The Prediction of Corporate Bankruptcy and Czech Economy's Financial Stability through Logit Analysis

Petr Jakubík
Petr Teplý

This paper can be downloaded at: http://ies.fsv.cuni.cz

The Prediction of Corporate Bankruptcy and Czech Economy's Financial Stability through Logit Analysis

Petr Jakubík*
Petr Teplý[#]

*Czech National Bank and
IES, Charles University Prague,
E-mail: Petr.Jakubik@cnb.cz

[#] EEIP, a.s and
IES, Charles University Prague,
E-mail: petr.teply@gmail.com

Abstract:

This article presents a financial scoring model estimated on Czech corporate accounting data. Seven financial indicators capable of explaining business failure at a 1-year prediction horizon are identified. Using the model estimated in this way, an aggregate indicator of the creditworthiness of the Czech corporate sector (named as JT index) is then constructed and its evolution over time is shown. This indicator aids the estimation of the risks of this sector going forward and broadens the existing analytical set-up used by the Czech National Bank for its financial stability analyses. The results suggest that the creditworthiness of the Czech corporate sector steadily improved between 2004 and 2006, but slightly deteriorated in 2007 what could be explained through global market turbulences.

Keywords: bankruptcy prediction, financial stability, logit analysis, corporate sector risk, JT index

JEL: G32, G33, G21, G28.

Acknowledgements

1. Introduction

Credit scoring methods are a standard part of financial institutions' risk management processes. They allow lenders to rate the creditworthiness of their potential debtors by estimating the probability of default[1], with the aim of maintaining a high-quality loan portfolio. The most common type of credit scoring used in banks for the legal entities segment is financial scoring. In this case, companies are rated using financial parameters derived from their accounting statements. The financial scoring process generates a score expressing the company's creditworthiness. This type of model can be applied analogously to aggregate economic data to construct a financial stability indicator based on the creditworthiness of the non-financial sector. From the credit risk assessment perspective, the indicator can be used to complement the sectoral macroeconomic models that have been estimated for the Czech economy and incorporated into the banking sector stress tests (Jakubík, 2007a).

This article begins with a literature review on credit scoring and bankruptcy prediction models. Section 3 looks briefly at the definition and estimation of scoring models with a primary focus on logit methodology. Section 4 discusses the corporate financial indicators that can be used as explanatory variables for business failure. Section 5 contains a description of the data used to estimate the model. The resulting estimated model is presented in section 6, and section 7 then applies the model to data for the entire sector to estimate a creditworthiness indicator for the non-financial corporations sector. The final section summarises the results.

2. Literature Overview

Although the history of credit came back 5,000 years, the history of credit scoring is much shorter (around 70 years). Credit scoring is essentially a way to identify different groups in population when one cannot see the characteristic that defines the group but only related ones (Thomas et al., 2002). In 1930s first studies on bankruptcy prediction emerged such as Smith and Winakor (1935) or Fitzpatrick (1932) who tried to find the sign of financial failure with comparative analysis of the ratios from frustrated companies and healthy companies. Fisher

---

[1] Default is generally defined as the failure of an obligor to meet its obligations arising under a loan agreement. The Basel Committee on Banking Supervision (2006) defines default as a situation where at least one of the following events has taken place. The first is the situation where the bank finds that the obligor is unlikely to pay its credit obligations in full, without recourse by the bank to actions such as realising security. The second is the situation where the obligor is past due more than 90 days on any of its obligations. In this article, default will mean the failure of the firm.

(1936) introduced the idea of discriminating between groups in a population; he focused, among others, on two species of iris by using measurements of the physical size of the plants. Following the Fischer's idea, Durham (1941) recognized that a similar method could be used for discrimination between good and bad loans. The Durham's work was done as a research project for the US National Bureau of Economic Research and was not used for any predictive purpose. However, no advanced statistical methods or computers were available for the researchers at that time. Hence financial ratios of healthy firms were compared with bankrupted firms and it was found that bankrupted firms reported poorer results than the healthy ones (Thomas et al., 2002).

Beaver (1966) applied a univariate model for discriminating between healthy and bankrupted ratios. He compared a list of ratios individually to for 79 failed firms and a matched sample for 79 healthy firms. Consequently, Beaver investigated how 30 financial ratios could predict the firm's bankruptcy and found that six financial ratios could discriminate well between healthy and bankrupted firms five years before the failure occurs. Although the Beaver's pioneer study presented a simple univariate model, it gave a solid base for future research in this field.

Altman (1968) created a multivariate discriminant model, which became one of the most used of all bankruptcy prediction models. He examined 33 healthy listed firms and 33 bankrupted listed firms[2] in the US manufacturing industry in the 1946-1965 period. Initially, Altman provided a multivariate discriminant analysis (MDA) on 22 financial ratios and constructed the Z-score model that consisted of 5 ratios. Based on the Z-Score he divided firms into three groups when predicting bankruptcy – healthy, bankrupted and the other firms.[3] The model proved to be extremely accurate in predicting bankruptcy (95%). However, this original model suffered several pitfalls such as it was applied on small listed firms and the US manufacturing industry. As a result, Altman expanded his model to larger firms (Altman, Haldeman, Narayanan, 1977), non-listed companies (Altman, 1983) and non-manufacturing companies (Altman, 1995)[4].

In 1970s several academics followed the works by Beaver and Altman for bankruptcy prediction, for example, Deakin (1972) tried to capture best of both models. Consequently, Wilcox (1971), Edmister (1972) and Libby (1975) further developed the models presented by Beaver and Altman.

As follows from the above, until during 1980s the MDA was dominant for bankruptcy prediction. However, method suffered from some assumptions that were violated in reality very often (e.g. the assumption of linearity and normality of the financial ratio distributions was problematic, particularly for the failing firms, or heteroscedasticity of residuals). To overcome some of the disadvantages of MDA and to provide higher prediction accuracy, the MDA was replaced by a logit regression method. The critique of the MDA can be found in Joy and Tollefson (1975), Altman and Eisenbeis (1978), Ohlson (1980) or Dimitras, Slowinski, Susmaga and Zopounidis (1999). In the Czech literature, credit scoring has been studied by, for example, Jakubík (2003). The first authors who used logit methodology for bankruptcy prediction were Santomero and Vinso (1977) and Martin (1977), who examined failures in the US banking sector. Ohlson (1980) applied it more generally also to non-banking firms or, for example, Wiginton (1980), used logit regression in his research. Zmijewski (1984) applied probit regression when predicting bankruptcy.

---

[2] The group of healthy firms matched with the group of bankrupted ones (in terms of size, industry etc.).
[3] The Z-score of these firms was in a gray area (or zone of ignorance), so one could not decide if the firm would be healthy or bankrupted.
[4] The model seems to be convenient also for emerging markets as documented e.g. by Sandin, Porporato (2003).

However, the logit methology suffers some problems such as the assumption that the cumulative distribution of the error term is logistic what does not always hold in reality. Hence in the following years, other methods for bankruptcy prediction have been developed such as classification trees (Breiman et al., 1984), semi-parameter models (Klein and Spady, 1993), neural networks (Zhang, et al., 1999) genetic algorithms (Back et al., 1996) hazard models (Shumway (1999) or Hillegeist et al. (2004) or generalized additive models (Berg, 2007).

3. Logit Methodology

Scoring models play a role in the decision whether or not to provide a loan. In practice, this is done by comparing information available on the client (obtained, for instance, from the client's loan application form or track record) against information on clients to whom loans have been granted in the past and whose quality is known. A predictive scoring model is estimated from the historical information on clients. By applying the model to known information on a potential obligor, one obtains the probability that the obligor will default. The decision is made by comparing the estimated probability of default against some threshold. A survey of these methods in the context of credit scoring can be found, for example, in Hand and Henley (1997) and Rosenberg and Gleit (1994).

A whole range of statistical methods can be used to construct scoring functions, among them linear regression, decision trees, neural networks and expert systems, hazard models (see above). In practice, however, logistic regression is one of the most commonly used methods.

The logit model comes from a simple linear regression that can be described through the following equation:

$$y_i = b_0 + \sum_{i=1}^{N} b_i x_i + e_i \tag{1}$$

where
y_i denotes the probability of default of the firm,

x_i represents the financial indicators of the firm,
b_i expresses the coefficients of the relevant scoring function indicators.

However, y_i (the probability of default) can go outside the interval <0,1>, hence linear regression is an inconvenient method to represent a probability function. To overcome this drawback, we need to normalized y_i into the 0 – 100 % range. Another problem of a linear regression lies in the assumption of homoscedasticity[5] that is often violated in reality. Hence linear models are not used in practice and non-linear models such as logit or probit models are preferred (Sironi, Resti, 2007)[6].

In the logit model, the linearity in equation (1) can be overcome through an exponential transformation (sometimes called as the logistic transformation):

$$y_i = f(w_i) = \frac{1}{1 + e^{-w_i}} \tag{2}$$

---

[5] Homoscedasticity means the constant variance of the residuals.
[6] For a detailed discussion on disadvantages of linear models see the discussion above.

where

y_i denotes the probability of default of the firm,

w_i represents the linear function of the financial indicators $w_i = b_0 + \sum_{i=1}^{N} b_i x_i$ in equation (1).

After providing some calculations we can get equation (3). In this case, it is assumed that the explanatory variables multiplied by the relevant coefficients are linearly related to the natural logarithm of the default rate (referred to as the logit – Mays, 2001).

$$\ln \frac{s}{1-s} = b_0 + \sum_{i=1}^{N} b_i x_i \,, \tag{3}$$

where

s represents the probability of default of the firm at the one-year forecast horizon,

x_i expresses the financial indicators of the firm,

bi denotes the coefficients of the relevant scoring function indicators.

This equation can then be used to derive the relationship for the probability of default. Hence the following relationship can be expressed using a logit curve (Ohlson,1980).

$$s = \frac{1}{1 + e^{-b_0 - \sum_{i=1}^{N} b_i x_i}} \tag{4}$$

In the case of financial scoring, financial indicators based on accounting data are considered as the explanatory variables. The coefficients of the function can be estimated using the maximum likelihood method (Baltagi, 2002). Owing to the large number of indicators that can be included in the model, stepwise regression is used to select the variables. This method involves testing various combinations of variables maximising the quality of the model. The model works with a binary dependent variable (0/1) and can be constructed for computation of either the probability of default or the probability of non-default, depending on the definition of the independent variable in the regression. If we denote a "bad firm" [7] with the value 1, the resulting score obtained from the model corresponds to the probability that the firm will default.[8]

If we assume that a large number of firms are used in order to estimate model (1), then according to the law of large numbers the variable s in equations (1) and (2) corresponds to the proportion of firms that default at the one-year forecast horizon. Assuming that model (2) is estimated on the set of firms to which the function will later be applied, the outcome of the model truly represents the probability of default. As the ratio of good to bad firms in the sample does not usually match the real situation, and given also that accounting data from various moments in time are taken into consideration, the outcome of the model cannot be interpreted as the probability of default. In this context, variable s is usually referred to as the score expressing the riskiness or creditworthiness of the firm.[9]

---

[7] A bad firm is defined here as a firm that defaults during the period under review but was a good firm prior to defaulting. A good firm means a firm that does not default during the period under review.

[8] Some studies, conversely, denote "good firms" with the number 1. In this case, the resulting score represents the probability that the firm will not default.

[9] The figure obtained can be converted to the probability of default with the aid of a suitable transformation. Either parametric or non-parametric estimates can be used for this purpose.

4. Financial indicators

The financial indicators used as the explanatory variables in model (2) can be broken down according to several perspectives – for example the perspectives of lenders, shareholders or state authorities. It is important to emphasise that there is no clear consensus either in theory or in practice on the ideal method for analysing the financial indicators. In the Czech literature, various authors present various breakdowns of relative indicators – see, for example, Blaha, Jindřichovská (2006) and Kislingerová (2007). There is a similar lack of unity in the foreign literature – see, for example, Damodaran (2002) and McKinsey et al. (2005).

Given the primary aim of our research, namely to construct a financial stability indicator based on the prediction of business failure, we chose 22 indicators and divided them into four main groups: liquidity indicators, solvency indicators, profitability indicators and activity indicators. The individual financial indicators are given in Table 1. For each indicator we also indicate its theoretical influence on business failure (positive or negative).

The liquidity indicators explore the firm's ability to meet its short-term liabilities (r_1, r_2, r_{15} and r_{19}) or to cover its long-term liabilities with long-term assets (r_{10}). Generally, higher liquidity implies a lower probability of default. Persisting problems with low liquidity usually indicate problems ahead with meeting long-term liabilities (i.e. declining solvency[10]), which in the extreme case can result in the company failure.

The solvency indicators describe the firm's ability to meet its long-term liabilities. Generally, a higher debt ratio (r_3, r_4 and r_{14}) and a longer debt repayment period (r_9) result in a higher probability of default. By contrast, an ability of the company to generate sufficient funds for debt repayment (r_5, r_6, r_{13} and r_{16}) and a higher proportion of internal funds (r_{17}) reduce this probability.

The profitability indicators explain how the company generates profit and the quantity of inputs it uses to do so. Generally, higher profitability implies a lower probability of default (r_7, r_8, r_{20} and r_{21}).

The activity indicators measure the efficiency of use of various inputs by the company. From the financial point of view, it would be ideal if the company generated sales/profit by using the minimum amount of resources. Generally, the lower the company's efficiency, the higher its probability of default (r_{11}, r_{12} and r_{22}). The *sales turnover ratio* (r_{18}) is constructed so that the value of the indicator rises – and the probability of default falls – as the volume of sales rises.

The potential influence of the individual indicators on corporate bankruptcy can be demonstrated on the following simplified example.[11] One classic symptom of declining solvency is when a company fails to make efficient use of inputs (its activity indicators deteriorate). Cash flows into the firm consequently shrink, leading to a decline in the firm's ability to meet its short-term liabilities (its liquidity indicators deteriorate). Over time, the company proves to be incapable of generating a profit (its profitability indicators deteriorate) to cover its short-term and long-term liabilities (its solvency indicators deteriorate). The firm's liabilities exceed its assets and it goes bankrupt.

To estimate model (1), the financial indicators obtained using the relationships given in Table 1 were further transformed into their relative order vis-à-vis the data sample used. In

---

[10] Liquidity is sometimes referred to as the short-term solvency of a company.
[11] In this simplified example we ignore alternative ways of restoring the firm to health (e.g. corporate restructuring, debt capitalisation and so on).

this way, each indicator value was transformed into a number lying in the interval (0,1). This simple transformation makes the model estimate more robust to outlying values of the indicators considered.

Table 1: Definitions of financial indicators

Ratio	Definition	Notation	Expected impact
Liquidity ratios			
Current ratio	$\dfrac{\text{current assets}}{\text{current liabilities}}$	r1	-
Quick ratio	$\dfrac{\text{cash+ST* receivables}}{\text{current liabilities}}$	r2	-
Cash ratio	$\dfrac{\text{working capital}}{\text{assets}}$	r19	-
Working capital	$\dfrac{\text{financial assets}}{\text{current liabilities}}$	r15	-
Capitalization ratio	$\dfrac{\text{fixed assets}}{\text{long-term liabilities}}$	r10	-
Solvency ratios			
Leverage I	$\dfrac{\text{debt}}{\text{equity}}$	r3	+
Leverage II	$\dfrac{\text{LT** debt+LT** bonds}}{\text{equity}}$	r4	+
Leverage III	$\dfrac{\text{debt}}{\text{assets}}$	r14	+
Debt payback period	$\dfrac{\text{LT** debt+ST* debt}}{\text{operating profit+interest expenses+depreciation}}$	r9	+
Interest coverage	$\dfrac{\text{operating profit+interest expenses}}{\text{interest expenses}}$	r5	-
Cash-flow I	$\dfrac{\text{net profit+depreciation}}{\text{(debt-reserves)/365}}$	r6	-
Cash-flow II	$\dfrac{\text{net profit+depreciation}}{\text{debt/365}}$	r13	-
No credit interval	$\dfrac{\text{money+ST* payables+LT** payables}}{\text{operating expenses}}$	r16	-
Retained earnings	$\dfrac{\text{retained earnings}}{\text{assets}}$	r17	-
Profitability ratios			
Gross profit margin	$\dfrac{\text{operating profit}}{\text{sales}}$	r7	-
Return on assets	$\dfrac{\text{operating profit}}{\text{assets}}$	r8	-
Return on equity	$\dfrac{\text{net profit}}{\text{equity}}$	r20	-
Net profit margin	$\dfrac{\text{net profit}}{\text{sales}}$	r21	-
Activity ratios			
Average receivable collection period	$\dfrac{\text{receivables}}{\text{sales/365}}$	r11	+
Inventory ratio	$\dfrac{\text{inventories}}{\text{sales/365}}$	r12	+
Sales turnover	$\dfrac{\text{sales}}{\text{assets}}$	r18	-
Payables ratio	$\dfrac{\text{ST* payables}}{\text{sales/365}}$	r22	+

* Short-term

** Long-term

Source: Authors

5. Data used

For our research we used the large database of the Czech Capital Information Agency (Česká kapitálová informační agentura, ČEKIA), which contains the accounting statements (balance sheets and profit-and-loss accounts) of selected Czech firms for the period 1993–2005. Of the total of 31,612 firms in the database, 932 went bankrupt. Since some of the accounting statements had been completed very sparsely, we focused on the records of firms whose main economic activity (NACE) was filled in, because for these firms most of the accounting items were filled in as well. In order to estimate the scoring function, from the firms that went bankrupt, we initially selected only those for which there was accounting data one year prior to the declaration of bankruptcy. There were 151 such firms.[12] Then, for the sample of firms that did not fail in the period under review we selected only those for which we had accounting statements for at least two consecutive years.[13] The data sample for the estimation of the model was constructed so as to best capture the true data structure.

Usually, however, a larger proportion of bad firms than exists in reality is included in the sample so that the good and bad firms can be distinguished using statistical methods. Sometimes a sample containing the same number of good and bad firms is used (Wezel, 2005). Generally, the good firms are chosen so as to be as similar as possible to the bad ones according to selected criteria, for example size as measured by assets, number of employees or sales.[14] We also randomly selected accounting periods for which statements were available for the immediately succeeding accounting period. In this way we made sure that the firm in question did not fail in the year following the period under review. In all, 606 good firms were ultimately selected using this procedure. The data sample thus contained a total of 757 firms, which were divided into two categories according to whether they went bankrupt in the period following the period for which the accounting data were selected for the company in question. According to the econometric literature, when the event of interest is rare, logistic regression underestimates the influence of the characteristics on the event, so an artificial sample is generated and the estimated values are further transformed so that they match the incidence in the population.[15]

Table 2 shows the breakdown of the data in the database on the selected data sample by accounting period and firm quality (good/bad). In the total data sample, moreover, there exists a set of firms for which we are unable to determine the quality in the given year (indeterminate firms). These are firms for which accounting statements for the following year are not available. Although the database contained accounting data for the period 1993–2005, in the final year it is no longer possible to determine the firm's quality. For this reason, the selected data sample does not cover 2005.

---

[12] We excluded from our analyses those firms which underwent composition. There were only nine such cases in the database. Unlike bankruptcy, composition is not associated with the dissolution of the legal entity (Jakubík, 2007b).

[13] To estimate the scoring function we need to have corporate accounting data for two consecutive years. The first period is used for estimating the function and the second for identifying the quality of the firm (failed, healthy). If no accounting data are available for the following period, we are unable to determine the quality of the company in question.

[14] A summary of the methods can be found, for example, in Heckman et al. (1997).

[15] For the estimation of the scoring function, an alternative sample constructed in the same way but with a new random selection for the good firms was used in the robustness tests – see section 5, where we discuss the results of the model.

Table 2: Breakdown of data sample by accounting period and firm quality[*]

	Total data				Used data sample		
	Total	Undefined firms	Bad firms	Good firms	Total	Bad firms	Good firms
1993	980	89	1	890	1	1	0
1994	1,824	53	0	1,771	4	0	4
1995	5,606	147	0	5,459	13	0	13
1996	7,023	1,032	9	5,982	53	9	44
1997	7,056	1,261	15	5,780	50	15	35
1998	6,802	1,028	12	5,762	48	11	37
1999	7,541	1,307	25	6,209	69	25	44
2000	7,377	3,094	18	4,265	62	17	45
2001	5,660	1,536	5	4,119	40	5	35
2002	7,869	956	8	6,905	57	8	49
2003	22,264	4,420	25	17,819	110	25	85
2004	18,989	18,490	35	464	250	35	215
Total[**]	98,991	33,413	153	65,425	757	151	606

Source: ČEKIA and authors' calculations
[*] A bad firm means a firm that went bankrupt at the one-year horizon, whereas a good firm for the given period means a firm that did not go bankrupt the following year.
[**] The "Total" row contains the number of observations for the given set of firms. On the full data sample this figure does not equal the total number of firms, because in the selection each company is monitored for several accounting periods.

Table 3 shows the breakdown of firms by size in the data sample. This is based on corporate assets and conforms to the European Commission categorisation.[16] Nonetheless, we should mention that the European Commission also offers other enterprise size categorisations (according, for example, to number of employees or sales).[17] The enterprise size definition chosen by us and used in Table 3 was based on the available data, which were part of the data source used. The source contained corporate assets, and not numbers of employees. Sales information did form part of the database, but had been filled in for only some companies, so it could not be used. Under the definition we used, micro-enterprises with assets not exceeding CZK 60 million have the largest representation in the data sample, while large enterprises with assets exceeding CZK 1,290 million have the lowest representation. However, large enterprises account for more than 80% of the aggregate assets of the firms represented in the sample.

---

[16] Commission Regulation (EC) No 70/2001 as amended by No 364/2004. The enterprise size boundaries were converted from EUR to CZK using the approximate exchange rate 1 EUR = 30 CZK.
[17] The Czech Statistical Office also uses a breakdown by number of employees.

Table 3: Description of data sample used

		Good firms			Bad firms		
Type	Assets (CZK million)	Number of firms	Share according to number of firms (%)	Share according to assets of firms (%)	Number of firms	Share according to number of firms (%)	Share according to assets of firms (%)
Micro firms	< 60	292	48.2%	0.8%	70	46.4%	1.0%
Small firms	61-300	138	22.8%	3.5%	36	23.8%	5.4%
Medium firms	301-1,290	90	14.9%	10.8%	24	15.9%	14.7%
Large firms	>1,291	86	14.2%	84.9%	21	13.9%	78.9%
Total	**-**	**606**	**100.0%**	**100.0%**	**151**	**100.0%**	**100.0%**

Source: ČEKIA and authors' calculations

6. Results of the model

The resulting model (3) confirmed the relationships between the liquidity, solvency, profitability and activity indicators and business failure. The best statistical properties were shown by the model containing seven statistically significant indicators (of the 22 considered in all). These included three solvency indicators (*financial leverage I* and *II* and *interest coverage*), two profitability indicators (*return on equity* and *gross profit margin*), one liquidity indicator (*cash ratio*) and one activity indicator (*inventory ratio*). The resulting model takes the following form:

$$score = \frac{1}{1 + e^{-(b_0 + b_1 r_3^* + b_2 r_4^* + b_3 r_5^* + b_4 r_7^* + b_5 r_{12}^* + b_6 r_{19}^* + b_7 r_{20}^*)}} \tag{3}$$

where

score expresses the risk of the firm, which is linked to the probability that the firm will go bankrupt at the one-year horizon,

r_i represents the individual financial indicators of the firm,

b_i denotes the coefficients of the relevant scoring function indicators,

* denotes the relative order operator in per cent, which returns the relative order of the value of a given indicator for a given firm vis-à-vis the full data sample used to estimate the model.[18]

As the model is based on the relative order of the indicators in the sample, the estimated coefficients of the function express their relative importance. The larger is the indicator's

---

[18] The relative order operator returns a number in the interval (0,1). It is analogous to seeking a quantile on the given data sample, except that the value for which we are seeking the position in the given sample is not part of the sample. In practice, we calculate the value of a given financial indicator, such as the *cash ratio*, and seek the two closest indicator values in the data sample between which the value sought lies. From the relative order of these two values we calculate the relative order for the sought value by linear interpolation. If, for example, the *cash ratio* takes the value 0.2, the relative order operator for it is calculated by linear interpolation of the relative order of the two closest values to 0.2 occurring in the data sample used for the estimation of the model, namely 0.1996 and 0.2015, whose relative orders are 0.5733 and 0.5746. We then obtain the resulting relative order value using the following relationship:

$$0.5733 \cdot \frac{0.2015 - 0.2}{0.2015 - 0.1996} + 0.5746 \cdot \frac{0.2 - 0.1996}{0.2015 - 0.1996} = 0.5736, \text{ i.e. } 0.2^* = 0.5736.$$

This means that in the original data sample on which the model was estimated, 57.36% of the values of this indicator are less than 0.2.

coefficient (in absolute terms), the larger is its weight in the scoring function.[19] From this perspective, *interest coverage*, *cash ratio* and *financial leverage I* appear to be the most important indicators (Table 4).

The estimated scoring model confirmed our expectations regarding the impact of the individual indicators on business failure. It is clear that a higher debt ratio increases the probability of default (see *financial leverage I* and *II*), whereas a higher ability to repay debts (see the *interest coverage*) reduces this probability. Likewise, higher profitability (see *gross profit margin* and *return on equity*) and higher liquidity (see *cash ratio*) increase the financial stability of the firm and reduce its probability of default. By contrast, lower efficiency (see *inventory ratio*) implies lower financial stability of the firm.

Table 4: Estimated scoring model

Variable	Type	Notation of ratio	Notation of coefficient	Coefficient	Standard Error	Siginificance
Constant	-	-	b_0	2.4192	0.9289	0.009207
Leverage I	Solvency	r3	b_1	2.5779	0.3788	0.000000
Leverage II	Solvency	r4	b_2	1.7863	0.5727	0.001813
Interest coverage	Solvency	r5	b_3	-3.4902	1.0005	0.000486
Gross profit margin	Profitability	r7	b_4	-2.4172	0.4802	0.000000
Inventory ratio	Actitvity	r12	b_5	1.7679	0.4033	0.000012
Cash ratio	Liquidity	r19	b_6	-3.3062	0.4246	0.000000
Return on equity	Profitability	r20	b_7	-2.2491	0.5621	0.000063

Source: Authors' calculations

Although the model confirmed some of the expected results, for example that solvency and liquidity ratios are the most important for predicting corporate bankruptcy, one surprising result is the importance of inventories, as contained in the *inventory turnover ratio* (i.e. the number of days a company has goods in stock in the form of inventories). The higher this indicator is, the longer goods lie in the company's store and the less saleable its inventories are.[20] One possible explanation for the importance of this indicator is the high stock of unsaleable inventories typical of businesses heading towards bankruptcy. This argument is supported by the fact that the *total liquidity* indicator, which includes inventories in current assets, proved to be insignificant. Conversely, the *cash ratio*, which does not include inventories in current assets at all, appears to be significant. This implies that the saleability of inventories – among other indicators – plays an important role in the prediction of corporate bankruptcy.

In comparison with other studies on predicting corporate bankruptcy we find similar results for two ratios - *leverage I* and c*ash ratio (*see Table 4*).* This result is not surprising as different authors studied various samples of firms in different periods using different methodologies. The only study listed in Table 5 and focused on Czech companies was provided by Neumaierova (2002), who examined financial statements of 2,000 Czech companies in the 1995-1998 period. We found three similar significant ratios as observed by

---

[19] The relative order operator applied to the individual financial indicators used in the scoring function ensures that the model is robust to extreme values.

[20] Nevertheless, we should point out that different industries display different *inventory ratios*. For example, this indicator, sometimes denoted as *average inventory processing period,* is high for ship manufacture, but very low for retail trade.

Neumaierova (2002) - *leverage I*, c*ash ratio* and *interest coverage*, what implicates the best fit of the mentioned studies. This fit makes sense as we examined similar companies in a similar period but using a different methodology.

Table 5: Comparison with other studies

Author(s) Year Methodology	This study 2008 Logit	Chi, Tang 2006 Logit	Neumaireová 2002 MDA*	Altman 1995 MDA*	Zmijewski 1984 Probit	Ohlson 1980 Logit	Beaver 1966 UM**
Leverage I	a	a	a	a	a	a	a
Leverage II	a						
Interest coverage	a		a				
Gross profit margin	a						
Inventory turnover	a						
Cash ratio	a	a	a	a	a	a	
Return on equity	a						

Notes: The operator "+ " indicates that a study has found a particular financial ratio significant (sometimes in a slightly modified form compared to this paper's defintion).
* Multivariate Discriminant Analysis
** Univariate model

The aim of the scoring model is to correctly separate good and bad firms. This property expresses the quality of the estimated function. To measure it, one can use the Gini coefficient, for example. The value of this coefficient should be as close as possible to 1, which would mean a 100% ability to separate firms in terms of their quality using the scoring function. The quality of the model can be demonstrated graphically by means of a histogram (Chart 1) or a Lorenz curve (Chart 2).

Chart 1: Histogram of estimated scoring function
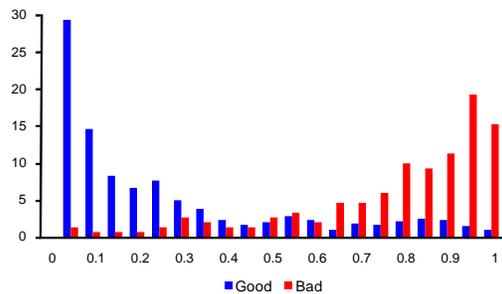
Source: Authors' calculations

Chart 2: Lorenz curve of estimated scoring function
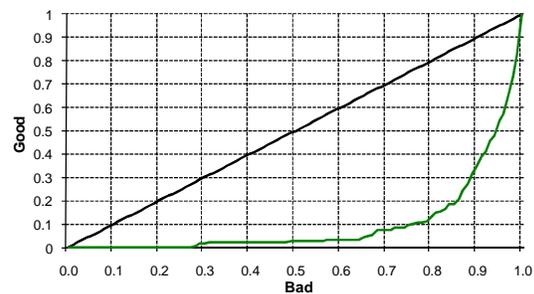
Source: Authors' calculations

Chart 1 shows the firm distribution of the data sample used according to score and according to whether bankruptcy occurred. The blue columns express the percentage of good firms and the red columns the percentage of bad firms for each score interval. The ideal situation would be if all the bankrupt firms were assigned a score of 1 and all the healthy ones a score of 0. This, however, does not happen in practice, as we are unable to observe the complete characteristics of the firm and so we are working with imperfect information. This implies that the function cannot fully separate the firms according to their quality. There is always a

set of bad firms that are classified as good ones, and vice versa. The aim is to keep such cases to a minimum.

Chart 2 depicts the cumulative distribution of the scores of good and bad firms. In the ideal case, guaranteeing a 100% rate of separation, this curve would take the form of a right angle. From the Lorenz curve one can compute the "Gini coefficient" as the ratio of the area enclosed by the green curve and the black diagonal and the total area below that diagonal. The generally accepted Gini coefficient for this type of model fluctuates above 60% depending on the data used and the purpose of the scoring (Mays, 2001). With a Gini coefficient of 80.41%, our estimated model satisfies the requirement of a sufficient rate of separation of the firms on the data sample used[21].

The estimation of the model for the alternative data sample, constructed according to the same rules as the sample used, confirmed that our estimate is sufficiently robust. The robustness of the model was also tested on another alternative data sample consisting of good clients selected entirely at random, and their representation according to the breakdown by assets was different from both the alternative and original data samples. In this case, a slight change was made to the model (two of the seven indicators were replaced with others[22]), but when the model was applied to aggregate data on financial corporations (discussed in section 5), similar results were obtained (the resulting score was different owing to a different ratio of good to bad clients in the sample, but the time profile of the score was similar). The quality of the model as measured by the Gini coefficient was also almost identical.

7.   Use of the model to assess the financial stability of the economy

Financial scoring is routinely used to assess the creditworthiness of individual firms. If we have aggregated data for the whole non-financial sector, we can imagine this sector as one large hypothetical firm with an aggregated balance sheet. Alternatively, given the use of relative indicators only, we can view the aggregated indicators as characteristics of the average firm in the sector. Assuming a degree of homogeneity, the estimated model can be applied to the aggregated indicators of non-financial corporations. If the situation in the sector takes a turn for the worse, the financial indicators of firms will deteriorate on average. This will be reflected in a falling score of the average representative firm. However, the scope and inhomogeneity of the sample of firms on which the model was estimated place some limitations on the model. We could get better results by decomposing the sample into several more homogeneous segments and then estimating the model for these groups of firms separately. In the ideal scenario, we would decompose the firms by size and area of economic activity. Owing to the small number of bad firms in the data source used, however, this is not possible.

An aggregated balance sheet can be obtained for Czech firms from the publicly available data of the Czech Statistical Office, which has data containing the economic results of non-financial corporations. This data is published in a sufficiently detailed structure (to enable the construction of the seven aforementioned indicators included in the model) only for corporations with 100 employees or more. The seven indicators obtained in this way (r_3, r_4, r_5, r_7, r_{12}, r_{19} and r_{20}) are substituted into equation (3) to give an aggregated score representing the level of risk of the entire sector.

---

[21] This result is comparable to Zmijewski (1984) recorded 76% accuracy of classification when employing probit regression.

[22] *Gross profit margin* and *interest coverage* were replaced with *retained earnings* and *cash flow*.

The resulting score was computed for 2004–2007.[23] The value of the creditworthiness indicator (the 1-score) or JT index[24] for 2004–2007 (see Chart 3) can be interpreted as the creditworthiness of the non-financial sector for the one-year prediction horizon. This indicator is related directly to the probability of default of the corporate sector. By contrast with the original data sample, the model is only applied to data on firms with 100 employees or more, but one can get some idea of the evolution of the corporate sector over time. Given the aforementioned limitation, the resulting score is probably underestimated and thus the creditworthiness is overestimated, owing to the higher level of risk of the small enterprises excluded from the aggregate data.

For financial stability purposes, however, the dynamics of this indicator over time are more important than its absolute level. The results suggest a steady improvement in the creditworthiness of the non-financial sector between 2004 and 2006 in line with the positive macroeconomic trend. Although there was a slight decline in *return on equity* and the *gross profit margin* in 2005, the positive development of the other five indicators (a falling debt ratio, rising liquidity, increasing interest coverage and a decreasing inventory turnover ratio) outweighed this effect and the resulting creditworthiness score increased slightly compared to 2004. A positive shift and a reduction in the risk of the sector occurred in particular in 2006, which saw improvements in five out of the seven financial indicators studied (the only deteriorations were recorded by *financial leverage I* and *II*). The biggest improvements were shown by *interest coverage* (a year-on-year improvement of 24.2% to 11.78), *cash ratio* (a year-on-year increase of 24.2% to 0.385) and *return on equity* (a year-on-year improvement of 10.7% to 0.105). The results of the model are consistent with the conclusions contained in the 2006 Financial Stability Report, according to which 2006 was an extraordinarily successful year for the large corporations sector and the outlook for 2007 was favourable (CNB, 2007). The constructed indicator offers a more comprehensive aggregate view of the riskiness of the sector as a whole going forward.

However, the JT index for 2007 is somewhat lower than that for 2006, but is still higher than that for 2005 (see Chart 3). This slight deterioration can be explained through global market turbulences in the year 2007 and expected slow down of the Czech economy. According to these results, the Czech corporate sector risk should show a modest increase in 2008. This expectation is driven chiefly by a higher debt ratio, lower interest coverage and a lower gross profit margin.

---

[23] Since some of the balance sheet items needed to calculate the necessary indicators were not monitored until after 2004, these values for 2004 were estimated from the available data for 2004 and 2005.
[24] The JT index is named after the authors of this study.

Chart 3: JT index for the Czech non-financial corporate sector in the 2004-2007 period
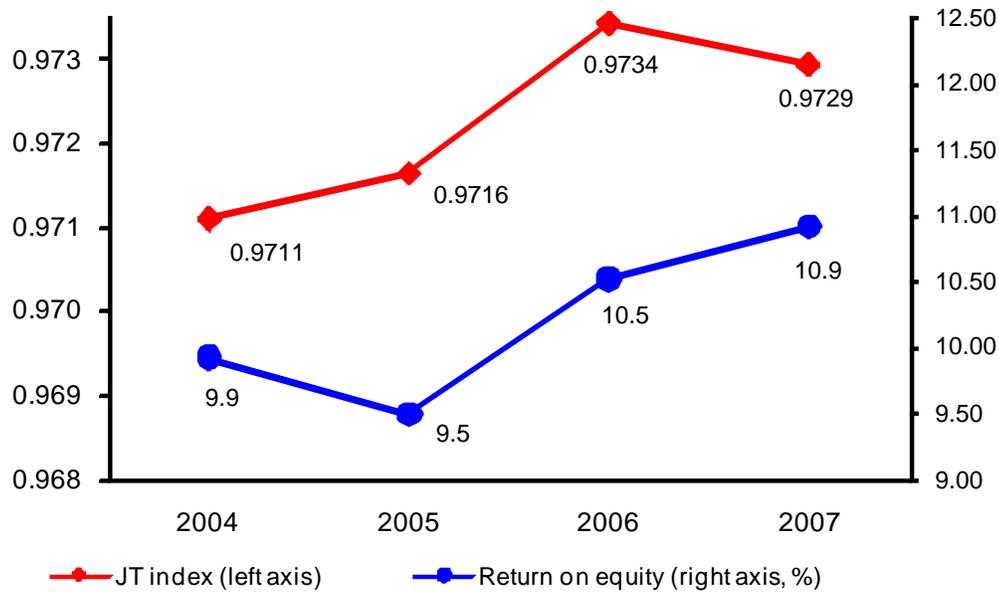
Source: Authors' calculations and Czech Statistical Office

8.   Conclusions

Financial scoring is a method used to assess the creditworthiness of obligors and thus is frequently used by lenders when deciding whether or not to provide credit products. This study showed that it is possible to use these traditional methods to monitor the financial stability of the corporate sector. Using accounting data on Czech firms, a scoring model based on seven financial indicators was estimated using logistic regression. By applying this model to the aggregate financial results of non-financial corporations, the scores of the Czech corporate sector as a whole – corresponding to its level of risk for the one-year prediction horizon – were calculated for 2004–2006. The results of our study suggest that the creditworthiness of the Czech non-financial corporate sector (JT index) improved between 2004 and 2006. However, the JT index for 2007 slightly deteriorated what could be explained through global market turbulences. This indicator will be incorporated into the quantitative system used by the Czech National Bank to assess financial stability. The calculated score will be used each year as auxiliary information for evaluating the probability of the corporate sector running into difficulties at the one-year prediction horizon.

References

Altman, E. I. (1968): *Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy*, Journal of Finance, Vol. 23, No. 4, pp. 589–609

Altman, E.I., Haldeman, R., Narayanan, P. (1977), *ZETA analysis: a new model to identify bankruptcy risk of corporations*, Journal of Banking and Finance, June, Vol. 1 pp.29-54.

Altman, E.I. and Einsenbeis,R.A. (1978): *Financial Applications of Discriminant Analysis: A Clarification*, Journal of Financial and Quantitative Analysis, March, pp. 185-195.

Altman, E.I. (1996): Corporate Bond and Commercial Loan Portfolio Analysis, The Wharton Financial Institutions Center, Wharton School, University of Pennsylvania, September 1996.

Altman, E.I. (1993): Corporate Financial Distress and Bankruptcy: A Complete Guide to Predicting & Avoiding Distress and Profiting from Bankruptcy, Second Edition. John Wiley & Sons.

Altman, E.I. (2000): *Predicting financial distress of companies: Revisiting the z-score and zeta models*, Update of Journal of Finance 1968 and Journal of Banking & Finance 1997.

Baltagi, B. D. (2002): Econometrics, Springer

Back, B., Laitinen, T., Sere, K., and van Wezel M. (1996): Choosing Bankruptcy Predictors Using Discriminant Analysis, Logit Analysis, and Genetic Algorithms, TUCS Technical Report No. 40, September

Basel Committee on Banking Supervision (2006): International Convergence of Capital Measurement and Capital Standards: A Revised Comprehensive Framework, Bank for International Settlements

Beaver, W. (1966): *Financial Ratios as Predictors of Failure*, Journal of Accounting Research, Vol. 4, pp. 71–102

Berg, D. (2007): *Bankruptcy prediction by generalized additive models*, Applied Stochastic Models in Business and Industry, Vol. 23, No. 2, pp. 129-143

Blaha, Z.S., Jindřichovská, I. (2006): Jak posoudit finanční zdraví firmy: finanční analýza pro investory: bankéře, brokery, manažery, podnikatele i drobné akcionáře, 3rd edition, Management press

Breiman L., Friedmann H.L., Olshen R. A., and Stone C.J. (1983): Classification and regression trees, Wadsworth International Group

Chye K. H., Chin T., Peng G. (2005): *Credit Scoring Using Data Mining Techniques*, Singapore Management Review, Volume 26 No 2, pp. 25-46.

CNB (2007): Financial Stability Report 2006, Czech National Bank

CNB (2008): Financial Stability Report 2007, Czech National Bank

Damodaran, A. (2002): Investment Valuation: Tools and Techniques for Determining the Value of Any Asset, 2nd Edition, John Wiley & Sons

Deakin E.B. (1972): *A Discriminant Analysis of Predictors of Business Failure*, Journal of Accounting Research, 10(1), pp.167–179

Dimitras, A. I., Slowinski, R., Susmaga, R., Zopounidis, C. (1999): *Business Failure Prediction Using Rough Sets*, European Journal of Operational Research, Vol. 114, No. 2, pp. 263–280

Edmister, R.O. (1972): *An Empirical Test of Financial Ratio Analysis for Small Business Failure Prediction*, The Journal of Financial and Quantitative Analysis, 7(2), pp. 1477–1493

Eisenbeis, R.A. (1977): *Pitfalls in the application of discriminant analysis in business and economics*, Journal of Finance, vol. 32, pp. 875–900

Fitzpatrick P. (1932): A comparison of ratios of successful industrial enterprises with those of failed firms*,* Certified Public Accountant, October, November, and December, 598-605, 656-662, and 727-731.

Hand, D., Henley, W. (1997): *Statistical Classification Methods in Consumer Credit Scoring: A Review.* Journal of the Royal Statistical Society, Vol. 160, No. 3, pp. 523–541

Heckman, J., Ichimura, H., Todd, P. (1997): *Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program*, Review of Economic Studies, Vol. 64, No. 4, pp. 605–654

Hillegeist, S.,Cram, D., Keating E. and Lundstedt K. (2004): *Assessing the Probability of bankruptcy,* Review of Accounting Studies 9(1), pp. 5–34.

Jakubík, P. (2003): *Úloha skóringu při řízení kreditního rizika*, Acta Oeconomica Pragensia – Finanční krize, VŠE Praha

Jakubík, P. (2007a): *Macroeconomic Environment and Credit Risk*, Czech Journal of Economics and Finance, Vol. 57, Nos. 1 and 2, pp. 60–78

Jakubík, P. (2007b): Exekuce, bankroty a jejich makroekonomické determinanty, IES Working Paper 2007/29

Joy, O.M. and Tollefson, J.O. (1975):*On Financial Application of Discriminant Analysis*, Journal of Financial and Quantitative Analysis, December, pp. 723-739.

Kislingerová, E. (2007): Manažerské finance, 2nd Edition, C. H. Beck

Klein, R., Spady, R. (1993): *An Efficient Semi-parametric Estimator of the Binary Choice Model*, Econometrica 61, pp.387-421

Kolari J., Caputo M.., Wagner D.(1996): *Trait Recognition: An Alternative Approach to Early Warnings Systems in Commercial Banking*, Journal of Business Finance and Accounting, Decemebr 1996, pp. 1415-34

Libby R. (1975): *Accounting Ratios and the Prediction of Failure: Some Behavioral Evidence*, Journal of Accounting Research, 13(1), pp.150–161

Chi, L., Tang, T. (2006): *Bankruptcy Prediction: Application of Logit Analysis in Export Credit Risks*, Australian Journal of Management, Vol. 31, No. 1

Martin, D. (1977): *Early warning of bank failure: A logit regression approach,* Journal of Banking and Finance, Vol. 1, No. 6, November, pp. 249-276.

Mays, E. (2001): Handbook of Credit Scoring, Glenlake Publishing Company, Ltd.

McKinsey et al. (2005): Valuation – Measuring and Managing the Value of Companies, 4th Edition, John Wiley & Sons

Ohlson J. (1980): *Financial ratios and the probabilistic prediction of bankruptcy*, Journal of Accounting Research, 18 (1), pp. 109-131

Rosenberg, E., Gleit, A. (1994): *Quantitative Methods in Credit Management: A Survey*, Operations Research, Vol. 42, No. 4, pp. 589–613

Santomero A. and Vinso J.D. (1977): *Estimating the Probability of Failure for Commercial Banks and the Banking* System, Journal of Banking and Finance.

Smith, R. F. and Winakor A. H. (1935): Changes in the Financial Structure of Unsuccessful Corporations, University of Illinois, Bureau of Business Research

Thomas, L.C., Edelman, D.B., Crook, J.N. (2002): Credit Scoring & Its Applications, Society for Industrial Mathematics, 1st edition

Wezel, T. (2005): Determinants of Foreign Direct Investment in Emerging Markets: An Empirical Study of FDI Flows from Germany and its Banking Sector, Peter Lang GmbH, Frankfurt am Main

Wiginton (1980): *A note on the comparison of logit and discriminant models of consumer credit behavior,* Journal of Finance and Quantitative Analysis, 15, 757-76

Wilcox J.W. (1971): *A Simple Theory of Financial Ratios as Predictors of Failures*, Journal of Accounting Research, 9(2), pp. 389–395

Zhang, G., Hu, M.Y., Patuwo, B.E. (1999): *Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis*, European Journal of Operational Research, vol. 116, pp 16–32.

Zmijewski M (1984): *Methodological issues related to the estimation of financial distress prediction models*, Journal of Accounting Research, 22, Supplement, pp. 59-82

IES Working Paper Series

2008

All papers can be downloaded at: http://ies.fsv.cuni.cz