




Article

Reliability and Accuracy of Alternative Default Prediction Models: Evidence from Slovakia

Daniela Rybárová¹, Helena Majdúchová¹, Peter Štetka^{1,*}  and Darina Luščíková²

¹ Faculty of Business Management, University of Economics in Bratislava, Dolnozemska cesta 1/B, 85235 Bratislava, Slovakia; daniela.rybarova@euba.sk (D.R.); helena.majduchova@euba.sk (H.M.)

² Independent Researcher, Dostojevského 12, 05801 Poprad, Slovakia; darinaluscikova@gmail.com

* Correspondence: peter.stetka@euba.sk

Abstract: The aim of this paper is to assess the reliability of alternative default prediction models in local conditions, with subsequent comparison with other generally known and globally disseminated default prediction models, such as Altman's Z-score, Quick Test, Creditworthiness Index, and Taffler's Model. The comparison was carried out on a sample of 90 companies operating in the Slovak Republic over a period of 3 years (2016, 2017, and 2018) with a narrower focus on three sectors: construction, retail, and tourism, using alternative default prediction models, such as CH-index, G-index, Binkert's Model, HGN2 Model, M-model, Gulka's Model, Hurtošová's Model, Model of Delina and Packová, and Binkert's Model. To verify the reliability of these models, tests of the significance of statistical hypotheses were used, such as type I and type II error. According to research results, the highest reliability and accuracy was achieved by an alternative local Model of Delina and Packová. The least reliable results within the list of models were reported by the most globally disseminated model, Altman's Z-score. Significant differences between sectors were identified.

Keywords: financial distress; financial health; default models; credit risk; error rate; reliability



Citation: Rybárová, Daniela, Helena Majdúchová, Peter Štetka, and Darina Luščíková. 2021. Reliability and Accuracy of Alternative Default Prediction Models: Evidence from Slovakia. *International Journal of Financial Studies* 9: 65. <https://doi.org/10.3390/ijfs9040065>

Academic Editors: Tihana Škrinjaric and Sabri Boubaker

Received: 31 August 2021

Accepted: 25 November 2021

Published: 30 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In the current turbulently changing and uncertain economic environment, demand on the reliability and validity of the tools used to predict companies' financial health and financial distress is growing. Using default prediction models to assess companies' financial situation, and to separate prosperous from non-prosperous companies, or solvent from insolvent ones, has become the standard. Default prediction models are primarily based on the information provided by financial statements, transformed into ratios. Financial statements are the basic, and sometimes the only, available source of information. According to [Zalai et al. \(2016\)](#), it is evident that the indicators of successful companies differ significantly from indicators of non-prosperous companies several years before bankruptcy, both at the absolute level and with regards to annual change. Default prediction models aim to classify companies as future prosperous (solvent) or non-prosperous (insolvent) companies ([Zalai et al. 2016](#)).

To assess the financial health or financial distress of companies, several well-established default prediction models are being used on a global scale. These well-known models, such as Altman Z-score ([Altman 1983](#)) or Quick Test ([Kralicek 1993](#)), are applied in different local conditions, regardless of the specifics of the environment in which they originated and the differences between the original business environment and the environment in which they are applied. Yet, several alternative models are emerging in different regions and localities, aiming to reflect the given market's specifics and thus increase the reliability and decrease the error rate of the company's financial health assessment, e.g., [Chrastinová \(1998\)](#), [Delina and Packová \(2013\)](#), [Gulka \(2016\)](#), and [Hurtošová \(2008\)](#). Authors of this research paper are focused on testing the reliability and error rate of such alternative models in local conditions and comparing this reliability with the reliability of the most

used standard prediction models in specific local conditions. When selecting a locality, the criterion of a high concentration of alternative models in the given locality was applied, i.e., a high number of alternative prediction models compared to the size of the market in which they are being applied. From the spectrum of European national markets, the Czech and Slovak Republics were chosen as transition economies, in which several default prediction models have been published. In both countries, it is possible to identify a similar structure of financial accounting, which allows the comparability of economic content of indicators applied in default prediction models, comparable capital structure of companies due to the absence of developed capital market, as well as similar legislation governing corporate bankruptcy. These attributes are similar in several transition economies (such as Poland, Hungary, Russia, Ukraine, the Baltic countries, etc.) because they resulted from the same economic system. The predominant legal form of businesses in the Slovak Republic are limited liability companies and joint stock companies which are required by law to book the share capital, which largely affects the equity structure. Most of these entities are required to prepare financial statements in accordance with the Accounting Act (NC 2002) and further detailed accounting procedures with a clear definition of the content of each accounting item, which differ to some extent from accounting items presented in other accounting standards, e.g., IFRS or US GAAP. A specific element included in Slovak legislation is the legal regulation of the company in crisis, which is defined as the ratio of equity to liabilities lower than 0.08 (NC 1991).

Eight alternative default prediction models were identified in Slovak market, while IN group models were taken from the Czech market and were tested on a sample of 90 companies from three different industries. The reliability and error rate of these models were evaluated and compared to the reliability and error rate of standard prediction models such as Altman's Z-score, Quick Test, Binkert's Model, Creditworthiness Index, and Taffler's Model. Research has shown high reliability of some alternative default prediction models, which were designed in local conditions. The highest reliability and accuracy was achieved by an alternative local Model of Delina and Packová. However, it should be emphasized that a large group of unreliable models within the group of alternative default prediction models were identified. This subset of models was excluded from further testing. The least reliable results within the final list of models were reported by the most globally disseminated model, Altman's Z-score. In contrast, the Quick Test, as one of globally well disseminated models, achieved a high ranking as it reported low error rates, for both type I and type II errors. While Index IN05, one of the most applied models in Slovak Republic, showed its ability to correctly classify non-prosperous companies, the overall rating of this model was low as it reported a high type II error rate. Significant differences in reliability were identified across the spectrum of alternative default prediction models, but also across economy sectors.

The following section focuses on a literature review of default prediction models and an overview of previous research in the field of evaluating the reliability of these models. The next section is dedicated to research design, where the research goals, methods, and procedures are described. The main research results are presented in Section 4, separately for each category of companies. The research results are supplemented by general findings. Partial analytical results are listed in appendices. These results are further discussed in Sections 5 and 6 draws conclusions.

2. Literature Review

The scientific literature provides a very wide range of default prediction models for assessing the financial health and financial distress of companies. The reliability of these models is the subject of research of many scientists. In the following section, the latest global scientific efforts to test the reliability of default prediction models are summarized. Given the main goal of this research, this paper is primarily focused on local models created in the Slovak and Czech Republic, which are not generally known. Subsequently, selected globally disseminated models are briefly described, the reliability of which was further

compared with alternative default prediction models. At the end of this section, results of previous studies are summarized.

In the CEE (Central and Eastern Europe) region, due to the geopolitical situation and the introduced economic system, the issue of default predictions started to be the subject of research starting in 1990s. Prusák (2018) analysed the level of advancement in these countries and came to conclusion that the most advanced research in this area is conducted in Visegrad countries (Czech Republic, Slovak Republic, Poland, Hungary), Estonia, and Russia. Kristóf and Virág (2020) conducted a comprehensive analysis based on 30 years of Hungarian empirical results and concluded that, considering the validity of a key theoretical finding that no bankruptcy prediction model might function independently of time, space, and economic environment, it is not recommended to apply bankruptcy models on Hungarian companies that were developed on a foreign sample. This conclusion was confirmed by Singh and Mishra (2016), based on their research on a sample of Indian companies. Korol (2019) applied logistic regression and multilayer perceptron to predict bankruptcy using tax arrears' information and came to conclusion that models created indicate that shortly before bankruptcy, tax arrears' models outrun the financial ratio-based models in terms of accuracy. The accuracy reduces when further periods before bankruptcy declaration are considered. The highest accuracy is obtained by using tax arrears and financial ratios simultaneously.

There are two parallel sources of default prediction models being used in the Slovak Republic, i.e., widely disseminated models and alternative models considering the specifics of the Slovak business environment. The following Table 1 presents a basic overview of default prediction models published by Slovak authors.

Table 1. List of local default prediction models.

Title	Year of Origin	Author	Method, Characteristics	Sector	Size, Legal Form	Number of Companies
CH-index ¹	1998	Chrastinová, Z.	Multidimensional discriminant analysis, bankruptcy model	agriculture	unspecified	unspecified
Model of Binkert ²	1999	Binkert, Ch.	Multidimensional discriminant analysis, financial situation forecasting model	production, services, trade	joint-stock companies	1550
Model of Hurtošová ³	2009	Hurtošová, J.	Logistic regression, rating model	non-financial SME's	excluded very small and very large companies	427 from the bank's client database
G-index ⁴	2002	Gurčík, L.	Multidimensional discriminant analysis	agriculture	unspecified	60
Model of Delina and Packová ⁵	2013	Delina, R.; Packová, M.	Regression analysis, bankruptcy model	unspecified	unspecified	1560, of which 103 bankrupt
HGN Model ⁶	2014	Hyránek E.; Grell, M.; Nagy, L.	Multidimensional discriminant analysis, creditworthy model	production, services, trade	regardless of the legal form	260

Table 1. Cont.

Title	Year of Origin	Author	Method, Characteristics	Sector	Size, Legal Form	Number of Companies
M-model ⁷	2014	Harumová, A.; Janisová, M.	logistic regression, scoring model	unspecified	small and medium-sized companies	53,206 company data for 3 years
Model of Gulka ⁸	2016	Gulka, M	logistic regression	all non-financial corporations	unspecified	unspecified

Source: Own processing according to ¹ Chrastinová 1998; ² Zalai 2000; ³ Hurtošová 2008; ⁴ Gurčík 2002; ⁵ Delina and Packová 2013; ⁶ Hyránek et al. 2018; ⁷ Harumová and Janisová 2014; ⁸ Gulka 2016).

A more detailed look at the construction of alternative default prediction models and the criteria used to evaluate the financial health and financial distress of companies is listed in Appendix A (Table A1).

The most frequently used default prediction models in Czech Republic are IN group models (Index IN95, Index IN99, Index IN01, and Index IN05), created by Neumaierová and Neumaier (2005), adapted specifically to the local conditions of the Czech market. The construction base of these models is similar to the Altman's Z-score. It contains standard ratios measuring the activity, profitability, financial leverage, and liquidity. Over the 50 years since the publication of the Z-score model for predicting firm financial distress and bankruptcy, many practitioner applications have emerged, not to mention the enormous number of scholarly works that have utilized the Altman model as a reliable and easily replicable benchmark (Altman 2002, 2018; Altman et al. 2017). The subject for testing was the latest version of these models, i.e., Index IN05.

Various Czech and Slovak agencies and credit risk advisors dealing with the issue of financial health and financial distress of companies operating in local conditions are usually using globally disseminated default prediction models, such as Altman's Z-score, Quick Test, Creditworthiness Index, Binkert's Model, and Taffler's Model. A more detailed look at the construction of globally disseminated default prediction models and the criteria used to evaluate the financial health and financial distress of companies is listed in Appendix B. These models, together with IN group models, were also the most frequently tested models in various publications discussed below.

Sušický (2011) investigated in his work the hypotheses that Czech bankruptcy models are not more successful than foreign bankruptcy models when applied to companies operating in local conditions, and whether it is possible to choose the so-called universal bankruptcy model that would be always successful in predicting the bankruptcy of a company, regardless the industry in which the company operates. This author tested Altman's Z-score, ZETA, IN99, IN01, IN05, and Taffler's Model on companies in five Czech economy sectors (agriculture, food industry, motor vehicles, metal structures, energy). The data analysis confirmed the hypothesis that Czech bankruptcy models are not more successful in predicting companies' financial health in local Czech conditions than foreign bankruptcy models. The hypothesis that there would be a bankruptcy model that would always be successful regardless the industry was not confirmed.

Delina and Packová (2013) verified in their research the explanatory value of Altman's model, Creditworthiness Index, and IN05 Index. This research was conducted on a sample of 1560 companies operating in Slovak conditions. The authors state that out of the analysed number of 1560 companies, 103 companies went bankrupt, representing 8.33% of the total sample. By testing accuracy, they found that Altman's Z-score showed the worst results in all categories (first-degree error, accuracy). Based on their research, Delina and Packová concluded that the results were not satisfactory, and these models are not suitable for testing in the Slovak business environment.

Machek (2014) analysed Kralick's Quick Test, Taffler's Model, IN99, and IN05, Altman's Z-score and the prediction ability of these models in the conditions of Czech compa-

nies with more than 10 employees for the period of 2007 to 2012. He found cases in which, based on the examined models' results, it was possible to correctly predict companies' default up to 5 years in advance. He found that IN05 and IN99 gave the best results, as did Altman's Z-score. The ability to predict default using Taffler's Model and Kralick's Quick Test was limited.

Čámská (2014) mapped individual models and approaches for predicting corporate bankruptcies, which were preferentially used, or which are currently used in the Czech Republic and similar transition economies, whether created in these economies or taken over from advanced economies. Testing of 40 models was carried out in three sectors: metal manufacturing and metalworking products, except machinery and equipment (CZ-NACE 25); machinery and related equipment manufacturing (CZ-NACE 28); and construction (CZ-NACE F). Čámská classified these models as successful and unsuccessful based on quantifying 1st and 2nd type errors and subsequently further evaluated the compliant models based on the ROC curve and its corresponding AuROC coefficient. Testing models in different sectors of the economy has shown different reliability within different sectors. Čámská and Klečka (2020) published the research results comparing the partial results on a same sample of companies from the time of the recession in 2012 to the results of 18 models in 2017, the time of expansion. The research results confirmed the expectation that companies generally achieve better financial scores in times of economic expansion, both in the group of solvent and insolvent companies. Therefore, it is unnecessary to insert macroeconomic indicators directly into the models. However, the information users should assess financial health in the context of overall economic conditions, including macroeconomic and industrial development, esp. the phases of the economic cycle.

Režňáková and Karas (2015) presented the research results of the Altman model's discriminative ability. They dealt with the applicability of models that originated in another environment or in a different period. Testing the Altman model's accuracy on a sample of data from industrial companies operating in Visegrad 4 countries showed that the original model works with lower statistical accuracy than stated by the author and with a high proportion of unclassified companies. The researchers also found that by re-evaluating the weight of each model's coefficients and the grey zone boundary based on conditions specific to each V4 country while maintaining the model variables, it is possible to increase the distinctive ability of the model, i.e., the ability to distinguish failing businesses from prosperous ones correctly. Režňáková and Karas also considered the differences in financial statements and the methodology of financial reporting in V4 countries to be the cause of the presented research results. These differences in the structure of financial statements are interrelated with key macroeconomic indicators, such as interest rate, level of tax rates, wage level, access to capital markets, etc.

A study of Gavurová et al. (2017), in which the authors focused on the evaluation of four bankruptcy prediction models in the Slovak business environment, also compared the most frequently used models. The following models were verified on a sample of 700 Slovak companies: Altman's Model, Ohlson's Model, IN01, and IN05 Index. According to authors, the uniqueness of their research lies in the validation and evaluation of the accuracy of the bankruptcy models at three levels: overall accuracy, the accuracy of the bankruptcy prediction and the accuracy of the prospecting company's forecast, thus ensuring greater objectivity in evaluating these models. Gavurová et al. (2017) considered IN05, which proved the highest prediction ability of bankruptcy, to be the most suitable model for local conditions.

Valášková et al. (2017) tested models created for the agricultural sector. The authors compared CH-index and G-index models, which were defined for Slovak agricultural companies, with Altman's Z-score. Only the G-index, which complied with the tested sample of agricultural companies, passed this test, as it was the only one to reach the limit value for assessing the classification ability and prediction accuracy of models (Valášková et al. 2017).

Bakeš and Valášková (2018) verified the prediction models created in Slovakia, namely: Ch-index, Binkert's Model, G-index, Hurtošová's Model, Delina and Packová's Model, and Gulka's Model—on a sample of 266 companies from a selected economic sector (the article does not specify which one). The authors used the criteria set by the Commercial Code (a company in crisis) and the Bankruptcy and Restructuring Act, as well as economic criteria—the same approach as Valášková et al. (2017) applied in previous research, to select companies and classify them as prosperous and non-prosperous. According to the above evaluation, Model proposed by Delina and Packová and Binkert's Model achieved an ideal classification ability, as they were able to determine the company's prosperity with an accuracy of 98.3%. Research of Kováčová and Kubala (2018), working at the same faculty, was carried out on the same basis. To define a non-prosperous company, they used the same basis as the authors in the last two types of research (described above). For testing, they selected models using discriminant analysis: Altman's Z-score (Altman 2018), Creditworthiness Index, IN05 Index, Taffler's Model (Taffler 1983), CH-index, G-index, and tested them in the conditions of the Slovak Republic. The models' verification was carried out on a sample of 400 prosperous companies and the same number of non-prosperous companies. The source of data was the database of financial statements reported in 2016. A different sample of the assessed companies (despite the same criteria of the initial classification of companies and the same research methodology) caused that in the case of the Ch-index and G-index, significantly different prediction abilities were found.

The following research from the same faculty and period can be assigned to both previous studies, as it used the same parameters to classify the input sample of companies and a similar methodology, in which errors of type I and II were assessed. The research was carried out by Siekelová et al. (2018) based on the data from financial statements for the accounting periods 2016 and 2017 on a sample of 500 companies, from the following sectors: wholesale and retail, repair of vehicles and motorcycles (125 companies), construction (101 companies), industrial production (83 companies), accommodation and food services (69 companies), information and communication (55 companies), and other (67 companies). Based on the defined conditions, out of a sample of 500 companies, 383 companies were classified as prosperous and 117 companies as non-prosperous. Following models were tested: Model of Jakubík and Teplý, Hurtošová's Model, Gulka's Model, and IN05 Index. The IN05 Index (88.40%), Hurtošová's Model (79.80%), and the Model of Jakubík and Teplý (79.60%) had the highest prediction ability, while the Gulka's Model (76.80%) had the lowest, but still acceptable, prediction ability. The study did not describe the results of investigated models' prediction ability for each industry separately.

One of the main determinants causing differences in the results of prediction models testing is an incomparable set of tested non-prosperous companies, created by differently set criteria for selecting these companies. In their research, many authors consider non-prosperous companies to be companies that are in bankruptcy based on the rules of the relevant legislation. Many authors considered in the comparative sample non-prosperous companies to be companies that were declared bankrupt, e.g., Čámská (2014), under Czech Act no. 182/2006 Coll. on Bankruptcy and its Resolution (Insolvency Act), as well as Delina and Packová (2013), Gulka (2016), Mihalovič (2018) under Slovak Act no. 7/2005 on bankruptcy and restructuring. When testing the models, Sušický (2011) considered not only the declaration of bankruptcy as a state when the company becomes insolvent as an essential base for classifying companies into prosperous and non-prosperous, but he also included companies in liquidation, companies in preliminary administration, companies in settlement and companies under sequestration (according to the Czech legislation). It is necessary to understand that the legislation governing the bankruptcy differs from country to country, which means that the companies' classification based only on legislative criteria without using relevant economic indicators necessarily leads to different results.

3. Research Design

The aim of this paper is to assess the reliability of alternative default prediction models in local conditions, with subsequent comparison with other generally known and globally disseminated default prediction models. The partial goals supporting the fulfilment of the main goal were:

- compare the results of locally originated default prediction models with globally accepted models,
- compare the explanatory value of these models among selected sectors,
- evaluate the applicability of examined models for predicting companies' financial distress, considering the sector of economy.

The explanatory value of the selected default prediction models was tested on a specific sample of 90 companies over the period of 3 years (2016, 2017, 2018). Due to the models' inability to consider the systematic risk and the occurrence of extraordinary events, such as the COVID-19 pandemic and the related structural changes in economy, the pre-pandemic research period was selected. No significant deviations were observed during the period considered. The FinStat and Credit Bureau Slovakia databases were used to select a sample of companies. The essential selection of the sample was carried out according to precisely defined criteria:

- SME's (small and medium-sized companies) defined in terms of categorising business entities according to the size following the European Commission's recommendation 2003/361/EC effective from January 2005, (less than 250 employees, and turnover less than EUR 50 million);
- companies accounting in the double-entry bookkeeping system following Act no. 431/2002 Coll. on Accounting and MEASURES of the Ministry of Finance of the Slovak Republic of 16 December 2002 no. 23054/2002-92 specifying details of accounting procedures and a general chart of accounts for businesses accounting in the double-entry bookkeeping system;
- companies with the legal form of a limited liability company.

Regarding the accuracy and purpose of the research, the first selection of sectors was conducted based on the general relevancy of the sector, especially the size of businesses operating in sector, employment, share of GDP, and a higher risk of insolvency and bankruptcy. The final selection was based on the following criteria: number of enterprises in the sector, the average share of liabilities, and the survival rate (the share of companies operating in the sector at least for 5 years on the total number of entities established at the initial period). The highest number of enterprises was recorded in the services sector (46.10%), retail sector (19%) and construction sector (17%). The highest share of liabilities in 2018 was recorded in retail sector (65.16%) and construction sector (54.80%). The lowest survival rate (SBA 2019) was recorded in construction sector (37.4%) and services sector (41.2%). Based on these results, three sectors were selected, namely construction sector, retail sector, and services. Services sector contains a wide range of business activities. It was therefore necessary to reduce the internal heterogeneity by focusing on a narrower range of business activities. Out of the services sector, subset of companies operating in tourism was selected.

To create a sample of companies, the following criteria were applied:

- Annual sales from 30 thousand up to 50 mil. EUR. The intention of setting the lower limit was to eliminate non-operating companies and very small companies. The upper limit represents the constraint for classifying a company as SME (small and medium-sized enterprise).
- Type of ownership—private domestic.
- Date of the business establishment no later than 31 December 2014. The period considered in this research was three accounting periods (2016, 2017, and 2018), during which the stability of the company was assessed, by analysing continuous development of sales and economic results. Start-ups could report very specific results

during the initial year. This factor could greatly distort the research results. Therefore, only companies operating more than 1 year were assessed. The other reason for setting this criterion is the definition of a company in difficulty (EC 2014a), according to which businesses that exist less than three years are excluded from this category. The research period (2016, 2017, and 2018) was not affected by the systematic risk, the source of which was the global pandemic.

- Number of employees up to 250. This criterion represents an upper limit for classifying a company as SME.

Within the focus and needs of research, companies meeting the above criteria were divided into three groups, reflecting the local legislation, namely non-prosperous companies, companies in difficulty, and prosperous companies. It was crucial to properly define companies in difficulty. For this purpose, following legislation was applied: (EC 2014a) Communication from the Commission—Guidelines on State aid for rescuing and restructuring non-financial undertakings in difficulty (2014/C 249/01), and (EC 2014b) Commission Regulation (EU) No 651/2014 of 17 June 2014 declaring certain categories of aid compatible with the internal market in application of Articles 107 and 108 of the Treaty Text with EEA relevance. Based on this legislation, non-prosperous companies are companies with negative equity, companies in difficulty are companies with equity less than half of a share capital as the amount of money invested by its owners in exchange for shares of ownership. Prosperous companies could be therefore defined as companies with the equity higher than a half of share capital.

Based on the above described procedure, prosperous from non-prosperous companies were separated. However, applying this procedure a target sample of prosperous companies has not been created yet. Defining prosperous companies as companies with equity higher than a half of their share capital cannot be considered as sufficient. According to Brealey et al. (2016), liquidity and profitability are the two basic financial goals of a company, which in mutual interaction demonstrate the solvency and the ability to achieve financial results. Therefore, the second level liquidity as the first indicator was selected. This indicator is calculated as the sum of financial accounts and short-term receivables divided by short-term liabilities. This indicator is also known as quick liquidity. It expresses the ability to cover short-term liabilities of the company with financial accounts and short-term receivables (Zalai et al. 2016). The median of this indicator ranged in 2018 from 0.88 in tourism sector to 1.32 in construction sector. The lower limit of the interval for the 2nd level liquidity was therefore set to 1.0. As the profitability indicator, the Return on Sales (ROS) was applied, calculated as EBIT divided by total sales. The acceptance interval was chosen at the level of the upper quartile for each sector separately (3.78% in tourism sector, 4.01% in construction sector, and 9.22% in retail sector). A random sample was made from such a narrowed sample of prosperous companies.

From the processed data of a total sample of 11,168 small and medium-sized enterprises operating in the tourism, construction, and retail sector, which were active in 2018 (reported sales higher than 30,000 EUR), up to 23.13% of companies reported negative equity. Most companies with negative equity occurred in 2018 in the tourism sector—up to 35.07%. The construction sector reported the lowest share of companies with negative equity—up to 15.70%. The finding of such a high share of companies with negative equity operating in these sectors is alarming. This finding emphasizes the need for reliable and accurate default prediction models that could be applied in local conditions. The research design respects the fact that companies, especially limited liability company, can be set up with almost no equity. The selected sample of non-prosperous companies includes only those companies, which reported positive equity in 2016 and 2017, but the declining trend in the value of this indicator resulted in negative equity in 2018. The fact that only those companies that were established by the end of 2014 at the latest were selected is not affected. A random sample was applied from such a narrowed sample of non-prosperous companies.

The structure of total sample gathered applying this criteria for the 2018 year is presented in Table 2.

Table 2. Share of prosperous and non-prosperous companies in selected sectors in 2018.

Economy Sectors	Business Categories	Initial Categorization (Legislative Criteria) in 2018 [n]	Final Categorization (Passing Additional Criteria) in 2018 [n]
Tourism sector	Non-prosperous companies	955	158
	Companies in difficulty	158	158
	Prosperous companies	1610	399
Construction sector	Non-prosperous companies	758	165
	Companies in difficulty	154	154
	Prosperous companies	3917	881
Retail sector	Non-prosperous companies	870	45
	Companies in difficulty	85	85
	Prosperous companies	2661	496
Total	Non-prosperous companies	2583	368
	Companies in difficulty	397	397
	Prosperous companies	8188	1776

Source: Own processing; based on data extracted from the FinStat database.

Selecting a proper sample of companies by applying above stated selection criteria, thus ensuring consistent sample within the category and significant differences between categories, was a prior interest. Resulting sample may not be large and therefore could be considered as a research limitation. 270 observations (90 companies \times 3 years) were applied, resulting in 3 510 estimations (13 models \times 270 observations).

The research methodology offers various ways of compiling a sample (Tomšik 2017). A multistage selection was applied, based on a hierarchical description of the elements of the base set. These elements were specified by gradual selections through higher selection units. Gradual selections were compiled using cluster selection. Clusters were designed to be homogeneous within and heterogeneous between groups in relation to testing the reliability and accuracy of default prediction models. The generalizability of research results is linked to this deliberate sample.

Based on the success rate of each default prediction model to correctly classify non-prosperous companies within the applied sample (1 minus type I Error rate), the confidence interval of the model for the base set was determined. If significance tests are available for general values of a parameter, then confidence intervals can be constructed by including in the 100% confidence region all those points for which the significance test of the null hypothesis that the true value is the given value is not rejected at a significance level of $(1 - p)$ (Cox and Hinkley 1974).

The research was focused on testing the reliability of default prediction models created in local conditions using statistical methods. The focus was on comparing alternative default prediction models which originated in local conditions with each other and further comparing these alternative models with well-known and globally used models.

The models' quality was verified using tests of the significance of statistical hypotheses and quantification of type I and II errors. Neyman and Pearson (1928) identified two sources of errors, the error of rejecting a hypothesis that should have not been rejected, and the error of failing to reject a hypothesis that should have been rejected. There are errors (I) we reject H_0 [i.e., the hypothesis to be tested] when it is true, and (II) we fail to reject H_0 when some alternative hypothesis H_1 is true. Neyman and Pearson (1933) call these two kinds of errors type I and type II error.

This method is often used to verify the explanatory value of scoring, credit risk, and bankruptcy models. The following test results occur when verifying statistical hypotheses (Kováčová and Kubala 2018):

- True Positive, TP—the test result matches the correct result; there is a positive match in this case. This means that the company was correctly classified as healthy.
- True Negative, TN—the test result matches the correct result; there is a negative match in this case. This means that the company was correctly classified as unhealthy.
- False Positive, FP—the company was incorrectly classified as healthy and belonged to the unhealthy group.
- False Negative, FN—the company was incorrectly classified as unhealthy and belonged to the healthy group.

The results of the model estimates are visually interpreted as shown in the Table 3.

Table 3. Contingency table showing the number of positive and negative tests and type I and II errors.

		Estimates (Results Provided by Model)	
		Healthy	Unhealthy
Reality	Healthy	True Positive (TP)	False Negative (FN)—type II error
	Unhealthy	False Positive (FP)—type I error	True Negative (TN)

Source: Own table processed according to (Lee et al. 2009).

The main measures used to evaluate default prediction models are accuracy and F1 score.

Accuracy is a measure of all correctly identified cases, most used when all classes are equally important. It is calculated as the ratio of correctly predicted classifications to the total set of test data. The numerical value of accuracy represents the proportion of true results (both true positive and true negative) in the selected population (Zhu et al. 2010).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

All tested observations are the sum of true positive, true negative, false positive, and false negative cases. It is expressed on a scale of 0 to 1, with higher meaning better. Accuracy adj., also known as Accuracy_{chance} or Cohen’s Kapa, was used to interpret the results, applying the following calculation.

$$Accuracy\ adj. = \frac{Accuracy - Pe}{(1 - Pe)}, \tag{2}$$

where (Cohen 1968) *Pe* is the probability of a chance detection for the given data set. *Pe* can be interpreted as normalizing the accuracy from the range [0,1] to [*p*_{chance} 1], where *p*_{chance} is the accuracy expected by random guess given a test subset (i.e., 1/M in an M-class classifier). *Pe* (Berthold et al. 2020) is the probability of the raters to agree by chance

$$pe = p_{pos} + p_{nes} \tag{3}$$

$$p_{pos} = \frac{TP + FN}{n} \times \frac{TP + FP}{n} \tag{4}$$

$$p_{nes} = \frac{TN + FP}{n} \times \frac{TN + FN}{n}, \tag{5}$$

when *n* is the number of instances in the data set.

Though accuracy provides a single simple number for diagnostic performance, it is often too simple and must be interpreted with considerable caution (Metz 1978).

The F1 score (Chicco and Jurman 2020) is the most used member of the parametric family of the F-measures, named after the parameter value $\beta = 1$. F1 score is defined as the harmonic mean of precision and recall. The calculation formula is as follows.

$$F1 = \frac{2 \times (\textit{precision} \times \textit{recall})}{\textit{precision} + \textit{recall}} \quad (6)$$

$$\textit{whereas for precision : } \textit{precision} = \frac{TP}{TP + FP}$$

$$\textit{and for recall : } \textit{recall} = \frac{TP}{TP + FN}$$

F1 ranges in $[0, 1]$, where the minimum is reached for $TP = 0$, that is, when all the positive samples are misclassified, and the maximum for $FN = FP = 0$, that is for perfect classification (Chicco and Jurman 2020).

A false positive test result means that the model identifies the bankrupt company as a healthy company, which in practical use causes risk and increased costs (bankruptcy, default, etc.) By taking completeness into account with the F1 score, the reported risk can be minimised.

The overall assessment of the models considers other statistical categories:

- The sensitivity (Steward 2019) of a test is also called the true positive rate (TPR) and is the proportion of samples that are genuinely positive that give a positive result using the test in question. The test that correctly identifies all positive samples is very sensitive. The test that detects only part of the positive samples would be deemed to have lower sensitivity as it is missing positives and giving higher a false negative rate (FNR). Referred to as type II errors, false negatives are the failure to reject a false null hypothesis (the null hypothesis being that the sample is negative).
- The specificity (Steward 2019) of a test, also referred to as the true negative rate (TNR), is the proportion of samples that test negative using the test in question that are genuinely negative. The test that identifies all healthy as being negative is very specific. The test that incorrectly identifies part of healthy panel as having the condition would be deemed to be less specific, having a higher false positive rate (FPR). Referred to as type I errors, false positives are the rejection of a true null hypothesis (the null hypothesis being that the sample is negative).

The test can be very specific without being sensitive, or it can be very sensitive without being specific. Both factors are equally important. A good test is a one that has both high sensitivity and specificity (Zhu et al. 2010). Sensitivity and specificity are conditions for usefulness. The higher the sensitivity and specificity, the smaller the number of false positive and false negative tests, the more useful the model.

Furthermore, this metrics (Trevethan 2017) should not be regarded as unquestionably valid and fixed attributes of a test: the values depend on how stringent the test is and the prevalence of the target condition in the sample of businesses used in the analysis. As a result of these complexities, it is sometimes necessary to examine the validity of measurement procedures within both the reference standard and it might also be necessary to question the stringency of the test and to ensure that there is a match between the samples that were used for assessing the test and the businesses subsequently being tested.

4. Results

On a sample of selected companies in the tourism, construction, and retail sectors, predictions using alternative default models were calculated, and companies were assigned the relevant classification zone using the original methodology developed by model's author. Results of Binkert's Model, Altman's (Revised) Z-score (Altman 2002), IN05 Index, Quick Test, Creditworthiness Index, and Taffler's Model were taken into research as presented by FinStat database and included in testing as well as alternative models—

Hurtošová's Model, Ch-index, G-index, Model of Delina and Packová, Model M, HGN model, and Gulka's Model.

The Accuracy Adj. model (Appendix C) and the success rate of the F1-score (Appendix D) were used as the primary measures for the assessment of default prediction models. The accuracy of the prediction adjusted for the effect of a random match (P_e) was calculated overall for all categories of business success. Accuracy adj. is expressed on a scale of 0 to 1, where a higher value means a better explanatory value of the model.

The average of the Accuracy adj. (Appendix C) for the entire monitored period from 2016 to 2018 reached satisfactory values only in the case of the Model of Delina and Packová. The models' prediction power is increasing towards the final period—2018. The overall average of the models is negatively affected mainly by worse results of predictions of non-prosperous companies, which in 2016 and 2017 were less reliable, as companies were included in the sample of non-prosperous companies based on financial results reported in 2018, and the condition for inclusion was that the companies did not have negative equity in 2016 and 2017.

The models' overall success rate was positive in 2018 for only four models: Model of Delina and Packová, Gulka's Model, Creditworthiness Index, and Quick Test, which achieved the correct classification of companies in more than 50% of cases. However, the overall success rate of less than 70% cannot be considered satisfactory; therefore, the Model of Delina and Packová and the Model of Gulka can be considered the only models generally applicable for testing companies in Slovak conditions. One year and two years before the decline, only Model of Delina and Packová reached more than a 50% success rate. In 2017 it reached 70%, and in 2016 it reached 53.3%, which, however, can no longer be considered a satisfactory result.

The F1-score (Appendix D) is a measure of the success of testing the predictive ability of models calculated from the results of Sensitivity and Precision for the target category of non-prosperous companies. Sensitivity and Precision are listed in Appendix E for each year and economic sector.

Using the F1-score (Appendix D), the Model of Delina and Packová was also evaluated as the most successful one, being successful for all three monitored periods (2016, 2017, and 2018). Model of Gulka was successful in 2018, and the Creditworthiness Index and G-Index also achieved satisfactory results.

When analysing the results of statistical research, it was found that some models are not able to distinguish non-prosperous companies from prosperous companies in certain circumstances. It is worth noting the zero values of Accuracy adj. and the F1 score for the Hurtošová's Model and CH-index. These values were also confirmed by testing it with an F1 success rate. Altman's Z-score reached surprisingly low values. According to our research, the HGN2 Model and M-model can be considered as models that are not suitable for general use for all considered categories.

Hurtošová's Model classified almost all companies as prosperous companies, and therefore it did not show any distinctive ability. The authors of the HGN2 model focused only on assessing and predicting the financial situation of prosperous companies. Non-prosperous companies were excluded in the process of structuring the model, which was one of the reasons for its inability to correctly classify non-prosperous companies. This model classified non-prosperous companies as prosperous companies. Furthermore, the M model showed its distinctive ability only in the assessment of prosperous companies and could not correctly classify non-prosperous companies. The CH-index classified all tested companies, including prosperous and non-prosperous companies, into the grey zone. In his study, Gurčík (2002) concluded that the grey zone of the CH-index is too broad, which reduces its explanatory power. Sušický (2011) also states in his study that according to Kopta (2006), the explanatory power of the CH-index was the lowest because this model correctly classified only 0.9% of companies in the analysed sample among non-prosperous companies. All other companies, which accounted for 99.1%, were included in the grey zone. Of the prosperous companies, the CH-index included 89.7% of companies in the

grey zone, i.e., only 9.3% were classified correctly. Řezbová (2001) recommended using the CH-index with caution and made several adjustments to the model. In our research, the original model was used, and the authors' objections were confirmed.

Unsatisfactory reliability results of the examined models were the reason why, in addition to the overall statistical analysis performed in the first phase of the research, where the total reliability and F1 score of the examined models were evaluated, another phase followed, in which reliability and error rate were evaluated separately for each category of companies and results of models for individual sectors were compared.

4.1. Testing Prosperous Companies

By comparing the results of models for given economy sectors, slight differences were found. Values less than 0.5 are highlighted as red numbers in following tables. A value of 0.5 represents the probability of random selection; therefore, this value and less is considered an incorrect result of the company's financial health assessment and the model does not have an adequate explanatory value for the given category.

The prediction power of models in the case of prosperous companies operating in tourism sector was slightly lower than for the construction and retail sectors, the results of which were almost identical. Table 4 presents the results of testing default prediction models on prosperous companies in tourism sector in the period from 2016 to 2018.

Table 4. Results of testing default prediction models on prosperous companies in tourism sector in 2016–2018.

Model	Number of Correct Tests			Accuracy			Misclassification Rate		
	2016	2017	2018	2016	2017	2018	2016	2017	2018
Model of Hurtošová	10	10	10	1.00	1.00	1.00	0.00	0.00	0.00
CH-index	0	0	0	0.00	0.00	0.00	1.00	1.00	1.00
G-index	5	4	6	0.50	0.40	0.60	0.50	0.60	0.40
Model of Delina and Packová	8	8	9	0.80	0.80	0.90	0.20	0.20	0.10
M-model	10	10	10	1.00	1.00	1.00	0.00	0.00	0.00
HGN model	10	9	8	1.00	0.90	0.80	0.00	0.10	0.20
Model of Gulka	8	10	10	0.80	1.00	1.00	0.20	0.00	0.00
Altman Z-score	4	5	5	0.40	0.50	0.50	0.60	0.50	0.50
Index IN05	5	7	7	0.50	0.70	0.70	0.50	0.30	0.30
Quick Test	7	5	8	0.70	0.50	0.80	0.30	0.50	0.20
Model of Binkert	1	2	1	0.10	0.20	0.10	0.90	0.80	0.90
Creditworthiness Index	9	9	9	0.90	0.90	0.90	0.10	0.10	0.10
Model of Taffler	9	9	9	0.90	0.90	0.90	0.10	0.10	0.10

Notes: Accuracy is a measure of all correctly identified cases. Misclassification rate is a measure of all incorrectly identified cases (1 minus Accuracy). If the Accuracy is equal to or less than 0.5 (highlighted as red number), the model does not have an adequate explanatory value for the given category. Source: Own table.

According to the presented data, the CH-index and Binkert's Model are entirely unsuitable for examining the financial situation of companies operating in tourism sector. The G-index and Z-score also did not achieve satisfactory results. By testing the results of Altman's Z-score, it was found that this model classified a large proportion of prosperous companies as companies attributed to the grey zone and subsequently all companies in difficulty as unhealthy companies. It is evident that the Altman's Z-score was set to be more strict than necessary, which is causing the misclassification of healthy companies. The results of the G-index predictions were not as unambiguous as in the Altman's Z-score. It classified companies in difficulty among all evaluation categories (healthy, grey zone, unhealthy), which indicates an overall ambiguity of G-index results. Altman's

model was consistent in its evaluation, in contrast to the G-index, and it could gain a higher score by simply redefining boundaries of categorisation zones. Hurtošová's Model, although correctly identifying all prosperous companies as healthy, cannot be considered as successful, as it considered almost all non-prosperous companies and companies in difficulty to be healthy. Therefore, it has no distinctive ability. Similarly, the M-model and the HGN2 model appear to be problematic in terms of interpretation, as they also consider most non-prosperous companies to be healthy. These models therefore cannot be applied without an initial assessment of the company based on criteria that would exclude non-prosperous companies. For this reason, their high reliability when assessing prosperous companies loses its informative quality in the context of the assessment of non-prosperous companies. Furthermore, Taffler's Model shows a high error rate in the case of non-prosperous companies.

For the above stated reasons, Hurtošová's Model, CH-index, HGN2 model, M-model, and Taffler's Model were excluded from the final evaluation of successful models. However, their results were kept in tables for comparison purposes.

Table 5 presents results of testing default prediction models on prosperous companies in construction sector in the period from 2016 to 2018.

Table 5. Results of testing default prediction models on prosperous companies in construction sector in 2016–2018.

Model	Number of Correct Tests			Accuracy			Misclassification Rate		
	2016	2017	2018	2016	2017	2018	2016	2017	2018
Model of Hurtošová	10	10	10	1.00	1.00	1.00	0.00	0.00	0.00
CH-index	0	0	0	0.00	0.00	0.00	1.00	1.00	1.00
G-index	3	6	6	0.30	0.60	0.60	0.70	0.40	0.40
Model of Delina and Packová	4	8	9	0.40	0.80	0.90	0.60	0.20	0.10
M-model	10	10	10	1.00	1.00	1.00	0.00	0.00	0.00
HGN model	9	10	10	0.90	1.00	1.00	0.10	0.00	0.00
Model of Gulka	8	10	10	0.80	1.00	1.00	0.20	0.00	0.00
Altman Z-score	5	7	9	0.50	0.70	0.90	0.50	0.30	0.10
Index IN05	9	9	10	0.90	0.90	1.00	0.10	0.10	0.00
Quick Test	4	9	9	0.40	0.90	0.90	0.60	0.10	0.10
Model of Binkert	0	4	0	0.00	0.40	0.00	1.00	0.60	1.00
Creditworthiness Index	8	9	10	0.80	0.90	1.00	0.20	0.10	0.00
Model of Taffler	9	10	10	0.90	1.00	1.00	0.10	0.00	0.00

Notes: Accuracy is a measure of all correctly identified cases. Misclassification rate is a measure of all incorrectly identified cases (1 minus Accuracy). If the Accuracy is equal to or less than 0.5 (highlighted as red number), the model does not have an adequate explanatory value for the given category. Source: Own table.

For the construction sector, the reliability of tested default prediction models was clearer. According to the results, only three models were classified as unreliable: CH-index, G-index, and Binkert's Model. Other models correctly identified prosperous businesses as healthy businesses with 100% success rate, the Quick Test achieved 90% success rate. Similar results were found in the retail sector.

Table 6 presents results of testing default prediction models on prosperous companies in retail sector in the period from 2016 to 2018.

Table 6. Results of testing default prediction models on prosperous companies in retail sector in 2016–2018.

Model	Number of Correct Tests			Accuracy			Misclassification Rate		
	2016	2017	2018	2016	2017	2018	2016	2017	2018
Model of Hurtošová	10	10	10	1.00	1.00	1.00	0.00	0.00	0.00
CH-index	0	0	0	0.00	0.00	0.00	1.00	1.00	1.00
G-index	3	5	3	0.30	0.50	0.30	0.70	0.50	0.70
Model of Delina and Packová	10	10	10	1.00	1.00	1.00	0.00	0.00	0.00
M-model	10	10	10	1.00	1.00	1.00	0.00	0.00	0.00
HGN model	10	10	10	1.00	1.00	1.00	0.00	0.00	0.00
Model of Gulka	10	10	10	1.00	1.00	1.00	0.00	0.00	0.00
Altman Z-score	9	9	10	0.90	0.90	1.00	0.10	0.10	0.00
Index IN05	10	9	10	1.00	0.90	1.00	0.00	0.10	0.00
Quick Test	9	9	9	0.90	0.90	0.90	0.10	0.10	0.10
Model of Binkert	3	2	2	0.30	0.20	0.20	0.70	0.80	0.80
Creditworthiness Index	10	10	10	1.00	1.00	1.00	0.00	0.00	0.00
Model of Taffler	10	10	10	1.00	1.00	1.00	0.00	0.00	0.00

Notes: Accuracy is a measure of all correctly identified cases. Misclassification rate is a measure of all incorrectly identified cases (1 minus Accuracy). If the Accuracy is equal to or less than 0.5 (highlighted as red number), the model does not have an adequate explanatory value for the given category. Source: Own table.

According to above stated results for the construction and retail sectors, the HGN2 and M-model failed to correctly classify non-prosperous companies.

4.2. Testing Companies in Difficulty

As mentioned above, the testing of companies in difficulty is specific, as these are legally defined as companies whose existence is at risk and are at risk of insolvency. Their classification as prosperous companies can therefore be considered incorrect. Yet, it would be appropriate to individually assess and evaluate the company's overall financial health to decide whether to attribute such a company to the grey zone or to classify it as a non-prosperous company.

In the case of companies in difficulty, two accuracy ratings were assessed, namely Accuracy 1 and Accuracy 2 (Appendix F). The reason for setting two different criteria for assessing reliability is the problematic ambiguity of the classification of companies in difficulty and, at the same time, the absence of a grey zone in some models. Accuracy 1 means the ratio of the number of firms considered as grey zone companies to the total number of firms in difficulty, and Accuracy 2 the ratio of the sum of companies attributed to the grey zone and companies classified as unhealthy to the total number of companies in difficulty. Accuracy 2 reflects companies' financial situation in difficulty more accurately. It should be emphasised that models with only two categorisation zones could deal with the company in only two ways and classify it as either healthy or unhealthy. These are the Hurtošová's Model, Model of Delina and Packová, the M-model, and the Gulka's Model.

In previous text, we identified four problematic models, which, according to the research results, are not suitable for assessing the whole spectrum of companies: Hurtošová's Model, CH-index, M-model, and Taffler's Model, even though in this specific category of companies in difficulty, the CH-index reached 100% reliability and the HGN2 model 50%.

In the tourism sector (Appendix F), Gulka's Model attributed companies in difficulty to the grey zone or classified them as unhealthy companies with the lowest success rate (20%). This model seems to be inapplicable for assessing companies defined as companies in difficulty because their classification as healthy companies cannot be considered as correct.

Most models for the Accuracy 2 measure rated companies in difficulty as unhealthy companies or companies attributed to the grey zone. Z-score and IN05 considered companies in difficulty to be unhealthy in almost all cases, while the Quick Test classified them as companies attributed to the grey zone. There was no significant difference in the numbers of unhealthy companies and companies attributed to the grey zone for the other models. The following models can be considered successful when testing companies in difficulty: G-index, Model of Delina and Packová, Altman's Z-score, IN05, Quick Test, Binkert's Model, and Creditworthiness Index.

In the construction sector (Appendix F), the models' success rate to classify companies in difficulty as unhealthy companies or companies attributed to the grey zone was lower, but not significantly. Only the G-index and the Quick Test achieved a 100% success rate for the Accuracy 2 measure. Model of Delina and Packová, Altman's Z-score (its accuracy increases to 90% towards the final period), the Creditworthiness Index, and the IN05 Index also achieved a satisfactory level of reliability. Gulka's Model ranked most companies as healthy in every period, therefore it cannot be considered reliable.

Testing results in the retail sector (Appendix F) were very similar to above mentioned results achieved in the construction sector. G-index and Quick Test achieved 100% success rate in the Accuracy 2 measure. In the final period, IN05 and the Creditworthiness Index achieved a 90% success rate. Model of Delina and Packová and Binkert's Model achieved still satisfactory results (70 and 80%). Gulka's Model success rate (50% to 60%) was too close to random selection and therefore cannot be considered satisfactory.

4.3. Testing Non-Prosperous Companies

The primary condition for selecting and defining a company as non-prosperous was that the company had negative equity in 2018, while in 2016 and 2017 the equity was positive, but with declining trend. According to how the research was set, the reliable model is the one predicting future decline and negative financial state in 2018.

Table 7 presents results of testing default prediction models on non-prosperous companies in tourism sector in the period from 2016 to 2018.

Table 7. Results of testing default prediction models on non-prosperous companies in tourism sector in 2016–2018.

Model	Number of Correct Tests			Reliability			Misclassification Rate		
	2016	2017	2018	2016	2017	2018	2016	2017	2018
Model of Hurtošová	0	0	0	0.00	0.00	0.00	1.00	1.00	1.00
CH-index	0	0	0	0.00	0.00	0.00	1.00	1.00	1.00
G-index	0	1	6	0.00	0.10	0.60	1.00	0.90	0.40
Model of Delina and Packová	8	8	10	0.80	0.80	1.00	0.20	0.20	0.00
M-model	0	1	1	0.00	0.10	0.10	1.00	0.90	0.90
HGN model	1	2	0	0.10	0.20	0.00	0.90	0.80	1.00
Model of Gulka	0	4	4	0.00	0.40	0.40	1.00	0.60	0.60
Altman Z-score	3	6	8	0.30	0.60	0.80	0.70	0.40	0.20
Index IN05	6	7	10	0.60	0.70	1.00	0.40	0.30	0.00
Quick Test	0	0	0	0.00	0.00	0.00	1.00	1.00	1.00
Model of Binkert	0	0	3	0.00	0.00	0.30	1.00	1.00	0.70
Creditworthiness Index	0	0	6	0.00	0.00	0.60	1.00	1.00	0.40
Model of Taffler	2	3	3	0.20	0.30	0.30	0.80	0.70	0.70

Notes: Accuracy is a measure of all correctly identified cases. Misclassification rate is a measure of all incorrectly identified cases (1 minus Accuracy). If the Accuracy is equal to or less than 0.5 (highlighted as red number), the model does not have an adequate explanatory value for the given category. Source: Own table.

Model of Delina and Packová and the IN05 Index were able to identify the decline of companies in the tourism sector with 100% success rate. Less successful was Altman's Z-score with 80% success rate, followed by the Creditworthiness Index and the Gulka's Model, which successfully identified a decline in 60% of cases in 2018. Given the sample specifics, it is natural that the best prediction results were achieved in 2018. Both models with the best results were able to identify the future financial distress of most companies in previous two periods. The Model of Delina and Packová was slightly more successful than IN05. Hurtošová's Model, M-model, HGN model, and Quick Test were unable to correctly identify non-prosperous companies as their success rate in 2018 was less than 10%. The shortcomings in prediction value of CH-index have already been described above and are not addressed in further interpretations.

Table 8 presents results of testing default prediction models on non-prosperous companies in construction sector in the period from 2016 to 2018.

Table 8. Results of testing default prediction models on non-prosperous companies in construction sector in 2016–2018.

Model	Number of Correct Tests			Reliability			Misclassification Rate		
	2016	2017	2018	2016	2017	2018	2016	2017	2018
Model of Hurtošová	0	0	1	0.00	0.00	0.10	1.00	1.00	0.90
CH-index	0	0	0	0.00	0.00	0.00	1.00	1.00	1.00
G-index	4	4	8	0.40	0.40	0.80	0.60	0.60	0.20
Model of Delina and Packová	9	10	10	0.90	1.00	1.00	0.10	0.00	0.00
M-model	2	0	0	0.20	0.00	0.00	0.80	1.00	1.00
HGN model	1	1	0	0.10	0.10	0.00	0.90	0.90	1.00
Model of Gulka	4	6	9	0.40	0.60	0.90	0.60	0.40	0.10
Altman Z-score	2	3	4	0.20	0.30	0.40	0.80	0.70	0.60
Index IN05	6	6	8	0.60	0.60	0.80	0.40	0.40	0.20
Quick Test	5	5	10	0.50	0.50	1.00	0.50	0.50	0.00
Model of Binkert	0	1	2	0.00	0.10	0.20	1.00	0.90	0.80
Creditworthiness Index	2	1	7	0.20	0.10	0.70	0.80	0.90	0.30
Model of Taffler	0	1	1	0.00	0.10	0.10	1.00	0.90	0.90

Notes: Accuracy is a measure of all correctly identified cases. Misclassification rate is a measure of all incorrectly identified cases (1 minus Accuracy). If the Accuracy is equal to or less than 0.5 (highlighted as red number), the model does not have an adequate explanatory value for the given category. Source: Own table.

In the construction sector, Model of Delina and Packová was again the most successful in classifying non-prosperous companies, even in previous two periods. The same result—100% success rate in 2018 was achieved also by the Quick Test, but which predicted future decline in only half of companies in previous periods. The following models were also successful in identifying the financial distress in the last period, when companies in the construction sector were overleveraged: Gulka's Model (90% success rate), G-index, IN05 Index (both 80% success rate), and Creditworthiness Index with a 70% success rate. Hurtošová's Model, M-model, HGN model, Taffler's Model, Binkert's Model, and Altman's Z-score were unable to correctly classify non-prosperous companies. Surprisingly, Altman's model successfully classified as unhealthy only four companies out of 10 in 2018.

Table 9 presents results of testing default prediction models on non-prosperous companies in retail sector in the period from 2016 to 2018.

Table 9. Results of testing default prediction models on non-prosperous companies in retail sector in 2016–2018.

Model	Number of Correct Tests			Reliability			Misclassification Rate		
	2016	2017	2018	2016	2017	2018	2016	2017	2018
Model of Hurtošová	0	0	0	0.00	0.00	0.00	1.00	1.00	1.00
CH-index	0	0	0	0.00	0.00	0.00	1.00	1.00	1.00
G-index	5	6	10	0.50	0.60	1.00	0.50	0.40	0.00
Model of Delina and Packová	7	7	10	0.70	0.70	1.00	0.30	0.30	0.00
M-model	0	0	1	0.00	0.00	0.10	1.00	1.00	0.90
HGN model	0	1	0	0.00	0.10	0.00	1.00	0.90	1.00
Model of Gulka	4	4	10	0.40	0.40	1.00	0.60	0.60	0.00
Altman Z-score	4	6	8	0.40	0.60	0.80	0.60	0.40	0.20
Index IN05	5	5	9	0.50	0.50	0.90	0.50	0.50	0.10
Quick Test	2	3	10	0.20	0.30	1.00	0.80	0.70	0.00
Model of Binkert	0	0	3	0.00	0.00	0.30	1.00	1.00	0.70
Creditworthiness Index	3	5	10	0.30	0.50	1.00	0.70	0.50	0.00
Model of Taffler	3	4	5	0.30	0.40	0.50	0.70	0.60	0.50

Notes: Accuracy is a measure of all correctly identified cases. Misclassification rate is a measure of all incorrectly identified cases (1 minus Accuracy). If the Accuracy is equal to or less than 0.5 (highlighted as red number), the model does not have an adequate explanatory value for the given category. Source: Own table.

In the retail sector, five models were able to identify unhealthy companies with a 100% success rate in 2018: G-index, Model of Delina and Packová, Gulka's Model, Quick Test, and Creditworthiness Index. On the contrary, Hurtošová's Model, CH-index, M-model, and HGN2 model were unable to identify non-prosperous companies also in this sector. Binkert's Model and Taffler's Model showed insufficient reliability.

4.4. General Findings

The basic criterion to exclude default prediction models and to classify them as unreliable is that the model cannot identify non-prosperous companies with more than 50% success rate. It is well-known that the negative equity may cause some calculation problems when calculating ratios. However, in 2016 and 2017 the book value of equity of companies included in the research sample of non-prosperous companies was positive, and achieved results were worse than in 2018. Therefore negative equity was not the only cause of models' failure to detect non-prosperous companies.

We are adding the list of default prediction models that were able to identify companies that were overleveraged in 2018 as non-prosperous companies:

- G-index
- Model of Delina and Packová
- Gulka's Model (not applicable in the tourism sector)
- Altman Z-score (not applicable in the construction sector)
- Index IN05
- Quick Test
- Creditworthiness Index

These models were successful in 2018; however, their success rate to predict financial distress in advance (one or two years before the event) was lower. Model of Delina and Packová was the only model that was able to identify most non-prosperous companies and classify them as unhealthy in all three periods and all sectors.

The assessed reliability of default prediction models enabled us to distinguish models with higher reliability from models with insufficient reliability. However, the models' success rate was set not only due to its high reliability but also due to its low error rate.

The error rate was given as a type I error in the case of non-prosperous companies, i.e., the number of non-prosperous companies incorrectly classified as healthy to all non-prosperous companies, and type II error in the category of prosperous companies, i.e., the number of prosperous companies incorrectly classified as unhealthy to the number of all prosperous companies. In general, default prediction models are scaling financial health or financial distress of companies into three categories, specifically healthy companies, unhealthy companies, and companies in the grey zone. Attributing company to the grey zone does not provide a clear prediction of its financial health. Different models are proposing different interpretations for this interspace. This category was considered in this research and dealt with as an auxiliary category for the final assessment purposes, based on type I and type II errors, and interpreted as non-assignment to non-prosperous companies. If a non-prosperous company was attributed to the grey zone, it was considered and treated as an incorrect result, while if a prosperous company was attributed to the grey zone, it was considered and treated as an appropriate classification. This procedure emphasizes the primary purpose of default prediction models and the requirement to correctly classify non-prosperous companies. It is based on a premise that risks associated with incorrect classification of non-prosperous companies outweigh the risks of classifying prosperous companies as the grey zone companies.

Table 10 presents the results of error rate assessment of default prediction models in tourism, construction and retail sector in 2018. The high success rate determined the final selection of suitable models in the case of non-prosperous companies, eliminating the type I error. The consequences of a type I error are more dangerous, because classifying a non-prosperous company as a healthy one brings many risks for the management, owners, investors, and creditors. The following table lists the models characterised by high reliability and low error rate, and presents synthesized results obtained in 2018. Data included in this table were transferred from Appendix E.

Table 10. Error rate assessment of default prediction models in tourism, construction and retail sector in 2018.

Model	Tourism		Construction		Retail		Average (I, II)	Rank (I, II)	Average (I)	Rank (I)
	Type I Error	Type II Error	Type I Error	Type II Error	Type I Error	Type II Error				
G-Index	0.40	0.20	0.20	0.10	0.00	0.15	0.175	5	0.200	4
Model of Delina and Packova	0.00	0.10	0.00	0.10	0.00	0.00	0.033	1	0.000	1
Model of Gulka	0.60	0.00	0.10	0.00	0.00	0.00	0.117	3	0.233	5
Altman Z-score	0.20	0.55	0.60	0.20	0.20	0.15	0.317	7	0.333	6
INDEX IN05	0.00	0.45	0.20	0.30	0.10	0.35	0.233	6	0.100	3
Quick Test	0.10	0.10	0.00	0.30	0.00	0.15	0.108	2	0.033	2
Creditworthiness Index	0.40	0.10	0.30	0.05	0.00	0.10	0.158	4	0.233	5

Notes: Rank (I, II) expresses the order of default prediction models in relation to the average value of type I and type II errors for all sectors. Value of 1 is assigned to the model with the lowest error rate. Rank (I) ranks default prediction models only according to the average type I error calculated for all sectors. Source: Own table.

Considering the importance of type I errors, the above stated table presents two rankings. Rank (I, II) is based on Average (I, II), which represents an overall average of type I and type II errors, while Rank (I) is based on Average (I), which represents the average of type I errors only. Comparing these two rankings, we can conclude the impact of type I and type II errors on overall ranking of a specific model.

From the comparison between the sectors, the differences in the models' error or success rates are apparent. It must be acknowledged that the cross-sectoral results could be distorted by the small number of companies within the sectors and that this claim should be verified on a larger sample of companies.

Due to the sample limitations, the confidence interval was further estimated, based on the classification of non-prosperous companies. The population was set to 368 non-prosperous companies, which represents the base set of all non-prosperous companies within all three sectors of economy passing the selection criteria, which were specified and further described in the methodology section. The sample applied was set to 30 non-prosperous companies, based on which the research was conducted. The confidence level was set to 95%. The results of confidence interval estimation are presented in Table 11.

Table 11. Confidence interval of correctly classifying non-prosperous companies for each default prediction model according to type I error rate, 2018.

Model	The Success Rate of Classifying Non-Prosperous Companies	Confidence Interval	Interpretation	Range
G-Index	80.00%	13.74	In 95% of cases, the model correctly classifies non-prosperous companies with the success rate interval from 93.74% to 66.26%.	0.94–0.66
Model of Delina and Packová	100.00%	3.42	In 95% of cases, the model correctly classifies non-prosperous companies with the success rate interval from 100% to 96.58%.	1–0.97
Model of Gulka	76.67%	14.45	In 95% of cases, the model correctly classifies non-prosperous companies with the success rate interval from 91.12% to 62.22%.	0.91–0.62
Altman Z-score	66.67%	16.15	In 95% of cases, the model correctly classifies non-prosperous companies with the success rate interval from 82.82% to 50.52%.	0.83–0.51
Index IN05	90.00%	10.3	In 95% of cases, the model correctly classifies non-prosperous companies with the success rate interval from 100% to 79.70%.	1.00–0.80
Quick Test	96.67%	5.86	In 95% of cases, the model correctly classifies non-prosperous companies with the success rate interval from 100% to 90.81%.	1.00–0.91
Creditworthiness Index	76.67%	14.45	In 95% of cases, the model correctly classifies non-prosperous companies with the success rate interval from 91.12% to 62.22%.	0.91–0.62

Source: Own table.

The above stated confidence interval of each model corresponds with the success rate of classifying non-prosperous companies (one minus type I Error rate). The narrowest range was estimated for the Model of Delina and Packová, followed by Quick Test. The broadest range was estimated for the Altman Z-score, followed by Creditworthiness Index and model of Gulka.

5. Discussion

Model of Delina and Packová was rated as the most reliable model with the lowest average error rate (both type I and type II errors), which was also confirmed by Accuracy adj. (Appendix C) and F1-score (Appendix D) measures. From the available literature, the only previous research testing the Model of Delina and Packová is known (Bakeš and Valášková 2018), where this model was compared with CH-index, Binkert's Model, Hurtošová's Model, and Gulka's Model. This model proved the best classification ability on a sample of 266 companies. However, the same result in the mentioned research was achieved by Binkert's Model, which was excluded from our final assessment due to its inability to identify non-prosperous companies. The reason for the different result in the case of Binkert's Model can be most likely considered the selection of a sample, which included and separately treated non-prosperous companies.

The second most reliable model, considering both, type I and type II errors, was ranked the Quick Test, a representative of globally well-disseminated default prediction models. While Index IN05, one of the most applied models in local conditions, did not prove its validity when considering both types of errors. Yet, it showed outstanding results and the ability to predict financial distress when relying on Type I Errors assessment. This Index is well-known and often used in local conditions, as it originated in Czech Republic, and was a research subject of several authors. Siekelová et al. (2018) evaluated it as the most successful model, compared to Model of Jakubík and Těplý, Hurtošová's Model, Gulka's Model, in 2018 on a sample of 500 companies. The research of Valášková et al. (2017) confirms these results and the Rank (I) listed in the table above. CH-index, G-index, and Altman Z-score were compared in the research, according to which the G-index achieved better results than the Altman Z-score, while the CH-index achieved the worst result. The comparison of Altman's Z-score and the IN05 Index in the study of Gavurová et al. (2017) turned out in favour of IN05 Index. The authors also considered the correct identification of non-prosperous companies as the most critical criterion in their assessment. In Sušický's dissertation (2011), higher reliability of Altman's Model was presented, compared to the IN05 Index; however, both achieved high reliability. High reliability reported by IN05 Index when relying on type I error rate is crucial, but a high rate of misclassifying other than non-prosperous companies should be considered in the final assessment of this model. The overall Rank (I, II) of the IN05 Index was therefore one of the lowest within the final list of models presented in Table 10.

The non-optimal assessment results of Altman Models in local conditions are worthy of consideration. It achieved the least reliable results within the final list of models. Yet, Altman's model is the most frequently used default prediction model in local conditions. However, in presented studies (Čámská 2014; Machek 2014), the reliability results did not reach highest levels, but it was still possible to consider those results as satisfactory. There are studies (Delina and Packová 2013; Režňáková and Karas 2015) in which Altman's Model failed, mainly due to the setting of boundaries of the grey zone and the inconsistency of its assessment. We also dealt with this issue and concluded that it is correct in the case of the two-criteria classification of companies between healthy and unhealthy to assign a grey zone to the positive identification, i.e., assign it to a healthy state of a business. In this case, the grey zone warns us that the company in one examined criterion does not achieve optimal results, but it cannot be considered an unhealthy company. Similarly, we dealt with the grey zone in other models. Altman's Model, given the results of research, attributed most prosperous companies to the grey zone, what was the primary "trigger" for this consideration.

The low explanatory value of the models, which were discarded as insufficiently reliable for general testing of all categories, mainly focusing on identifying non-prosperous companies, whose detection was of the highest importance, was also confirmed in the research of other authors. The CH-index was marked as an unsatisfactory model in the study of Čámská (2014), Gurčík (2002), Valášková et al. (2017), and others. All authors agreed on the model's inability to detect failing businesses, i.e., type I error.

In Gulka's Model, high reliability was demonstrated, which is shown in the previous table as the overall Rank (I, II), and this finding is also confirmed by the measures of Accuracy adj. and F1-score. However, in the individual assessment of the classification of companies in difficulty, this model did not show a sufficient distinctiveness and classified companies in difficulty mostly as healthy companies, which can cause a risk for information users.

The Taffler's Model was very often tested in domestic studies and rated as one of the models with the lowest reliability (Kováčová and Kubala 2018; Machek 2014; Sušický 2011), which was also confirmed by the results of our study.

It must be acknowledged that this research deserves a larger sample of companies focusing on more sectors to determine which model is the most suitable for each sector. According to research results, the Model of Delina and Packová appears to be the most reliable generally applicable model. It was able to reliably identify companies that are prosperous and non-prosperous within all considered sectors. Other models show some differences in reliability between sectors. Gulka's Model, due to its type I error, is unsuitable for the tourism sector, while in the other two sectors, it achieved reliable results. Altman's Z-score showed a high error rate in the construction sector. It's important to mention that while some default prediction models, such as the Model of Delina and Packová and M-model, were constructed for the general use, regardless of the economy sector, some other tested models were specifically designed for a particular sector of economy, such as CH-Index and G-Index, which were meant to predict the financial health or financial distress in the sector of agriculture. There's also a group of default prediction models, which were designed for broader applications, such as the Model of Binkert and HGN2 model, but still contain partial sector-specificity. When interpreting the research results presented in this article, it is highly recommended to consider the primary sector for which the model was constructed.

6. Conclusions

All models are wrong, but some are useful (Box 1987). This statement fits for models of predicting financial distress. The research results have shown that some models could be classified as unreliable because of their inability to detect impending decline. This research identified also a group of models which were incorrectly penalizing healthy companies. Only few models were delivering reliable results, and could be used in real conditions to support qualified economic decisions. We consider results of this research to be directly applicable when selecting and choosing a suitable model for financial distress assessment.

Research has shown high reliability of some alternative default prediction models, which were designed in local conditions. The highest reliability and accuracy was achieved by an alternative local Model of Delina and Packová. The least reliable results within the final list of models were reported by the most globally disseminated model Altman's Z-score. In contrast, the Quick Test, as one of globally well disseminated models, achieved high ranking as it reported low error rates, both type I and type II errors. While Index IN05, one of the most applied models in Slovak Republic, showed its ability to correctly classify non-prosperous companies, the overall rating of this model was low as it reported high type II error rate. However, significant differences in reliability were identified across the spectrum of alternative default prediction models, but also across economy sectors. It should be emphasized that a large group of unreliable models within the group of alternative default prediction models was identified. This subset of models was excluded from further testing.

This research confirmed doubts about the reliability of some models, which were proposed by several studies quoted in this paper, e.g., CH-index and Tafflers' model. Different approaches applied by different scientists to defining the sample and the methodology for evaluating the reliability resulted in a low degree of comparability of these studies. Authors consider the methodology proposed in this paper as a possible benchmark for further research.

Based on the research results, it cannot be generally stated that a higher reliability and accuracy was achieved by alternative default prediction models, which originated in local conditions, compared to globally disseminated default prediction models. The highest overall ranking was achieved by the local alternative model and the least reliable results were recorded by the most globally disseminated model. But the score between these two models provides ambiguous results when comparing these two categories of default prediction models. It is reasonable to assume similar findings in other local markets. Further research territorially targeting other markets is therefore necessary.

When choosing a proper model to predict the financial distress in real conditions, based on above presented research results, three factors should be considered, i.e., the sector-specificity of a selected model, whether to rely on the overall Rank (I, II) or to prefer the type I error rate, and the target period as the distance of time for which the model should be reliable.

It should be noted that this research has some limitations. When interpreting the research results, the sector-specificity of assessed models must be considered. The subjects of testing were also default prediction models, which were designed for specific industries, e.g., agriculture. These models were included in the testing in order to verify their reliability for other sectors. The research sample size could be considered as another limitation of this research. However, it should be noted that this sample consists of companies, which passed the selection criteria described in the methodology. It could be therefore defined as a homogeneous sample, which ensures internal consistency of each category and differences between categories. Though, for the further research, it would be appropriate to expand the sample. It is also necessary to realize that the tested models do not reflect systematic risk and the occurrence of extraordinary events, such as the COVID-19 pandemic and the related structural changes in economy. The research was therefore conducted in a pre-pandemic period (2016, 2017, and 2018).

Author Contributions: D.R.: Term, Conceptualization, Methodology, Validation, Resources. H.M.: Supervision, Project administration, Funding acquisition. P.Š.: Visualization, Writing—Review & Editing, Validation. D.L.: Software, Data Curation, Formal Analysis, Investigation, Writing—Original draft preparation. All authors have read and agreed to the published version of the manuscript.

Funding: This research and the APC was funded by the Ministry of Education, Science, Research and Sports of the Slovak Republic, VEGA No. 1/0340/19, project share is 50%; and VEGA No. 1/0708/20, project share is 50%.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analyzed in this study. This data can be found here: [<https://finstat.sk/databaza-financnych-udajov>, accessed on 16 September 2020].

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix A. Alternative Models

Table A1. Parameters of local default prediction models.

Model's Construction			
Title	Classification Criteria		Indicators Used
	Prosperous	Non-Prosperous	
$CH\ index = 0.37 * x_1 + 0.25 * x_2 + 0.21 * x_3 - 0.10 * x_4 - 0.07 * x_5$			
CH-index ¹	unspecified	unspecified	$x_1 = \text{EAT/assets}$
			$x_2 = \text{EAT/revenues}$
			$x_3 = \text{current assets/current liabilities}$
			$x_4 = \text{current liabilities/revenues}$
			$x_5 = \text{third-party sources/assets}$
$D_S^{1\ p0\ 3} = 0.180 * U_1^{t-2} + 0.147 * U_{40}^{t-1} + 0.237 * U_{49}^{t-1} + 0.377 * U_{63}^{t-1} + 0.514 * U_{13}^t + 0.505 * U_{29}^t + 0.271 * U_{30}^t + 0.207 * U_9^{t-1}$ where: superscript represents the period of reporting (t represents the current period), subscript identifies the indicator in the initial set of indicators (its order)			
Model of Binkert ²	current assets/current liabilities > 1.5 return on sales > 5%	current assets/current liabilities < 1.0 negative return on sales	$U_1^{t-2} = \text{current assets/current liabilities (indicators reported in the oldest period)}$
			$U_{40}^{t-1} = \text{equity/tangible assets (indicators reported in the 2nd period)}$
			$U_{49}^{t-1} = \text{EAT/sales (indicators reported in the 2nd period)}$
			$U_{63}^{t-1} = \text{revenues/added value (indicators reported in the 2nd period)}$
			$U_{13}^t = \text{(total assets of the current year/total assets of the previous year) minus 1, (indicators reported in the most current period)}$
			$U_{29}^t = \text{(equity of the current year/equity of the previous year) minus 1, (indicators reported in the most current period)}$
			$U_{30}^t = \text{total liabilities of the current year/total liabilities of the previous year) minus 1, (indicators reported in the most current period)}$
$logit(\beta) = \ln\left(\frac{p}{1-p}\right) = -1.6889 + 0.00337 * UK27 - 4.4075 * UK55 + 1.4058 * UK57 - 0.0165 * UK67$			
Model of Hurtošová ³	(good) late payment of less than 30 days	(bad) late payment of more than 60 days	$UK27 = (\text{Inventory/sales}) \times 365$
			$UK55 = \text{depreciation/operating costs}$
			$UK57 = \text{interest expense/cost of financial activity}$
			$UK67 = \text{equity/total capital}$
$G = 3.412 * x_1 + 2.26 * x_2 + 3.277 * x_3 + 3.149 * x_4 - 2.063 * x_5$			
G-index ⁴	for 3 years in a row made a profit, in the last period return on equity > 8%	loss in the last three years	$x_1 = \text{EAT/total liabilities}$
			$x_2 = \text{EBT/total liabilities}$
			$x_3 = \text{EBT/revenues}$
			$x_4 = \text{cash flow/total liabilities}$ cash flow = EAT + depreciation + time – shifted operating income and expenses
			$x_5 = \text{inventories/revenues}$

Table A1. Cont.

Model's Construction			
Title	Classification Criteria		Indicators Used
	Prosperous	Non-Prosperous	
			$P = 2.836 - 0.0001331 * \frac{(\text{financial assets} - \text{short-term liabilities})}{\text{operating costs} - \text{depreciation}} + 0.05077 * A_2 + 0.8027 * A_3 - 0.0006079 * A_4 + 0.00009913 * B_1 + 0.0005512 * B_2 - 0.5996 * B_3 - 0.0005996 * B_4$
Model of Delina and Packová ⁵	those for which bankruptcy has not been declared	bankrupt companies	$A_2 = \text{retained earnings}/\text{total capital}$
			$A_3 = \text{EBIT}/\text{total capital}$
			$A_4 = \text{share capital}/\text{total liabilities}$
			$B_1 = \text{cash flow}/\text{third-party capital}$ cash flow = EAT + depreciation + time-shifted operating income and expenses
			$B_2 = \text{total capital}/\text{debt capital}$
			$B_3 = \text{EBT}/\text{total capital}$
			$B_4 = \text{EBT}/\text{total output}$
			$HGN2 = (x_1 + x_2 + x_3) - (y_1 + y_2 + y_3)$
HGN2 model ⁶	profitable companies	negative economic result, extreme values of current debt	$x_1 = \text{EAT}/\text{equity}$
			$x_2 = (\text{EAT} + \text{depreciation})/\text{sales}$
			$x_3 = \text{sales}/\text{assets}$
			$y_1 = \text{short-term receivables}/\text{sales}$
			$y_2 = \text{liabilities}/(\text{EAT} + \text{depreciation})$
			$M = -1.32 + 0.40 * x_1 + 0.19 * x_2 + 1.25 * x_3 + 1.93 * x_4 + 1.00 * x_5 + 0.92 * x_6$
M-model ⁷	unspecified	unspecified	$x_1 = (\text{commercial liabilities}/(\text{revenues from sales of goods} + \text{revenues from sales of own products and services})) \times 360$
			$x_2 = (\text{EBITDA})/(\text{revenues from the sale of goods} + \text{revenues from the sale of own products and services}) \times 100$
			$x_3 = (\text{EBT}/\text{total assets}) \times 100$
			$x_4 = (\text{current assets}/\text{current liabilities}) \times 100$
			$x_5 = (\text{sales of goods} + \text{sales of own products and services})/\text{total assets}$
			$x_6 = (\text{liabilities} + \text{accruals})/(\text{depreciation and adjustments to intangible fixed assets and tangible fixed assets} + \text{EAT})$
			$p = \frac{e^{0.0216 - 0.6131 * X1 - 0.0068 * X2 - 0.0293 * X3 - 0.0011 * X4 + 0.0240 * X5 + 0.0317 * X6 - 1.0663 * X7}}{1 + e^{0.0216 - 0.6131 * X1 - 0.0068 * X2 - 0.0293 * X3 - 0.0011 * X4 + 0.0240 * X5 + 0.0317 * X6 - 1.0663 * X7}}$
Model of Gulka ⁸	those for which bankruptcy has not been declared	bankrupt companies	$X1 = \text{financial accounts}/(\text{short-term liabilities} + \text{short-term financial assistance} + \text{current bank loans})$
			$X2 = \text{total sales}/(\text{current assets} - \text{current liabilities})$
			$X3 = \text{total financial accounts}/\text{assets}$
			$X4 = \text{total equity}/\text{assets}$
			$X5 = (\text{bank loans} + \text{short-term financial assistance})/\text{total assets}$
			$X6 = (\text{social security liabilities} + \text{tax liabilities and subsidies})/\text{total assets}$
			$X7 = \text{EBITDA}/\text{total assets}$

Source: Own processing according to ¹ Chrastinová 1998; ² Zalai 2000; ³ Hurtošová 2008; ⁴ Gurčík 2002; ⁵ Delina and Packová 2013; ⁶ Hyránek et al. 2018; ⁷ Harumová and Janisová 2014; ⁸ Gulka 2016).

Appendix B. Globally Disseminated Models

Table A2. Parameters of globally disseminated default prediction models.

Model	Number of Original Classification Zones	Classification Zones According to FinStat Database			
		Number of Zones Applied	Prosperous Companies	Grey Zone	Non-Prosperous Companies
Altman's Z-score ¹	3	3	$Z > 2.9$	$1.2 \leq Z \leq 2.9$	$Z < 1.2$
Formula			$Z = 0.7177 * x_1 + 0.847 * x_2 + 3.107 * x_3 + 0.42 * x_4 + 0.998 * x_5$		
Indicators used			where $x_1 = \frac{\text{current assets} - \text{current liabilities}}{\text{total capital}}$; $x_2 = \frac{\text{retained earnings}}{\text{total capital}}$; $x_3 = \frac{\text{EBIT}}{\text{total capital}}$; $x_4 = \frac{\text{equity}}{\text{debt capital}}$; $x_5 = \frac{\text{sales}}{\text{total capital}}$		
INDEX IN05 ²	3	3	$\text{IN05} > 1.6$	$1.6 < \text{IN05} < 0.9$	$\text{IN05} < 0.9$
Formula			$\text{IN05} = 0.13 * x_1 + 0.04 * x_2 + 3.97 * x_3 + 0.21 * x_4 + 0.09 * x_5$		
Indicators used			where $x_1 = \frac{\text{total assets}}{\text{liabilities}}$; $x_2 = \frac{\text{EBIT}}{\text{interest expense}}$; $x_3 = \frac{\text{EBIT}}{\text{total assets}}$; $x_4 = \frac{\text{total revenues}}{\text{total assets}}$; $x_5 = \frac{\text{current assets}}{\text{current liabilities}}$		
QUICK test ³	5	4	4–7	8–15	16–20
Formula			$Qt = x_1 + x_2 + x_3 + x_4$		
Indicators used			where $x_1 = \frac{\text{equity}}{\text{total capital}}$; $x_2 = \frac{\text{debt capital} - \text{financial accounts}}{\text{EAT} + \text{depreciation}}$; $x_3 = \frac{\text{EAT} + \text{depreciation}}{\text{sales}}$; $x_4 = \frac{\text{EBIT}}{\text{total capital}}$		
Creditworthiness Index ³	8	3	$\text{CI} > 1$	$-1 \leq \text{CI} \leq 1$	$\text{CI} < -1$
Formula			$\text{CI} = 1.5 * x_1 + 0.08 * x_2 + 10 * x_3 + 5 * x_4 + 0.3 * x_5 + 0.1 * x_5$		
Indicators used			where $x_1 = \frac{\text{EAT} + \text{depreciation}}{\text{debt capital}}$; $x_2 = \frac{\text{total assets}}{\text{debt capital}}$; $x_3 = \frac{\text{EBIT}}{\text{total assets}}$; $x_4 = \frac{\text{EBIT}}{\text{sales}}$; $x_5 = \frac{\text{sales}}{\text{total assets}}$		
Taffler's Model ⁴	3	3	$T > 0.3$	$0.2 \leq T \leq 0.3$	$T < 0.2$
Formula			$T = 0.53 * x_1 + 0.13 * x_2 + 0.18 * x_3 + 0.16 * x_4$		
Indicators used			where $x_1 = \frac{\text{EBIT}}{\text{current liabilities}}$; $x_2 = \frac{\text{current assets}}{\text{debt}}$; $x_3 = \frac{\text{current liabilities}}{\text{total assets}}$; $x_4 = \frac{\text{sales}}{\text{total assets}}$		

Source: Processed according to (¹ Altman 1983; ² Neumaierová and Neumaier 2005; ³ Kralicek 1993; ⁴ Taffler 1983).

Appendix C. Accuracy Adj.

Table A3. Results of testing models for each sector using the Accuracy Adj. measure.

Sector	Tourism Sector			Construction Sector			Retail Sector		
	2018	2017	2016	2018	2017	2016	2018	2017	2016
Model of Hurtošová	0.000	0.000	0.000	0.100	0.000	0.000	0.000	0.000	0.000
CH-index	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
G-Index	0.350	0.050	0.050	0.600	0.350	0.300	0.500	0.400	0.100
Model of Delina and Packová	0.900	0.600	0.600	0.900	0.800	0.300	1.000	0.700	0.700
M-model	0.100	0.100	0.000	0.000	0.000	0.200	0.100	0.000	0.000
HGN2 Model	0.000	0.100	0.150	0.000	0.050	0.000	0.000	0.050	0.000
Model of Gulka	0.400	0.400	-0.200	0.900	0.600	0.200	1.000	0.400	0.400
Altman's Z-score	0.150	0.050	-0.150	0.400	0.250	0.050	0.550	0.500	0.300
INDEX IN05	0.350	0.250	0.050	0.600	0.350	0.400	0.550	0.350	0.400
Quick Test	0.750	0.300	0.500	0.650	0.650	0.250	0.800	0.500	0.400
Model of Binkert	0.150	0.100	0.000	-0.100	0.150	-0.100	0.000	-0.100	-0.050
Creditworthiness Index	0.600	0.250	0.250	0.700	0.300	0.350	0.800	0.450	0.400
Taffler's Model	0.200	0.200	0.100	0.100	0.100	0.050	0.250	0.200	0.200

Source: Own processing.

Table A4. Average results of testing models using the Accuracy Adj. measure.

Model	Average 2018	Average 2017	Average 2016	Average (Total)
Model of Hurtošová	0.033	0.000	0.000	0.011
CH-index	0.000	0.000	0.000	0.000
G-Index	0.483	0.267	0.150	0.300
Model of Delina and Packová	0.933	0.700	0.533	0.722
M-model	0.067	0.033	0.067	0.056
HGN2 Model	0.000	0.067	0.050	0.039
Model of Gulka	0.767	0.467	0.133	0.456
Altman's Z-score	0.367	0.267	0.067	0.233
INDEX IN05	0.500	0.317	0.283	0.367
Quick Test	0.733	0.483	0.383	0.533
Model of Binkert	0.017	0.050	−0.050	0.006
Creditworthiness Index	0.700	0.333	0.333	0.456
Taffler's Model	0.183	0.167	0.117	0.156

Source: Own processing.

Appendix D. F1 Score**Table A5.** Results of testing models for each sector using the F1 Score measure.

Sector	Tourism Sector			Construction Sector			Retail Sector		
	2018	2017	2016	2018	2017	2016	2018	2017	2016
Model of Hurtošová	0.000	0.000	0.000	0.182	0.000	0.000	0.000	0.000	0.000
CH-index	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
G-Index	0.521	0.111	0.000	0.662	0.344	0.396	0.480	0.319	0.196
Model of Delina and Packová	0.888	0.694	0.694	0.789	0.652	0.353	1.000	0.824	0.824
M-model	0.182	0.182	0.000	0.000	0.000	0.333	0.182	0.000	0.000
HGN2 Model	0.000	0.192	0.101	0.000	0.135	0.139	0.000	0.182	0.000
Model of Gulka	0.571	0.571	0.000	0.947	0.750	0.324	1.000	0.571	0.571
Altman's Z-score	0.428	0.339	0.185	0.303	0.269	0.188	0.403	0.394	0.224
INDEX 05	0.558	0.404	0.339	0.438	0.319	0.319	0.259	0.196	0.228
Quick Test	0.799	0.363	0.462	0.519	0.553	0.328	0.480	0.219	0.119
Model of Binkert	0.462	0.000	0.000	0.333	0.182	0.000	0.219	0.000	0.000
Creditworthiness Index	0.615	0.000	0.000	0.697	0.116	0.333	0.581	0.273	0.219
Taffler's Model	0.237	0.237	0.135	0.099	0.116	0.000	0.450	0.377	0.462

Source: Own processing.

Table A6. Average results of testing models using the F1 Score measure.

Model	Average 2018	Average 2017	Average 2016	Average (Total)
Model of Hurtošová	0.061	0.000	0.000	0.020
CH-index	0.000	0.000	0.000	0.000
G-Index	0.554	0.258	0.197	0.336
Model of Delina and Packová	0.892	0.723	0.623	0.746

Table A6. Cont.

Model	Average 2018	Average 2017	Average 2016	Average (Total)
M-model	0.121	0.061	0.111	0.098
HGN2 Model	0.000	0.170	0.080	0.083
Model of Gulka	0.840	0.631	0.299	0.590
Altman's Z-score	0.378	0.334	0.199	0.304
INDEX 05	0.418	0.306	0.295	0.340
Quick Test	0.599	0.378	0.303	0.427
Model of Binkert	0.338	0.061	0.000	0.133
Creditworthiness Index	0.631	0.130	0.184	0.315
Taffler's Model	0.262	0.243	0.199	0.235

Source: Own processing.

Appendix E. Sensitivity and Precision

Table A7. Sensitivity and Precision results of alternative models, 2018.

Characteristics	Model of Hurtošová	CH-Index	G-Index	Model of Delina and Packová	M-Model	HGN2 Model	Model of Gulka
Tourism sector							
Accuracy	0.500	0.333	0.567	0.950	0.550	0.333	0.700
Pe	0.500	0.333	0.333	0.500	0.500	0.333	0.500
Accuracy Adj.	0.000	0.000	0.350	0.900	0.100	0.000	0.400
Sensitivity	0.000	0.000	0.600	1.000	0.100	0.000	0.400
Sp	1.000	1.000	0.800	0.900	1.000	0.750	1.000
Error Type I	1.000	1.000	0.400	0.000	0.900	1.000	0.600
Error Type II	0.000	0.000	0.200	0.100	0.000	0.250	0.000
Pr	0.284	0.221	0.221	0.284	0.284	0.221	0.284
Precision	0.000	0.000	0.460	0.798	1.000	0.000	1.000
F1 Score	0.000	0.000	0.521	0.888	0.182	0.000	0.571
Construction sector							
Accuracy	0.550	0.333	0.733	0.950	0.500	0.333	0.950
Pe	0.500	0.333	0.333	0.500	0.500	0.333	0.500
Accuracy Adj.	0.100	0.000	0.600	0.900	0.000	0.000	0.900
Sensitivity	0.100	0.000	0.800	1.000	0.000	0.000	0.900
Sp	1.000	1.000	0.900	0.900	1.000	0.778	1.000
Error Type I	0.900	1.000	0.200	0.000	1.000	1.000	0.100
Error Type II	0.000	0.000	0.100	0.100	0.000	0.222	0.000
Pr	0.158	0.139	0.139	0.158	0.158	0.139	0.158
Precision	1.000	0.000	0.564	0.652	0.000	0.000	1.000
F1 Score	0.182	0.000	0.662	0.789	0.000	0.000	0.947

Table A7. Cont.

Characteristics	Model of Hurtošová	CH-Index	G-Index	Model of Delina and Pačková	M-Model	HGN2 Model	Model of Gulka
Retail sector							
Accuracy	0.500	0.333	0.667	1.000	0.550	0.333	1.000
Pe	0.500	0.333	0.333	0.500	0.500	0.333	0.500
Accuracy Adj.	0.000	0.000	0.500	1.000	0.100	0.000	1.000
Sensitivity	0.000	0.000	1.000	1.000	0.100	0.000	1.000
Sp	1.000	1.000	0.850	1.000	1.000	0.789	1.000
Error Type I	1.000	1.000	0.000	0.000	0.900	1.000	0.00
Error Type II	0.000	0.000	0.150	0.000	0.000	0.211	0.000
Pr	0.083	0.065	0.065	0.083	0.083	0.065	0.284
Precision	0.000	0.000	0.316	1.000	1.000	0.000	1.000
F1 Score	0.000	0.000	0.480	1.000	0.182	0.000	1.000

Source: Own processing.

Table A8. Sensitivity and Precision results of globally disseminated models, 2018.

Characteristics	Altman's Z-Score	INDEX IN05	Quick Test	Model of Binkert	Creditworthiness Index	Taffler's Model
Tourism sector						
Accuracy	0.433	0.567	0.833	0.433	0.733	0.467
Pe	0.333	0.333	0.333	0.333	0.333	0.333
Accuracy Adj.	0.150	0.350	0.750	0.150	0.600	0.200
Sensitivity	0.800	1.000	0.900	0.300	0.600	0.300
Sp	0.450	0.550	0.900	1.000	0.900	0.650
Error Type I	0.200	0.000	0.100	0.700	0.400	0.700
Error Type II	0.550	0.450	0.100	0.000	0.100	0.350
Pr	0.221	0.221	0.221	0.221	0.221	0.221
Precision	0.292	0.387	0.719	1.000	0.630	0.196
F1 Score	0.428	0.558	0.799	0.462	0.615	0.237
Construction sector						
Accuracy	0.600	0.733	0.767	0.267	0.800	0.400
Pe	0.333	0.333	0.333	0.333	0.333	0.333
Accuracy Adj.	0.400	0.600	0.650	−0.100	0.700	0.100
Sensitivity	0.400	0.800	1.000	0.200	0.700	0.100
Sp	0.800	0.700	0.700	1.000	0.950	0.850
Error Type I	0.600	0.200	0.000	0.800	0.300	0.900
Error Type II	0.200	0.300	0.300	0.000	0.050	0.150
Pr	0.139	0.139	0.139	0.139	0.139	0.139
Precision	0.244	0.301	0.350	1.000	0.694	0.097
F1 Score	0.303	0.438	0.519	0.333	0.697	0.099

Table A8. Cont.

Characteristics	Altman's Z-Score	INDEX IN05	Quick Test	Model of Binkert	Creditworthiness Index	Taffler's Model
Retail sector						
Accuracy	0.700	0.700	0.867	0.333	0.867	0.500
Pe	0.333	0.333	0.333	0.333	0.333	0.333
Accuracy Adj.	0.550	0.550	0.800	0.000	0.800	0.250
Sensitivity	0.800	0.900	1.000	0.300	1.000	0.500
Sp	0.850	0.650	0.850	0.900	0.900	0.950
Error Type I	0.200	0.100	0.000	0.700	0.000	0.500
Error Type II	0.150	0.350	0.150	0.100	0.100	0.050
Pr	0.065	0.065	0.065	0.065	0.065	0.065
Precision	0.270	0.151	0.316	0.172	0.409	0.409
F1 Score	0.403	0.259	0.480	0.219	0.581	0.450

Source: Own processing.

Appendix F. Accuracy of Testing Companies in Difficulty

Table A9. Accuracy of testing companies in difficulty.

Model	2016		2017		2018	
	Accuracy 1	Accuracy 2	Accuracy 1	Accuracy 2	Accuracy 1	Accuracy 2
Tourism sector						
Model of Hurtošová	N/A	0.00	N/A	0.00	N/A	0.00
CH-index	1.00	1.00	1.00	1.00	1.00	1.00
G-Index	0.60	1.00	0.60	1.00	0.50	0.90
Model of Delina and Packová	N/A	0.80	N/A	0.80	N/A	0.70
M-model	N/A	0.00	N/A	0.10	N/A	0.00
HGN2 Model	0.20	0.70	0.10	0.60	0.20	0.70
Model of Gulka	N/A	0.20	N/A	0.30	N/A	0.10
Altman's Z-score	0.00	1.00	0.00	1.00	0.00	1.00
INDEX IN05	0.00	1.00	0.10	1.00	0.00	0.80
QUICK Test	1.00	1.00	0.80	1.00	0.80	1.00
Model of Binkert	0.90	0.90	1.00	1.00	0.90	0.90
Creditworthiness Index	0.60	1.00	0.60	1.00	0.70	0.90
Taffler's Model	0.10	1.00	0.20	0.90	0.20	0.90

Table A9. Cont.

Model	2016		2017		2018	
	Accuracy 1	Accuracy 2	Accuracy 1	Accuracy 2	Accuracy 1	Accuracy 2
Construction sector						
Model of Hurtošová	N/A	0.00	N/A	0.10	N/A	0.10
CH-index	1.00	1.00	1.00	1.00	1.00	1.00
G-Index	0.90	1.00	0.70	1.00	0.80	1.00
Model of Delina and Packová	N/A	0.90	N/A	0.90	N/A	0.90
M-model	N/A	0.00	N/A	0.10	N/A	0.00
HGN2 Model	0.10	0.20	0.40	0.50	0.20	0.60
Model of Gulka	N/A	0.20	N/A	0.30	N/A	0.20
Altman's Z-score	0.40	0.70	0.50	0.80	0.50	0.90
INDEX IN05	0.30	0.90	0.20	0.90	0.00	0.80
QUICK Test	0.60	1.00	0.90	1.00	0.10	1.00
Model of Binkert	0.80	0.80	0.80	0.80	0.60	0.60
Creditworthiness Index	0.70	0.70	0.60	0.80	0.70	0.80
Taffler's Model	0.20	0.40	0.10	0.30	0.10	0.40
Retail Sector						
Model of Hurtošová	N/A	0.00	N/A	0.00	N/A	0.00
CH-index	1.00	1.00	1.00	1.00	1.00	1.00
G-Index	0.40	0.90	0.70	1.00	0.70	1.00
Model of Delina and Packová	N/A	0.70	N/A	0.70	N/A	0.80
M-model	N/A	0.00	N/A	0.00	N/A	0.00
HGN2 Model	0.00	0.40	0.20	0.20	0.10	0.50
Model of Gulka	N/A	0.60	N/A	0.60	N/A	0.50
Altman's Z-score	0.30	0.60	0.50	0.70	0.30	0.60
INDEX IN05	0.30	0.70	0.30	0.70	0.20	0.90
QUICK Test	0.70	1.00	0.80	1.00	0.70	1.00
Model of Binkert	0.60	0.80	0.60	0.70	0.50	0.70
Creditworthiness Index	0.50	0.70	0.40	0.70	0.70	0.90
Taffler's Model	0.10	0.10	0.00	0.10	0.00	0.10

Source: Own processing.

References

- Altman, Edward I. 1983. *Corporate Financial Distress; A Complete Guide to Predicting, Avoiding, and Dealing with Bankruptcy*. New York: Wiley.
- Altman, Edward I. 2002. Revisiting Credit Scoring Models in a Basel II Environment. Prepared for "Credit Rating: Methodologies, Rationale, and Default Risk". London Risk Books. Available online: <https://web.archive.org/web/20060918190436/http://www.stern.nyu.edu/fin/workpapers/papers2002/pdf/wpa02041.pdf> (accessed on 15 July 2021).
- Altman, Edward I. 2018. Applications of Distress Prediction Models: What Have We Learned After 50 Years from the Z-Score Models? *International Journal of Financial Studies* 6: 70. [CrossRef]
- Altman, Edward I., Małgorzata Iwanicz-Drozdowska, Erkki K. Laitinen, and Arto Suvas. 2017. Financial Distress Prediction in an International Context: A Review and Empirical Analysis of Altman's Z-Score Model. *Journal of International Financial Management & Accounting* 28: 131–71.

- Bakeš, Vladimír, and Katarína Valášková. 2018. Application and Verification of Slovak Prediction Models in Conditions of National Economy. *Podniková ekonomika a manažment* 1: 3–15.
- Berthold, Michael R., Christian Borgelt, Frank Höppner, Frank Klawonn, and Rosaria Silipo. 2020. *Guide to Intelligent Data Science: How to Intelligently Make Use of Real Data*, 2nd ed. Cham: Springer.
- Box, George E. P. 1987. Science and Statistics. *Journal of the American Statistical Association* 356: 791–99. [CrossRef]
- Brealey, Richard A., Stewart C. Myers, and Franklin Allen. 2016. *Principles of Corporate Finance*, 12th ed. New York: McGraw-Hill.
- Čámská, Dagmar, and Jiří Klečka. 2020. Comparison of Prediction Models Applied in Economic Recession and Expansion. *Journal of Risk and Financial Management* 13: 52. [CrossRef]
- Čámská, Dagmar. 2014. Modely predikcie finanční tísne s důrazem na tranzitivní ekonomiky. Ph.D. Thesis, Vysoká škola ekonomická v Praze, Prague, Czech Republic.
- Chicco, Davide, and Giuseppe Jurman. 2020. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics* 21: 6. [CrossRef] [PubMed]
- Chrastinová, Zuzana. 1998. *Metódy hodnotenia ekonomickej bonity a predikcie finančnej situácie poľnohospodárskych podnikov*, 1st ed. Bratislava: VÚEPP.
- Cohen, Jacob. 1968. Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological Bulletin* 70: 213–20. [CrossRef]
- Cox, David Roxbee, and David Victor Hinkley. 1974. *Theoretical Statistics*, 1st ed. London: Chapman and Hall/CRC.
- Delina, Radoslav, and Miroslava Packová. 2013. Prediction bankruptcy models validation in Slovak business environment. *E + M Ekonomie a Management* 16: 101–12.
- EC—European Commission. 2014a. Communication from the Commission—Guidelines on State aid for Rescuing and Restructuring Non-Financial Undertakings in Difficulty (2014/C 249/01). Available online: <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex%3A52014XC0731%2801%29> (accessed on 12 October 2020).
- EC—European Commission. 2014b. Commission Regulation (EU) No 651/2014 of 17 June 2014 Declaring Certain Categories of Aid Compatible with the Internal Market in Application of Articles 107 and 108 of the Treaty. Available online: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32014R0651> (accessed on 12 October 2020).
- Gavurová, Beata, Miroslava Packová, Mária Mišanková, and Ľuboš Smrčka. 2017. Predictive potential and risks of selected bankruptcy prediction models in the Slovak business environment. *Journal of Business Economics and Management* 18: 1156–73. [CrossRef]
- Gulka, Martin. 2016. Predictive Model of Corporate Failure in the Slovak Business Environment. *Biatic* 24: 5–9.
- Gurčík, Ľubomír. 2002. G-index—The financial situation prognosis method of agricultural enterprises. *Zemědělská ekonomika* 48: 373–78. [CrossRef]
- Harumová, Anna, and Marianna Janisová. 2014. Rating Slovak enterprises by scoring functions. *Journal of Economics* 62: 522–39.
- Hurtošová, Jana. 2008. Postup vývoja ratingového modelu. *Trendy Ekonomiky a Management* 2: 34–40.
- Hyránek, Eduard, Michal Grell, Ladislav Nagy, and Ivona Ďurinová. 2018. The Economic-mathematical Nature of the HGN Model Concept as a Tool for Measuring Performance of Enterprises. *Journal of Economics* 66: 309–25.
- Kopta, Daniel. 2006. Využití ukazatelů finančního zdraví při hodnocení zemědělských podniků. In *Hradecké ekonomické dny 2006. Podnikání a rozvoj regionu*. Hradec Králové, 7th and 8th February 2006. Hradec Králové: Gaudeamus.
- Korol, Tomasz. 2019. Dynamic Bankruptcy Prediction Models for European Enterprises. *Journal of Risk and Financial Management* 12: 185. [CrossRef]
- Kováčová, Mária, and Pavol Kubala. 2018. Verifikácia predikčných modelov vytvorených na báze diskriminačnej analýzy v podnikoch SR. *Podniková ekonomika a manažment* 1: 52–64.
- Kralicek, Peter. 1993. *Základy finančního hospodaření: Bilance, účet zisků a ztrát, cash-flow, základy kalkulace, finanční plánování, systémy včasného varování*. Praha: Linde.
- Kristóf, Tomás, and Miklós Virág. 2020. A Comprehensive Review of Corporate Bankruptcy Prediction in Hungary. *Journal of Risk and Financial Management* 13: 35. [CrossRef]
- Lee, Alice C., John C. Lee, and Cheng-Few Lee. 2009. *Financial analysis, Planning and Forecasting, Theory and Application*, 2nd ed. Singapore: World Scientific Publishing Co. Pte. Ltd.
- Machek, Ondřej. 2014. Long-term Predictive Ability of Bankruptcy Models in the Czech Republic: Evidence from 2007–2012. *Central European Business Review* 3: 14–17. [CrossRef]
- Metz, Charles E. 1978. Basic principles of ROC analysis. *Seminars in Nuclear Medicine* 8: 283–98. [CrossRef]
- Mihalovič, Matúš. 2018. Využitie skóringových modelov pri predikcii úpadku ekonomických subjektov v Slovenskej republike. *Politická ekonomie* 66: 689–708. [CrossRef]
- National Council of the Slovak Republic. 1991. Zákon č. 513/1991 Zb. Obchodný zákonník. Available online: <https://www.slov-lex.sk/pravne-predpisy/SK/ZZ/1991/513/20201001> (accessed on 10 December 2020).
- National Council of the Slovak Republic. 2002. Zákon č. 431/2002 Z.z. Zákon o účtovníctve. Available online: <https://www.slov-lex.sk/pravne-predpisy/SK/ZZ/2002/431/20210101> (accessed on 10 December 2020).
- Neumaierová, Inka, and Ivan Neumaier. 2005. Index IN05. In *Evropské Finanční Systémy*. Brno: Masarykova Univerzita.
- Neyman, Jerzy, and Egon Sharpe Pearson. 1928. On the Use and Interpretation of Certain Test Criteria for Purposes of Statistical Inference Part I. *Biometrika* 20A: 175–240.

- Neyman, Jerzy, and Egon Sharpe Pearson. 1933. The testing of statistical hypotheses in relation to probabilities a priori. *Mathematical Proceedings of the Cambridge Philosophical Society* 29: 492–510. [CrossRef]
- Prusák, Blažej. 2018. Review of Research into Enterprise Bankruptcy Prediction in Selected Central and Eastern European Countries. *International Journal of Financial Studies* 6: 60. [CrossRef]
- Řezbová, Helena. 2001. Kritéria hodnocení podnikatelských subjektů v zemědělství. Ph.D. Thesis, Česká zemědělská univerzita v Praze, Prague, Czech Republic.
- Režňáková, Mária, and Michal Karas. 2015. The Prediction Capabilities of Bankruptcy Models in a Different Environment: An example of the Altman Model under the Conditions in the Visegrad Group Countries. *Journal of Economics* 63: 617–33.
- Siekelová, Anna, Tomáš Kliešтик, and Peter Adamko. 2018. Predictive ability of chosen bankruptcy models: A case study of Slovak republic. *Economics and Culture* 15: 105–14. [CrossRef]
- Singh, Bhanu Pratap, and Alok Kumar Mishra. 2016. Re-estimation and comparisons of alternative accounting based bankruptcy prediction models for Indian companies. *Financial Innovation* 2: 6. [CrossRef]
- Slovak Business Agency. 2019. Analýza ukazovateľov finančnej výkonnosti MSP účtujúcich v sústave podvojného účtovníctva v roku 2018. Available online: http://www.sbagency.sk/sites/default/files/analiza_ukazovatelov_financnej_vykonnosti_msp_2018_final.pdf (accessed on 5 November 2020).
- Steward, Karen. 2019. Sensitivity vs. Specificity. Technology Networks. Analysis & Separations. Available online: <https://www.technologynetworks.com/analysis/articles/sensitivity-vs-specificity-318222> (accessed on 12 October 2021).
- Sušický, Ján. 2011. Využitelnost bankrotních modelů a jejich aplikace v podmínkách České republiky. Ph.D. Thesis, Česká zemědělská univerzita v Praze, Czech Republic, Praha.
- Taffler, Richard J. 1983. The Assessment of Company Solvency and Performance Using a Statistical Model. *Accounting and Business Research* 52: 295–308. [CrossRef]
- Tomšík, Robert. 2017. *Kvantitatívny výskum v pedagogických vedách. Úvod do metodológie a štatistického spracovania*. Nitra: Pedagogická fakulta, Univerzita Konštantína Filozofa v Nitre.
- Trevethan, Robert. 2017. Sensitivity, Specificity, and Predictive Values: Foundations, Pliabilities, and Pitfalls in Research and Practice. *Frontiers in Public Health* 5: 307. [CrossRef] [PubMed]
- Valášková, Katarína, Lucia Švábová, and Marek Ďurica. 2017. Verification of prediction models in conditions of the Slovak agricultural resort. *Economics Management Innovation* 9: 30–38.
- Zalai, Karol, Dávid Arnold, Jana Šnircová, Elena Moravčíková, Jana Hurtošová, and Daniela Tučníková. 2016. *Finančno-ekonomická analýza podniku*, 9th ed. Bratislava: Sprint 2.
- Zalai, Karol. 2000. *Finančno-ekonomická analýza podniku*, 3rd ed. Bratislava: Sprint vfra.
- Zhu, Wen, Nancy Zeng, and Ning Wang. 2010. Sensitivity, Specificity, Accuracy, Associated Confidence Interval and ROC Analysis with Practical SAS Implementations. Paper present at NESUG Proceedings: Health Care and Life Sciences, Baltimore, MD, USA, November 14–17. Available online: <http://www.nesug.org/Proceedings/nesug10/hl/hl07.pdf> (accessed on 7 July 2021).