Distribuční vlastnosti finančních ukazatelů: případ českých konkurzních údajů

The Distributional Properties of Financial Ratios: The Case of Czech Bankruptcy Data

Michal Karas, Mária Režňáková

Abstract:
Purpose of the article: The purpose of this paper is to analyse the distributional properties of financial data, suitable for building a bankruptcy forecast model, in the sense of normality deviation and the existence of outliers.
Methodology/methods: In praxis, financial data in the form of financial ratios is very often not normally distributed. A Shapiro-Wilk's procedure was used to test normality (Shapiro, Wilk, 1965) and a Box-Cox transformation (Box, Cox, 1964) for normalizing financial ratios**.**
Scientific aim: We would like to contributed to the previous pieces of research in following ways. Firstly, by analysing a greater range of accounting ratios or indicators (*i.e.* 44), secondly, by focusing on data of a different character (data suitable for building a bankruptcy forecast model), thirdly, by explaining cases in which the parameter λ is not possible to estimate, and finally fourthly, identifying a possible cause of transformation failure in achieving normality of financial ratios.
Findings: Before the transformation none of the analysed financial ratios met the condition of one-dimensional normality, not even on the 1-% level. After transformation, the condition of one-dimensional normality was met, at the 1-% level, by 34% of the analysed financial ratios. The same condition, but at the 5 or 10-% level, was met by 27% of the analysed financial ratios. The parameter λ was not possible to estimate in the case of 18% of financial ratios.
Conclusions: The condition of normality for untransformed Czech bankruptcy data seems almost as impossible to fulfil. This conclusion implies the use of non-parametric methods, such as artificial neural networks. However, the comparison of the parametric method's performance using untransformed or transformed data is the subject of further research.

Keywords: bankruptcy, financial ratios, outlier detection, normality, data transformation

JEL Classification: G320, G330

Introduction

Beaver (1966) and Altman (1968) were the first to come up with the idea of financial ratios being used to sense the risk of bankruptcy. Many similar models have been built since (Deakin, 1972; Altman, 1977; Ohlson, 1980; Zmijewski, 1984; Shumway, 1999, and others). At present, many authors are endeavouring to find a more perfect classification algorithm. Niemann *et al.* (2008) believe that the choice of classification algorithm offers little leeway for improving the precision of rating models. The remaining potential to increase the precision of a model includes methods of variable choice and methods supporting the statistical significance of predictors. Moreover, there are studies (Grice, Dugan, 2001; Wu, Gaunt, Gray, 2010; Niemann *et al.* 2008) showing that the precision of a bankruptcy model is significantly degraded if used in a field, period, and/or business environment different from that in which the learning data were observed.

Therefore, it is generally not a good idea to use models favoured in the literature believing that they and their predictors will work well even in domestic conditions.

On the other hand, deriving a new model seems to be an effective solution.

The purpose of this paper is to analyse the distributional properties of bankruptcy data, in the sense of normality deviation and the existence of outliers.

Furthermore, we aim to explore the effectiveness of using Box-Cox data transformation to help in achieving approximately normal distribution.

1. The research sample

The research sample consists of 207 Czech industrial companies.

As this paper focuses on data suitable for building a bankruptcy model, the research sample consists only partly of data on active -companies (175 of 207). The rest of the sample is comprised of on bankrupt firms (32 of 207). Only companies with complete financial statements were considered. This approach was chosen for the analysis to include a maximum number of potential predictors. The sample data included financial statements submitted one year prior to the bankruptcy. As Beaver-Altman's matched-pairs approach, that is comparing only companies of identical sizes, was not used on purpose, the observed sample includes companies of different sizes. The reason is the following: the company size as such may itself be a significant bankruptcy indicator in the first place (see Ohlson, 1980; Peel & Peel, 1987). Second, as bankruptcy is a rare occurrence[1], this matching may influence the sample size and, thus, the number of the degrees of freedom (Taffler, 1982).

Table 1. The sample characteristics.

	N	Aver.	Min,	Max.	1% quan.	99% quan.	Std. Dev.
Total Assets (Active)	175	7 353 172	267 425	138 464 258	355 760	68 275 976	14 671 915
Sales (Active)	175	6 508 839	352 117	102 159 712	439 028	44 167 126	10 897 050
Total Assets (Bankrupt)	32	487 055	13 077	3 162 368	13 077	3 162 368	784 593
Sales (Bankrupt)	32	550 966	19 514	3 250 613	19 514	3 250 613	909 766

Source: Our own analysis of data from the Amadeus database.

The period of interest is the time period 2007–2010.

Each analysed company is represented by 44 financial ratios. These financial ratios were used in previous bankruptcy prediction studies (Beaver, 1966; Altman 1968; Deakin, 1972; Ohlson, 1980; Ding *et al.*, 2008; Wang, Lee 2008; Niemann et al, 2008; Beaver, 2005; Tseng, Hu, 2010; Psillaki, Tsolas, Margaritis, 2009). Unlike the research performed by Nikkinen and Sahlstrőm (2004), who test 12 ratios (8 accounting, 4 market-based) in ten countries, we test 44 accounting ratios in one country, *i.e.* the Czech Republic.

Furthermore, the research performed by Nikkinen and Sahlstrőm (2004), focuses only on active companies.

The ratios analysed[2] are listed in the following table 2.

1.1 The theoretical cause of non-normality of financial ratios

Deviation from normal distribution may be caused by a lack of proportionality between the numerator (Y) and denominator (X) of financial ratios (Barnes, 1982, 1987; Nikkinen, Sahlstrőm, 2004). In the case studied, proportionality means, that: "*the relationship between the two variables is linear and the*

Table 2. List of analysed financial ratios.

1	CA/TA	Current assets/total assets	23	OI/AC	Oper. income (loss)/average capital
2	CD/S	Current debt/sales	24	OP/OR	(Oper. revenue – oper. cost)/oper. Rev.
3	CF/S	Cash flow/sales	25	OR/CA	Oper. revenue/current assets
4	CF/TA	Cash flow/total assets	26	OR/CL	Oper. revenue/current liabilities
5	CF/TD	Cash flow/total debt	27	OR/FA	Oper. revenue/fixed assets
6	CR	Current ratio	28	OR/LTL	Oper. revenue/long-term liabilities
7	DR	Debt ratio	29	OR/TA	Oper. revenue/total assets
8	E/TA	EBIT/total assets	30	OR/TL	Oper. revenue/total liabilities
9	EBIT (E-vol)	EBIT (3-year volatility)	31	PM	profit margin (3-year average)
10	EBIT/Int.	EBIT/interest	32	QA/S	Quick assets/sales
11	EBITDA/Int.	EBITDA/interest	33	QA/TA	Quick assets/total assets
12	EBITDA/TL	EBITDA/total liabilities	34	RE/TA	Retained earnings/total assets
13	EBT/OR	Income/loss before tax/oper. rev.	35	S	Log of sales
14	EQ	log of equity	36	S/TA	Sales/total assets
15	FA/LTL	Fixed assets/long-term liabilities	37	TA	Total assets
16	NI/AC	Net income (loss)/average capital	38	TD/EDA	Total debt/EBITDA
17	NI/CA	Net income/current assets	39	TL/TA	Total liabilities/total assets
18	NI/FA	Net income/fixed assets	40	WC/OE	Working capital/operating cost
19	NI/OR	Net income/oper. revenue	41	WC/S	Working capital/sales
20	NI/TA	Net income/total assets	42	WC/TA	Working capital/total assets
21	NI-change	Ohlson change	43	Tan.A/Tot.A	Tangible assets/total assets
22	OC/OR	Oper. cost/oper. revenue	44	Int.A/Tot.A	Intangible assets/total assets

Source: Beaver, 1966; Altman ,1968; Deakin, 1972; Ohlson, 1980; Ding et al., 2008; Wang, Lee, 2008; Niemann et al, 2008; Beaver, 2005; Tseng, Hu, 2010; Psillaki, Tsolas, Margaritis, 2009.

constant is zero" (Whittington, 1980). The relation between Y and X can be described by using linear regression as an alternative to using ratio, this means:

$$Y = \alpha + \beta \cdot X + \varepsilon \,, \tag{1}$$

where:

$$\varepsilon \sim \mathrm{N}\left(0, \sigma^2\right), \tag{2}$$

and

α, β are regression parameters,

ε is residual.

By dividing *X* we obtain the following formula:

$$\frac{Y}{X} = \frac{\alpha}{X} + \beta + \frac{\varepsilon}{X} \,, \tag{3}$$

if $$\alpha = 0 \,, \tag{4}$$

then $$\frac{Y}{X} = \beta + \frac{\varepsilon}{X} \,. \tag{5}$$

And the ratio of *Y*/*X* represents an unbiased estimate of the β parameter (Barnes, 1982).

The nonzero value of α is supposed to be a cause of β parameter bias. The use of linear regression represents a more robust alternative method to using financial ratios in financial analysis (Barnes, 1982).

Nikkinen and Sahlstrőm (2004) offer an alternative explanation of this problem, claiming that: *"the non-normality of financial ratios may also be caused by their definitions. Some financial ratio, such as quick ratio and current ratio, are limited to be greater than zero, and some ratios, such as equity to total capital ratio, have an upper limit of 100%."* We further refer to this type of ratios as bounded. The problem of bounded ratios is analysed in more detail in a paper by McLeay and Omar (2000). Zimmerman (1994, 1995 and 1998) was concerned with the influence of non-normality and outliers on the precision of parametric (t-test) and non-parametric testing (Mann-Whitney-Wilcoxon U-test). He found that non-normality and the existence of extreme outliers also influences the results of non-parametric tests, in terms of the second-type error.

Furthermore, we aim to explain the ability of financial ratios to achieve extreme values, *i.e.* outliers. We believe that, the cause of this problem could also be found in the definition of financial ratios, as bounded ratios cannot reach extreme values.

2. Methods

2.1 Testing normal distribution

A Shapiro-Wilk's procedure was used to test normality (Shapiro, Wilk, 1965). This test is especially suitable for small-sized samples (Nikkinen, Sahlstrőm, 2004; Meloun, Militký, 1994; Hebák *et al.*, 2004). The Shapiro-Wilk's procedure tests the null hypothesis that a sample $x_1, x_2, \dots, x_n$ came from a normally distributed population. The test statistic is (Hebák *et al.*, 2004):

$$SW = \frac{\left[\sum_{i=1}^{n} a_i x_{(i)}\right]^2}{Q(x)}, \tag{6}$$

where

$$Q(x) = \left(x_i - \bar{x}\right)^2, \tag{7}$$

and

$x_{(i)}$ are order statistics,
a_i are constants specially derived by Shapiro and Wilk for the purposes of this test, these constant are tabulated.

2.2 The transformation of the data

Data transformation techniques are frequently used in cases in which an approximate normal distribution of data is needed. Deakin (1976) was the first to find, that a logarithm or root data transformation could be used, in certain cases, to achieve normality of data. Watson (1990) extends this idea by suggesting a use of Box-Cox data transformation (Box, Cox, 1964).

Nikkinen and Sahlstrőm (2004) explore the suitability of Box-Cox data transformation for normalizing financial ratios based on different accounting concepts. They found, that using this transformation results in substantial approximation to normality.

In the event that non-normality is proved, the Box-Cox transformation (Box, Cox, 1964) will be used to achieve an approximately normal distribution of the data.

This is a form of power transformation designed by Box and Cox (Box, Cox, 1964).

The y representing a non-normal financial ratio can be transformed to a normally distributed variable with mean μ and variance σ^2 by using the following transformation formula:

$$y^{(\lambda)} = \begin{cases} \dfrac{(y+\lambda_2)^{\lambda_1} - 1}{\lambda_1} & ; \ \lambda_1 \neq 0 \\ \ln(y+\lambda_2) & ; \ \lambda_1 = 0 \end{cases}. \tag{8}$$

The parameter λ_1 can be estimated by maximizing the log-likelihood function (Nikkinen, Sahlstrőm, 2004):

$$\ell = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{n}\left(y_t^{(\lambda)} - \mu\right)^2 + $$
$$+(\lambda - 1)\sum_{t=1}^{n} y_t. \tag{9}$$

In the case of a negative value of the financial ratio (y) a positive constant λ_2 is added to ensure positivity of the variable $y+\lambda_2$ to be transformed.

Here the indicators of sales (S), total assets (TA), and equity (EQ), originally designed as logarithms, are considered non-logarithm values. The logarithm of a value as such is a special case of Box-Cox transformation for $\lambda_{1,2}=0$ (see equation 5). The value of λ_1 taken to be the maximum likely estimate, its value need not be assumed. In some cases, the value of the parameter may diverge or, if strongly non-normal, the transformation may not achieve normality at all within the preset value of the Shapiro-Wilk's test.

2.3 The outlier's problem

Outliers can be seen, as noted by, Škapa (2011), as: "*an observation that is very different from the rest of the data*".

It has been proved that outliers do influence both parametric and non-parametric tests (see Zimmerman, 1994, 1995, 1998).

When setting up a bankruptcy model, outliers are often *winsorized* (Shumway, 1999; Wu, Gaunt, Gray, 2010) or even removed (Mileris, Boguslauskas, 2011).

For outliers detection we used the Grubbs test (see Grubbs, 1969), which tests the null hypothesis, that there are no outliers in the data sample. The Grubbs test statistic is the largest absolute deviation from the sample mean in units of the sample standard deviation and can be written as follows (Grubbs, 1969):

$$G = \frac{\max\limits_{i=1,2,\dots,N}\left|Y_i - \bar{Y}\right|}{s}, \tag{10}$$

where $\bar{Y}$ and s denotes the sample mean and the standard deviation, respectively.

To demonstrate the influence of outliers, the mean of each analysed ratio is compared to its 5-% winsorized mean (see table 3a or 3b).

The winsorized mean is given by the formula (Meloun, Militký, 1994):

$$\bar{x}_w(\vartheta) = \frac{1}{n}\left[(M+1)\left(x_{(M+1)} + x_{(n-M)}\right) + \sum_{i=M+2}^{n-M-1} x_{(i)}\right], \tag{11}$$

Table 3a. Descriptive statistics of research sample.

	Mean	Wins. mean	Grubbs T.	p-value	Min.	Max.	Std. Dev.
CA/TA**	0.5396	0.5414	2.08303	1,00000	0.0337	0.997	0.2429
CD/S*	0.3846	0.3658	6.20669	0,00000	0.0792	2.84	0.2922
CF/S	0.0765	0.081	5.60777	0.000001	–0.7775	0.5516	0.1523
CF/TA	0.0621	0.0744	9.14075	0,00000	–1.4268	0.4024	0.1629
CF/TD	0.2237	0.2097	7.58014	0,00000	–0.6335	2.23	0.3639
CR*	1.47	1.92	6.46265	0,00000	0.2459	7.85	0.9398
DR*	0.5483	0.5297	6.16629	0,00000	0.1037	2.46	0.2929
E/TA	0.0327	0.042	7.37413	0,00000	–1.0425	0.3866	0.1458
EBIT(3-vol)	271916.3	189360.5	10.04225	0,00000	435	7949341	764512
EBIT/Int.	6858	62	9.64423	0,00000	–2937	505043	51656
EBITDA/Int.	7957	110	9.48666	0,00000	–692	576293	59909
EBITDA/TL	0.2433	0.2372	6.76295	0,00000	–1.9357	2.69	0.3643
EBT/OR	0.0283	0.0362	9.33891	0,00000	–1.311	0.438	0.1434
EQ	3839644	2733991	9.45778	0,00000	–2854651	100673936	10238589
FA/LTL*	430.6695	15.8119	11.07392	0,00000	0.0232	43061	3850
Int.A/Tot.A**	0.013	0.0072	11.00193	0,00000	–0.2123	0.7364	0.0657
NI/AC	8.17	0.6487	13.56418	0,00000	–14.4024	962.1945	70.3266
NI/CA	0.0614	0.0627	7.24234	0,00000	–2.4695	1.82	0.3495
NI/FA	–0.0889	0.163	14.13158	0,00000	–56.1132	4.11	3.45
NI/OR	0.0207	0.0292	9.67387	0,00000	–1.3133	0.3782	0.1379
NI/TA	0.0194	0.0319	9.22443	0,00000	–1.476	0.3596	0.1621
NI-change	–0.063	–0.063	1.81434	1,00000	–1	1	0.5859
OC/OR	1.032	1.0403	9.48436	0,00000	–0.2154	1.69	0.1315

Source: Our own analysis of data from the Amadeus database.

Table 3b. Descriptive statistic of research sample.

	Mean	Wins. mean	Grubbs T.	p-value	Min.	Max.	Std. Dev.
OC/OR	1.032	1.0403	9.48436	0,00000	–0.2154	1.69	0.1315
OI/AC	106.2629	18.9703	9.27543	0,00000	–0.5117	5578.859	590.0101
OP/OR	0.032	0.0403	9.48436	0,00000	–1.2154	0.3069	0.1315
OR/CA	3.1376	3.0325	4.63889	0.000396	0.3758	12.0265	1.62
OR/CL	4.0263	3.939	4.50259	0.000815	0.4443	13.5861	2.1232
OR/FA	18.9458	6.1668	13.8955	0,00000	0.1575	1997.0673	142.357
OR/LTL	1045.81	98.74	8.62112	0,00000	0.4486	58065	6614
OR/TA	1.06	1.17	6.43	0,00000	0.1497	8.75	1.0762
OR/TL	3.1336	3.0718	5.25162	0.000011	0.2986	12.799	1.04
PM	0.0131	0.0358	13.19982	0,00000	–4.0439	0.438	0.3073
QA/S*	0.2423	0.233	4.51476	0.000765	–0.1965	1.57	0.2267
QA/TA**	0.3848	0.3847	2.39386	1,00000	0.0204	0.9063	0.2178
RE/TA	0.1978	0.226	8.06648	0,00000	–2.7693	0.807	0.3678
S*	5587815	4748118	9.06	0,00000	19514	102159712	10251141
S/TA*	1.32	1.19	6.91778	0,00000	0.0694	8.088	0.9663
TA*	6291743	4836953	9.63686	0,00000	13077	138464258	13715308
Tan.A/Tot.A**	0.382	0.378	2.81101	0.948021	0.0022	1.0627	0.2421
TD/EDA	6.57	6.0982	9.16227	0,00000	–64.6618	161.1914	16.9113
TL/TA*	0.5482	0.5296	6.16902	0,00000	0.1037	2.46	0.2928
WC/OE	0.0799	0.0918	7.39807	0,00000	–1.9058	1.0063	0.2684
WC/S	0.0979	0.1054	6.26882	0,00000	–1.5473	1.0531	0.2624
WC/TA	0.1049	0.1137	7.25891	0,00000	–1.7676	0.7635	0.258

Source: Our own analysis of data from the Amadeus database.

where

$$M = \text{int}(\vartheta n/100), \qquad (12)$$

and

ϑ stands for the percentage of "cut-off" order statistics,
$x_{(i)}$ is the i-th order statistics,
n is the sample size.

3. Research results (findings)

Descriptive statistics of the analysed financial ratios are listed in the following table 1. A ratio or indicator limited by its nature to be greater than zero is denoted by (*). A ratio/indicator limited with both lower and upper boundaries is denoted by (**).

The cause can be seen in the previously mentioned ratio definition. The ratios bound to be greater than zero, but lower than 100%, *i.e.* CA/TA, QA/TA and Tan. A/ Tot. A, cannot theoretically reach any outlier value.

Two exceptions were found by analysing Grubbs test results. The first exception is the ratio of intangible assets to total assets (Int. A/ Tot. A). This ratio is bounded in the same way as the previously mentioned ratios are, but, according to the results of the Grubbs test, at least one outlier value in the calculated values is detected. The second exception is represented by the *NI-change* indicator, which measures the Ohlson's change of net income (NI). The *NI-change* was calculated in following way:

$$NI - change = \frac{NI_t - NI_{t-1}}{|NI_t| + |NI_{t-1}|}. \qquad (12)$$

As there are no theoretical reasons to expect a bounded value of net income (*NI*), we can assume this ratio to be unbounded. Even though this ratio is of an unbounded character, it does not exhibit outliers values.

The Ohlson's approach to net income could mean an alternative solution to the outlier problem.

3.1 Results of Shapiro-Wilk's normality test

The results of the SW test of non-transformed financial ratios are listed in the following table 4a or 4b.

As can be seen from table 4a or 4b, none of the non-transformed ratios reached normality at least at

Table 4a. SW normality test results of non-transformed data.

Ratio	SW	p-value
CA/TA**	0.97494	0.00094
CD/S*	0.73854	0.00000
CF/S	0.84343	0.00000
CF/TA	0.60931	0.00000
CF/TD	0.77269	0.00000
CR*	0.80014	0.00000
DR*	0.81964	0.00000
E/TA	0.70822	0.00000
EBIT(3-vol)	0.33750	0.00000
EBIT/Int.	0.11587	0.00000
EBITDA/Int.	0.11457	0.00000
EBITDA/TL	0.76464	0.00000
EBT/OR	0.69612	0.00000
EQ	0.37021	0.00000
FA/LTL*	0.08686	0.00000
Int. A/Tot. A**	0.23855	0.00000
NI/AC	0.09843	0.00000
NI/CA	0.76740	0.00000
NI/FA	0.10708	0.00000

Source: Our own analysis of data from the Amadeus database.

Table 4b. Shapiro-Wilk's normality test results of non-transformed data.

Ratio	SW	p-value
NI/OR	0.65464	0.00000
NI/TA	0.59918	0.00000
NI-change	0.95209	0.00000
OC/OR	0.67806	0.00000
OI/AC	0.1634	0.00000
OP/OR	0.67806	0.00000
OR/CA	0.84513	0.00000
OR/CL	0.92851	0.00000
OR/FA	0.08884	0.00000
OR/LTL	0.1405	0.00000
OR/TA	0.79354	0.00000
OR/TL	0.92234	0.00000
PM	0.28459	0.00000
QA/S*	0.89543	0.00000
QA/TA**	0.92687	0.00000
RE/TA	0.68565	0.00000
S*	0.46353	0.00000
S/TA*	0.82285	0.00000
TA*	0.42715	0.00000
Tan. A/Tot. A**	0.97039	0.00024
TD/EDA	0.63474	0.00000
TL/TA*	0.81934	0.00000
WC/OE	0.74871	0.00000
WC/S	0.86486	0.00000
WC/TA	0.87421	0.00000

Source: Our own analysis of data from the Amadeus database.

Table 5. The maximum likelihood estimates of Box-Cox transformation parameters.

Ratio	λ_1	λ_2	LCL	UCL	Ratio	λ_1	λ_2	LCL	UCL
CA/TA**	0.8799	0.9663	0.0435	1.37	OC/OR	5.000*	1.54		
CD/S*	–3.2487	0.9208	–4.0638	–2.4869	OI/AC	–0.39	1.17	–0.4845	–0.3015
CF/S	1.96	1.75	0.9666	2.08	OP/OR	5.000*	2.54		
CF/TA	5.000*	2.68			OR/CA	–0.2318	0.6242	–0.4822	0.0163
CF/TD	–0.8357	1.35	–1.269	–0.4094	OR/CL	0.1863	0.5557	–0.0603	0.4336
CR*	–0.5932	0.7541	–0.9118	–0.2846	OR/FA	–0.4434	0.8425	–0.5733	–0.3238
DR*	–0.9379	0.8963	–1.5098	–0.412	OR/LTL	–0.1743	0.5514	–0.2367	–0.1153
E/TA	5	2.0425			OR/TA	–0.5687	0.8503	–0.8775	–0.2708
EBIT(3-vol)	0.0275	0	–0.0352	0.0903	OR/TL	0.1341	0.7014	–0.1138	0.3817
EBIT/Int.	0.0207	2938	–0.0091	0.0517	PM	5.000*	5.0439		
EBITDA/Int.	–0.0752	693	–0.1051	–0.0446	QA/S*	–1.456	1.65	–2.1604	–0.768
EBITDA/TL	0.839	2.57	0.5222	1.1806	QA/TA**	–1.6071	0.9796	–2.5073	–0.7145
EBT/OR	4.08	2.311	3.42	5.28	RE/TA	4.95	3.93	3.75	5.92
EQ	0.197	2854652	0.158	0.2423	S*	0.1486	0	0.0809	0.2174
FA/LTL*	–0.2834	0.9768	–0.3667	–0.2078	S/TA*	–0.4949	0.9306	–0.8215	–0.1794
Int. A/Tot. A**	–4.2689	1.23	–5.0365	–3.5242	TA*	0.0765	0	0.0109	0.1431
NI/AC	–0.2346	15.4024	–0.2994	–0.1696	Tan. A/Tot. A**	0.0049	0.9978	–0.7925	0.7908
NI/CA	2.0262	3.95	1.89	2.07	TD/EDA	0.8511	65.6618	0.7241	0.9874
NI/FA	5.000*	57.1132			TL/TA	–0.938	0.8963	–1.5098	–0.4122
NI/OR	5.000*	2.33			WC/OE	3.0683	2.58	2.09	3.74
NI/TA	5.000*	2.476			WC/S	2.08	2.73	1.61	2.22
NI-change	0.7582	2	0.3475	1.1736	WC/TA	3.82	2.76	2.91	4.77

Source: Our own analysis of data from the Amadeus database.

Table 6. Spearman's correlation coefficient.

	W1	L	SB	KB	SA	KA	SR	KR	W2
W1	1	–0.0856	**–0.36161**	**–0.97093**	0.02319	**–0.46605**	0.11382	0.23828	0.18084
L	–0.0856	1	**–0.75516**	0.17176	**0.30584**	**0.4072**	**0.53817**	0.23015	0.03181
SB	**–0.36161**	**–0.75516**	1	0.26159	–0.08584	–0.15856	**–0.46399**	–0.27422	–0.08774
KB	**–0.97093**	0.17176	0.26159	1	–0.04214	**0.54249**	–0.05356	–0.20008	–0.10677
SA	0.02319	**0.30584**	–0.08584	–0.04214	1	0.12911	0.23834	0.03672	0.17048
KA	**–0.46605**	**0.4072**	–0.15856	**0.54249**	0.12911	1	**0.50726**	**0.40925**	–0.11981
SR	0.11382	**0.53817**	**–0.46399**	–0.05356	0.23834	**0.50726**	1	**0.80348**	–0.08281
KR	0.23828	0.23015	–0.27422	–0.20008	0.03672	**0.40925**	**0.80348**	1	–0.14462
W 2	0.18084	0.03181	–0.08774	–0.10677	0.17048	–0.11981	–0.08281	–0.14462	1

Source: Our own analysis of data from the Amadeus database.

the 1% level of the SW test. As the non-normality of the ratios had been proved, the transformation of the data was needed to achieve approximate normality. Table 5 shows the details of Box-Cox transformation, such as the maximum likelihood estimates of parameter λ. The results of the SW test of transformed financial ratios are shown in table 7.

As can be seen from table 5, the parameter λ_1 (denoted in the table as L) was not estimated in the case of 8 ratios (18.1% of all ratios). In searching for the cause of this effect, the following possible causes were explored: value of SW statistics of non-transformed ratios (SW 1), value of SW statistics of transformed ratios (SW 2), skewness before transformation (SB), skewness after transformation (AB), kurtosis before transformation (KB), skewness reduction (SR), kurtosis reduction (KR). A non-parametric Spearman's correlation coefficient was used to find the cause. Statistically significant correlations at the 5% level are shown in bold face (see table 6).

The strongest correlation can be found between λ_1 and the skewness before transformation. This could mean that values of skewness before transformation are a cause of not estimating parameter λ_1.

Table 7. Shapiro-Wilk's normality test results of transformed data.

Ratio	SW	p-value	Ratio	SW	p-value
CA/TA**	0.97530	0.00105	OC/OR	0.88269	0.000000
CD/S*	0.98377	0.01752	OI/AC	0.99191	0.307440
CF/S	0.8516	0.00000	OP/OR	0.88269	0.000000
CF/TA	0.90394	0.00000	OR/CA	0.99599	0.868010
CF/TD	0.91921	0.00000	OR/CL	0.99622	0.894580
CR*	0,99493	0.71642	OR/FA	0.99219	0.336720
DR*	0,96877	0.00015	OR/LTL	0.99382	0.578700
E/TA	0,90409	0.00000	OR/TA	0.99624	0.897280
EBIT(3-vol)	0,99614	0.88503	OR/TL	0.9952	0.758930
EBIT/Int.	0,22135	0.00000	PM	0.77717	0.000000
EBITDA/Int.	0,31618	0.00000	QA/S*	0.98884	0.106820
EBITDA/TL	0,76437	0.00000	QA/TA**	0.97750	0.002120
EBT/OR	0,87393	0.00000	RE/TA	0.97304	0.000530
EQ	0,68301	0.00000	S*	0.95281	0.000000
FA/LTL*	0,99239	0.38960	S/TA*	0.99662	0.934320
Int. A/Tot. A**	0,34896	0.00000	TA*	0.98603	0.039050
NI/AC	0,31079	0.00000	Tan. A/Tot. A**	0.97405	0.000710
NI/CA	0,80617	0.00000	TD/EDA	0.65200	0.000000
NI/FA	0.41065	0.00000	TL/TA	0.96865	0.000140
NI/OR	0.86194	0.00000	WC/OE	0.87771	0.000000
NI/TA	0.89175	0.00000	WC/S	0.91385	0.000000
NI-change	0.95193	0.00000	WC/TA	0.98291	0.012980

Source: Our own analysis of data from the Amadeus database.

The stronger the skewness before transformation is, the higher the λ_1 coefficient is. As the skewness reaches a higher value than approximately 6[3], the parameter λ_1 bounded by an upper limit of 5 cannot be estimated.

Moreover, the kurtosis before transformation does not seem to affect the possibility of estimating the λ_1 parameter, as there is no statistically significant correlation between them.

As is shown in table 7, the Box-Cox transformation led to normality being achieved at the 1% level of the SW test in 15 cases (34.09%) and at the 5 or 10% level in 12 cases (27.27%).

The values of Spearman's correlation coefficient from table 6 could again be used to find a possible cause of this effect. There is no statistically significant correlation between the value of the SW test statistic after transformation and other effects (skewness, kurtosis and others). However, the lack of proportionality, in the sense mentioned by Whittington (1980), was not explored.

Other details such as the values of skewness, kurtosis and its reduction[4], are included in the appendix. From this table it is clear, that the use of Box-Cox transformation, led to the expected result (see Nikkinen, Sahlstrőm, 2004). The average skewness was reduced more (*i.e.* by 81.76%) than the average kurtosis (*i.e.* by 67.26%). After transformation, the average skewness was not statistically different from zero, but the average kurtosis was statistically different from zero[5], even after transformation.

4. Discussion

Our research differs from previous pieces of research in that it focusing on data suitable for building a bankruptcy prediction model, *i.e.* our research sample consists of data on both active and bankrupt companies. As statistically significant difference can be found, between the ratios of bankrupt companies and the ratios of active ones (see Beaver, 1966 or Altman, 1968), the results achieved may differ from studies using only data on active-companies. In the research conducted, the Box-Cox transformation, completely removed the non-normality induced by skewness. These results are consistent with the results of other works (see Nikkinen, Sahlstrőm, 2004). However, we found two cases of bounded ratios in which Box-Cox transformation led to an increase of the skewness or kurtosis instead of decreasing skewness and kurtosis. The ratio's ability to achieve its limits (bounds) may represent a possible cause.

According to Nikkinen and Sahlstrőm (2004), a positive skewness is exhibited if a ratio is limited so as to be greater than zero[6]. A ratio[7] with both a lower and upper limit has slightly negative skewness and profitability ratios have no clear pattern.

Our research into ratios with a lower limit of zero, confirmed the result achieved by Nikkinen and Sahlstrőm (2004). Unlike them, we used different ratios/indicators, such as CA/TA, CD/S, FA/LTL, S and TA. Moreover, we found, that these ratios/indicators exhibit a strong positive kurtosis.

We focus more on ratios bounded both by a lower and an upper limit (and found that they can be divided into two subgroups, based on the lower limit character, *i.e.* if the ratio can achieve a zero value (*e.g.* Int.A/Tot.A, Tan.A/Tot.A or TL/TA) or can only approach to the zero value, but not reach it (*e.g.* CA/TA and QA/TA).

The first group can be characterized by a positive skewness and kurtosis, the second group exhibits a slight positive skewness and a slight negative kurtosis. Although there was only a slight skewness and kurtosis of second-group untransformed ratios, the transformation significantly increased the skewness (in the case of CA/TA) or the kurtosis (in the case of QA/TA), and as a result the transformed ratios were not normally distributed.

Furthermore, we found, that the outlier problem may possibly be solved by an alternative transformation approach. This approach represents an analogy of Ohlson's change of net income (NI-change). The usefulness of this approach is subject to further research.

Conclusions

Before transformation, none of the analysed financial ratios met the condition of one-dimensional normality, not even on the 1% level. After transformation, the condition of one-dimensional normality was met, at the 1% level, by 34% of the analysed financial ratios. The same condition, but at the 5 or 10% level, was met by 27% of the analysed financial ratios. The parameter λ was not possible to estimate in the case of 18% of financial ratios. The value of skewness before transformation seems to be responsible for this effect, as the values of skewness before transformation are strongly correlated with the values of parameter λ estimates. The critical value of skewness before transformation seems to be between 6.03 and 6.76. However, the values of kurtosis before transformation have no statistically significant effect on parameter λ estimates.

We contributed to the previous pieces of research in four ways. Firstly, by analysing a greater range of accounting ratios or indicators (*i.e.* 44), secondly, by focusing on data of a different character (data suitable for building a bankruptcy forecast model), thirdly, by explaining cases in which the parameter λ is not possible to estimate, and finally fourthly, identifying a possible cause of transformation failure in achieving normality of financial ratios.

The condition of normality for untransformed Czech bankruptcy data seems almost as impossible to fulfil. This conclusion implies the use of non-parametric methods, such as artificial neural networks. However, the comparison of the parametric method's performance using untransformed or transformed data is the subject of further research.

Notes

1. In the Czech Republic from 2006 to 2010, the number of wound-up joint-stock companies ranged between 2 and 2.6% (Felcman, 2010).
2. In this way, 53 potential predictors were obtained with 44 potential predictors being calculated from the data available. Mostly those indicators were not determined using capital market data as the shares of none of the bankrupt sample companies were marketable
3. The skewness of the EQ indicator before transformation reached at value of 6.039 with an estimated value of $\lambda 1$ of 4.799, the skewness of EBIT (3-vol) before transformation reached a value of 6.7638 and the parameter $\lambda 1$ was not possible to estimate within the limited values.
4. The reduction (*e.g.* of skewness) was calculated in the following way: [(SA – SB)/SB]*100%
5. A t-test was used to test significance. The t-statistic value for skewness was –0.45054 with a p--value of 0.653454, and in the case of kurtosis the t-static value was 3.49140 with a p-value of 0.0007.
6. Nikkinen and Sahlstrőm (2004) tested these ratios limited to be greater than zero: debt ratio, current ratio, quick ratio and inventory turnover.
7. Nikkinen and Sahlstrőm (2004) only tested the ratio of equity to total capital. However, this ratio was not included in our analysis.

Appendix: Skewness, kurtosis before and after transformation.

Ratio	Before transfor.		After transfor.		Skewness reduction [%]	Kurtosis reduction [%]
	Skewness	Kurtosis	Skewness	Kurtosis		
CA/TA**	0.0115	–0.9249	–0.0206	–0.9138	78.23	–1.2
CD/S*	2.95	10.0791	0.206	–0.6819	–92.37	–93.23
CF/S	–0.3516	6.88	0.3304	4.35	–6.02	–27.2
CF/TA	–5.0346	40.2493	–0.3244	4.13	–93.56	–87.87
CF/TD	2.911	16.8716	–0.134	4.0438	–95.4	–76.03
CR*	2.75	10.1629	0.0076	0.2742	–99.69	–97.3
DR*	2.37	12.598	–0.0128	0.3279	–99.47	–97.4
E/TA	–3.4041	20.869	0.1986	4.93	–94.17	–79.54
EBIT(3-vol)	6.39	56.6589	0.064	0.1583	–99.05	–99.72
EBIT/Int.	8.41	71.5388	0.3542	43.0318	–95.75	–39.85
EBITDA/Int.	8.375	71.8039	–0.8018	40.3027	–90.43	–43.87
EBITDA/TL	1.06	16.1207	0.682	16.5271	–42.72	2.52
EBT/OR	–3.9767	37.1918	0.5939	4.74	–85.07	–86.59
EQ	6.0396	45.6792	1.0899	11.71	–81.95	–74.66
FA/LTL*	10.0404	102.1164	0.0396	–0.1404	–99.61	–99.86
Int. A/Tot. A**	8.1752	82.7984	–2.1407	55.686	–73.82	–32.75
NI/AC	12.99	166.7747	–1.261	46.3262	–89.89	–72.22
NI/CA	–1.389	15.2192	0.6804	7.09	–51.02	–51.37
NI/FA	–13.8212	196.2908	–5.7374	74.4404	–58.49	–62.08
NI/OR	–4.5455	42.8521	0.5137	4.47	–88.7	–88.41
NI/TA	–5.1012	40.7577	–0.5544	4.957	–89.13	–87.84
NI-change	0.1045	–0.6591	–0.0441	–0.6841	–57.76	3.8
OC/OR	–4.314	39.7079	0.3972	3.74	–90.79	–90.64
OI/AC	8.0177	68.7679	0.0925	0.0258	–98.85	–99.96
OP/OR	–4.314	39.7079	0.3972	3.74	–90.79	–90.64
OR/CA	1.33	4.94	0.0265	0.0673	–98.58	–98.57
OR/CL	1.02	2.75	0.0268	–0.0609	–97.79	–97.72
OR/FA	13.2725	183.5863	0.1074	–0.4806	–99.19	–99.74
OR/LTL	7.19	59.941	0.044	–0.1634	–99.43	–99.73
OR/TA	2.06	10.0421	0.0245	–0.0086	–99.03	–99.91
OR/TL	1.02	3.18	0.02	–0.2159	–98.43	–93.21
PM	–11.3205	148.9661	–2.0368	17.5286	–82.01	–88.23
QA/S*	1.68	4.0397	–0.0179	0.7796	–98.84	–80.7
QA/TA**	0.7712	–0.3058	0.089	–0.7383	–88.47	141.42
RE/TA	–4.27	30.9661	0.2325	1.0808	–94.56	–96.51
S*	5.69	40.4779	0.1711	1.03	–96.82	–97.06
S/TA*	2.63	11.37	0.0268	–0.0657	–98.9	–99.43
TA*	5.795	45.3315	0.1128	0.5716	–98.05	–98.74
Tan. A/Tot. A**	0.3218	–0.6115	0.0169	–0.893	–94.75	46.05
TD/EDA	3.1576	36.8033	1.15	28.7106	–43.58	–21.99
TL/TA**	2.67	12.43	–0.013	0.331	–99.46	–97.38
WC/OE	–2.8237	21.1032	0.7456	7.49	–73.59	–63.06
WC/S	–1.3018	10.15	0.4996	5.068	–61.62	–51.46
WC/TA	–2.1112	13.3211	0.2525	1.77	–88.04	–89.36
Average	*1.22*	*42.0177*	*–0.0744*	*9.0005*	*–81.76*	*–67.26*

Source: Our own analysis of data from the Amadeus database.

Acknowledgment

This paper was supported by grant FP-S-12-1 Efficient Management of Enterprises with Regard to Development in Global Markets from the Internal Grant Agency at Brno University of Technology.

References

Aziz, M., Dar, H. (2006). Predicting corporate bankruptcy: where we stand? *Corporate Governance*, 6(1), pp. 18–33.

Altman, E. I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance*, 23(4), pp. 589–609.

Altman, E. I, Haldeman, R. G., Narayanan, P. (1977). ZETA Analysis. A new model to identify bankruptcy risk of corporations. *Journal of Banking and Finance*, 1(1), pp. 22–54.

Barnes, P. (1982). Methodological implications of non-normally distributed financial ratios. *Journal of Business Finance and Accounting*, 9(1), pp. 51–62.

Barnes, P. (1987). The analysis and use of financial ratios: *A review article. Journal of Business Finance and Accounting*, 14(4), pp. 449–461.

Beaver, W. H. (1966). Financial Ratios as Predictors of Failure. *Journal of Accounting Research*, 4(1), pp. 71–111.

Box, G. E. P., Cox, D. R. (1964). An Analysis of Transformations, *Journal of the Royal Statistical Society. Series B (Methodological*), 26(2), pp. 211–252.

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

Ding, Y., Song, X., Zen, Y. (2008). Forecasting financial condition of Chinese listed companies based on support vector machine. *Expert Systems with Applications*, 34(4), pp. 3081–3089.

Gorczyńska, M. (2011). The Dilemmas over Credit Policy Management in a Company. *Trends of Economics and Management*, 5(8), pp. 97–106.

Felcman, M (2010). Vývoj počtu a struktury zaniklých ekonomických subjektů v období 2006–2010. [The development of number and structure of extinct economic subjects in the period of 2006–2010].[online] Volume 2011 (3). Available: http://www.ekonomikaamanagement.cz/cz/clanek-vyvoj-poctu-a-struktury-zaniklych-subjektu-v-obdobi-200-2010.html.

Grice, J. S., Dugan, M. T. (2001). The limitations of bankruptcy prediction models: Some cautions for the researchers. *Review of Quantitative Finance and Accounting*, 17(2), pp. 151–166.

Grubbs, F. E. (1969). Procedures for Detecting Outlying Observations in Samples. *Technometrics*, 11(1), pp. 1–21.

Hebák, P., Hustopecký, J., Jarošová, E., Pecáková, I. (2004). Vícerozměrné statistické metody (1), Informatorium.

McLeay, S., Omar, A. (2000). The sensitivity of prediction models to the non-normality of bounded and unbounded financial ratios. *British Accounting Review*, 32(2), pp. 213–230.

Mileris, R., Boguslaukas, V. (2011). Credit Risk Estimation Model Development Process: Main Steps and Model Improvement, *Inzinerine Ekonomika-Engineering Economics*, 22(2), pp. 126–133.

Niemann, M., Schmidt, J. H., Neukirchen, M. (2008). Improving performance of corporate rating prediction models by reducing financial ratio heterogeneity. *Journal of Banking & Finance*, 32(3), pp. 434–446.

Nikkinen, J., Sahlstrőm, P. (2004). Distributional properties and transformation of financial ratios: The impact of the accounting environment. *Advances in International Accounting*, 17(1), pp. 85–101.

Meloun, M., Militký, J. (1994). Statistické zpracování experimentálních dat, Plus.

Ohlson, J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, 18(1), pp. 109–131.

Peel, M. J., Peel, D. A. (1987). Some further empirical evidence on predicting private company failure. *Accounting and Business Research*, 18(69), pp. 57–66.

Psillaki, M. Tsolas, I. T., Margaritis, M. (2010). Evaluation of credit risk based on firm performance, *European Journal of Operational Research*, 201(3), pp. 873–881.

Shapiro, S. S., Wilk, M. B. (1965). An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, 52(3/4), pp. 591–611.

Shumway, T. (1999). Forecasting Bankruptcy More Accurately: A Simple Hazard Model. *Journal of Business*, 74(1), pp. 101–124.

Škapa, S. (2011). Risk Measurement of Equity Markets and Private Investor Behaviour. *Trends of Economics and Management*, 5(8), pp. 85–96.

Taffler, R. J. (1982). Forecasting company failure in the UK using discriminant analysis and financial ratio data. *Journal of the Royal Statistical Society*, 145(3), pp. 342–358.

Tseng, F. M., Hu, Y. C. (2010). Comparing four bankruptcy prediction models: Logit, quadratic interval logit, neural and fuzzy neural networks. *Expert Systems with*

Applications, 37(3), pp. 1846–1853.
Whittington, G. (1980). Some basic properties of accounting ratios. *Journal of Business Finance and Accounting*, 7(2), pp. 219–232.
Wang, Y. J., Lee, H. S. (2008). A clustering method to identify representative financial ratios, *Information Sciences*, 178(4), pp. 1087–1097.
Watson, C. J. (1990). Multivariate distributional properties, outliers and transformation of financial ratios. *The Accounting Review*, 65(3), pp. 682–695.
Wu, Y., Gaunt, C., Gray, S., (2010). A comparison of alternative bankruptcy prediction models. *Journal of Contemporary Accounting & Economics*, 6(1), pp. 34–45.
Zimmerman, D. W. (1998). Invalidation of parametric and nonparametric statistical tests by concurrent violation of two assumptions. *Journal of experimental education*, 67(1), pp. 55–69.
Zimmerman, D. W. (1994). A note on the influence of outliers on parametric and nonparametric tests, *The journal of general psychology*, 121(4), pp. 391–401.
Zimmerman, D. W. (1995). Increasing the power of nonparametric tests by detecting and down weighting outliers, *Journal of experimental education*, 64(1), pp. 71–79.
Zmijewski, M. E. (1984). Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research*, 22(1), pp. 59–82.

Doručeno redakci: 30. 10. 2012
Recenzováno: 20. 11. 2012
Schváleno k publikování: 6. 3. 2013

Ing. Michal Karas
Prof. Ing. Mária Režňáková, CSc.
Brno University of Technology,
Faculty of Business and Management,
Department of Finance,
Kolejní 2906/4,612 00 Brno,
Czech Republic
tel.: +420 541114 3787
e-mail: karas@fbm.vutbr.cz
e-mail reznakova@fbm.vutbr.cz