and successful organisations operating in this sector for a considerable time. Organisations that are less than six years old (in the range of 1-5 years) account for 38.3%. A substantial proportion of 30.6% is accounted for by organisations that have been in existence for over a decade, specifically those between the ages of 11 and 17 years. The age distribution of organisations in the industry is, therefore, both diverse and reflective of the stability, experience and dynamism of the retail sector as a whole.

In the subsequent phase, the median values of financial ratios for Russian organisations were analysed. These ratios were calculated on the basis of financial statements for 2020, and the results were then contextualised within the framework of the TEAs under consideration. This approach was adopted in order to identify any differences between the various industries (see Table 5).

The liquidity of organisations from different industries was assessed based on the analysis of the following ratios: current liquidity, quick liquidity and absolute liquidity.

The current liquidity ratio was found to take the maximum value for organisations in the field of retail trade, equalling 1.93, whilst the minimum value was found to be characteristic of oil and natural gas extraction, equalling 0.97. Concurrently, the median estimates for all analysed TEAs exhibited homogeneity, as evidenced by the coefficient of variation not exceeding 33%.

Organisations providing pollution remediation and other services related to waste management have the maximum level of quick liquidity (the value is 1.23). The minimum level of this ratio is 0.53, and is characteristic of organisations in the field of tobacco production. The analysed series of values of the quick liquidity ratio is also homogeneous.

The absolute liquidity ratio, in contrast to other indicators within the same group, exhibits significant variation across the specified TEAs. This observation is substantiated by the high value of the coefficient of variation, which stands at 84%. Concurrently, the maximum value is indicative of organisations providing services in the field of elimination of pollution consequences and other services related to waste disposal (0.3), while the minimum value corresponds to tobacco products manufacturers (0.02).

Table 5 – Median values of financial ratios by TEA [Compiled by the author]

Financial ratio	Type of economic activity						Coefficient of variation
	Oil and natural gas production	Mining of metal ores	Production of tobacco products	Provision of electricity, gas and steam; air conditioning	Provision of pollution control and other waste management services	Retail trade, except for trade in motor vehicles and motorcycles	
Liquidity ratios							
Current liquidity ratio	0,97	1,45	1,05	1,06	1,88	1,93	**31%**
Quick liquid ratio	0,75	0,97	0,53	0,95	1,23	0,75	**28%**
Absolute liquidity ratio	0,04	0,13	0,02	0,08	0,30	0,15	**84%**
Financial stability ratios							
Debt ratio	0,56	0,47	0,22	0,22	0,01	0,00	**93%**
Interest coverage ratio	7,07	8,89	14,88	5,23	3,73	1,22	**69%**
Business activity ratios							
Accounts receivable turnover ratio	3,56	6,76	3,66	4,08	3,44	17,58	**85%**
Inventory turnover ratio	16,35	5,95	2,67	26,03	10,15	5,96	**77%**
Asset turnover ratio	0,42	1,25	0,97	1,38	1,63	2,81	**57%**
Accounts Payable Turnover Ratio	2,64	3,91	2,67	3,34	2,56	8,82	**61%**
Profitability ratios							
Net sales profitability ratio	17%	13%	11%	7%	5%	6%	**49%**
Return on equity ratio	18%	53%	88%	26%	44%	36%	**56%**
Return on assets ratio	7%	15%	14%	9%	13%	14%	**28%**

The financial stability of organisations from different sectors of the economy was analysed using two ratios: debt and interest coverage. It was established that both indicators are heterogeneous, a conclusion that is substantiated by the high values of the coefficient of variation, which stood at 93% and 69% respectively. Furthermore, the minimum values of the financial ratios are observed to be characteristic of organisations in the retail trade sector, for which the median of the debt ratio is equal to 0, and the interest coverage ratio is 1.22. Conversely, the maximum median value of the debt ratio is observed among organisations involved in oil and natural gas production. The median value of the latter indicator for oil and natural gas producing organisations is 0.56. Conversely, the maximum value of the interest coverage ratio is observed in organisations producing tobacco products (14.88). It is evident that the financial structures of diverse industry sectors vary considerably, a finding that is substantiated by the heterogeneity of the median values of financial stability ratios by industry type.

The level of business activity by type of activity was assessed using four turnover ratios: accounts receivable, inventories, total assets, and accounts payable.

For all four indicators, the coefficient of variation exceeds 33%, which indicates significant differences in the levels of business activity among the organisations engaged in various TEAs. It is noteworthy that the maximum values for the turnover ratios of accounts receivable, assets and accounts payable are indicative of organisations operating within the retail trade sector. The values for the aforementioned indicators were 17.58, 2.81 and 8.82, respectively.

Conversely, organisations providing pollution remediation services and other services related to waste management exhibited the lowest values for receivables and payables turnover ratios, at 3.44 and 2.56 respectively. The lowest recorded inventory turnover ratio was observed among organisations involved in tobacco product manufacturing (2.67), while the lowest asset turnover ratio was recorded among oil and natural gas extraction companies (0.42).

The profitability of Russian organisations was assessed using three ratios: net return on sales, return on equity and return on assets.

The maximum level of net profitability of sales was observed for organisations in the field of oil and natural gas extraction (17%), while the minimum was recorded for organisations providing services in the field of elimination of pollution consequences and other services related to waste disposal (5%). Concurrently, tobacco product manufacturers demonstrated the maximum value of return on equity, which was 88%, while the minimum value for the same indicator was 18% for oil and natural gas extraction organisations. The latter TEA is distinguished by the minimum value of the return on assets coefficient, which stood at 7%, while the maximum value was recorded in metal ore mining companies at 15%. It is evident that the median values of the coefficients of net return on sales and return on equity vary considerably across different TEAs, a finding corroborated by the substantial values of the coefficient of variation, which stood at 49% and 56%, respectively. It is notable that only the numerical series with median estimates of the return on assets ratio is homogeneous for all TEAs under study.

The subsequent phase of the study will entail an examination of alterations in a specific financial indicator across diverse groups, with a focus on the dynamics experienced during the period 2012-2020.

Figure 7 presents a graph showing the change in the median value of the current liquidity ratio by TEA.

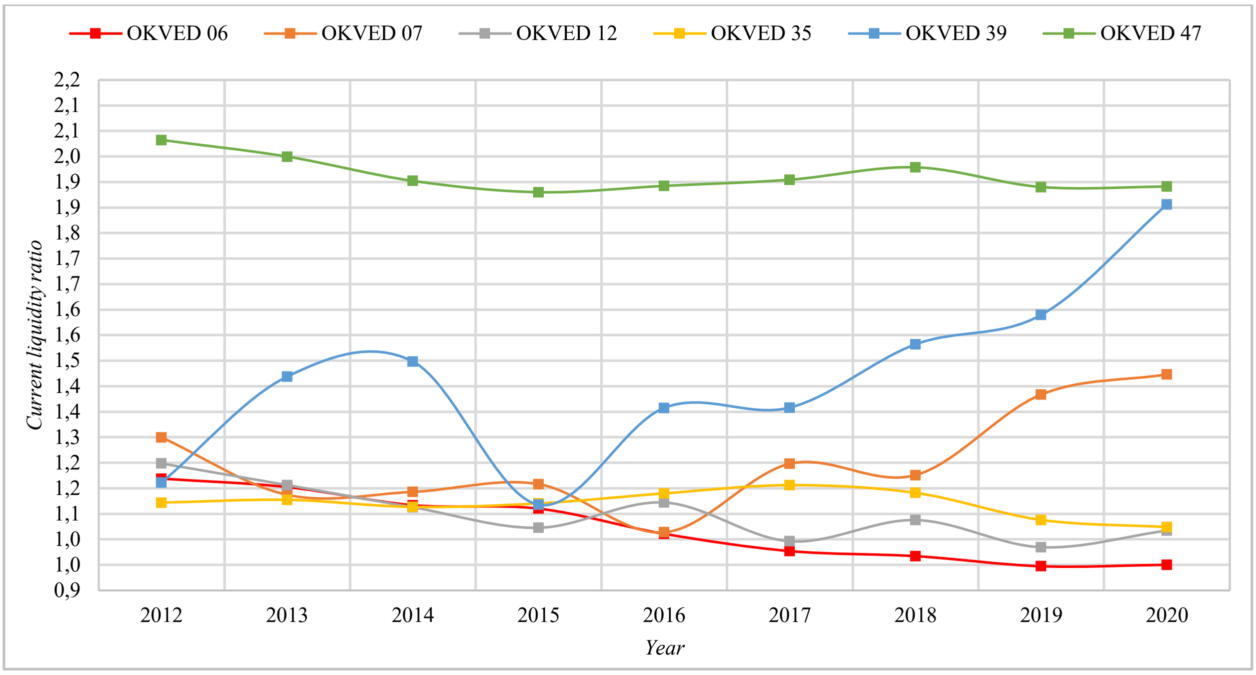


Figure 7 – Dynamics of changes in the median value of the current liquidity ratio by TEA for 2012-2020. [Compiled by the author]

It is evident from Figure 7 that the median value of the current liquidity ratio for the oil and natural gas extraction sector (OKVED 06) reached its maximum in 2012 at 1.18, and its minimum in 2019 at 0.96. The current ratio has been in a state of gradual decline since 2012, with a median decline of approximately 19 per cent from 2012 to 2020, and a time series median of 1.06.

For the metal ores mining sector (OKVED 07), the minimum level of the liquidity ratio was 1.05 in 2016, the maximum was 1.45 in 2020, with the average value for all periods being 1.23. It is noteworthy that, commencing in 2016, the indicator within the industry has undergone rapid growth.

Minor fluctuations in the current liquidity ratio are also evident among tobacco product manufacturers (OKVED 12). The mean value is 1.09, the minimum was 1.01 in 2019, and the maximum was 1.22 in 2012. A comparable pattern of fluctuation in the indicator under scrutiny is discernible in the provision of electricity, gas and steam; air conditioning (OKVED 35), where the minimum level was recorded at 1.06 in 2020, the maximum at 1.17 in 2017, and the mean value for the periods under consideration was determined to be 1.12.

TEA ‘Provision of services in the field of pollution remediation and other services related to waste disposal’ (OKVED 39) is characterised by the maximum level of heterogeneity of the time series containing information on median values of the liquidity ratio: from 2012 to 2014, the indicator increased from 1.17 to 1.48; in 2015, it decreased again to the level of 1.12; starting from 2015, it grew rapidly and in 2020 amounted to 1.88. The average value of the indicator for all periods was 1.44.

The most stable time series among all considered TEAs turned out to be for retail trade (OKVED 47): the coefficient of variation was 2%, with the indicator varying in the range from 1.91 to 2.05.

As demonstrated in Figure 8, a graphical representation is provided which illustrates the dynamics of the median values of the debt-to-equity ratio by TEA.

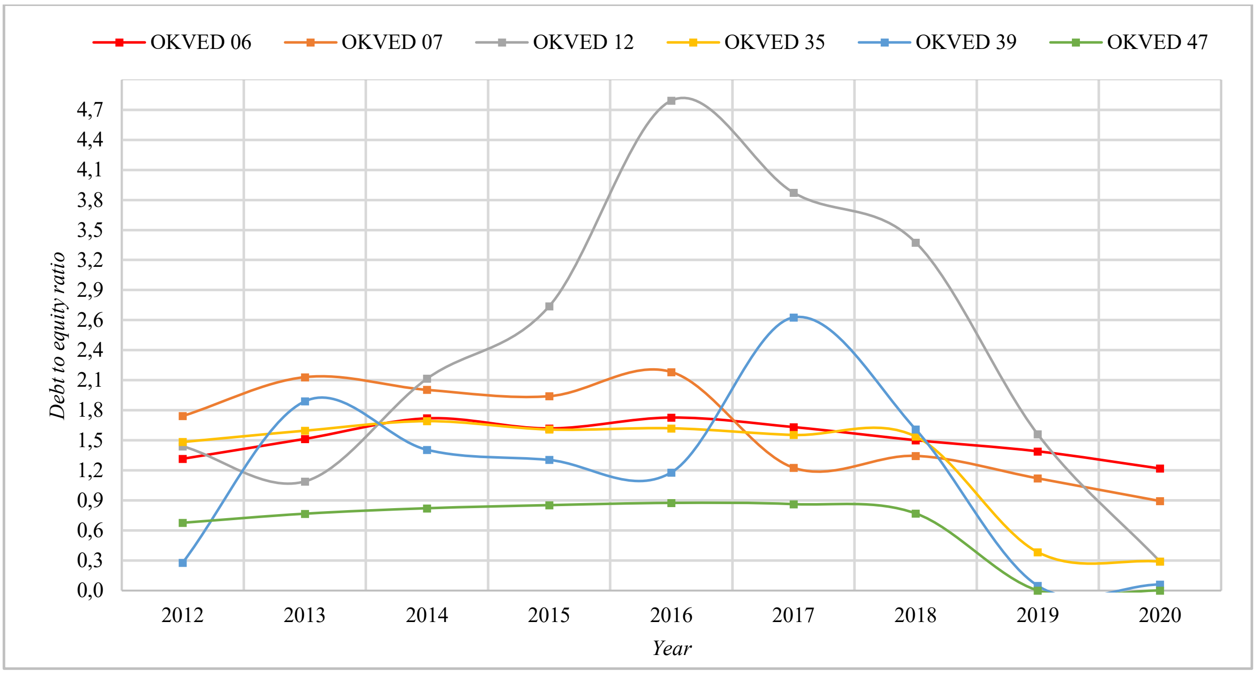


Figure 8 – Dynamics of change in the median value of the debt-to-equity ratio by TEA for 2012-2020 [Compiled by the author]

It is evident from the graph that the time series data are homogeneous for a mere two TEAs: extraction of oil and natural gas (OKVED 06) and extraction of metal ores (OKVED 07). For the first TEA, the minimum value recorded was 1.19 in 2020, the maximum value was 1.69 in 2014 and 2016, and the average for all periods was 1.49. For the second TEA, the minimum value was recorded as 0.88 in 2020, the maximum value was 2.14 in 2016, and the average for all periods was 1.59.

For manufacturers of tobacco products (OKVED 12), a significant change in the dynamics of the indicator is characterised, which can be divided into two intervals: annual growth from 2013 to 2016 (from 1.07 to 4.8); and rapid decline from 2016 to 2020 (from 4.8 to 0.28).

It is also possible to distinguish between the TEA 'Supply of electricity, gas and steam; air conditioning' (OKVED 35) and 'Retail trade, except trade in motor vehicles and motorbikes' (OKVED 47) as a separate group, given that they demonstrate similar dynamics in relation to the indicator under study. For these TEAs, there is an insignificant fluctuation in the coefficient from 2012 to 2018: the range of changes for the first TEA is from 1.45 to 1.66; for the second TEA, from 0.66 to 0.86. In 2019, a precipitous decline in the median value was observed for both activities: for the first TEA, the minimum was recorded in 2020 and amounted to 0.28; for the second TEA, in 2019 with an indicator value of 0.

The change in the median debt-to-equity ratio by year in the pollution remediation and other waste management services sector (OKVED 39) can be divided into four periods: a sharp increase in the ratio from 0.27 to 1.85 in 2012-2013; a decline from 1.85 to 1.16 from 2013 to 2016; a sharp increase to 2.67 in 2017; and a significant decline from 2.67 to 0.06 in 2017-2020.

The dynamics of median values of the receivables turnover ratio for 2012-2020 by TEA is presented in Figure 9.

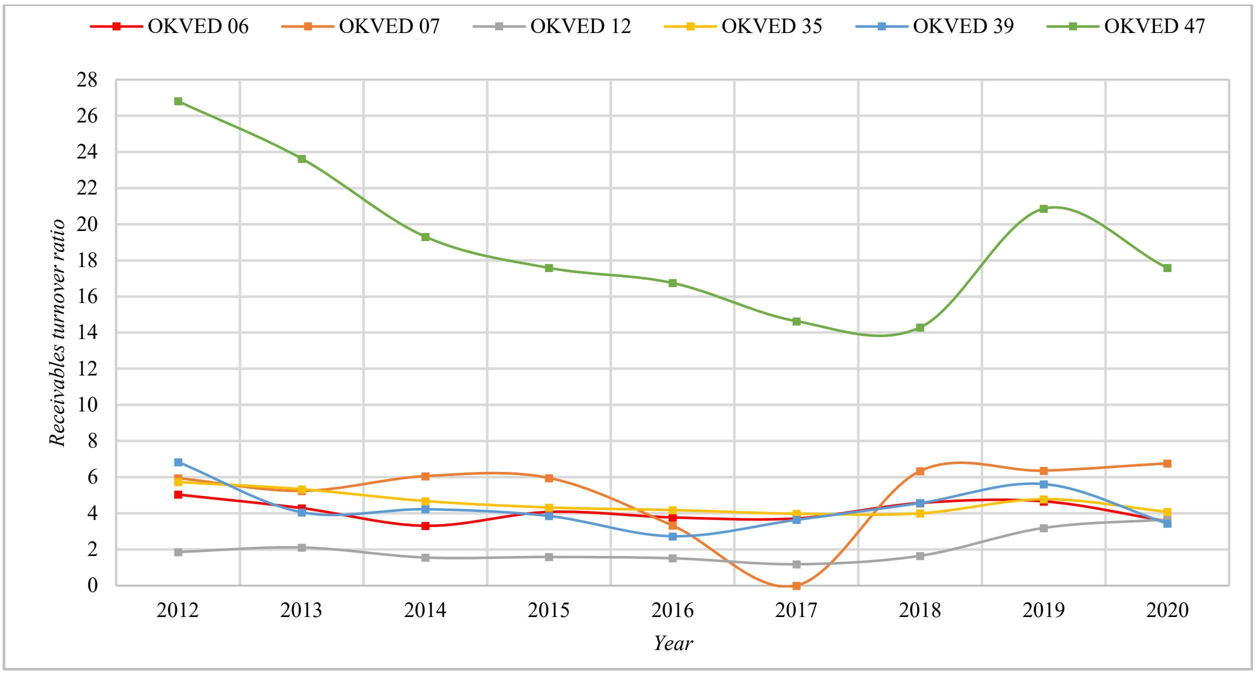


Figure 9 – Dynamics of change in the median value of the receivables turnover ratio by TEA for 2012-2020. [Compiled by the author]

For the activities 'Extraction of oil and natural gas' (OKVED 06), 'Provision of electricity, gas and steam; air conditioning' (OKVED 35) and "Provision of pollution remediation services and other services related to waste disposal" (OKVED 39), the time data do not change significantly throughout the periods under consideration and range from 2.73 to 6.84, with average estimates for the first TEA – 4.11, the second – 4.56, the third – 4.33.

In the sector of metal ores mining (OKVED 07), the time data are heterogeneous, a state of affairs that is confirmed by a high coefficient of variation equal to 42%. This TEA is characterised by a stable value of the coefficient from 2012 to 2015, ranging from 5.24 to 6.06. However, a notable decline was observed in 2016, with a decrease from 5.95 to 3.32. This was followed by an increase in 2017 to 5.95, and subsequently, a period of stability with values ranging from 6.34 to 6.76 from 2018 to 2020.

The receivables turnover ratio for Tobacco Products Manufacturing (OKVED 12) remained consistent from 2012 to 2018, fluctuating between 1.51 and 2.11, with the exception of 2017 when the indicator experienced a notable decline to 1.18 from

1.51. In 2019, the ratio increased to 3.19 from 1.64 and continued to increase to a maximum level of 3.66 in 2020.

The highest level of the receivables turnover ratio throughout all analysed periods was observed in retail trade (OKVED 47), with an average value of 19.05. Despite a decrease in the ratio from 26.81 to 14.29 between 2012 and 2018, there was a sharp increase to 20.86 in 2019 and a decrease again to 17.58 in 2020.

The median values of the net profitability of sales coefficient for 2012-2020 for the TEAs under consideration exhibited variation (Fig. 10).

For instance, within the oil and natural gas extraction sector (OKVED 06), the coefficient values consistently exhibited the highest levels across all periods, ranging from 15% to 21%, in comparison to other TEAs.

In the field of metal ores mining (OKVED 07), with the average value of the coefficient of 10%, the coefficient of variation was 37%, indicating a high level of variability of statistical estimates. During the period 2012-2015, the indicator ranged from 6-8%, in 2016 it increased to its maximum of 16%, and in 2017-2020 it varied between 9-14%.

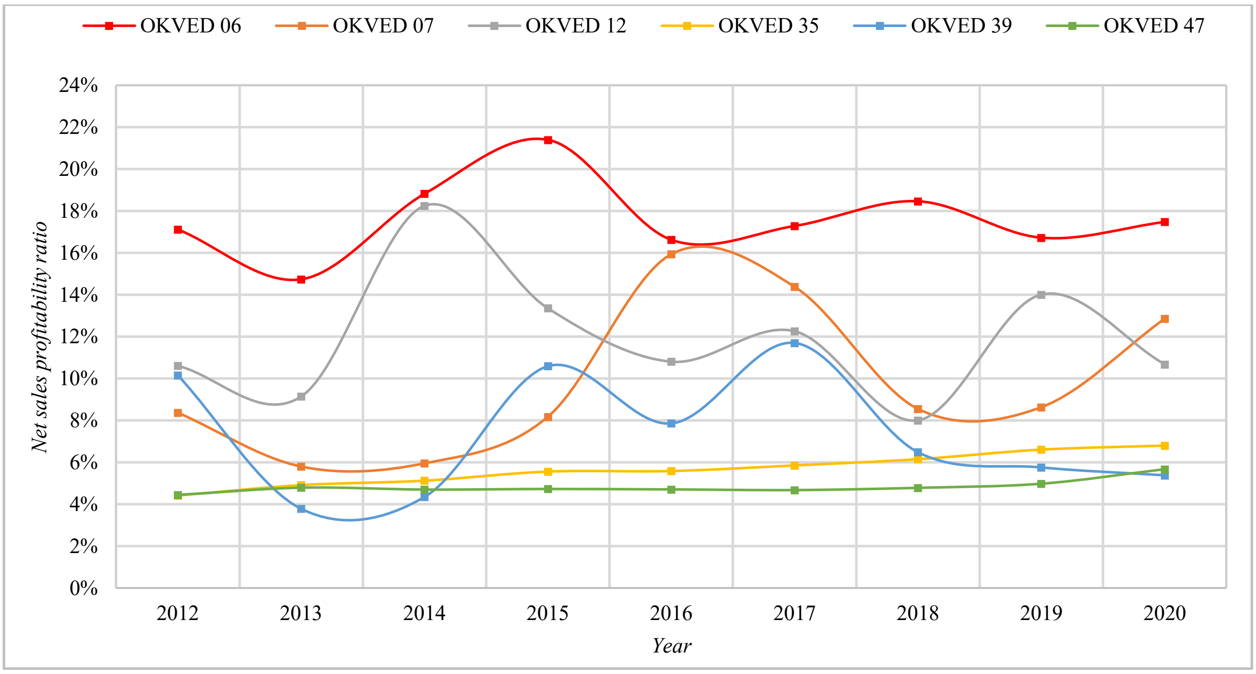


Figure 10 – Dynamics of change in the median value of the net profitability of sales ratio by TEA for 2012-2020 [Compiled by the author]

In the production of tobacco products (OKVED 12), the median value for all periods except 2014 was 8-14%. However, in 2014, there was a marked increase in the coefficient, reaching 18%. The mean value for all periods under consideration is thus 12%.

The dynamics of change in the net profitability of sales ratio for the types of activity 'Provision of electricity, gas and steam; air conditioning' (OKVED 35) and 'Retail trade, except for trade in motor vehicles and motorbikes' (OKVED 47) are similar, with values ranging from 4-7% across all periods under consideration. Concurrently, the mean value for the initial TEA was 6%, while the secondary TEA exhibited a mean value of 5%.

High variability of values is observed in the sector of provision of pollution remediation services and other services related to waste management (OKVED 39). In 2013, the median value decreased to 4 per cent from 10 per cent compared to 2012, in 2014 it remained at the 2013 level, in 2015-2017 it ranged from 8 per cent to 12 per cent, and in 2018-2020 it held at 5-6 per cent.

The analysis confirmed the existence of discrepancies in the performance of organisations engaging in diverse activities by means of a comparative analysis of financial and non-financial indicators. Initially, distinctions in the distribution of organisations by age were identifiable, thereby signifying the heterogeneity in the maturity levels of individual markets, the presence of competitors and new market entrants, and the demand level. Secondly, median values of financial ratios were utilised to confirm disparities in liquidity, financial stability, business activity and profitability between different TEAs. Furthermore, it was determined that the dynamics of median values of financial ratios for the period 2012-2020 exhibited significant variation between TEAs, which may be attributed to the stability levels of various industries due to the influence of both exogenous and endogenous factors.

Therefore, given the disparities between TEAs obtained during the course of the study, the development of conventional financial models should be undertaken within the context of TEAs. Furthermore, in order to ensure the generalisability of models, it is necessary to analyse the temporal stability of individual financial indicators of an industry and then select only those industries for which the models will retain a high predictive power after a significant time period from the moment of their development.

2.3 Justification of the choice of clustering method as an element of machine learning for the development of typical financial models

The term 'machine learning' first appeared in the scientific literature in the middle of the 20th century. Nevertheless, the precise origins and initial utilisation of this concept remain ambiguous, as it evolved progressively and was employed by various researchers in diverse contexts.

The development of the term can be traced back to the following time periods.

The 1950s. In 1959, the American scientist Arthur Samuel was among the first to introduce the term 'machine learning' as a process by which a computer learns and behaves in a way that was not originally programmed to do. The researcher's work focused on the development of a draughts programme capable of enhancing its game strategy through experience [Samuel, 1959].

The 1960s and 1970s witnessed the emergence of machine learning as a distinct scientific discipline. During this period, machine learning began to be developed as an independent scientific direction in its own right. Researchers began to create machine learning algorithms, including k-nearest neighbours and support vectors [Cover, Hart, 1967; Vapnik, Chervonenkis, 1971].

The 1980s and 1990s witnessed the emergence of machine learning as a distinct scientific discipline. During this period, machine learning was used to solve applied problems such as pattern recognition, natural language processing, and medical diagnosis. The following algorithms are of particular significance in this period: the development of backpropagation [Rumelhart et al., 1986] represents a significant milestone in the evolution of neural networks, while the introduction of AdaBoost [Freund & Schapire, 1997] emerged as a pivotal approach in ensemble learning. Additionally, the utilisation of Hidden Markov Models (HMM) [Rabiner, 1989] in speech recognition and natural language processing has been instrumental in the analysis of sequential data.

The 2000s and subsequent decades. Machine learning has become increasingly common and important in today's world due to the increase in computing power and availability of big data. This period has also seen the emergence of novel methodologies, including deep learning, which have profoundly impacted the field of machine learning. Significant advancements in the field of deep learning include the use of Convolutional Neural Networks (CNNs) for image classification [Krizhevsky, Sutskever, Hinton, 2012], and the development of models capable of playing video games using reinforcement learning techniques [Mnih, et al., 2013); development of neural network architecture for machine translation and natural language processing tasks, which has become an important method in the field of sequential data processing [Sutskever, Vinyals, Le, 2014]; development of ResNet deep neural network architecture, which has become a key element in computer vision [He, Zhang, Ren, Sun, 2016]; development of Transformer architecture, which introduced an attention mechanism and became a standard for sequential data processing, including machine translation and natural language processing [Vaswani, et al., 2017].

Advances in machine learning are also actively used to solve scientific and applied problems in economics and finance. The introduction of machine learning algorithms has made it possible to effectively model complex financial phenomena, forecast market changes and optimise investment strategies. The main directions and results of using machine learning algorithms in economics and finance include:

1. Forecasting financial time series.

Machine learning techniques have been identified as having significant potential in the field of finance, particularly with regard to the prediction of time series, including stock prices, exchange rates and interest rates. Machine learning algorithms, such as recurrent neural networks (RNNs) and convolutional neural networks (CNNs), have the capacity to analyse multiple factors, including historical data, news events and social media, in order to make more accurate predictions [Sezer, Ozbayoglu, 2018]. Specifically, recurrent neural network (RNN) models have the capacity to discern dependencies in sequential data, utilising these insights to forecast future stock prices [Selvin S. et al., 2017].

2. Credit scoring.

It is evident that financial institutions have successfully implemented machine learning algorithms to assess customer creditworthiness and manage risk. Machine learning models, including gradient boosting and random forest algorithms, have demonstrated proficiency in analysing voluminous data sets and identifying customers with a high propensity for bankruptcy [Karminsky, Burekhin, 2019; Liu, Fan, Xia, 2022].

3. Portfolio Optimisation.

Portfolio optimisation is one of the key challenges for investors. Machine learning algorithms can determine the optimal mix of assets by considering various factors, including expected return and risk, subject to constraints. These methods allow creating more efficient investment strategies [Chen, et al., 2021].

4. Text data analysis.

The utilisation of natural language processing (NLP) algorithms for the analysis of textual data has emerged as a pivotal instrument within the domains of accounting, auditing and financial analysis. Models of this class have been shown to facilitate the analysis of news articles, financial reports and social media, with a view to identifying trends and market sentiment [Fisher, et al., 2016].

5. Forecasting macroeconomic indicators

Machine learning has also found application in forecasting economic indicators such as GDP, inflation and unemployment. Modelling and forecasting of economic phenomena based on the analysis of large amounts of macroeconomic data is used in solving applied tasks for government agencies and various financial institutions [Coulombe, et al., 2021].

Machine learning is a process whereby data is input into a system and algorithms are applied to process it. This results in a model being trained, which is then used to predict values within acceptable ranges. Subsequent to the arrival of new data, the model undergoes training, incorporating new dependencies, thereby enhancing prediction accuracy and versatility.

In this study, individual machine learning algorithms were considered, which can be divided into two groups: supervised learning and unsupervised learning.

Supervised learning

Supervised learning involves the use of labelled training data to construct a model and its subsequent application to predict new, previously unseen data. The fundamental property of labelled data is that its input and output values are known in advance. In this scenario, the operator is privy to the correct solution for the input parameters, while the algorithm identifies internal natural dependencies in the data, learns from the marked-up data, and makes predictions. In the event that the outcomes of these predictions fail to meet the operator's expectations, adjustments are made to the algorithm's characteristics, and the model is retrained. This iterative process persists until the operator receives high scores on quality metrics following model testing.

The classification of algorithms for learning with a teacher can be divided into two groups, depending on the type of tasks: algorithms for solving the regression problem and algorithms for solving the classification problem. The former is employed to predict a real variable, while the latter is used to predict a discrete response, according to which it is determined to which class or category an object belongs. Consequently, the BFMs described in the first chapter belong to the group of machine learning algorithms with a teacher for solving the classification problem. This is because, depending on the value of the composite bankruptcy ratio for a particular organisation, one of two classes (for most models) is assigned: the area of financial insolvency or the area of financial stability.

There are several types of classification.

1. Binary classification. In this classification, the model divides the data into two classes.
2. Multi-class classification is another type of classification in which the model divides the data into multiple classes. In this scenario, the model partitions the data into multiple classes, exceeding two categories. Each observation is assigned to one of several classes.

We will consider one of the popular models of binary classification – logistic regression.

Logistic regression is an algorithm employed for binary classification, the purpose of which is to predict the probability of an observation belonging to one of two classes. Fundamentally, the model incorporates a sigmoid function that is provided with a linear classifier. This method is capable of not only making class predictions, but also estimating the confidence of the model (probability) in these predictions. The general form of the model can be represented as the following formula:

$$y = sigmoid(z) = \frac{1}{(1 + e^{-z})}, \tag{15}$$

$$z = w_0 + w_1 * x_k + \cdots + w_k * x_k, \tag{16}$$

where y is probability that the observation belongs to the target class, z is linear classifier, $x_1, \ldots, x_л$–independent variables (attributes), $w_0, \ldots, w_k$ – model coefficients, which are tuned during training.

In order to ascertain the most efficacious logistic regression model, it is imperative to solve the problem of finding the minimum of the loss function, which is known as cross-entropy.

$$logloss = -\frac{1}{N} * \sum_{i=1}^{N} \left(y_i * log\left(\hat{y_i}\right) + (1 - y_i) * log\left(1 - \hat{y_i}\right)\right), \tag{17}$$

where N – number of observations, y_i – фactual value of the target variable for the i-th observation, $\hat{y}_i$ – predicted value of the target variable for the i-th observation.

In this study, we explore alternative approaches to the classification problem that do not fall under the umbrella of linear models.

Ensemble models (ensembles) are machine learning methods with a teacher that combine multiple underlying models together to improve generalisability and predictive quality. Ensembles use multiple models to improve prediction quality, stability and generalisability. These models find application in both classification and regression tasks.

The Random Forest algorithm, which is based on creating a composition of a large number of decision trees, will be used here to provide a more detailed examination of ensemble models.

The main features of the method under consideration include:

1. The Random Forest combines weak baseline algorithms, which themselves may not be very accurate as they are prone to overfitting and are highly sample dependent. However, the Random Forest methodology is predicated on the construction of each tree independently. This has the effect of creating variance in predictions, and, as a result, enhancing prediction quality and generalisability through the averaging of predictions from all trees in the composition.

2. The Random Forest algorithm functions in a black box capacity, providing no interpretable explanations regarding the manner in which the data is organised or the dependencies it has discovered. Researchers utilising this algorithm primarily focus on the accuracy of the predictions made by these models.

3. The 'Random Forest algorithm has been demonstrated to resolve issues pertaining to both regression (the process of constructing regression decision trees to predict a continuous target variable) and classification (the process of constructing classification decision trees to predict a discrete target variable).

In the process of constructing learning models with the guidance of a supervisor, there is a significant risk of encountering the issue of overtraining. This is characterised by the model attaining elevated scores on quality indicators during the training phase, yet subsequently achieving low scores when evaluated on a separate test sample. This phenomenon can be attributed to the operator's inclination to calibrate the model to ensure that during the training process, the algorithm incorporates all dependencies, including those that may be considered to be noise dependencies, so that when the model is used on the test sample, the trained model may not be able to detect real (natural) dependencies in the data.

The problem of overtraining is also relevant for models in economics and finance, as it can lead to significant socio-economic consequences when applied in practice.

In this section, we will consider the difficulties faced by economic researchers in developing and using machine learning models that are related to overlearning.

1. Data volatility. In the fields of economics and finance, models are frequently employed to analyse and forecast time series, which can be characterised as noisy and contain numerous anomalies due to unforeseen events associated with economic crises and recessions, political decisions, natural disasters, technological changes, and changes in consumer behaviour. This renders them particularly susceptible to overfitting.

2. Insufficient training sample size. The availability of economic and financial data is often limited or non-existent. Consequently, the development of models on a limited sample can increase the risk of overfitting, particularly when complex algorithms are employed.

3. Model complexity. In the pursuit of enhancing model quality, researchers may employ an excessive number of attributes influencing the dependent variable to obtain estimates that approximate the values from the test sample. This can result in erroneous predictions when utilising novel data not encountered by the model.

Unsupervised learning

Machine learning algorithms that are unsupervised are capable of identifying patterns in data without the need for operator input. These algorithms are capable of processing large data sets independently and deriving conclusions from them. The unsupervised machine learning model attempts to organise unstructured data in various ways, either by ordering it by some criterion or by class. Subsequent to the arrival of new data, the model is re-trained and learns new patterns, thereby increasing the accuracy of decisions.

The classification of unsupervised learning algorithms is predicated on the particularities of the problems to be solved, and can be divided into three groups: clustering, dimensionality reduction and recommender systems.

In the present study, the focus was exclusively on clustering algorithms.

Clustering algorithms solve the following problems:

– automatic search for similar objects;

– identification of anomalies – isolated objects;

– more detailed analysis of clusters, which involves building models for each cluster rather than a general model for all objects.

Groups of clustering methods include:

1. Clustering based on prototypes (Prototype-based methods). Prototype-based clustering methods allow to obtain a strict division of objects into clusters. In this case, each cluster is characterised by a basic element, for example, the k-means method corresponds to the centre of mass (centroid) of objects from this cluster.

2. Hierarchical clustering (Hierarchical methods). Hierarchical clustering methods allow to obtain a hierarchy of clusters in two ways: by means of agglomerative and divisive algorithms. The agglomerative algorithm considers each object as a separate cluster at the first step, then combines the two closest clusters at each step and stops when the only cluster remains. Divisive algorithm at the first step considers that all objects belong to the same cluster, then at each step splits one of the clusters into two parts and stops when all clusters consist of a single object.

3. Density-based clustering (Density-based methods). The employment of density-based clustering methods facilitates the identification of clusters of arbitrary shape. A cluster is defined as a region characterised by a high density of objects. The application of such algorithms is in the identification of isolated objects that are outliers in statistical analysis.

4. Probabilistic model-based methods (Probabilistic clustering). This approach to data clustering involves the application of probabilistic models and the assumption that the data are generated using probabilistic processes. The objective of this class of models is to identify hidden clusters by maximising the likelihood of the data.

5. Grid clustering (Grid-based methods). A data clustering method that divides the data space into a grid of cells and then aggregates the objects within a cell into clusters.

6. Spectral-based clustering (Spectral-based methods). This method of data clustering is based on the analysis of the spectrum (eigenvalues and eigenvectors) of the similarity matrix between data points. This method identifies cluster structures in the data based on the properties of the eigenvectors of the similarity matrix.

In this paper, the k-means method was used to identify typical financial models.

The following section will consider the peculiarities of the k-means clustering algorithm.

Problem statement

– let there be a set of objects X = $\{x_1, x_2, ...x_i,..., x_n\}$,

– each object x_i has its own set of characteristics (in the case of balance sheet BFMs – values of financial indicators of accounting statements): x_i= $\{x_{i1}, x_{i2}, ...x_{ij},..., x_{im}\}$,;

– to determine the measure of similarity of objects x_i and $x_i^{'}$ Euclidean distance was used, calculated by the formula:

$$L(x_i, x_i') = \sqrt{\sum_{j=1}^{m}\left(x_{ij} - x_{ij}'\right)^2}, \tag{18}$$

– each cluster is characterised by a basic element: $C_k \leftrightarrow$ centroid μ_k. Centroid Calculation Formula:

$$\mu_k = \frac{\sum_{x_i \in C_k} x_i}{|C_k|}, \tag{19}$$

where C_k – k-th cluster; x_i – i-th object belonging to the k-th cluster; μ_k – centroid of the k-th cluster;

– object x_i belongs to the cluster C_k if and only if the distance from the centroid of this cluster to the object under consideration $L(\mu_k, x_i)$ is the smallest among all distances between the object and centroids of all clusters;

– the objects will be clustered in such a way as to minimise the function L(C).

$$L(C) = \sum_{j=1}^{k} \sum_{x_i \in C_j} \left\| x_i - \mu_j \right\|^2 \tag{20}$$

Algorithm sequence

Let the objects X and the number of clusters k be given to the input of the system. After which the steps are performed sequentially:

1. Initialisation of centroids: $\mu_1, \mu_2, \dots, \mu_k$.
2. Cluster update: objects are assigned to the nearest centroid.
3. Update centroids: recalculate the position of centroids using the formula:

$$\mu_j = \frac{\sum_{x_i \in C_j} x_i}{|C_j|} \tag{21}$$

Actions 2 and 3 continue until the stopping rule is triggered or the centroids stay in the same place.

To determine the number of clusters, we used the elbow method, which involves calculating the value of the criterion for different k values:

$$L^{(k)}(C) = \sum_{j=1}^{k} \sum_{x_i \in C_j} \left\| x_i - \mu_j \right\|^2 \tag{22}$$

Before running the k-means algorithm, it is necessary to carry out the data standardisation procedure: centring and normalisation.

In the following section, we proceed to examine a number of specific instances wherein the aforementioned machine learning algorithms are employed to solve problems in the area of financial insolvency assessment, including bankruptcy prediction.

Thus, in foreign and domestic studies devoted to the development of balance sheet BFMs, classification algorithms, including linear and ensemble models, are used to solve this class of problems (see Chapter 1). The peculiarity of these models is the use of pre-labelled data - scientists initially know whether each organisation in the sample belongs to one of two groups - financial insolvency or financial stability.

It is important to note that the testing of the predictive ability of modern domestic linear models yielded low scores according to the author's quality metric (see paragraph 1.3). The findings suggest that the overfitting problem is pertinent to these models, as classification accuracy was high on the training and test samples, yet the predictive ability of BFMs was revealed to be low when utilising the new data generated within the scope of this study. The general reasons for model overfitting noted in this paragraph also apply to domestic linear BFMs.

1. Non-stationary time series with values of individual financial ratios considered for 6 TEAs for 2012-2020 in paragraph 2.2;

2. A lack of training sample size has been attributed to the utilisation of specialised RAS, which are subject to limitations when importing data (see paragraph 1.2). These limitations are governed by the stipulated terms and conditions of the user agreement;

3. Complexity of individual models that include many different financial indicators used as explanatory variables. The use of a large number of attributes improves the predictive accuracy of the model on the training data. This is because the model tends to account for all dependencies, including spurious ones, which subsequently leads to erroneous predictions when new data are used. In addition, some models use as independent variables logarithms of financial indicators, the values of which are negative for the majority of financially insolvent organisations (see paragraph 1.2).

The existence of non-linear dependence of the composite bankruptcy ratio on individual financial indicators has been demonstrated in several research papers. These papers also justify the use of ensemble algorithms to create author's BFMs. Despite the enhanced predictive capability of this class of models, they are afflicted by the same retraining issues as linear models. Furthermore, this class of models represents a set of decision trees whose scores are averaged to obtain the final model score. This characteristic makes it impossible for other stakeholders to use BFMs, for whom the model is perceived as a black box that does not provide interpretable explanations about the structure of the data and the relationships between them.

Another feature of both linear and ensemble balance sheet BFMs is the use of individual financial ratios derived from accounting statements as explanatory variables, which leads to a number of limitations. Firstly, the coefficient analysis involves working with separate elements of accounting statements, which does not allow for a comprehensive assessment of the financial position of the organisation. Secondly, the purpose of classification algorithms does not imply the allocation of typical financial models in the form of averaged accounting statements, as the solution of the latter problem is provided through the use of clustering algorithms.

In light of the identified limitations of BFMs developed using classification algorithms, this study proposes a fundamentally different approach to assessing financial insolvency. This approach is predicated on the assumption of the existence of several typical financial models for each industry and the identification of such a typical financial model that is typical of the majority of financially insolvent organisations. The proposed machine learning method is to be implemented in accordance with the following sequence of actions:

1. The subsequent stage of the process involves the identification of similar objects so that they can be grouped into clusters. The basic element, which is typical financial statements, is then assigned to the cluster.

2. The allocation of all financially insolvent entities into clusters is based on maximum similarity to the typical financial models identified in step 1.

3. The identification of clusters with a high frequency of financially insolvent organisations in step 2 will be taken as an indication of a high risk of financial insolvency for any other organisation in the study that belongs to this cluster.

The implementation of the aforementioned actions is possible using the k-means clustering algorithm and the Euclidean distance to determine the measure of similarity of objects. The outcome of the clustering process is the identification of a centroid as the basic element of the cluster [Kovalev et al., 2022].

The proposed approach facilitates the utilisation of unit values from the balance sheet sections as variables. This is also referred to as the financial model of the organisation. The centroid of the individual cluster is thus representative of the n-dimensional estimate of the typical financial model. The study presents typical financial models using five measures of balance sheet sections: current and non-current assets, equity, long-term and short-term liabilities. It is also important to note that one of the advantages of the k-means clustering method is its scalability. This means that in the future the number of measures can be increased by detailing the sections of the balance sheet to individual items (indicators).

Another advantage of the k-means algorithm with predetermined centroids (typical financial models) is its ease of use, since an interested third-party user, in order to determine whether the organisation under study belongs to one of the clusters, will need to provide an n-dimensional estimate of the organisation's financial model and calculate the distance from the organisation's financial model to each typical financial model. The financial model that yields the minimum distance from the n-dimensional estimate of the organisation's financial model is designated as the cluster number.

The issue of non-stationarity in financial indicators can be addressed by analysing the change in the coordinates of comparable centroids. In instances where deviations are deemed insignificant, the utilisation of centroids for further analysis, including the assessment of financial insolvency, becomes a viable option. Significant deviations require the identification and systematisation of economic and other factors influencing such shifts. However, the present study does not address this latter task.

Chapter conclusions

The features of open government data (data sources) as well as related statistical codes are systematised and described. A two-level grouping scheme has been developed for data sources and statistical codes.

In order to combine data from different sources to form a consolidated database, the mechanisms of their interfacing were proposed. These mechanisms included identification codes, interfacing keys developed by the Ministry of Economic Development of the Russian Federation, and comparable codes from different classifiers. The formation of data slices from the consolidated database enabled the resolution of several complex applied tasks, including the establishment of typical financial models.

The findings of the comparative analysis of the TEA by age distribution of organisations and median values of individual financial ratios for 2020 substantiated the necessity for sectoral segmentation of organisations in order to evaluate financial insolvency. Furthermore, an analysis of the dynamics of the median values of financial ratios employed as predictors of BFMs for the period 2012-2020 revealed the heterogeneity of the obtained estimates. This finding serves to confirm the absence of temporal stability in BFMs based on financial ratios, which are subject to a high level of fluctuations during the analysed time interval.

The utilisation of the k-means clustering algorithm to establish typical financial models, which are the centroids of the selected clusters, is justified. Consequently, each centroid signifies an n-dimensional estimate of a prototypical financial model. The scalability of the k-means algorithm is advantageous in this context, as it allows for the increase in the dimensionality (number of attributes) of estimates to the level of detailed balance sheet items.

CHAPTER 3. DEVELOPMENT OF TYPICAL FINANCIAL MODELS AND THEIR APPLICATION TO THE ASSESSMENT OF FINANCIAL INSOLVENCY AND BEYOND

3.1 Formation of typical financial models based on the use of clustering algorithms

The identified limitations of the coefficient analysis based on the use of individual fragments of accounting statements; the difference in the sets of explanatory variables presented as financial ratios in the BFMs considered; the limited sample size for which the BFMs were developed; the high level of fluctuations in the industry median values of financial ratios, which explains the temporary volatility of the BFMs - these are the key factors that had a negative impact on the predictive power of the existing balance sheet BFMs developed based on supervised machine learning algorithms to solve the classification task.

In order to surmount the identified issues, it is proposed that typical financial models be formed in the context of TEA by means of the k-means clustering algorithm. This will enable the determination of the centroid (i.e. the typical financial model) for each cluster, the coordinates of which will reflect the n-dimensional average assessment of the balance sheet. In order to ascertain the temporal stability of a typical financial model, the level of change in the values of its indicators over several periods was assessed.

The following stages are realised within the framework of forming typical financial models:

1. Collection and processing of initial data presented in the form of a large array of accounting statements data.

2. Clustering of organisations according to the balance sheet data. Determination of centroid coordinates and interval values of specific indicators of the balance sheet for all allocated clusters by TEA.

3. Determination of financial models stable over time in the context of TEA.

Stage 1: Collection and processing of source data

The array with the accounting statements of Russian organisations for 2018 in the form of a csv file presented on the official website of Rosstat was considered as the initial data [45] (see Fig. 3). The size of the initial sample exceeds 2 million organisations.

TIN was used as the identification attribute of the organisation, and the first 4 digits of the MTA code from the OKVED were used as the industry affiliation (see Figure 2).

For automated processing of a large array of accounting statements data, codes of items established by statistical authorities in the form of five-digit numerical values, which correlate with the names of reporting items, were used[46]. The code format is as follows:

$$FSNNP,$$

where F – report form number, S – report section number, NN – item number of the reporting section, P – reporting period (3 – current period, 4 – previous period).

The data processing procedure can be delineated in five successive stages:

1. Initially, a selection of the array of data of accounting statements of organisations is made from the initial sample according to the following criteria: the total assets (Balance Currency) is equal to the total liabilities (Balance Currency); non-current assets (NCA), current assets (CA), equity capital (EC), long-term liabilities (LTL) and short-term liabilities (STL) are not less than zero. The number of observations was reduced to 1.2 million.

2. Selection of active organisations by the criterion - revenue (R) and balance currency (BC) greater than zero. The array of data of organisations' accounting statements decreased to 1.1 million.

---

[45] Accounting (financial) statements of enterprises and organisations for 2018. URL: https://rosstat.gov.ru/opendata/7708234640-7708234640bdboo2018

[46] Order of the Ministry of Finance of Russia dated 02.07.2010 N 66n (as amended on 19.04.2019) "On the forms of financial statements of organisations" (Registered in the Ministry of Justice of Russia on 02.08.2010 N 18023) (as amended and supplemented, entered into force with the reporting for 2020). URL: https://www.consultant.ru/document/cons_doc_LAW_103394/b990bf4a13bd23fda86e0bba50c462a174c0d123/

3. Division of the aggregate set of organisations formed in item 2 into industry subsets. 625 industry sets with accounting statements of organisations are obtained.

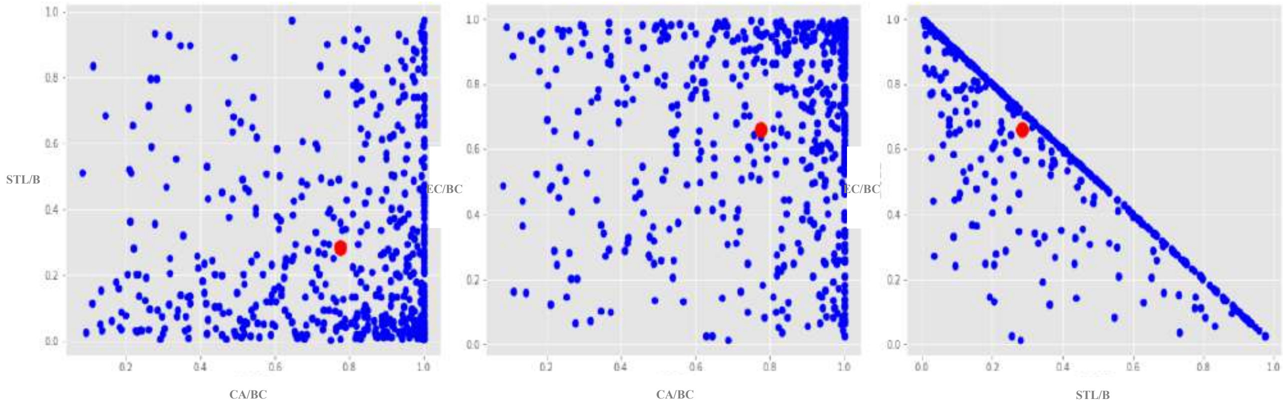
4. Determination of a statistically justified sample size and exclusion from further analysis of industry sets that do not meet the calculated criterion. As a result, 244 industry sets of accounting statements were obtained, the aggregate number of which totalled 800 thousand.

5. The financial model of each organisation is to be assessed by means of a five-dimensional calculation, formed from industry sets. This calculation takes the form of a system of specific values of balance sheet sections in the balance sheet currency:

– NCA / BC – ratio of non-current assets to total assets;
– CA / BC – ratio of current assets to total assets;
– STL / BC – ratio of current liabilities to total liabilities;
– LTL / BC – ratio of long-term liabilities to total liabilities;
– EC / BC – ratio of equity to total liabilities.

Stage 2: Clustering of organisations according to balance sheet data

Utilising the TEA 'Marine Fishing' (OKVED 03.11) as a case study, the following scatter diagram is employed to analyse the distribution of organisations in relation to the prepared estimates of individual indicators of the balance sheet: CA / BC and STL / BC (Fig. 11a); CA / BC and EC / BC (Fig. 11b); STL / BC and EC / BC (Fig. 11c). The utilisation of blue markers is intended to denote the organisations under study, whilst red markers are employed to indicate average estimates of specific balance sheet indicators before clustering.



(a) (b) (c)

Figure 11 – Scatter diagram of organisations on the example of ‘Marine Fishing’ (OKVED 03.11) in the coordinate system: a) CA/BC and STL/BC; b) CA/BC and EC/ BC; c) STL/ BC and EC/ BC [Compiled by the author].

The analysis of the presented image revealed that the data are characterised by significant heterogeneity. Consequently, the utilisation of average estimates is inappropriate for solving applied problems, as they do not carry sufficient value.

Let us apply the k-means clustering algorithm to divide organisations into groups in such a way that elements within a group are similar, while elements from different groups differ. In this case, the coordinates of the centroid of an individual cluster will reflect the average estimates of the indicators of the typical financial model, and the maximum distance between the elements of this cluster - the permissible intervals within which the estimates of the indicators of the typical financial model can vary.

Prior to the implementation of the aforementioned algorithm, the data underwent standardisation, which entailed centring and normalisation.

Two hypotheses were postulated and evaluated during the course of the research:

1. The centroids of clusters formed on the basis of indicators reflecting the scale of activity, namely revenue and balance sheet currency, determine the values of indicators of typical financial models.

2. The selection of clusters for a particular TEA is influenced solely by the specific values of the balance sheet sections, and not by the scale of activity.

Hypothesis 1. The first hypothesis posits that, as a consequence of the clustering of organisations by scale of activity, the centroids of the clusters will reflect typical financial models.

As demonstrated in Figure 12, the outcome of the clustering of organisations with MTA 'Marine Fishing' (OKVED 03.11) is shown according to the values of revenue and balance currency. Green markers denote small organisations, purple markers indicate medium-sized organisations, yellow markers represent large organisations, and red markers indicate the centroids of the listed clusters.

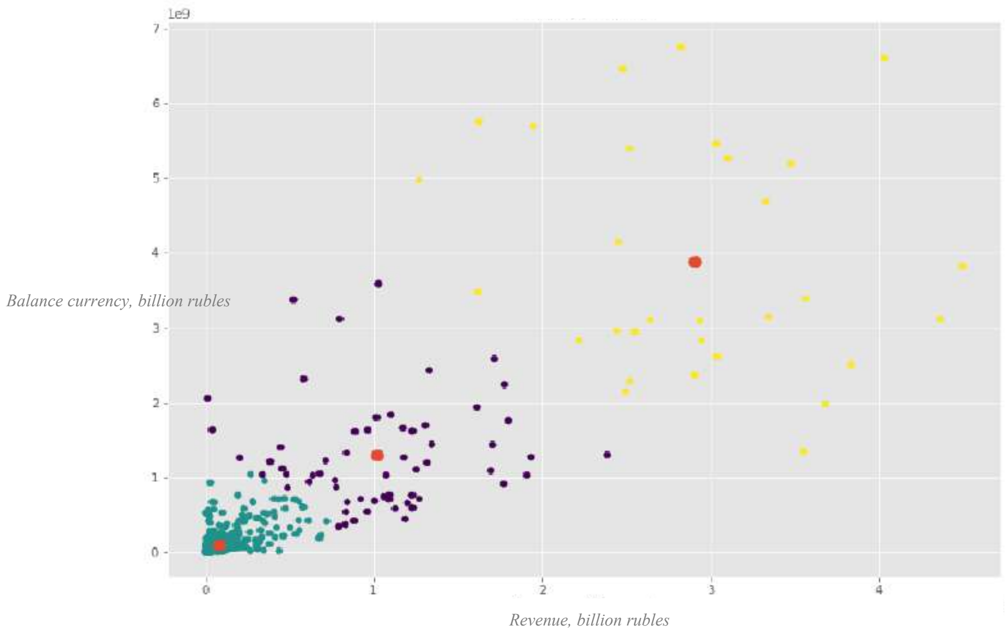
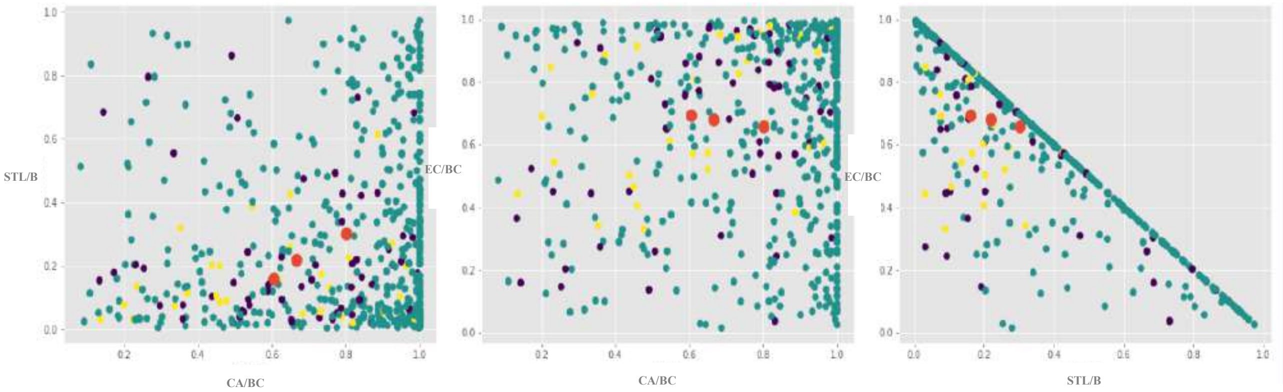


Figure 12 – Clustering of organisations with MTA ‘Marine Fishing’ (OKVED 03.11) by scale [Compiled by the author]

The subsequent step involves the application of the obtained colour labels to the organisations, with the aim of reflecting their belonging to one or another cluster. This is followed by a consideration of the distribution of data in the system of specific values of the balance sheet sections (Fig. 13).

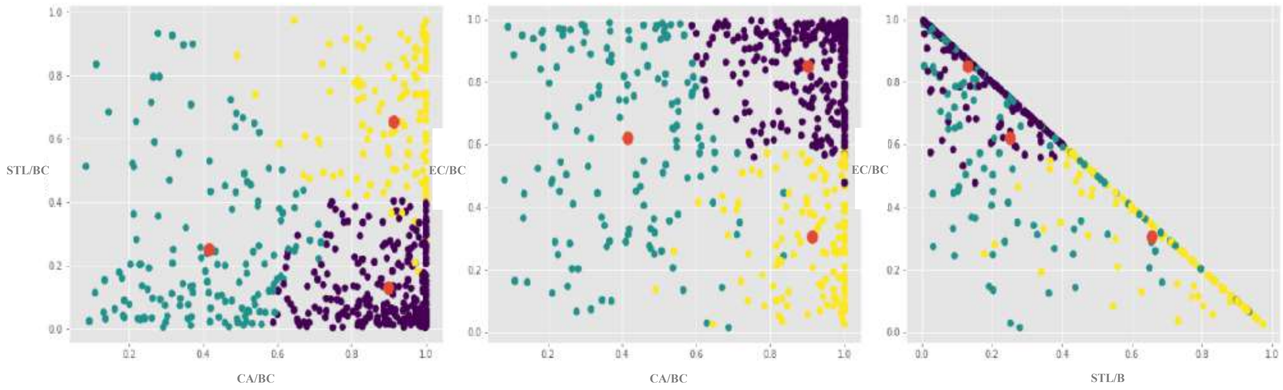


(a) (b) (c)

Figure 13 – Scatter diagram of organisations after clustering by scale of activity on the example of TEA ‘Marine Fishing’ (OKVED 03.11) in the coordinate system: a) CA/BC and STL/BC; b) CA/BC and EC/BC; c) STL/BC and EC/BC [Compiled by the author].

Following a thorough analysis of Figure 13, it is evident that organisations belonging to different clusters, reflecting the scale of activity, are mixed in the system of indicators of the balance sheet sections, which does not allow for the formation of different typical financial models, the values of which would be clearly different depending on the scale of activity. Therefore, the division of organisations by the scale of activity does not affect the formation of a system of different typical financial models in the context of TEA.

Hypothesis 2. Let us divide all organisations (without segmentation by scale) with MTA ‘Marine Fishing’ (OKVED 03.11) into clusters by specific values of balance sheet sections. The obtained results are presented in Figure 14.



(a) (b) (c)

Figure 14 – Scatter diagram of organisations after clustering by specific values of balance sheet sections on the example of TEA ‘Marine Fishing’ (OKVED 03.11) in the coordinate system: a) CA/BC and STL/BC; b) CA/BC and EC/BC; c) STL/BC and EC/BC [Compiled by the author].

Following the implementation of the k-means clustering algorithm, the data is segmented into three distinct clusters. For each of these clusters, a centroid is then determined. The coordinates of the centroids of the selected clusters are proposed for use as n-dimensional estimates of typical financial models, and the maximum distances between the objects of one cluster are suggested as intervals reflecting the permissible change in the values of indicators of a typical financial model. The results of the clustering process for the TEA 'Marine Fishing' (OKVED 03.11) are presented in Table 6.

Table 6 – Values of indicators of typical financial models for TEA ‘Marine Fishing’ (OKVED 03.11) [Compiled by the author].

Group of indicators	**Indicator**	**Cluster name**		
		Cluster 0	**Cluster 1**	**Cluster 2**
Average values	*NCA / BC*	58%	10%	9%
	CA / BC	42%	90%	91%
	EC / BC	62%	85%	30%
	LTL / BC	13%	2%	4%
	EC / BC	25%	13%	66%

Continuation of table 6

Group of indicators	Indicator	Cluster name		
		Cluster 0	**Cluster 1**	**Cluster 2**
Interval values	*NCA / BC*	29-88%	0-37%	0-35%
	CA / BC	12-71%	63-100%	65-100%
	EC / BC	9-98%	59-100%	4-57%
	LTL / BC	0-66%	0-20%	0-36%
	EC / BC	1-86%	1-39%	34-95%
Share of organisations		**27%**	**50%**	**23%**

In light of the results obtained, the following characteristics can be identified for each cluster within the considered TEA ' Marine Fishing' (OKVED 03.11):

– cluster 0 (green markers) - includes organisations with the share of non-current assets (58%) slightly exceeding the share of current assets (42%) in the balance sheet currency, with equity (62%) being the main source of asset financing, followed by short-term and long-term liabilities - 25% and 13%, respectively. The share of organisations with this typical financial model is 27%;

– cluster 1 (purple markers) - includes organisations whose main share of property is current assets (90%), with equity capital being the main source of asset financing (85%). The share of organisations with this typical financial model is 50%;

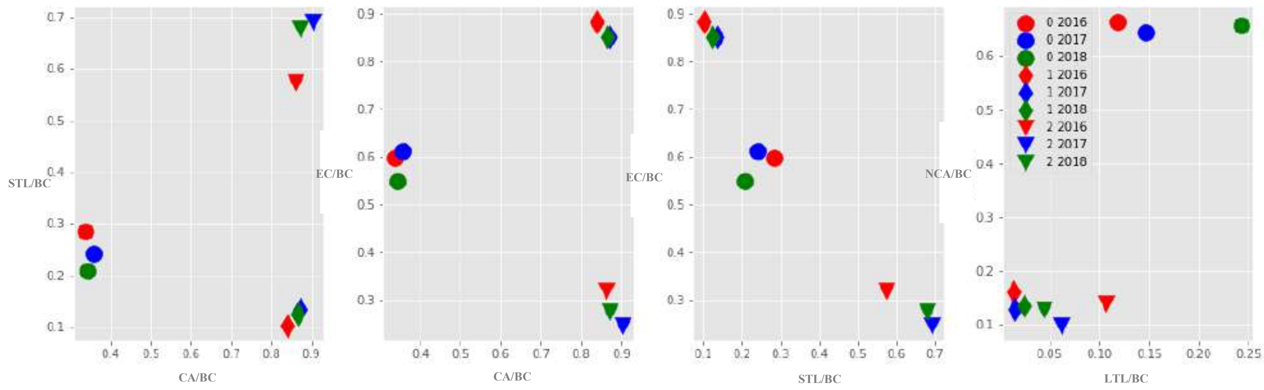
– cluster 2 (yellow markers) - includes organisations whose main share of property is current assets (91%), with short-term liabilities being the main source of asset financing (66%), followed by equity and long-term liabilities - 30% and 4%, respectively. The share of organisations with this typical financial model is 23%.

Stage 3: Identification of typical financial models that are sustainable over time

In the context of the TEA, it is necessary to identify those financial models whose values of indicators have not changed significantly over several years. This class of models will be further used to assess the financial insolvency of organisations and beyond. To assess the degree of change in the values of indicators of a particular typical financial model over different periods, the variation coefficient was used.

Therefore, with reference to the constructed clusters based on the balance sheet data for 2018, the same algorithm was utilised to form clusters based on the data for 2016 and 2017 by TEA. To determine comparable clusters for different periods, the search for minimum distances between the corresponding centroids was automated, which made it possible to form triples of similar objects of the same cluster for different years, where each object represents the centroid of a particular cluster in a particular year. For each such set of objects, the average values of their coordinates were calculated, reflecting the specific values of balance sheet sections in total assets and liabilities. Figure 15 presents 4 scatter plots of centroids of the selected clusters for 2016-2018 for TEA ‘Marine Fishing’ (OKVED 03.11) for the following pairs of indicators: CA/BC and STL/BC, CA/BC and EC/BC, STL/BC and EC/BC, LTL/BC and NCA/BC. The shape of the marker on the graphs determines the ordinal number of the cluster: circle – cluster 0, rhombus – cluster 1, triangle – cluster 2. The colour of the marker determines the period to which the centroid corresponds: red – 2016, blue – 2017, green – 2018. Then an example of a trio of similar objects are the centroids of cluster 1 for 2016-2018 (represented in the graph as rhombuses of three colours).

As can be seen from Figure 15, for each trio of centroids it is possible to calculate the average specific values of balance sheet sections for organisations in the marine fisheries sector: for cluster 0, non-current and current assets in total assets make up 66% and 34%, respectively, equity, long-term and current liabilities in total liabilities are 59%, 17% and 24%, respectively; for cluster 1, the share of non-current and current assets in total assets is 14% and 86%, respectively, the share of equity, long-term and current liabilities in total liabilities is 86%, 2% and 12%, respectively; for cluster 2, the share of non-current and current assets in total assets is 12% and 88%, respectively, the share of equity, long-term and current liabilities is 28%, 7% and 65%, respectively.



(a) (b) (c) (d)

Figure 15 – Results of clustering of balance sheets for 2016-2018 of organisations with MTA 'Marine Fishing' (OKVED 03.11) in the coordinate system: a) CA/BC and STL/BC; b) CA/BC and EC/BC; c) STL/BC and EC/BC; d) LTL/BC and NCA/BC [Compiled by the author].

Table 7 presents the obtained results of the calculation of average estimates of indicators of typical financial models, based on the data of the 2016-2018 financial statements for 13 TEAs.

Table 7 – Average estimates of indicators of typical financial models in the context of TEA and clusters according to the accounting data for 2016-2018 [Compiled by the author].

OKVED	Cluster number	NCA/BC	CA/BC	STL/BC	LTL/BC	EC/BC
02.20 – Logging	0	10%	90%	79%	2%	19%
	1	8%	92%	18%	3%	79%
	2	63%	37%	47%	13%	40%
03.11 – Marine fishing	0	66%	34%	59%	17%	24%
	1	14%	86%	86%	2%	12%
	2	12%	88%	28%	7%	65%
10.20 – Processing and canning of fish, crustaceans and mollusks	0	8%	92%	17%	2%	81%
	1	20%	80%	22%	51%	27%
	2	61%	39%	58%	5%	37%
	3	10%	90%	78%	1%	20%
11.07 – Manufacture of soft drinks; manufacture of mineral waters and other bottled drinking waters	0	12%	88%	18%	4%	78%
	1	43%	57%	16%	60%	25%
	2	65%	35%	59%	6%	35%
	3	11%	89%	78%	2%	20%

Continuation of table 7

OKVED	Cluster number	NCA/BC	CA/BC	STL/BC	LTL/BC	EC/BC
22.21 – Manufacture of plastic plates, strips, pipes and profiles	0	9%	91%	16%	2%	82%
	1	34%	66%	19%	52%	30%
	2	12%	88%	73%	2%	26%
	3	61%	39%	61%	5%	34%
25.11 – Manufacture of building metal structures, products and their parts	0	56%	44%	45%	11%	44%
	1	5%	95%	15%	3%	82%
	2	6%	94%	74%	2%	24%
25.12 – Production of metal doors and windows	0	4%	96%	20%	2%	78%
	1	66%	34%	47%	9%	44%
	2	6%	94%	74%	7%	19%
33.15 – Repair and maintenance of ships and boats	0	64%	36%	67%	2%	31%
	1	3%	97%	18%	0%	81%
	2	2%	98%	81%	0%	19%
	3	8%	92%	78%	1%	21%
47.75 – Retail trade in cosmetic and personal care products in specialised shops	0	3%	97%	24%	1%	75%
	1	20%	80%	26%	55%	20%
	2	3%	97%	83%	1%	16%
	3	68%	32%	60%	5%	36%
52.10 – Warehousing and storage activities	0	73%	27%	67%	3%	30%
	1	54%	46%	20%	58%	22%
	2	11%	89%	81%	1%	18%
	3	11%	89%	18%	2%	80%
52.22 – Transport support activities	0	3%	97%	24%	1%	75%
	1	46%	54%	18%	58%	24%
	2	5%	95%	83%	0%	17%
	3	64%	36%	69%	2%	29%
61.10 – Activities in the field of communications based on wire technologies	0	8%	92%	21%	1%	78%
	1	7%	93%	80%	1%	19%
	2	43%	57%	21%	57%	23%
	3	65%	35%	67%	2%	31%
81.21 – General cleaning activities for buildings	0	16%	84%	85%	0%	15%
	1	22%	78%	17%	2%	81%
	2	0%	100%	93%	0%	7%
	3	2%	98%	85%	1%	15%
	4	1%	99%	20%	0%	80%

Among the sets of typical financial models constructed in the context of the TEA and presented in Table 7, it is necessary to identify models that are stable over time. The coefficient of variation has been used as a criterion for assessing stability over time, calculated on a time series consisting of the values of each indicator of a typical financial model over 3 years. If the coefficient of variation of the time series is less than 33%, the sample is considered homogeneous, otherwise the data set analysed is characterised by a high degree of dispersion. At the same time, a particular typical financial model is stable if the time series of values of all indicators of this typical financial model for 3 years are homogeneous. The results of testing typical financial models for their stability over time are presented in Table 8. Unstable models are marked in the table with the symbol 'U', stable models with ’S'.

It is also important to note that the homogeneity condition is not met for some triples of centroids. Let us take the example of the TEA ' Marine Fishing' (OKVED 03.11): for cluster 0, the coefficient of variation of the indicator reflecting the share of long-term liabilities in total liabilities exceeds the threshold of 33% and is 37%, which indicates a high degree of data dispersion and, consequently, the inappropriateness of using the typical financial model corresponding to this cluster, averaged over several periods, for the purpose of assessing the financial insolvency of organisations and beyond.

Table 8 – Estimation of temporal stability of typical financial models in the context of TEA [Compiled by the author]

OKVED	**Cluster number**				
	0	**1**	**2**	**3**	**4**
02.20 – Logging	U	S	S		
03.11 – Marine fishing	S	S	U		
10.20 – Processing and canning of fish, crustaceans and mollusks	S	S	U	U	
11.07 – Manufacture of soft drinks; manufacture of mineral waters and other bottled drinking waters	U	S	U	S	
22.21 – Manufacture of plastic plates, strips, pipes and profiles	S	S	S	S	
25.11 – Manufacture of building metal structures, products and their parts	S	S	U		
25.12 – Production of metal doors and windows	U	U	U		
33.15 – Repair and maintenance of ships and boats	U	U	S	U	
47.75 – Retail trade in cosmetic and personal care products in specialised shops	S	U	S	U	
52.10 – Warehousing and storage activities	S	S	S	S	
52.22 – Transport support activities	U	S	U	S	
61.10 – Activities in the field of communications based on wire technologies	S	S	S	U	
81.21 – General cleaning activities for buildings	S	U	U	U	U

It is suggested to use the identified time-stable typical financial models to assess the financial insolvency of organisations and beyond, as the values of indicators of such models do not change significantly from year to year. Within the framework of the present study, it is noted that there are significant fluctuations in the values of indicators of unstable financial models, in the case of which it is recommended to be extremely cautious. The analysis of the nature of significant fluctuations is the subject of future research [Kovalev, Moldobaev, 2021].

Suppose the system receives as input the financial statements of a given organisation for which the MTA is known. Let us consider the sequence of actions to identify a typical financial model that is stable over time and maximally similar to the balance sheet of this organisation (hereafter referred to as the instruction):

1. Search in the directory with sets of typical financial models for all TEAs for a suitable TEA equal to the organisation's MTA (an example of the directory is presented in the form of Table 7). If the required TEA is included in the generated directory, it is possible to determine the most similar financial model for the organisation.

2. Transformation of the balance sheet of the organisation under study into a reporting format (financial model of the organisation) containing five relative indicators: NCA/BC, CA/BC, EC/BC, LTL/BC, STL/BC.

3. Determine the measure of similarity of the organisation's financial model from step 2 to each typical financial model from step 1 using Euclidean distance.

4. Selecting the most similar generic financial model from step 3 for which the distance between the objects (the organisation's financial model and the selected typical financial model) is minimal.

5. If the typical financial model identified in paragraph 4 is sustainable (see table 8), its continued use for financial insolvency assessment purposes and beyond is recommended; otherwise, the subsequent use of an unsustainable model should be extremely cautious.

3.2 Application of typical financial models to assess the financial insolvency of organisations

Taking into account the sets of typical financial models, including those that are stable over time, developed within the framework of TEA, the next step is to identify such models that are typical for the majority of financially insolvent organisations. Then, if the correlation of the values of individual financial statement indicators of any studied organisation with the values of indicators of the average sustainable model of financially insolvent organisations confirms their maximum similarity, we will consider that the studied organisation is at risk of financial insolvency.

In order to identify typical financial models characteristic of insolvent organisations, for each insolvent organisation in the sample its typical financial model is identified using the instructions described in paragraph 3.1. of the Instruction. Having identified a typical financial model for each insolvent organisation, which is at the same time the centroid of the cluster, the distribution of the financially insolvent organisations into clusters is carried out and the share of financially insolvent organisations is calculated ($FFI_{i,j}$) using the formula:

$$FFI_{i,j} = \frac{N_{i,j}}{N_j} \times 100\%, \tag{23}$$

where $N_{i,j}$ is the number of financially insolvent organisations with j-th TEA from the i-th cluster; N_j is total number of financially insolvent organisations with j-th TEA.

The results of distribution of financially insolvent organisations by clusters by TEA using the indicator $FFI_{i,j}$ are presented in Table 9. The cells for which the index $FFI_{i,j}$ exceeds 60% are shown in grey, indicating that out of the set of typical financial models for the j-th TEA, the i-th model is typical for the majority of financially insolvent organisations. Therefore, if the financial statements of an organisation studied with the j-th TEA are as similar as possible to the typical financial model of the i-th cluster of the j-th TEA, then this organisation faces the risk of financial insolvency.

Table 9 – Relative distribution of financially insolvent organisations into clusters using the index $FFI_{i,j}$ by TEAs [Compiled by the author]

OKVED	**Cluster number**				
	0	**1**	**2**	**3**	**4**
02.20 – Logging	0%	67%	33%		
03.11 – Marine fishing	33%	0%	67%		
10.20 – Processing and canning of fish, crustaceans and mollusks	62%	17%	18%	2%	
11.07 – Manufacture of soft drinks; manufacture of mineral waters and other bottled drinking waters	0%	100%	0%	0%	
22.21 – Manufacture of plastic plates, strips, pipes and profiles	70%	20%	10%	0%	
25.11 – Manufacture of building metal structures, products and their parts	9%	85%	6%		
25.12 – Production of metal doors and windows	86%	9%	5%		
33.15 – Repair and maintenance of ships and boats	0%	96%	0%	4%	
47.75 – Retail trade in cosmetic and personal care products in specialised shops	70%	8%	9%	13%	
52.10 – Warehousing and storage activities	9%	15%	9%	67%	
52.22 – Transport support activities	65%	16%	9%	10%	
61.10 – Activities in the field of communications based on wire technologies	100%	0%	0%	0%	
81.21 – General cleaning activities for buildings	0%	38%	0%	0%	62%

It is important to note that not all typical financial models corresponding to clusters with high $FFI_{i,j}$ index are stable over time. In order to identify sustainable models that can be used as a tool to assess the financial insolvency of organisations, we combine the research results presented in Tables 8 and 9. The classification of typical financial models of insolvent organisations by temporal stability is presented in Table 10. The green colour reflects typical financial models of insolvent organisations that meet the criterion of time stability. Consequently, it is recommended to use such models in relation to any organisation with a TEA from the table in order to assess its financial insolvency. The red colour represents typical financial models that are characteristic of insolvent organisations, but these models are not sustainable, so there is a risk of a significant change in the estimates of the indicators of the typical financial model in future periods.

Table 10 – The classification of typical financial models of insolvent organisations by temporal stability [Compiled by the author]

OKVED	Cluster number				
	0	**1**	**2**	**3**	**4**
02.20 – Logging	U	S	S		
03.11 – Marine fishing	S	S	U		
10.20 – Processing and canning of fish, crustaceans and mollusks	S	S	U	U	
11.07 – Manufacture of soft drinks; manufacture of mineral waters and other bottled drinking waters	U	S	U	S	
22.21 – Manufacture of plastic plates, strips, pipes and profiles	S	S	S	S	
25.11 – Manufacture of building metal structures, products and their parts	S	S	U		
25.12 – Production of metal doors and windows	U	U	U		
33.15 – Repair and maintenance of ships and boats	U	U	S	U	
47.75 – Retail trade in cosmetic and personal care products in specialised shops	S	U	S	U	
52.10 – Warehousing and storage activities	S	S	S	S	
52.22 – Transport support activities	U	S	U	S	
61.10 – Activities in the field of communications based on wire technologies	S	S	S	U	
81.21 – General cleaning activities for buildings	S	U	U	U	U

Temporal stability of typical financial models characteristic for insolvent organisations is observed for the following TEAs: ‘Logging’ (OKVED 02.20), ‘Processing and canning of fish, crustaceans and mollusks’ (OKVED 10.20), ‘Manufacture of soft drinks; manufacture of mineral waters and other bottled drinking waters’ (OKVED 11.07), ‘Manufacture of plastic plates, strips, pipes and profiles’ (OKVED 22. 21), ‘Manufacture of building metal structures, products and their parts’ (OKVED 25.11), ‘Retail trade in cosmetic and personal care products in specialised shops’ (OKVED 47.75), ‘Warehousing and storage activities’ (OKVED 52.10), ‘Activities in the field of communications based on wire technologies’ (OKVED 61.10).

Let us look in more detail at the obtained sustainable typical financial models of insolvent organisations by assessing the values of their indicators and derived financial ratios that allow to describe the organisation from the point of view of its liquidity and financial stability. We will also highlight the characteristics of each model of financially insolvent organisations in comparison with other typical financial models from the set corresponding to a given TEA.

Table 11 presents the values of liquidity and financial stability ratios calculated on the basis of the indicators of time-stable typical financial models of insolvent organisations (see tables 7 and 10).

Table 11 –Values of liquidity and financial stability indicators according to typical financial models' data of insolvent organisations by TEA [Compiled by the author]

Name and formula for calculating the indicator [47]	**OKVED code**							
	02.20	**10.20**	**11.07**	**22.21**	**25.11**	**47.75**	**52.10**	**61.10**
Current liquidity ratio *= (CA/BC) / (STL/BC)*	1,16	1,14	2,28	1,11	1,16	1,29	1,11	1,18
Ratio of own working capital *= [(CA/BC) – (STL/BC)] / (CA/BC)*	0,14	0,12	0,56	0,1	0,14	0,23	0,1	0,15
Autonomy ratio *= (EC / BC) / [(EC/BC) + (LTL/BC) + (STL/BC)]*	0,18	0,17	0,16	0,16	0,15	0,24	0,18	0,21
Financial dependency ratio *= [(LTL/BC) + (STL/BC)] / [(EC/BC) + (LTL/BC) + (STL/BC)]*	0,82	0,83	0,85	0,84	0,85	0,76	0,82	0,79
Financial leverage ratio *= [(LTL/BC) + (STL/BC)] / (EC/BC)*	4,56	4,88	5,31	5,25	5,67	3,17	4,56	3,76
Coefficient of maneuverability of own working capital *= [(CA/BC) – (STL/BC)] / (EC/BC)*	0,72	0,65	2	0,56	0,87	0,92	0,5	0,67
Ratio of immobilized to mobile assets *= (NCA/BC) / (CA/BC)*	0,09	0,09	0,75	0,1	0,05	0,03	0,12	0,09
Ratio of mobile to immobilized assets *= (CA/BC) / (NCA/BC)*	11,5	11,5	1,33	10,11	19	32,33	8,09	11,5

[47] The description of the symbolic designations of the indicators of typical financial models is presented in paragraph 3.1.

In financially insolvent organisations with MTA ‘Logging’ (OKVED 02.20) the share of non-current and current assets in total assets is 8% and 92% respectively, while the share of equity, long-term and short-term liabilities in total liabilities is 18%, 3% and 79% respectively.

The current liquidity ratio is equal to 1.16. The share of own current assets in current assets is 14%. The financial leverage ratio is 4.56. Current assets exceed non-current assets 11.5 times. Having compared this typical financial model of insolvent organisations with other models within the TEA under consideration, the following features can be noticed: despite the similar structure of assets with one of the other models, the latter has a high level of equity capital equal to 79%, which determines a low level of financial dependence; the second other model is characterised by the excess of non-current assets over current assets, a large share of long-term sources of financing equal to 13%, and an autonomy ratio of 47%.

For financially insolvent organisations with MTA ‘Processing and canning of fish, crustaceans and mollusks’ (OKVED 10.20) the following structure of the average balance sheet is typical: 92% of current assets and 8% of non-current assets in total assets; 17% of equity, 2% of long-term liabilities and 81% of short-term liabilities in total liabilities. It can be seen that the values of the indicators of the two typical financial models of insolvent organisations with the above described MTAs are exactly the same for the asset sections and slightly different for the liability sections. The revealed similarity of the models also determines the same average estimates of relative liquidity and financial stability indicators.

The main differences between the typical financial model of insolvent organisations and the three other models for the current TEA include: for the first other model - current assets exceed non-current assets 4 times, the main source of financing is long-term liabilities, the share of which in the composition of liabilities is 51%, and approximately equal ratio of equity capital and short-term liabilities, the shares of which are 22% and 27% respectively; for the second other model – non-current assets exceed current assets by 1.56 times – assets are financed to a greater extent by equity and current liabilities with shares of 58% and 37% respectively; for the third other model – significant differences are observed only in the structure of liabilities – the main source of financing is equity, which accounts for 78% of total liabilities.

The structure of the average balance sheet of financially insolvent organisations with MTA ‘Manufacture of soft drinks; manufacture of mineral waters and other bottled drinking waters’ (OKVED 11.07) is described by centroid coordinates, according to which non-current and current assets in total assets make up 43% and 57% respectively, equity capital, long-term and short-term liabilities in total liabilities – 15%, 60% and 25% respectively. Using the obtained estimates of the model indicators, the following financial ratios are calculated: current liquidity ratio is 2.28, financial leverage ratio is 5.31, manoeuvrability ratio of own current assets is 2, ratio of mobile and immobilised assets is 1.33. The main differences between the typical financial model of insolvent organisations and the two other models for the TEA under consideration are: for the first other model current assets exceed non-current assets by 7.33 times, most of the liabilities are represented by short-term liabilities – 78%, while almost no long-term borrowed sources are attracted, the share of which is 4%; for the second other model the ratio of immobilised assets to mobile assets is 1.86, assets are mostly financed by equity (59%) and short-term liabilities (35%); for the third other model the ratio of current assets to non-current assets is 8.09, in the structure of liabilities a large share falls on equity, equal to 78% of total liabilities.

The following structure of the average balance sheet is typical for financially insolvent organisations with MTA ‘Manufacture of plastic plates, strips, pipes and profiles’ (OKVED 22.21): current and non-current assets in total assets make 91% and 9% respectively; equity capital, long-term and short-term liabilities in liabilities are 16%, 2% and 82% respectively. Based on the obtained values of indicators of the average balance sheet sections, let us estimate the relative financial indicators: current liquidity ratio is 1.11, current assets ratio is 0.1, financial dependence ratio is 0.84, financial leverage ratio is 5.25, mobile to immobilised assets ratio is 10.11.

The considered typical financial model of insolvent organisations differs from three others from the same TEA by such criteria as: for the first other model current assets exceed non-current assets twice, current liquidity ratio is 2.2, a large share in liabilities are long-term and short-term liabilities equal to 52% and 30% respectively. For the second other model, the ratio of mobile and immobilised assets is 7.33, the main source of asset financing is equity (73%). For the third other model non-current assets exceed current assets 1.56 times, the manoeuvrability ratio of own current assets is 0.08, equity and short-term liabilities are the main sources of financing, the shares of which in the composition of liabilities are 61% and 34%, respectively.

The following structure of the average balance sheet is typical for financially insolvent organisations with MTA ‘Manufacture of building metal structures, products and their parts’ (OKVED 25.11): current and non-current assets in the composition of assets make up 95% and 5% respectively; equity capital, long-term and short-term liabilities in the composition of liabilities are 15%, 3% and 82% respectively. The obtained average assessments of the balance sheet sections determine the values of relative liquidity and financial stability indicators: current liquidity ratio is 1.16, own current assets ratio is 0.14, financial dependence ratio is 0.85, financial leverage ratio is 5.67, ratio of mobile assets to immobilised assets is 19. For this TEA, in addition to the typical financial model of insolvent organisations described above, two other models are identified, the characteristics of which are as follows: for the first other model, the excess of non-current assets over current assets is 1.27 times, the main sources of financing are equity and short-term liabilities, the shares of which are approximately equal – 45% and 44% respectively. For the second model differences are only observed in the structure of liabilities. For this cluster the share of equity is 74%, which influences the low value of the financial dependence ratio (26%) and the high value of the current liquidity ratio (3.92).

The structure of the average balance sheet of financially insolvent organisations with MTA ‘Retail trade in cosmetic and personal care products in specialised shops’ (OKVED 47.75) is as follows: the share of current and non-current assets in total assets is 97% and 3% respectively; equity, long-term and short-term liabilities in total liabilities are equal to 24%, 1% and 75% respectively. Let's calculate relative indicators of liquidity and financial stability using the average assessments of balance sheet sections obtained: current liquidity ratio is 1.29, own current assets ratio is 0.23, financial dependence ratio is 0.76, financial leverage ratio is 3.17, ratio of mobile assets to fixed assets is 32.33. Three other typical financial models within the TEA under consideration differ significantly from the model of financially insolvent organisations by the following parameters: for the first other model current assets exceed non-current assets 4 times, the main source of financing is long-term liabilities (55%), high level of current liquidity ratio (4) and own working capital ratio (0.75). For the second other model differences are observed only in the structure of the liabilities: the main source of financing is equity (83%), which determines a high level of the autonomy ratio and a low level of the financial dependence ratio. For the third other model the ratio of immobilised assets to mobile assets is 2.13, the main source of financing is equity (60%), then short-term liabilities (36%).

The following structure of the average balance sheet is typical for financially insolvent organisations with MTA 'Warehousing and storage activities' (OKVED 52.10): 89% of current assets and 11% of non-current assets in total assets; 18% of equity, 2% of long-term liabilities and 80% of short-term liabilities in total liabilities. Using the obtained estimates of the sections of the average balance sheet, let us determine the relative liquidity and financial stability indicators: the current liquidity ratio is 1.11, the ratio of provision with own working capital is 0.1; the financial dependence ratio is 0.82, the financial leverage ratio is 4.56, the coefficient of manoeuvrability of own working capital is 0.5, the ratio of mobile assets to immobilised assets is 8.09. In the context of the TEA under consideration, three other typical financial models have also been identified, the characteristics of which (in comparison with the model of insolvent organisations) are as follows: for the first other model - the prevalence of non-current assets over current assets is 2.7 times, the main sources of financing are equity and short-term liabilities, the shares of which in total liabilities are 67% and 30%, respectively. For the second other model the ratio of current and non-current assets is approximately equal, with their shares accounting for 46% and 54%, respectively. The main source of financing is long-term liabilities (58%), followed by short-term liabilities (22%) and equity (20%). For the third other model with a similar asset structure, differences are observed in the composition of the sources of financing. The majority of them are equity (81%), which positively affects the financial autonomy ratio and reduces the level of financial dependence, while the current liquidity ratio is 4.94 and the ratio of provision with own working capital is 0.8.

For financially insolvent organisations with MTA ‘Activities in the field of communications based on wire technologies’ (OKVED 61.10) the average estimates of the balance sheet sections are: current and non-current assets in total assets are 92% and 8% respectively; equity, long-term and short-term liabilities are 21%, 1% and 78% respectively. Based on the results obtained, we calculate the relative financial indicators: current liquidity ratio is 1.18, own working capital ratio is 0.15, financial dependence ratio is 0.79, financial leverage ratio is 3.76, own working capital manoeuvrability ratio is 0.67, mobile assets to immobilised assets ratio is 11.5.

Within the framework of the TEA under consideration, three other typical financial models were also identified, the main differences of which (compared to the model of insolvent organisations) are: for the first other model - a high level of the current liquidity ratio equal to 4.89, a high autonomy ratio and a low value of the financial dependence ratio due to the predominance of equity as the main source of asset financing (80%). For the second other model the ratio of mobile assets to immobile assets is 1.33, the current liquidity ratio is 2.48, a large share in the structure of liabilities falls on long-term (57%) and short-term liabilities (23%); for the third other model - non-current assets exceed current assets 1.86 times, the main source of financing is equity, the share of which in the structure of liabilities is 67%.

As a result of the analysis of sustainable typical financial models of insolvent organisations in the context of 8 TEAs, it can be noted that for all TEAs, except for

'Manufacture of soft drinks; production of mineral water and other bottled drinking water' (OKVED 11. 07), the average balance sheets have a similar structure of assets and liabilities with the following average estimates: the share of current and non-current assets in total assets is 7% and 93% respectively; the share of equity, long-term and short-term liabilities is 18%, 2% and 80% respectively. As can be seen from the above statistics, financially insolvent organisations are characterised by a significant excess of mobile assets over immobilised assets and a high level of short-term liabilities in liabilities. And only in the case of organisations with the MTA 'Manufacture of soft drinks; production of mineral water and other bottled drinking water' (OKVED 11.07), we consider that there is a risk of their financial insolvency if current assets slightly exceed non-current assets and the main source of financing assets are long-term liabilities, which leads to a decrease in the autonomy ratio.

It is important to note that the use of a typical financial model of insolvent entities without taking into account other typical financial models for a particular TEA makes no practical sense, since establishing a link between the financial model of an organisation under study and the model of insolvent entities requires a prior determination of the degree of similarity of the financial model of that organisation to all typical financial models within a particular TEA and a subsequent comparison of those degrees of similarity. Then, to assess the financial insolvency of a particular organisation, it is necessary to: determine its MTA; assess the similarity of its financial model with all typical financial models within a specific TEA, equal to the TEA of the organisation under study, using the Euclidean distance; select the most similar typical financial model. And if the most similar typical financial model is stable over time and characterised by a high frequency of financially insolvent organisations (as assessed by the index $FFI_{i,j}$), only then can it be said that the organisation under study is at risk of financial insolvency [Kovalev, Moldobaev, 2021].

3.3 Other uses of typical financial models beyond the assessment of financial insolvency

Let us assume that the application of the method of financial insolvency assessment using typical financial models described in paragraph 3.2 to the organisation under study has resulted in the maximum similarity of its balance sheet to the average model of insolvent organisations. In this context, let us consider the possibility of further application of typical financial models in order to improve the financial position of this insolvent organisation by implementing the following management decisions: merger and acquisition (hereinafter - M&A); introduction of innovative technologies. It is important to note that traditional approaches to assessing the effectiveness of the listed management decisions allow for the possibility of using financial statement indicators. At the same time, it is necessary to take into account the fact that, in practice, some actions taken as part of management decisions to improve the financial position of the organisation are not ultimately reflected in the financial statements. Therefore, the application of the proposals described below to modify the existing approaches to solving applied economic problems requires a preliminary assessment of the impact of a particular decision on the indicators of the accounting statements [Moldobaev, 2022].

Evaluating the effectiveness of M&A deals

In order to develop business, improve competitiveness and maximise profits, organisations make various investment decisions, including investment in research and development, investment in modernisation of production facilities, creation of new structural units, as well as acquisition of or merger with existing other organisations (M&A) [Ivanova, 2013; Miller, 2013]. The latter type of decision has the important advantage of reducing the cost and time of expensive research and development. On the one hand, a successful M&A strategy accelerates the development of the organisation. On the other hand, unsuccessful transactions can lead to significant financial losses and, as a result, serious financial difficulties.

Therefore, it is important to evaluate the effectiveness of M&A transactions in a timely manner in order to reduce the risk of losses and maximise the overall financial outcome of this strategy. Traditionally, all methods of assessing the effectiveness of M&A transactions are divided into two groups: prospective, used to calculate the optimal price of the target before its purchase; and retrospective, used to assess the performance of organisations after their integration. Prospective valuation methods include: comparative approach - valuation on the basis of organisation analogues; cost approach – valuation from the position of the organisation's existing assets; income approach – forecasting cash flows and bringing them to the current date using a discount rate. Among the retrospective methods there are: accounting approach – comparison of financial indicators before and after the transaction; market approach – analysis of the profitability of the shares of the organisation before and after the transaction; combined approach – assessment of the correlation between financial indicators and profitability of shares [Shemanueva, 2017].

This study proposes to extend the above methods of M&A performance assessment by using the previously identified typical financial models. Thus, if an organisation is at risk of financial insolvency, determined by assessing the similarity of the organisation's financial model and the typical model of insolvent organisations, it is proposed to consider acquiring another organisation with the same TEA in order to improve its financial position. Then, as a result of the combination of the two entities, the consolidated financial statements will represent a financial model that is as close as possible to the typical model of financially healthy entities. However, the acquired entity may not necessarily be financially efficient because the combination of the financial models of the two entities would change the structure of the assets and liabilities of the single entity. In addition, the acquiring entity may use different remuneration options to implement such a strategy, which, in turn, will also affect the structure of the consolidated financial model.

Let's take the example of how the structure of the consolidated balance sheet (BS) changes when a parent company (PC) acquires 100% of the shares of a subsidiary company (SC). In this case, the consolidated financial statements (CFS) will reflect the assets and liabilities of the entities as a single economic entity, so we will use the rules of consolidation of financial statements, which are defined in IFRS 10[48]. Assume that the fair value of the net assets of the acquiree at the acquisition date equals the book value and is equal to the consideration transferred (10,000 y.e.). Then goodwill at the acquisition date would be zero.

Figure 16 shows the scheme of formation of the consolidated balance sheet depending on the method of acquisition of the subsidiary: at the expense of non-current assets (NCA), current assets (CA), equity – shares issued (EC), formation of long-term liabilities (LTL) and short-term liabilities (STL). In the first stage, at the acquisition date, the acquiring company recognises an investment in non-current assets in its individual financial statements (IFS). The investment in the subsidiary recognised in the IFS is then eliminated on consolidation and the remaining assets of the parent and subsidiary are added together. Liabilities are also summarised in the consolidated financial statements, with equity at the time of consolidation being equal to the equity of the parent company. In the final stage, the resulting consolidated balance sheets with absolute values of indicators are transformed into a format with relative values of balance sheet sections. The final transformed consolidated balance sheets represent financial models whose values depend on the individual financial models of the acquiring and acquired entities and the type of consideration transferred.

[48] International Financial Reporting Standard (IFRS) 10 "Consolidated Financial Statements" (put into effect in the territory of the Russian Federation by Order of the Ministry of Finance of Russia dated 28.12.2015 N 217n) (as amended on 27.06.2016)

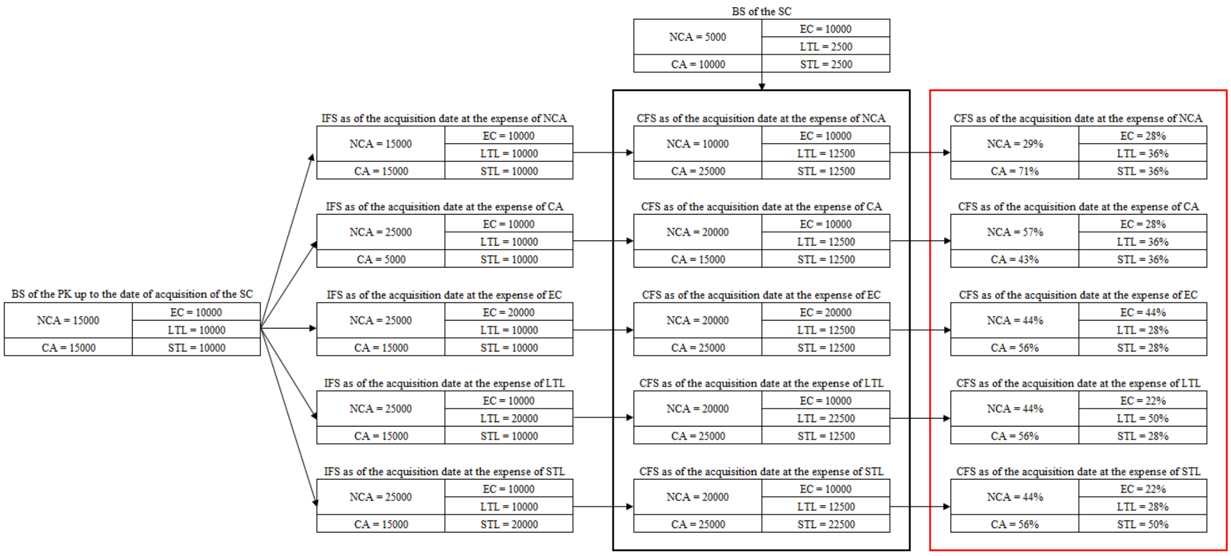


Figure 16 – Scheme of consolidated balance sheet formation depending on the type of compensation transferred for the acquisition of a subsidiary by a parent organisation [Compiled by the author].

Among the identified financial models of the merged organisation, there will be a model that has maximum similarity with the typical financial model of financially healthy organisations within a particular TEA. Therefore, if the acquisition of an organisation that will improve the financial position of the group requires lower costs compared to internal changes in business processes, it is recommended to consider such an organisation as a target for an M&A transaction.

Evaluate the efficiency of the implementation of innovative technologies

Innovative technologies provide exponential growth in the performance of organisations and are a necessary condition for scaling a business in the current economic environment. That is why the decision to introduce technologies into the main and auxiliary business processes of a financially insolvent organisation can lead to its financial recovery and subsequent rapid development. This thesis is confirmed by Russian researchers in their scientific works.

For example, V.V. Spinitz and L.Y. Spitsyna found a highly significant positive impact of financial performance and sustainability indicators on the net return on assets of enterprises of high-tech industries in Russia during the crisis. In addition, the researchers were able to confirm the hypothesis of a parabolic dependence between the size of the company and the net return on assets [Spitsyn, Spitsyna, 2020].

N.M. Kuznetsova established the interdependence between the competitive position of the organisation, its efficiency and innovation potential with the help of the developed model of occupational risk management on the basis of innovation implementation. The researcher considered the impact of innovation on the reduction of occupational risks in the field of occupational safety, health and environmental protection. In this study, the economic impact of the introduction of innovative technologies is determined by savings in the form of reduced compensation payments for injuries to workers and other factors [Kuznetsova, 2012].

A.V. Agalakova and G.I. Khrapovitsky see innovation activity as one of the key factors in ensuring long-term competitiveness and, as a result, improving the financial position of the organisation. Therefore, innovation is an integral part of the organisation's strategy. At the same time, the researchers believe that the level of innovation activity depends on the size of the business, the specifics of the industry and the current position of the organisation in the industry [Agalakova, Khrapovitsky, 2012].

E.V. Azimina notes that innovations ensure long-term business efficiency despite the negative background of macroeconomic factors. At the same time, she proposes a conventional division of innovations into 3 groups: design, technological and organisational. For each of these groups, the factors influencing the achievement of the target indicator of long-term efficiency - return on capital employed (ROCE) have been established. The author's decomposition of the company's innovation infrastructure and the identification of success factors for the implementation of the innovation strategy confirm the hypothesis of a close relationship between the level of the organisation's innovation potential and its performance indicators [Azimina, 2014].

From the analysis of the national studies that have been devoted to the assessment of the impact of innovative technologies on the performance of the organisation, it can be observed that the specifics of the industry and the size of the organisation have an impact on the development of the organisation through the adoption of technology.

Next, let us consider individual economic sectors from the point of view of the peculiarities of the distribution of all intellectual property objects that make up the innovative potential of the industry. As part of the analysis of the impact of innovative technologies on the efficiency of organisations' activities, the following industries with a significant impact on the national economy were considered: information and communication activities; mining; research and development activities; food industry; pharmaceutical industry; chemical industry; other high-tech industries. The last group includes production of computers, electronic and optical products and production of electrical equipment. The technological potential of the industry was assessed through a comparative analysis of various intellectual property objects, including databases (DB), inventions (I), utility models (UM), computer programmes (CP), industrial designs (ID), trade secrets (know-how) (TS), selection achievements (SA), integrated circuit topologies (ICT).

The total number of intellectual property objects in the 7 economic sectors under consideration as of the beginning of 2022 was 26,970 units. Of these, inventions account for 42%, computer programs for 32%, utility models for 10%, designs for 6%, industrial designs for 5%, databases for 4% and integrated circuit topologies for 1% (Figure 17)[49].

[49] Open data of the Federal Service for Intellectual Property. URL: https://rospatent.gov.ru/opendata

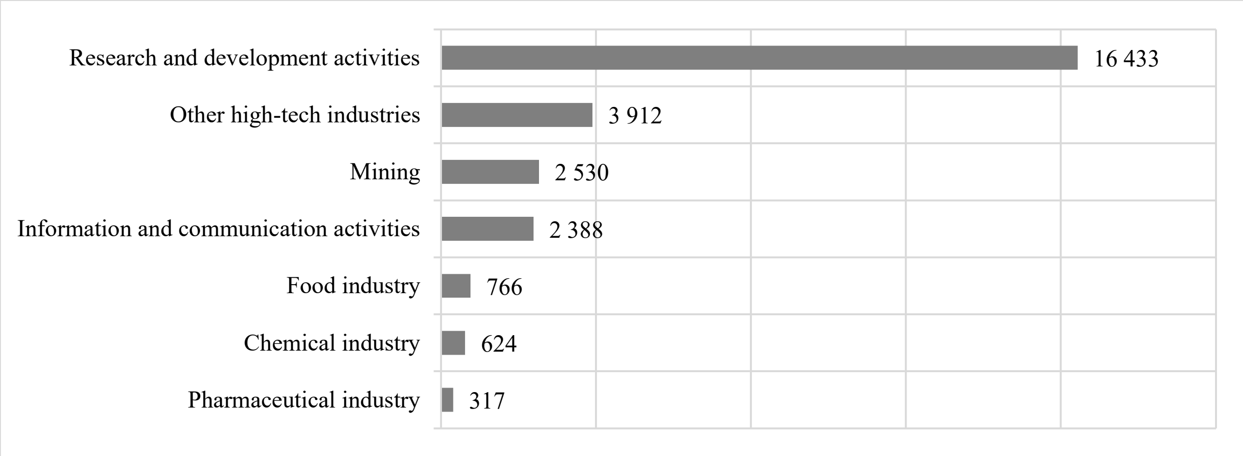


Figure 17 – Number of intellectual property objects by economic sectors, units. [Compiled by the author according to the Federal Service for Intellectual Property].

The distribution of different intellectual property items by economic sector is presented in Table 12.

Table 12 – Distribution of individual intellectual property objects in the context of industries [Compiled by the author according to the Federal Service for Intellectual Property]

Industry name	**Name of the intellectual property object**								**Total**
	DB	**I**	**UM**	**CP**	**ID**	**TS**	**SA**	**ICT**	
Other high-tech industries	31	1 348	1 000	1 072	363	2	1	95	**3 912**
Information and communication activities	45	914	39	1 365	16	3	1	5	**2 388**
Mining	117	1 065	295	1 028	22	3	0	0	**2 530**
Research and development activities	900	7 036	1 263	4 977	350	2	1 596	309	**16 433**
Food industry	11	97	10	104	539	2	3	0	**766**
Pharmaceutical industry	1	296	3	1	15	1	0	0	**317**
Chemical industry	4	454	48	97	20	1	0	0	**624**
Total	**1 109**	**11 210**	**2 658**	**8 644**	**1 325**	**14**	**1 601**	**409**	**26 970**

The analysis of the structure of intellectual property objects shows that inventions and computer programs account for the majority of objects, while low shares are characteristic of databases, integrated circuit topologies and trade secrets. The distribution of intellectual property objects differs significantly in the economic sectors considered. For example, a significant share of inventions among other intellectual property objects is found in research and development, mining and other high-tech industries. Computer programs are most common in R&D, information and communication and other high-tech industries, while industrial designs are most common in in the food industry.

In addition, there is a significant difference in the level of innovative development of the sectors considered, which is confirmed by the heterogeneity of the number of intellectual property objects in the context of individual sectors. In terms of the number of intellectual property objects, the most numerous are R&D activities (61%), other high-tech industries (15%), mining (9%) and information and communication activities (9%). On the other hand, the food industry (3%), the chemical industry (2%) and the pharmaceutical industry (1%) are among the important economic sectors that require special attention from the state and require high levels of investment to create, absorb and develop innovations.

The impact of intangible assets, including intellectual property, on the efficiency of organisations in high-tech industries has also been analysed.

At the same time, in order to determine the size of each sector, the number of organisations in each sector was first calculated. The total number of organisations in information and communication activities was 16,358 (of which 96% were small enterprises, 2% were medium-sized enterprises and 2% were large enterprises), in mining – 930 (of which 60% were small enterprises, 10% were medium-sized enterprises and 30% were large enterprises) and in research and development – 7,283 (of which 97% were small enterprises, 2% were medium-sized enterprises and 1% were large enterprises), in the food industry – 10,805 (of which 87% were small enterprises, 7% were medium-sized enterprises and 6% were large enterprises), in the pharmaceutical industry – 562 (of which 75% were small enterprises, 12% were medium-sized enterprises and 13% were large enterprises), in the chemical industry – 2,712 (of which 87% were small enterprises, 6% were medium-sized enterprises and 7% were large enterprises), in other high-tech activities – 5,095 (of which 90% were small enterprises, 6% were medium-sized enterprises and 4% were large enterprises)[50].

It is also important to note the peculiarity that the research revealed in all the organisations in the sectors studied – an extremely low share of intangible assets, which include intellectual property objects, in the total property complex of the organisations studied. Thus, according to the accounting data, the share of intangible assets did not exceed 4% on average, regardless of the economic sector and size of the organisations studied.

There was also a significant increase in the profitability of intangible assets according to the size of the organisation. On average, the profitability of intangible assets for small enterprises was 7%, for medium enterprises – 224%, and for large enterprises – 446%. This fact indicates a non-linear increase in the efficiency of Russian organisations with an increase in the intensity of the use of intellectual property objects.

The interim analysis allowed us to identify the economic sectors with the maximum and minimum levels of innovation activity. The number of intellectual property objects in the most technologically advanced industry (R&D) exceeds the number of objects in the pharmaceutical industry, with the minimum level of innovation potential, by almost 52 times. This regularity confirms the heterogeneity of technology distribution among economic sectors.

A stable unidirectional relationship between the share of intangible assets in the total property complex of the organisation and the profitability of assets taking into account the size of the business has also been established. This fact confirms the hypothesis about the impact of technology on the exponential growth of

[50] Integration of information from the Unified State Register of Legal Entities and the Unified State Register of Individual Entrepreneurs into the information systems of interested parties. URL: https://www.nalog.gov.ru/rn77/service/egrip2/egrip_vzayim/

organisations' performance and business scaling in the current economic environment.

The methodological approaches currently developed to assess the innovation potential of organisations are fragmented and often do not take into account the specificities of the innovative development of economic sectors and the size of companies.

Then decision-making on the introduction of innovative technologies into business processes of an insolvent organisation for the purpose of its financial recovery with the use of typical financial models has a number of peculiarities. On the one hand, the formed typical financial models take into account the industry specifics, which is an important factor in assessing the impact of innovation on the performance of an organisation from a certain economic sector. On the other hand, the 5-dimensional estimates of typical financial models proposed in the framework of this study do not imply detailing the balance sheet indicators to individual items, including intangible assets. In addition, the generated models do not take into account the scale of operations, which has an impact on the profitability of the organisation when introducing innovative technologies. Both limitations can be overcome by taking into account the previously described advantages of clustering algorithms, including their scalability. Thus, using the developed consolidated database integrating financial and non-financial sources (see Figure 3), it is proposed to increase the dimensionality of typical financial models by including the share of intangible assets in the balance sheet currency and further dividing all clusters by scale. Then, for new typical financial models, including models of insolvent organisations, the average valuation of intangible assets in total assets will be available, taking into account TEA and business size. Therefore, when deciding on the introduction of technologies for the purpose of financial recovery of the organisation, it is necessary to assess the importance of intangible assets in typical financial models, typical for financially healthy organisations, and to develop a step-by-step plan for the introduction of technologies in the organisation under study, as a result of which the latter will be able to achieve such an optimal structure of assets and liabilities that will allow a non-linear increase in performance indicators.

The proposed methodology will also make it possible to identify organisations in need of additional funding in the form of state subsidies, which have a high potential for development if innovative technologies are introduced. At the same time, state organisations authorised to finance enterprises within the framework of state support will be able to provide targeted budget funds and thus increase the efficiency of the use of budget funds.

Chapter conclusions

Using the k-means clustering algorithm, groups (clusters) of organisations with similar financial models by TEA were identified. For each cluster a centroid was established, the coordinates of which represent the n-dimensional estimation of a typical financial model. In addition, all typical financial models were tested for their temporal stability using the coefficient of variation, which was applied to the time series of values of all model indicators for 3 years. Instructions for finding a typical financial model that is as similar as possible to the model of the organisation under study have also been developed.

From the set of generated typical financial models, the models characteristic of financially insolvent organisations that are stable over time are identified using the index $FFI_{i,j}$ that reflects the share of insolvent organisations with the j-th TEA belonging to the i-th cluster in the total number of financially insolvent organisations with the j-th TEA. The similarity of the financial model of the organisation under study with the sustainable model of insolvent organisations signals the existence of the threat of insolvency.

The features of time-stable typical financial models of insolvent organisations compared to other models from the same set of models by TEA using individual model indicators and related financial ratios are also highlighted.

The use of typical financial models in making management decisions aimed at improving the financial position of the organisation through M&A deals or the implementation of innovative technologies is substantiated. For the latter decision, the requirements for modifying the typical financial models developed in this study are presented.

CONCLUSION

To summarise this thesis work, the following results can be formulated.

The author's characterisation of the limitations of known bankruptcy forecasting models as a tool in the assessment of financial insolvency is given.

Foreign and domestic scientific publications devoted to the development of authors' balance sheet BFMs have been analysed. All models are divided into two groups depending on the statistical methods used for their development: single-criteria models, which use the normative value of a single financial ratio (hereinafter – the first group of BFMs); multi-criteria models, which are based on the use of a system of financial indicators, the values of which determine the final assessment of the composite bankruptcy ratio (hereinafter – the second group of BFMs).

The following general limitations are highlighted for all domestic BFMs considered:

1. Use of financial ratios and other normalised accounting indicators. Financial ratios are characterised by the variability of approaches to their calculation, which allows interested parties to choose such analytical approaches that more optimistically reflect the financial condition of the organisation. Normalisation of reporting indicators implies the use of a logarithmic function, the area of definition of which is the set of all positive real numbers, while the profitability indicators of the majority of financially insolvent organisations are negative.

2. Limited size of the training sample formed from specialised RAS rather than from primary sources - open data of public authorities. The volume of data uploaded from RAS is regulated by the terms of the user agreement and in practice is much lower compared to the volume of the general population.

3. Industry specification of models using the OKVED directory developed by the Ministry of Economic Development of the Russian Federation. Given the hierarchical structure of the elements of this directory, researchers define the boundaries of the industry by different levels of hierarchy: sections (a set of classes), classes, subclasses or groups. Organisations are also segmented by industry on the

basis of their MTA, although an organisation may have a strategy of business diversification. Then, for other types of activities, the actual values of revenue and gross profit may account for a significant share of the organisation's total turnover.

4. The need to combine data from different sources using special automated tools to form a consolidated database containing financial indicators of organisations and their current bankruptcy status, which can then be used to test the predictive ability of the existing BFMs. For example, if the considered BFMs are used for a large number of organisations, it is necessary to expand the indicators of their financial statements formed from the data registers of Rosstat and FTS with information on the bankruptcy status and the stage of bankruptcy stored on the website of Fedresurs or in the electronic Card Library of Arbitration Cases.

In domestic studies, the first group of BFMs is based on the specification of normative values of financial ratios established by legislation, taking into account industry specifics and the size of the company, using such mathematical approaches as the development of classification trees and related ensemble machine learning models to solve the classification problem, as well as the determination of quartiles of an ordered series with the values of a single financial ratio. For this class of models, the following characteristics have been identified as having a negative impact on their predictive ability:

1. Drawing a conclusion about the organisation's susceptibility to insolvency on the basis of a single financial ratio that can only describe certain aspects of the organisation's activities. This leaves out other areas of activity that have an impact on financial insolvency.

2. The diversity of normative values of one and the same financial ratio in different scientific studies, caused by the variability of statistical methods for assessing normative values and approaches to the formation of the empirical base.

The second group of BFMs is based on establishing a relationship between the composite bankruptcy ratio and a set of financial ratios used as explanatory variables of the model. In turn, the models of this group are divided into subgroups depending on the possibility of their formulaic representation: classification linear models; classification ensemble models. Despite the high predictive ability of ensemble models proved by other researchers, their main limitation remains the impossibility to present BFMs data in the form of a formula, which makes it difficult for third-party stakeholders to use them. This circumstance is caused by the fact that these models are based on the construction of a set of basic algorithms (decision trees) that take into account the non-linear dependence of the composite bankruptcy ratio on financial indicators, with subsequent averaging of their forecasts.

A critique of the coefficient analysis of statements from the point of view of its use in assessing the financial insolvency of organisations is offered.

The peculiarity of coefficient analysis is the use of individual indicators of financial statements for the subsequent calculation of relative indicators, which are traditionally divided into groups depending on the analysed aspect of the financial position of the organisation: liquidity and solvency; financial stability; profitability; business activity. Due to the diversity of financial ratios, and taking into account the recommendations for overcoming the problem of retraining models that include a large number of characteristics, researchers must define the criteria of financial insolvency at the initial stage of model development, that is, select the most significant explanatory variables, on the values of which the final estimate of the composite bankruptcy ratio depends to the greatest extent.

Due to the variability of approaches to the definition of significant features, the 4 foreign and 3 domestic linear balance BFMs considered in this study use different sets of explanatory variables, which indicates a lack of understanding of which aspects of an organisation's financial position have a greater impact on its financial insolvency.

Thus, in E. Altman's model for non-public organisations, profitability and turnover of total assets are the main factors that contribute most to the assessment of the composite bankruptcy ratio. R. Lis' model, on the other hand, attaches great importance to the profitability of total assets, the share of working capital in total assets and the ratio of retained earnings to total assets. The final value of the composite bankruptcy ratio in R. Tuffler and G. Tishaw's model depends more on the profitability of short-term liabilities. In G. Springate's model, the profitability of total assets makes the largest contribution to the final value of the composite bankruptcy ratio. The comparison of domestic BFMs also reveals a divergence in the composition of the most significant features. In addition, in some models the composition of explanatory variables depends on the economic sector. For example, in the course of comparing the modules of weighting coefficients of A.V. Kazakov and A.V. Kolyshkin models, it was revealed that for agriculture and services the composite bankruptcy ratio depends to a greater extent on the absolute liquidity ratio, the weighting coefficient of which is many times higher than the weighting coefficients for other attributes; in construction, the absolute value of total liabilities makes the greatest contribution to the final assessment of the model; in trade – the absolute liquidity ratio and the absolute value of total liabilities.

In addition, the analysis of the dynamics of the median values of the financial ratios used as explanatory variables of BFMs in the context of 6 TEAs for 2012-2020 confirmed their heterogeneity, which indicates a low level of temporary stability of the BFMs considered in this study. It also revealed the impossibility of calculating certain characteristics of domestic BFMs on the basis of data from the financial statements of financially insolvent organisations, in which there are zero values of indicators used in the denominator in the calculation of this or that coefficient, or negative values of profit and retained earnings from the balance sheet, to which it is impossible to apply a logarithmic function to normalise the data.

To confirm the limitations described above in the use of financial ratios as explanatory variables of BFMs, we assessed the predictive ability of the most popular foreign and domestic BFMs on the basis of a sample of financially insolvent organisations using the author's quality metric (QM), which made it possible to assess the accuracy of classification of debtor organisations, taking into account the number of organisations with accounting statements that are unsuitable for calculating certain features of the model.

Foreign BFMs showed a high level of forecast accuracy for financially insolvent organisations. For domestic BFMs, the values of the quality metric were extremely low, which is primarily due to the inclusion of attributes that cannot be calculated from the data of accounting statements of debtor organisations. The testing also confirmed the hypothesis about the short-term forecasting horizon of foreign BFMs, as the value of the quality metric rapidly decreases as the bankruptcy forecasting period increases.

The advantages of multi-criteria assessments in diagnosing the financial insolvency of organisations using cluster analysis are described and the requirements to data sets necessary for the formation of such assessments are justified.

As a result of the analysis of known linear and non-linear BFMs, the limitations that have a negative impact on their applicability when using new data have been identified. The considered BFMs are developed on the basis of machine learning algorithms to solve the classification problem involving the use of pre-labelled data: in the case of BFMs, scientists initially know whether an organisation belongs to one of two classes – the area of financial insolvency or the area of financial stability. Thus, linear BFMs have low values of the author's quality metric, which confirms their weak predictive ability and their inappropriateness to the peculiarities of accounting reports of insolvent organisations. Nonlinear BFMs (ensemble models) are extremely difficult to apply by other interested parties, for whom the model looks like a black box, which does not explain the data structure and connections between them. Also, a common feature of all the BFMs under consideration is the use of financial ratios as explanatory variables, which represent a comparison of individual fragments of accounting statements, which deprives the possibility to comprehensively assess the financial position of the organisation, taking into account all the indicators of accounting statements and the relationships between them.

To overcome the identified limitations of the existing BFMs, an alternative approach to the assessment of financial insolvency based on the formation of typical financial models, which are a set of indicators in the form of specific values of individual balance sheet items, has been implemented. Typical financial models are developed for each industry using a large array of data from the accounting statements of Russian organisations. Since the solution of this problem requires dividing all organisations into clusters based on their accounting data and determining a prototype (centroid) – a typical financial model for each cluster, the k-means clustering algorithm was used as a machine learning method, and the Euclid distance was used to determine the measure of similarity of objects.

Then separately taken centroid of the cluster represents n-dimensional estimation of the typical financial model. In the framework of this study, the number of dimensions is equal to 5 and is determined by the number of sections of the balance sheet (non-current and current assets, equity, long-term and short-term liabilities). Given one of the positive features of the k-means clustering algorithm in the form of its scalability, the number of indicators of the typical financial model can be increased to include all available balance sheet items, which will provide a comprehensive view of the financial position of the organisation through its balance sheet as a perfect financial model.

Also, the models developed on the basis of cluster analysis are easy to use, as an outside interested party to determine whether an arbitrary organisation belongs to one of the clusters will need to present the balance sheet of this organisation in the form of an n-dimensional estimate of the financial model and determine the measure of similarity of the organisation's financial model with all typical financial models within the TEA using the Euclidean distance. The typical financial model for which the minimum distance from the n-dimensional estimate of the organisation's financial model is obtained will determine the cluster number.

The problem of temporal stability of models is solved by using centroids for different periods of time identified as a result of clustering. In the case of insignificant changes in the coordinates of centroids of the same cluster, we can talk about the temporal stability of the average typical financial model corresponding to these centroids, which is subsequently used in the assessment of financial insolvency of organisations.

A positive aspect of machine learning models is also their ability to learn on new data. For this purpose, a consolidated database has been developed, which includes many different disparate sources. All investigated sources, represented as open government data, and related identification and classification codes of government statistics are described and systematised. Thus, the collected data are divided into two groups: those describing financial and economic activities of Russian commercial organisations; those containing information on the movement of inventory.

When processing open government data, their difference in terms of format, structure and content was established. In order to integrate disparate data, typical scenarios of their merging were identified: using identifiers of organisations; using interface keys developed by the Ministry of Economic Development of the Russian Federation; by bitwise comparison of codes from different classifiers.

The possibility of forming typical financial models of organisations taking into account industry specifics on the basis of using algorithms of clustering of large data arrays is proved.

The initial sample of data with accounting statements of more than 2 million organisations was used within the framework of forming typical financial models. Organisations that did not meet the following conditions were excluded from the initial sample: assets equal to liabilities; values of all balance sheet items not less than zero; revenue and balance sheet currency greater than zero. In addition, all organisations were divided into industry sets. Of all the industry groups, those that met the minimum sample size criterion were selected. As a result of the data processing, the size of the final sample used to build typical financial models was 800 thousand organisations, distributed over 244 industry sets. For each organisation from the final sample, a financial model was also calculated, the indicators of which are specific values of the balance sheet sections.

The paper tests the hypothesis about the influence of the scale of activity on the values of indicators of typical financial models using the example of organisations with the MTA 'Marine Fishing' (MTA 03.11). As a result of the clustering of organisations by the indicators reflecting the scale of activity (in terms of revenue and balance sheet currency), centroids were formed, the values of which reflect the average estimates of indicators of financial models. It is important to note that the centroids obtained, when distributed in the system of specific values of the balance sheet sections, are close to each other. This indicates the similarity of the averaged financial models of organisations with different scale of activity. Consequently, the scale of activity does not affect the values of typical financial models of organisations from the same industry.

As a result of clustering of organisations with the MTA ‘Marine Fishing’ (OKVED 03.11) by specific values of the balance sheet sections (without segmentation by the scale of activity), 3 clusters were formed, each of which corresponds to a centroid with coordinates reflecting the n-dimensional assessment of the typical financial model. The values of indicators of 3 typical financial models within the TEA under consideration differ significantly, which indicates the diversity of organisations by their property complex and sources of financing. Consequently, it is inexpedient to create a single typical financial model for all organisations of the industry due to the heterogeneity of the values of the indicators of their financial models.

Given the revealed non-stationarity of median values of individual financial ratios for 2012-2020, the time-stable typical financial models were determined using the coefficient of variation. Thus, for each cluster the coordinates of centroids (n-dimensional estimates of typical financial models) for several years were calculated. In the case of homogeneity of values of all indicators of typical financial models built for different periods and corresponding to one cluster, the typical financial model averaged over several periods is proposed to be considered as stable over time.

Time-stable typical financial models in the context of the TEA are further used as tools for assessing the financial insolvency of organisations and beyond. At the same time, the use of unstable models should be extremely cautious due to the high risk of changes in estimates of model indicators. Analysing the nature of significant changes in the values of indicators of unstable typical financial models is an object of prospective research.

Modern domestic studies dedicated to the development of BFMs focus on the industry specificity of the models, but the need for such an approach is not substantiated. In the framework of this work, we selected 6 TEAs, for which we calculated the distribution of organisations by age, carried out a comparative analysis of the median values of the financial ratios for 2020, and analysed the dynamics of the median values of the financial ratios for 2012-2020. Thus, differences in the distribution of organisations by age have been established, which indicates the presence of heterogeneity in the maturity levels of individual markets, the intensity of competition, and the presence of leaders and outsiders in the market. The heterogeneity of median industry values of financial ratios calculated on the basis of the accounting statements of Russian organisations for 2020 has also been confirmed. In addition, it was found that the dynamics of the median values of individual financial ratios for 2012-2020 varies significantly between TEAs, which is probably due to the different degree of industries' resistance to the impact of various exogenous and endogenous factors.

The possibilities of assessment of financial insolvency of a commercial organisation on the basis of correlation of the data of its financial statements with the average model of financially insolvent organisations with correction for the industry specifics of its activity are shown.

In order to identify typical financial models characteristic of insolvent organisations, the distribution of these organisations into clusters is performed using the Euclidean distance calculated using the values of indicators of the financial model of each insolvent organisation and previously formed typical financial models (centroids) from the set within a particular TEA. Then the centroid of the cluster in which financially insolvent organisations are most frequently encountered reflects the n-dimensional estimate of the typical model of insolvent organisations. To calculate the frequency of occurrence of insolvent organisations in the cluster, the author's index is introduced $FFI_{i,j}$, reflecting the share of insolvent organisations with the j-th TEA belonging to the i-th cluster in the total number of financially insolvent organisations with the j-th TEA.

Having identified typical financial models with a high index $FFI_{i,j}$,, at the next stage their time stability was assessed for these models. Thus, for 8 out of 13 TEAs under consideration, it was possible to identify time-stable typical financial models characteristic of insolvent organisations. For each such model, the features described using the values of financial model indicators and related financial ratios reflecting liquidity and financial stability were identified. It was found that 7 out of 8 stable financial models of insolvent organisations from different TEAs have a similar structure of the balance sheet with the following average estimates: the share of current and non-current assets in total assets - 7% and 93% respectively; the share of equity, long-term and short-term liabilities - 18%, 2% and 80% respectively. As follows from the presented data, financially insolvent organisations are characterised by: a significant excess of mobile assets over immobilised assets; a high level of short-term liabilities in the structure of liabilities. An exception is observed with regard to the typical financial model of insolvent organisations with MTA ‘Manufacture of soft drinks; production of mineral water and other bottled drinking water’ (OKVED 11.07). Insolvent organisations from this industry are characterised by an insignificant excess of current assets over non-current assets, with long-term liabilities being the main source of asset financing, which determines a low level of the autonomy ratio.

The general rule for assessing the financial insolvency of an arbitrary organisation under study includes the following sequence of actions:

1. To evaluate the similarity of the financial model of the organisation under study to all typical financial models within a specific TEA equal to the MTA of the organisation under study, using Euclidean distance.

2. To select a typical financial model that is as similar as possible and for which the Euclidean distance between the objects (the organisation's financial model and the selected typical financial model) is minimal.

3. And if the most similar typical financial model is both stable over time and characterised by a high level of index $FFI_{i,j}$, then we believe that the organisation under review is at risk of financial insolvency.

The directions of use of typical financial models beyond the tasks of assessment of financial insolvency of organisations have been determined.

The paper considers the possibilities of applying typical financial models in relation to an insolvent organisation in order to justify management decisions aimed at its financial recovery in one of the following ways: merger and acquisition (hereinafter - M&A); introduction of innovative technologies.

For example, if an organisation is at risk of financial insolvency, determined by correlating its financial statements with an average model of insolvent organisations, it is suggested that the acquisition of another organisation in the same industry should be considered for its financial recovery. Such a merger should be based on the desire to obtain such a structure of the property complex and sources of its financing of the group of organisations that would be as close as possible to the typical model of financially healthy organisations. Therefore, if as a result of applying the proposed methodology an organisation is found, whose acquisition would provide an optimal structure of assets and liabilities, and this transaction would require lower costs in comparison with internal reengineering of business processes, it is recommended to consider this organisation as a target of the M&A transaction.

This study reveals a stable unidirectional relationship between the share of intangible assets in the total property complex of organisations and the profitability of the assets, taking into account the size of the company (average profitability of intangible assets for small companies – 7%, medium companies – 224%, large companies – 446%), which confirms the importance of introducing innovative technologies that provide exponential growth in the performance of the organisation. Consequently, the introduction of technology into the business processes of an insolvent organisation can lead to its rapid financial recovery and subsequent rapid development.

In addition, the paper reveals a significant difference in the distribution of intellectual property objects by industries, which confirms the need for sectoral specification of typical financial models in case of their use in the justification of management decisions on the introduction of innovative technologies.

The 5-dimensional estimates of the typical financial models proposed in this study do not include intangible assets. This limitation can be overcome by including intangible assets as a proportion of total assets in the new financial models. An insolvent organisation for which a recalculated financial model is known would then be able to identify a model that includes an average estimate of the proportion of intangible assets of financially successful organisations to guide its technology adoption.

At the same time, it is important to consider that not all measures to improve the financial position of the organisation will affect the final values of the financial statement indicators. Therefore, before applying the approaches described above to solve applied economic problems using typical financial models, it is necessary to assess how and to what extent this or that measure will affect the final values of the financial statement indicators.

LIST OF REFERENCES

1. "OK 028-2012. All-Russian Classifier of Organisational and Legal Forms" (approved by Order of Rosstandart dated 16.10.2012 N 505-st) (as amended on 14.03.2023) (together with "Explanations to OKOPF positions") [Electronic resource]. URL: https://www.consultant.ru/document/cons_doc_LAW_139192/ (date of access: 24.04.2024).

2. "OK 029-2014 (KDES Rev. 2). All-Russian Classifier of Types of Economic Activity" (approved by Order of Rosstandart dated 31.01.2014 N 14-st) (as amended on 30.11.2023) [Electronic resource]. URL: https://www.consultant.ru/document/cons_doc_LAW_163320/ (date of access: 24.04.2024).

3. "OK 034-2014 (KPES 2008). All-Russian classifier of products by types of economic activity (OKPD 2)" (approved by Order of Rosstandart dated 31.01.2014 N 14-st) (as amended on 04.02.2022) [Electronic resource]. URL: https://www.consultant.ru/document/cons_doc_LAW_163703/ (date of access: 24.04.2024).

4. A monthly updated complete database of accounting (financial) statements of organisations throughout Russia [Electronic resource]. URL: https://bo.nalog.ru/ (date of access: 26.03.2024).

5. Accounting (financial) statements of enterprises and organisations for 2018 [Electronic resource]. URL: https://rosstat.gov.ru/opendata/7708234640-7708234640bdboo2018 (date of access: 17.02.2020).

6. *Agalakova A. V., Khrapovitsky G. I.* The role of innovations in the process of forming an enterprise strategy // Siberian Aerospace Journal. – 2012. – №. 3 (43). – P. 132-135.

7. *Akulova N. G., Ryakhovsky D. I.* Problems of the quality of financial analysis in procedures applied in bankruptcy cases // Strategic decisions and risk management. – 2014. – №. 6 (87). – P. 74-83.

8. *Altman E. I.* Financial ratios, discriminant analysis and the prediction of corporate bankruptcy // The Journal of Finance. 1968. – V. 4. – P. 589–609.

9. *Altman E. I.* Why businesses fail //Journal of Business Strategy. – 1983. – V. 3. – №. 4. – P. 15-21.

10. *Altman E. I., Sabato G.* Modelling credit risk for SMEs: Evidence from the US market //Abacus. – 2007. – V. 43. – №. 3. – P. 332-357.

11. *Azimina E. V.* The role of innovations in managing long-term business efficiency // Bulletin of the St. Petersburg State University of Economics. – 2014. – №. 6 (90). – P. 107-113.

12. *Bakunova T. V., Koltsova T. A., Moldobaev T. Sh.* Features of accounting for operating and financial leases under IFRS // Accounting and Statistics. – 2019. – №. 1 (53). – P. 10-17.

13. *Bakunova T. V., Prisyazhny A. V., Detkov A. A., Moldobaev T. Sh., Taubaev A. A.* Spectral-scoring method for analyzing financial indicators of commercial enterprises // AIP Conference Proceedings. – AIP Publishing, 2019. – V. 2116. – №. 1.

14. Bankruptcies of organisations - Federal Resource statistics for 2019 [Electronic resource]. URL: https://fedresurs.ru/news/7b3c8884-b159-4ee7-b5fb-7770d9d941da?attempt=2 (date of access: 11.05.2024).

15. *Batrasova A. D., Konovalova T. V., Komarov P. I.* Clustering as a method for studying the financial stability of IT organisations // Management accounting. – 2022. – №. 1-2. – P. 177-182.

16. *Beaver W. H.* Financial ratios as predictors of failure //Journal of accounting research. – 1966. – P. 71-111.

17. *Blasiak K.* Big Data; A Management Revolution: The emerging role of big data in businesses. – 2014.

18. *Boyko I.P., Kazakov A.V., Kolyshkin A.V.* Methods of forecasting insolvency: problems and prospects // Russian entrepreneurship. – 2017. – V. 18. – № 8. – P. 1313-1326.

19. *Budarin A. D., Perepechkina E. G.* State regulation of the tobacco industry in Russia // Legal regulation of socio-economic processes: history and development prospects: collection of articles from the II All-Russian scientific and practical conference, April 24, 2018, Astrakhan. – Directmedia, 2018. – P. 39.

20. *Butakova N. M., Globa S. B., Zyablikov D. V.* Innovative development of infrastructure as a factor of economic growth // Bulletin of the Altai Academy of Economics and Law. – 2022. – №. 11. – P. 202-207.

21. *Butters R. A., Sacks D. W., Seo B. Why do retail prices fall during seasonal demand peaks? //The RAND Journal of Economics. – 2025. – V. 56. – №. 1. – P. 35-54.*

22. *Chen W., Zhang H., Mehlawat M. K., Jia L.* Mean–variance portfolio optimization using machine learning-based stock price prediction //Applied Soft Computing. – 2021. – V. 100. – P. 106943.

23. *Chudson W. A.* A Survey of Corporate Financial Structure // NBER Chapters in “The Pattern of Corporate Financial Structure: A Cross-Section View of Manufacturing, Mining, Trade, and Construction”. National Bureau of Economic Research. 1945. P. 1–16.

24. *Clifford L.* Big data: How do your data grow //Nature. – 2008. – V. 455. – №. 7209. – P. 28-29.

25. *Coulombe P. G., Leroux M., Stevanovic D., Surprenant S.* Macroeconomic data transformations matter //International Journal of Forecasting. – 2021. – V. 37. – №. 4. – P. 1338-1354.

26. *Cover T., Hart P.* Nearest neighbor pattern classification //IEEE transactions on information theory. – 1967. – V. 13. – №. 1. – P. 21-27.

27. Decision of the Council of the Eurasian Economic Commission of 14.09.2021 N 80 (as amended on 27.12.2023) "On approval of the unified Commodity Nomenclature of Foreign Economic Activity of the Eurasian Economic Union and the Single Customs Tariff of the Eurasian Economic Union, as well as on amending and recognizing as invalid certain decisions of the Council of the Eurasian Economic Commission" (as amended and supplemented, entered into force on

08.02.2024) [Electronic resource]. URL: https://www.consultant.ru/document/cons_doc_LAW_397176/ (date of access: 26.03.2024).

28. *Demeshev B.B., Tikhonova A.S.* Dynamics of the predictive power of bankruptcy models for medium and small Russian wholesale and retail trade organisations // Corporate Finance. – 2014. – V. 8. – №. 3. – P. 4-22.

29. *Dyagel O. Yu., Engelhardt E. O.* Diagnostics of the probability of bankruptcy of organisations: essence, tasks and comparative characteristics of methods // Economic analysis: theory and practice. – 2008. – №. 13. – P. 49-57.

30. *Edmister, R. O.* 1972. An Empirical Test of Financial Ratio Analysis for Small Business Failure Prediction. Journal of Financial and Quantitative Analysis 7: 1477–93.

31. Electronic justice [Electronic resource]. URL: https://kad.arbitr.ru/ (date of access: 27.03.2024).

32. *Erokhin V. L.* Northern Sea Route: cabotage and international transit in 2013-2020 // Marketing and logistics. – 2021. – №. 3 (35). – P. 15.

33. Export and import of the Russian Federation by goods [Electronic resource]. URL: https://customs.gov.ru/statistic (date of access: 14.02.2022).

34. *Fadeev A. N., Fadeeva I. S.* Logistics in the electric power industry // Logistics systems in the global economy. – 2020. – №. 10. – P. 300-303.

35. Federal Law "On Insolvency (Bankruptcy)" dated 26.10.2002 N 127-FZ [Electronic resource]. URL: http://www.consultant.ru/document/cons_doc_LAW_39331/ (date of access: 10.05.2020).

36. Federal State Statistics Service. Provision of accounting data upon user requests [Electronic resource]. URL: https://www.gks.ru/accounting_report (date of access: 20.02.2020).

37. *Fedorova E. A., Chukhlantseva M. A., Chekrizov D. V.* Normative values of financial stability coefficients: features of types of economic activity // Management Sciences. – 2017. – №. 2. – P. 44-55.

38. *Fedorova E. A., Khrustova L. E., Chekrizov D. V.* Industry specifics of application of enterprise bankruptcy forecasting models // Strategic decisions and risk management. – 2018. – № 1 (104). – P. 64–71.

39. *Fedorova E. A., Musienko S. O., Fedorov F. Yu.* Forecasting bankruptcy of small and medium-sized businesses in Russia // Finance and Credit. – 2018. – V. 24. – №. 11 (779). – P. 2537-2552.

40. *Fedorova E. A., Timofeev Ya. V.* Standards for financial stability of Russian enterprises: industry specific features // Corporate Finance. – 2015. – V. 9. – №. 1. – P. 38-47.

41. *Fedorova E.A., Dovzhenko S.E., Fedorov F.Yu.* Models for forecasting bankruptcy of Russian enterprises: industry specific features // "Problems of Forecasting". – 2016. – №3. – P. 32–40.

42. *Fisher I. E., Garnsey M. R., Hughes M. E.* Natural language processing in accounting, auditing and finance: A synthesis of the literature with a roadmap for future research //Intelligent Systems in Accounting, Finance and Management. – 2016. – V. 23. – №. 3. – P. 157-214.

43. *Freund Y., Schapire R. E.* A decision-theoretic generalization of on-line learning and an application to boosting //Journal of computer and system sciences. – 1997. – V. 55. – №. 1. – P. 119-139.

44. *Generalova N.V., Sokolova N.A.* Polymorphism of interpretations of economic information in an integrated accounting system // Finance and Business. – 2013. – № 2. – P. 159-168.

45. *Generalova N.V., Sokolova N.A.* Variability of interpretation of accounting information // Bulletin of St. Petersburg University. – Ser. 5. – 2012. – V. 1. – P. 175–181.

46. *Gil-Alana L. A., Gupta R.* Persistence and cycles in historical oil price data //Energy Economics. – 2014. – V. 45. – P. 511-516.

47. *Gnilitskaya E. V.* Main trends in the development of the oil and gas sector of Russia in the context of globalization of the world economy // Russian entrepreneurship. – 2002. – №. 1. – P. 2.

48. *Gorbunova O. I., Kanitskaya L. V.* Environmental management in oil and gas organisations of Russia: environmental responsibility rating // Bulletin of the Baikal State University. – 2017. – V. 27. – №. 3. – P. 366-371.

49. *Grishina V. T.* Modern marketing and new opportunities // Current problems of commerce, logistics and marketing in the context of transformation of the modern economy. – 2022. – P. 54-57.

50. *He K., Zhang X., Ren S., Sun J.* Deep residual learning for image recognition //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2016. – P. 770-778.

51. Head of the Analytical Center under the Government: SMEV has accumulated a technological lag [Electronic resource]. URL: https://www.cnews.ru/news/top/2018-11-20_glava_proektnogo_ofisa_tsifrovoj_ekonomiki_raskritikoval (date of access: 10.04.2019).

52. How bankruptcy and recognition of a debtor as bankrupt proceed: features of the procedure [Electronic resource]. URL: https://pravobez.ru/news/kak-prohodit-bankrotstvo-i-priznanie-dolzhnika-bankrotom-osobennosti-procedury.html (date of access: 25.03.2020).

53. *Ilysheva N. N., Savostina O. V., Moldobaev T. Sh.* Features of accounting for contracts in construction organisations according to the requirements of IFRS // Discussion. – 2018. – №. 3 (88). – P. 110-119.

54. Information on the amounts of debt for taxes, fees and insurance premiums, penalties and fines to the budgets of the budget system of the Russian Federation [Electronic resource]. URL: https://data.nalog.ru/opendata/7707329152-debtam/structure-20181201.xsd (date of access: 12.04.2024).

55. Information on the amounts of taxes and fees paid by the organisation in the calendar year preceding the year of posting the said information on the Internet in accordance with paragraph 1.1 of Article 102 of the Tax Code of the Russian Federation (for each tax and fee), excluding the amounts of taxes (fees) paid in connection with the import of goods into the customs territory of the EAEU, the

amounts of taxes paid by the tax agent, and the amounts of insurance premiums [Electronic resource]. URL: https://file.nalog.ru/opendata/7707329152-paytax/structure-20180110.xsd (date of access: 10.04.2024).

56. Information on the average number of employees of the organisation [Electronic resource]. URL: https://file.nalog.ru/opendata/7707329152-sshr2019/structure-20200408.xsd (date of access: 11.04.2024).

57. Integration of information from the Unified State Register of Legal Entities and the Unified State Register of Individual Entrepreneurs into the information systems of interested parties [Electronic resource]. URL: https://www.nalog.gov.ru/rn77/service/egrip2/egrip_vzayim/ (date of access: 21.02.2024).

58. *Ivanova T. V.* Development strategies of integrated economic structures // Bulletin of Omsk University. Series "Economics". – 2013. – №. 3. – P. 65-69.

59. *Kapil G., Agrawal A., Khan R. A.* A study of big data characteristics //2016 International Conference on Communication and Electronics Systems (ICCES). – IEEE, 2016. – P. 1-4.

60. *Kaplunov D. R., Yukov V. A.* Energy saving in underground copper ore mining processes // Mining information and analytical bulletin (scientific and technical journal). – 2016. – №. 4. – P. 5-17.

61. *Karminsky A. M., Burekhin R. N.* Comparative analysis of methods for forecasting bankruptcies of Russian construction organisations // Business Informatics. – 2019. – V. 13. – №. 3. – P. 52-66.

62. *Kazakov A. V., Kolyshkin A. V.* Development of bankruptcy forecasting models in modern Russian conditions // Bulletin of St. Petersburg University. Economics. – 2018. – V. 34. – Ser. 2. – P. 241–266.

63. *Kologermanskaya E. M.* Legal regulation of the use of renewable energy sources in the Russian Federation and foreign countries: diss. - St. Petersburg State University, 2020.

64. *Kolokolova O. V.* Assessment of the probability of bankruptcy of borrowing enterprises based on cluster analysis // Economic analysis: theory and practice. – 2007. – №. 18. – P. 53-60.

65. *Kolyshkin A. V., Gilenko E. V., Dovzhenko S. E., Zhilkin S. A., Choe S. E.* Forecasting financial insolvency of enterprises // Bulletin of St. Petersburg University. – Ser. 5. – 2014. – V. 2. – P. 122–142.

66. *Kovalenkova O. G., Reznikov S. N., Mikhailyuk M. V.* Transformation of business models of food delivery in the system of logistics management of online trade // SCIENCE AND EDUCATION: ECONOMY AND ECONOMICS; ENTREPRENEURSHIP; LAW AND MANAGEMENT. – 2023. – P. 18.

67. *Kovalev V. V.* Accounting balance in historical retrospect // Bulletin of St. Petersburg University. Economics. – 2011. – №. 3. – P. 85-99.

68. *Kovalev V. V., Kovalev Vit. V.* Corporate finance. – 2019.

69. *Kovalev V. V., Moldobaev T. Sh.* Testing foreign and domestic models for forecasting bankruptcies at Russian enterprises // Development of territories. – 2021. – №. 3 (25). – P. 10-19.

70. *Kovalev V. V., Moldobaev T. Sh., Molitvin M. N., Suyazov V. V.* Analysis of the effectiveness of support programs for Russian universities (2010-2020) // Bulletin of St. Petersburg University. Economics. – 2022. – V. 38. – №. 2. – P. 208-234.

71. *Krizhevsky A., Sutskever I., Hinton G. E.* Imagenet classification with deep convolutional neural networks //Advances in neural information processing systems. – 2012. – V. 25.

72. *Kuznetsova M. N.* Increasing the economic efficiency of an enterprise through innovation // Bulletin of the Voronezh State Technical University. – 2012. – V. 8. – №. 4. – P. 80-83.

73. *Leonov D. Yu., Voronov V. I.* Logistics in the fuel and energy complex (using gas exports as an example) // Internauka. – 2017. – №. 4-2. – P. 21-24.

74. *Liu W., Fan H., Xia M.* Credit scoring based on tree-enhanced gradient boosting decision trees //Expert Systems with Applications. – 2022. – P. 189. – V. 116034.

75. *Lvova N.A.* Financial diagnostics of Russian enterprises using the Altman model for developed and emerging markets // Financial analytics: problems and solutions. – 2015. – №. 7 (241). – P. 37-45.

76. *Lvova O. A.* Possibilities and limitations of using bankruptcy diagnostic models to prevent insolvency // Bulletin of Moscow University. Series 6. Economics. – 2021. – №. 4. – P. 73-94.

77. *Martins L.* Challenges and opportunities for a successful mining industry in the future // Boletín geológico y Minero. – 2019. – P. 130. – №. 1. – V. 99-121.

78. *Medvedev M.A., Detkov A.A., Moldobaev T.Sh.* Analysis of the competitive environment in industry markets using information systems //AIP Conference Proceedings. – AIP Publishing, 2019. – V. 2172. – №. 1.

79. *Migunova Yu. V.* Quality of life of the population through the prism of consumer preferences // Bulletin of the Ufa Scientific Center of the Russian Academy of Sciences. – 2020. – №. 3. – P. 88-92.

80. *Miller A. E.* Research of alternative variants of integration interactions of production structures // Bulletin of Omsk University. Series "Economics". – 2013. – №. 3. – P. 105-113.

81. *Mnih V., Kavukcuoglu K., Silver D., Graves A., Antonoglou I., Wierstra D., Riedmiller M*. Playing atari with deep reinforcement learning //arXiv preprint arXiv:1312.5602. – 2013.

82. *Moldobaev T. Sh.* The impact of innovations on the efficiency of enterprises: industry and regional analysis // Economy and Entrepreneurship – 2022. – № 11 (148). – P. 494-498.

83. *Munirov E. D., Gimadiev S. I., Minulina O. V.* Features of demand analysis in the electricity market // Theoretical and applied aspects in the field of humanities. – 2023. – P. 118-119.

84. *Munshi A. Yu., Aleksandrova L. Yu., Munshi Sh. M.* Development of retail trade in Russia in the context of its digitalization // Economics, entrepreneurship and law. – 2022. – V. 12. – №. 5. – P. 1599-1612.

85. *Novikova N. G., Shchepina S. V.* Factors influencing the efficiency of inventory management in the context of managing the competitiveness of a trading organisation // Bulletin of the Baikal State University. – 2023. – V. 33. – №. 1. – P. 74-83.

86. *Odom M., Sharda R. A.* Neur al Network Model for Bankruptcy Prediction // Proceedings of the International Joint Conference on Neural Networks. San Diego. – 1990. 17–21 June. – V. II. – P. 163–171.

87. *Ohlson J.* Financial ratios and the probabilistic prediction of bankruptcy. Journal of Accounting Research –1980. – V. 18. – P. 109–131.

88. Open Data in the Russian Federation. Open Data Bulletin, July 2015 [Electronic resource]. URL: https://ac.gov.ru/files/publication/a/5572.pdf (date of access: 18.04.2023).

89. Open data of the Federal Service for Intellectual Property [Electronic resource]. URL: https://rospatent.gov.ru/opendata (date of access: 02.06.2023)

90. Open Data Portal of the Russian Federation [Electronic resource]. URL: https://data.gov.ru/ (date of access: 17.04.2024).

91. Order of Rosstat dated 29.03.2017 N 211 "On approval of the Regulation on the All-Russian Classifier of Enterprises and Organisations (OKPO) and classifiers related to it" [Electronic resource]. URL: https://www.consultant.ru/document/cons_doc_LAW_215015/ (date of access: 23.01.2024).

92. Order of the Federal Tax Service of Russia dated 29.06.2012 N MMB-7-6/435@ "On approval of the Procedure and conditions for assigning, using, and changing the taxpayer identification number" (Registered in the Ministry of Justice of Russia on 14.08.2012 N 25183) [Electronic resource]. URL: https://www.consultant.ru/document/cons_doc_LAW_134082/ (date of access: 17.01.2024).

93. Order of the Ministry of Finance of Russia dated 02.07.2010 N 66n (as amended on 19.04.2019) "On the forms of financial statements of organisations" (Registered in the Ministry of Justice of Russia on 02.08.2010 N 18023) (as amended and supplemented, entered into force with the reporting for 2020) [Electronic resource]. URL: https://www.consultant.ru/document/cons_doc_LAW_103394/b990bf4a13bd23fda86e0bba50c462a174c0d123/ (date of access: 15.01.2024)

94. Order of the Ministry of Finance of the Russian Federation dated 30.10.2017 N 165н (as amended on 19.12.2022) "On approval of the Procedure for maintaining the Unified State Register of Legal Entities and the Unified State Register of Individual Entrepreneurs, making corrections to information included in the records of the Unified State Register of Legal Entities and the Unified State Register of Individual Entrepreneurs on electronic media that do not correspond to the information contained in the documents on the basis of which such records are made (correction of a technical error), and on recognizing as invalid the order of the Ministry of Finance of the Russian Federation dated 18 February 2015 N 25н" (Registered in the Ministry of Justice of Russia on 16.01.2018 N 49645) [Electronic resource]. URL: https://www.consultant.ru/document/cons_doc_LAW_288080/ (date of access: 18.01.2024).

95. *Petrov A. M.* Advantages and disadvantages of coefficient analysis // Economic sciences. – 2020. – №. 182. – P. 144-145.

96. *Petrova A. V.* Factors influencing the formation of the enterprise’s product range // Economics and business: theory and practice. – 2021. – №. 4-2. – P. 74-77.

97. *Pidyashova O. P., Kravchenko T. E., Tereshchenko T. A.* Statistical analysis of the structure of population expenditures // Bulletin of the Academy of Knowledge. – 2019. – №. 3 (32). – P. 209-217.

98. *Popenkova D.K., Stukalova I.B.* Transformation of retail trade in the context of digitalization // Bulletin of the Altai Academy of Economics and Law. – 2022. – №. 2-1. – P. 89.

99. *Powers D. M. W.* Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation //arXiv preprint arXiv:2010.16061. – 2020.

100. Production of main types of products in physical terms since 2017 (current data in accordance with OKPD2) [Electronic resource]. URL: https://www.fedstat.ru/indicator/57783 (date of access: 17.01.2024).

101. Pyatov M. L. Evolution of accounting methodology within the framework of the balance sheet model of the firm // Bulletin of St. Petersburg University. Economics. – 2014. – №. 4. – P. 56-75.

102. *Rabiner L. R.* A tutorial on hidden Markov models and selected applications in speech recognition //Proceedings of the IEEE. – 1989. – V. 77. – №. 2. – P. 257-286.

103. *Razumov P. E.* Classification of enterprises by bankruptcy level using the cluster analysis method // Socio-economic and humanitarian journal of Krasnoyarsk State Agrarian University. – 2015. – №. 1. – P. 24-27.

104. Resolution of the Government of the Russian Federation of 08.02.2017 N 145 (as amended on 27.03.2023) "On approval of the Rules for the formation and maintenance in the unified information system in the field of procurement of a catalog of goods, works, services to meet state and municipal needs and the Rules for the use of the catalog of goods, works, services to meet state and municipal needs" [Electronic resource]. URL: https://www.consultant.ru/document/cons_doc_LAW_212534/ (date of access: 21.03.2024).

105. Resolution of the Government of the Russian Federation of 25.06.2003 N 367 "On approval of the Rules for conducting financial analysis by an arbitration manager" [Electronic resource]. URL: http://www.consultant.ru/document/cons_doc_LAW_42901/ (date of access: 11.05.2023).

106. Resolution of the State Standard of Russia dated 30.03.1999 N 97 (as amended on 22.09.2023) "On the adoption and implementation of All-Russian

classifiers" (together with "OK 027-99. All-Russian classifier of forms of ownership") (date of introduction 01.01.2000) [Electronic resource]. URL: https://www.consultant.ru/document/cons_doc_LAW_26587/ (date of access: 10.01.2024).

107. Review: Cunning and stagnation have prolonged bankruptcies [Electronic resource]. URL: https://fedresurs.ru/news/ccfc7b66-7065-4eef-9bc2-a2426e0e895d?attempt=1 (date of access: 12.04.2020).

108. Review: The number of corporate bankruptcies in Russia in 2018 decreased to a natural level [Electronic resource]. URL: https://fedresurs.ru/news/5e75d843-000b-4ce1-9925-9e02d89fa0e7?attempt=1 (date of access: 12.04.2020).

109. *Rumelhart D. E., Hinton G. E., Williams R. J.* Learning representations by back-propagating errors //nature. – 1986. – V. 323. – №. 6088. – P. 533-536.

110. *Saeed N., Husamaldin L.* Big data characteristics (V's) in industry //Iraqi Journal of Industrial Research. – 2021. – V. 8. – №. 1. – P. 1-9.

111. *Salchengerger L. M., Cinar E. M., Lash N. A.* Neural networks: A New Tool for Predicting Th rift Failures // Decision Sciences. 1992. N 23 (4). P. 899–916.

112. *Salomatin V. A., Salomatina E. V., Kalashnikov S. V.* Development of the tobacco industry in the context of external challenges // Economics, entrepreneurship and law. – 2021. – V. 11. – №. 5. – P. 1305-1316.

113. *Samuel A. L.* Some studies in machine learning using the game of checkers //IBM Journal of research and development. – 1959. – V. 3. – №. 3. – P. 210-229.

114. *Saraev A. A.* Restoration of the environment on the territory of former military bases in the Arctic // ARCTIC: innovative technologies, personnel, tourism. – 2018. – P. 491-495.

115. *Schwartz E. A., Knizhnikov A. Yu., Pakhalov A. M., Sheresheva M. Yu.* Assessment of environmental responsibility of oil and gas organisations operating

in Russia: a rating approach // Bulletin of Moscow University. Series 6. Economics. – 2015. – №. 5. – P. 46-67.

116. *Selvin S., Vinayakumar R., Gopalakrishnan E. A., Menon V. K., Soman K. P.* Stock price prediction using LSTM, RNN and CNN-sliding window model //2017 international conference on advances in computing, communications and informatics (icacci). – IEEE, 2017. – P. 1643-1647.

117. *Sezer O. B., Ozbayoglu A. M.* Algorithmic financial trading with deep convolutional neural networks: Time series to image conversion approach //Applied Soft Computing. – 2018. – T. 70. – P. 525-538.

118. *Sharokhina S.* System of performance indicators for logistics management of inventory of retail enterprises // University Bulletin. – 2023. – №. 2. – P. 138-145.

119. *Shemanueva A.P.* Strategic alliances in modern business // Service in Russia and abroad. – 2007. – №. 3. – P. 37.

120. Shipped (transferred) products in kind since 2017 (current data in accordance with OKPD2) [Electronic resource]. URL: https://www.fedstat.ru/indicator/57786 (date of access: 16.01.2024).

121. *Sokolov M. M.* Prospects for oil consumption in the world and the dynamics of world prices for it // Bulletin of the Institute of Economics of the Russian Academy of Sciences. – 2019. – №. 4. – P. 108-124.

122. *Sokolova N.A.* Problems of determining doubtful accounts receivable based on solvency analysis // Finance and Business. 2011. № 2. P. 174-184.

123. *Sorokin K. V., Popova D. S.* Review of international classifications used in enterprise statistics for organizing accounting by separate divisions // International Journal of Humanities and Natural Sciences. – 2023. – №. 12-4 (87). – P. 199-202.

124. *Spitsyn V. V., Spitsyna L. Yu.* Stability, efficiency and effectiveness of enterprises in high-tech industries in times of crisis: econometric modeling // Bulletin of NSUEM. – 2020. – №. 2. – P. 160-176.

125. *Springate, G. L.* Predicting the possibility of failure in a Canadian firm // Unpublished MBA project, Simon Fraser University. 1978.

126. State Register of Accredited Branches and Representative Offices of Foreign Legal Entities (RAFP) [Electronic resource]. URL: https://data.nalog.ru/opendata/7707329152-rafp/structure-16032022.xsd (date of access: 27.03.2024).

127. *Surinov A.E.* Digital economy: challenges for Russian statistics // Questions of statistics. – 2018. – T. 25. – №. 3. – P. 3-14.

128. *Sutskever I., Vinyals O., Le Q. V.* Sequence to sequence learning with neural networks //Advances in neural information processing systems. – 2014. – V. 27.

129. *Syrovetsky V. A.* Modern development of transport and logistics infrastructure of the regions of the Far North, Siberia and the Far East // Logistics and supply chain management. – 2020. – P. 151-157.

130. Taffler R. J., Tisshaw H. Going, Going, Gone – Four Factors which Predict // Accountancy. 1977. Vol. 3. P. 50–54.

131. Transition keys between the Commodity Nomenclature of Foreign Economic Activity of the Eurasian Economic Union TN VED EAEU and the All-Russian Classifier of Products by Types of Economic Activity OK 034-2014 OKPD 2 [Electronic resource]. URL: https://economy.gov.ru/material/file/8fe3bac6d1fec0a3b943272e28212592/%D0%A2%D0%9D%D0%92%D0%AD%D0%94_%D0%9E%D0%9A%D0%9F%D0%942_20_07_2023.xlsx (date of access: 29.01.2024).

132. Unified Federal Register of Bankruptcy Information [Electronic resource]. URL: https://bankrot.fedresurs.ru/DebtorsSearch.aspx (date of access: 28.03.2024).

133. Unified register of small and medium-sized businesses [Electronic resource]. URL: https://file.nalog.ru/opendata/7707329152-rsmp/structure-10062023.xsd (date of access: 27.03.2024).

134. *Vapnik V. N., Chervonenkis A. Ya.* On the uniform convergence of relative frequencies of events to their probabilities // Probability Theory and Applications. – 1971. – V. 16. – №. 2. – P. 264-280.

135. *Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A. N., Polosukhin I.* Attention is all you need //arXiv preprint arXiv: 1706.03762. – 2017.

136. *Vyakina I. V., Garanikova L. F.* Capital intensity and investment resources of the extractive industry of the Russian Federation // Bulletin of the Tver State Technical University. Series: Social Sciences and Humanities. – 2015. – №. 2. – P. 129-138.

137. Weekly average consumer prices (tariffs) for individual goods and services [Electronic resource]. URL: https://fedstat.ru/indicator/37426 (date of access: 21.02.2024).

138. *Wixom B., Ariyachandra T., Douglas D., Goul M., Gupta B., Iyer L., Turetken, O.* The current state of business intelligence in academia: The arrival of big data //Communications of the Association for information Systems. – 2014. – V. 34. – №. 1. – P. 1.

139. *Yao L. M., Gabdulkhakova O. I., Anisimova E. V.* Social and environmental responsibility of enterprises engaged in the extraction and processing of hydrocarbon raw materials // Bulletin of the Kazan Technological University. – 2011. – №. 24. – P. 189-195.

140. *Zabortseva T. I., Kuznetsova A. N., Violin S. I., Sysoeva N. M.* Potential of the "green" economy in the socio-economic development of the Irkutsk region // Geography and natural resources. – 2017. – №. 4. – P. 154-161.

141. *Zharsky D. P., Kurazhenkov M. S., Akifev A. A.* Assessment of the financial condition of a debtor organisation taking into account industry specifics // Property relations in the Russian Federation. – 2023. – №. 5 (260). – P. 26-34.

142. *Zhukinsky A. A.* The impact of breakthrough technologies on the oil and gas industry // Novelty. Experiment. Traditions. – 2022. – T. 8. – №. 3 (19). – P. 29-34.