UNIVERSITY OF TARTU
Faculty of Social Sciences
School of Economics and Business Administration

Kerttu Lääne

CORPORATE BANKRUPTCY PREDICTION OF ESTONIAN FIRMS

Master's thesis

Supervisor: Oliver Lukason (PhD)

Tartu 2017

Name and signature of supervisor…………………………………………………….

Allowed for defence on..

(*date*)

I have written this master's thesis independently. All viewpoints of other authors, literary sources and data from elsewhere used for writing this paper have been referenced.

....

(signature of author)

Table of contents

Abstract

The aim of this research paper is to build a universal corporate bankruptcy prediction model based on Estonian firms' data. In Estonia only sector specific models have previously been built and therefore this paper is the first to cover many sectors in one multisector model. In this thesis bankruptcies that occurred between 2012 and 2016 have been analyzed. As the statistical method the author employ logistic regression. The final model includes three predictive variables, namely equity to total assets, net income to total revenues and total revenues to total assets, which all are statistically significant and model classification accuracy is 81.5%.

Keywords: bankruptcy, bankruptcy prediction, multisector model, financial ratios, Estonian firms, logit model

1. Introduction

Over the last five decades the reason why and when companies fail or go bankrupt has been an important research topic amongst researchers and practitioners. Bankruptcy has an effect on an entire company as well as on those who are involved with it - such as investors, managers, workers or even suppliers. Therefore having a compact and accurate bankruptcy prediction model is important for avoiding major losses.

Beaver (1966) was a pioneer who used financial ratios in corporate failure prediction. He recognized the importance of cash flow ratios and used univariate discriminant analysis method. Shortly after Beaver (1966), Altman (1968) introduced multiple discriminant analysis (MDA) to company failure prediction modelling. Later Ohlson (1980) proposed the logit model, which is also used in the current thesis. For a long time only statistical techniques, such as MDA, logit and probit models were used, due to huge development in technology, new intelligent techniques have been applied.

Most of the bankruptcy prediction models have been industry specific. For example Altman (1968) used manufacturing firms in his model. Fewer models are for a general application. Previous master theses about bankruptcy prediction, which have used Estonian firms' data, have been industry specific: Lukason (2006) chose commercial, Grünberg (2013) industrial and Onno (2015) Estonian road transportation companies. For credit companies, banks and investors it is crucial to understand the financial situation of the firms that they are interested to invest in or to lend money to. Wrong decisions can have a large scale impact on all parties.

A universal model that is not industry specific could be applied among all companies not depending on their sector. Moreover, many firms are active in different industries, and thus in those cases it is not possible to use industry specific models. The purpose of this thesis is to build a bankruptcy prediction model that would be applicable irrespective of the industry the firm is active in. For building this model all firms' annual reports are included regardless of the firm's field of activity. Bankruptcies that occurred between 2012 and 2016 and annual reports data between 2011 and 2015 are used.

After data scrubbing and matching the annual reports data with the bankrupt firms, 325 companies were included to the bankrupt firms' sample. 1712 non-bankrupt firms' annual reports data for 2011-2015 were available after data cleaning and the 3 random samples were generated using paired sample technique.

The author calculated 15 financial ratios. Lastly, three control variables, such as firm's age, size and sector, were also tested. The final model included three financial ratios equity to total assets, net income to total revenues and total revenues to total assets. Achieved classification accuracy for the model was quite high comparing to the previous research done in this field.

The structure of the paper is the following. The second chapter gives an overview about the bankruptcy definition and how it is used in different countries. It also covers theories of bankruptcies, failure processes and previous literature about logit method usage in the bankruptcy prediction models. The third chapter describes the data and different criteria that are used to choose firms to the final sample. Also, it is explained how the applied method works and what kind of variables were used in modelling. The fourth chapter includes the information about model composition. Moreover, the logit model is tested with control variables such as firm's age, size and sector to see whether in case of different firms the model financial ratios behave similarly among the firms with different age, size and sector. Afterwards the results are interpreted and compared to the previous models used in other master theses. Additionally, Altman *et al.* (2016) model is tested with the given dataset. Finally, future research directions are provided.

2. Literature review

2.1 Bankruptcy prediction

Since the late 1960s business failure prediction within corporate finance has developed into a wide research domain and based on different modelling techniques many various corporate failure prediction models have been developed (Balcaen, Ooghe 2006). There is not only one definition for the word "failure" but it has obtained several meanings in previous studies. Beaver (1966:71) wrote in his paper that: "Failure is defined as the inability of a firm to pay its financial obligations as they mature". The majority of the studies have leaned on some recorded event as a substitute measure of the failure. The two most common cases where the data have been available for are formal bankruptcy proceedings and the discontinuance of the activity in a business (Watson, Everett 1996).

Additionally, two further definitions have been proposed: failure of "make a go of it" by Cochran (1981) and termination in order to prevent additional losses by Ulmer and Nielsen (1947). On the other hand, Dimitras *et al*. (1996) noted that generally failure is a situation where company is not able to pay suppliers, stock shareholders or lenders and as well the case where the company is bankrupt based on law or a bill is overdrawn. There are multiple definitions but the author presented the most common ones. It is also important to point out that failure definition depends on the research area. Bankruptcy as the definition of failure has been mostly used in accounting and finance literature (Tamari 1966, Altman 1968, Kumar and Ravi 2007, Pretorius 2009).

As stated above, there are many different approaches for defining "failure" and therefore the authors of bankruptcy specific literature do not have a common understanding. Bankruptcy prediction research papers can be divided into two: the ones that have used the phrase "bankruptcy prediction" and others that have used "failure prediction". Most of the authors have used "failure prediction" but they actually mean "bankruptcy prediction" by it (Balcaen, Ooghe 2006). Some studies that have used failure but actually mean bankruptcy are Zavgren (1985), Hambrick and D'Aveni (1988), Daubie and Meskens (2002), Charitou *et al.* (2004) and Bhandari and Iyer (2013).

Insolvency legislation varies among European countries. Mainly there are three types of full insolvency proceedings: liquidation, reorganisation/restructuring and discharge/debt relief (which means that debts will be deleted and firms can have a new start). In many of the European Union countries the liquidation proceeding dominates (in Estonia as well) (Bariatti, Van Galen 2014).

Countries are divided as debtor friendly or creditor friendly. In debtor friendly countries such as France reorganization is common and rescuing the company is more important. On the other hand, in creditor friendly countries like the United Kingdom liquidation proceeding dominates (Davydenko, Franks 2008).

In this master thesis failure is defined according to the Bankruptcy Act of Estonia (Bankruptcy Act, 2003). According to the Bankruptcy Act of Estonia “Bankruptcy means the insolvency of a debtor declared by a court ruling“. Firstly, it states that: “A debtor is insolvent if the debtor is unable to satisfy the claims of the creditors and such inability, due to the debtor's financial situation, is not temporary.” Secondly, it states that: “A debtor who is a legal person is insolvent also if the assets of the debtor are insufficient for covering the obligations thereof and, due to the debtor's financial situation, such insufficiency is not temporary“.

2.2 Theories of bankruptcy

There is not one supported theory of bankruptcy in the failure literature that would explain how firms go bankrupt (Balcaen, Ooghe 2006). Therefore, most of the bankruptcy prediction studies are focused on finding empirically the best predictors of bankruptcy, without relying on any theory. However, there is some theoretical evidence available. Firstly, there is a stream of bankruptcy literature that is based on Beaver (1966) studies, where "the firm is like a reservoir of liquid assets, which is supplied by inflows and drained by outflows" (Beaver 1966:80). Beaver (1966) was the first who acknowledged the using of cash flow for financial failure prediction was important. In his analysis four ratios out of 30 were different cash flow ratios (Beaver 1966):

1) cash flow to sales;
2) cash flow to total assets;
3) cash flow to new worth;
4) cash flow to total debt.

The main idea of this stream is that an organization is like a cash reserve and different actions and activities can increase or decrease this reserve (*Ibid*). Giacomino and Mielke (1988), Gentry (1984), Bernard, Stober (1989), Zeller, Figlewicz (1990) and Carslaw, Mills (1991) and many others followed Beaver (1966). Their studies showed that cash inflows and outflows are highly connected with many organization activities and the above mentioned researchers used cash flow for predicting the non-bankrupt and bankrupt firms.

Secondly, Scott (1981) described a simple bankruptcy theory. The main idea for bankruptcy prediction was that when a firm's current year earnings are less than the debt obligations or are negative, it is more likely that the firm goes bankrupt. Also, it was added that the firm is assumed to go bankrupt when the sum of the expected value of equity and firm's current year's earnings is less than zero. Beaver (1966) prioritized cash flow variables, whereas Scott (1981) prioritized earning variables in bankruptcy prediction (Lukason 2016).

Thirdly, there is a literature stream that is based on option theory to describe the probability of bankruptcy (Lukason 2016). Together Black and Scholes (1973) and

Merton (1974) found a new option approach – "firm's equity can be viewed as a call option on the value of the firm's asset" (Hillegeist *et al.* 2004:6). The call option is not used and the bankrupt firm is handed over to its debtholders when the assets' value is lower than its strike price (*Ibid.*). The same paper found that it is possible to improve the forecasting accuracy by using previous accounting variables and adding market variables to them. It is also important to keep in mind that bankruptcies occur because of what happened in the company for some time before it goes bankrupt (Laitinen 1993). The bankruptcy process that happens in the company is an interesting topic for researches and it has been researched from different angles. Argenti (1976) had three failure trajectories:

1) Type 1 trajectory includes small businesses that fail in the beginning of trading and they never achieve sustainable profit. One of the reasons they fail can be poor management skills or unsuitable personalities in the management;
2) Type 2 trajectory includes firms which experience fast growth and sudden downfall;
3) Type 3 trajectory is formed by old and large firms that at some point lost their market share and passed through many stages of the failure process because of the new contestants.

Similarly to Argenti (1976), three different failure process patterns were introduced by D'Aveni (1989), who used financial and non-financial variables for building up the trajectories. Firstly, there is a sudden decline. It is described by a sudden collapse which is followed by bankruptcy. It can happen with impulsive firms. When companies have a slow decline and bankruptcy is not so sudden, this process is called a gradual decline. Lastly, there is a lingering that is characterized by firms that decline fast or gradually but the bankruptcy does not take place immediately, only after the decline is completed but maybe several years later (*Ibid.*).

Similarly to the previous researchers, Laitinen (1991) distinguished also three types of different failure processes: chronic, revenue financing (gradual) and acute failure processes. To represent the failure process six financial variables were used: return on investment (ROI), rate of growth in total assets, net sales to total assets, cash flow to net sales (CFR), capital assets ratio and current ratio. For chronic failure process the firm's financial ratios are poor for several years before failure and they usually have been earning losses for many years. For example, CFR is low and has been showing warning signs already four years before the firm fails. Also total debt to total assets and current ratios have low values 4-6 years before the firm fails. The second process is gradual failure process where ROI and CFR are negative and show some warnings signs two years before the failure. Nevertheless, it is still possible that the financial variables do not indicate any problems one or two years before the failure but they can rapidly deteriorate within the following years. In that case this is called acute failure process (Laitinen 1991). When financial ratios do not indicate any problems, other variables should be applied in failure prediction that could help to anticipate the problems in the firms.

Similarly to Argenti (1976) and D'Aveni (1989), Probst and Raisch (2005) defined two syndromes of the organizational failure logic, which can be used to divide firms in a crisis: Burnout Syndrome and Premature Aging Syndrome. The first one describes firms which have remarkable growth in five years before the failure and in the end have a sudden downfall. The second one, premature aging syndrome describes firms which stagnate consistently in the previous years before failure (Probst, Raisch 2005).

Usually the failure process includes signs that can predict the negative outcome. These signs differ from each other and may appear at various times before the eventual failure. Hambrick and D'Aveni (1988) studied large private sector corporations and their failures and found that for those companies the signs of difficulties appeared already 10 years before. Laitinen (1991) stated that sufficiency of the revenue financing is the main element that signals the failure process. The level of revenue financing and need for outside financing have a negative relationship: the higher the level of revenue financing, the less outside financing is needed for the firm. This also leads to fewer financial obligations for the company. For measuring the sufficiency of revenue financing the cash flow ratios can be used as well. Internal rate of returns, rate of growth and the rate of revenue accumulation are three parameters that affect the cash flow ratio. Thus, the causes for low revenue financing can be too fast growth, too low profitability, slow accumulation of revenues or the combinations of those three parameters (*Ibid*.).

2.3 Overview of previous bankruptcy prediction models

Failure prediction with financial ratios has been studied for many decades and different financial ratios and various methods have been used in order to predict bankruptcy and obtain as accurate model as possible (Keener 2013). Most of the bankruptcy prediction models have used a paired sample technique which means that number of bankrupt firms and non-bankrupt firms is equal. The financial ratios are calculated based on financial information that was available before the firms went bankrupt (Scott 1981).

Since Beaver's (1966) pioneering work on corporate bankruptcy prediction models with the usage of financial ratios and Altman's (1968) model, almost all of the following studies mainly focused on using the financial ratios. In recent times the non-financial variables such as employees, customers, age, firm size and payment defaults have been included to the models (Pervan, Kuvek 2013). Von Stein and Ziegler (1984) examined also the changes in the management of an endangered company that would give an earlier warning sign.

The first one who used financial ratios in the bankruptcy prediction model was Beaver (1966). Beaver (1966) used univariate discriminant analysis and tested 30 different financial ratios from which six statistically significant ratios were chosen after the calculations. Those six ratios were total debt to total assets, cash flow to total debt, net income to total assets, current ratios, working capital to total assets and the no-credit interval. Cash flow to total debt ratio had the strongest ability to predict the failure and the classification accuracy a year before the failure was 90% (*Ibid*.).

Shortly after Beaver (1966), Altman (1968) used a new statistical multivariate analysis technique that is called multiple discriminant analysis (MDA) for bankruptcy prediction. Even 50 years later Altman's Z-Score is still seen as a great indicator to predict bankruptcy (Lawrence *et al.* 2015). Altman (1968) chose five financial variables and the model classified bankruptcy correctly for 95% of the firms. So both Beaver (1966) and Altman (1968) had models with high prediction accuracy.

A logit model was proposed by Ohlson (1980) who was a pioneer of the logit analysis for business failure prediction. Ohlson (1980) did not agree with discriminant analysis because of the requirement of identical covariance-variance matrices for both failed and non-failed groups and also because with MDA there was a requirement of normally distributed predictors (Klieštik *et al.* 2015). Ohlson (1980) applied logistic regression in a larger sample and it did not involve pair-matching – 105 bankrupt firms and 2058 non-bankrupt firms' data were used. It was found that four statistically significant factors (that included 9 different variables in total) for identifying the probability of failure are the size of the company, measures of performance, measures of financial structure and measures of current liquidity. The logit regression model had one year before the failure high classification accuracy – 96% of the firms were classified correctly (*Ibid*.).

Logit analysis was widely used after Ohlson's (1980) research. Table 1 includes some of the bankruptcy prediction multisector models where logistic regression method was applied, with the exception of sector specific model by Zavgren (1985) who used manufacturing firms. Some of the authors who have used logistic regression are for example Zavgren (1985), Casey and Bartczak (1985), Aziz *et al.* (1988), Platt and Platt (1990), Gilbert *et al.* (1990), Pindado *et al*. (2008), Altman *et al*. (2016). Mostly the data sets that have been used in models are 4-6 years long – Ohlson (1980), Zavgren (1985), Platt, Platt (1990), Altman *et al*. (2016), but there are also researches where the used time horizon is up to 12 and 13 years as Casey and Bartczak (1985), Aziz *et al*. (1988) have used data from 1971-1982 and Pindado *et al.* (2008) used data from 1990-2002 (see Table 1).

Mostly all of those models mentioned above are multisector models and have been developed for general application across the industries. Only Zavgren (1985) used manufacturing firms in the model. Altman *et al*. (2016) used large international sample of firms from 28 European Union countries and 3 non-EU countries as well (see Table 1).

The smallest number of significant variables used in their models was three ratios that were profitability, financial expenses and retained earnings in Pindado *et al*. (2008) model. Altman's *et al.* (2016) model included 4 different financial ratios, Gilbert *et al*. (1990) and Aziz *et al.* (1988) had 6 financial ratios, Zavgren (1985) had 7 ratios, and 9 ratios were included in Casey and Bartczak's (1985) model. Most popular variables among those models were related to cash flows. Casey and Bartczak (1985) used 3 operating cash flow ratios, Aziz *et al.* (1988) used cash flow from operations, Gilbert *et al.* (1990) used cash flow from operations to total liabilities and also cash flow from operations to current liabilities and Platt and Platt (1990) used cash flow to sales. Ratios that included retained earnings information were used by Gilbert *et al.* 1990 (earnings before interest and taxes to total assets), Pindado *et al.* 2008 (retained earnings) and Altman *et al.* 2016 (retained earnings to total assets) (see Table 1).

Only two of those models in Table 1 used equal-sized matched samples. It means that each bankrupt firm is paired up with non-bankrupt firm, thus, the total number of bankrupt and non-bankrupt firms in the model is equal. Matched samples technique was used by Zavgren (1985) and Aziz *et al*. (1988) that had accordingly 45 and 49 paired firms in their samples. Others did not apply this technique and used the data that were available. Samples vary from 290 (60 failed and 230 non-failed) firms in case of Casey and Bartczak (1985) model to 18160 (17439 normal and 721 financial distress) firms as was used by Pindado *et al*. (2008).

One of the most important things about bankruptcy prediction models is their accuracy. The higher the model classification rates, the more accurate results in predicting bankruptcy for firms may be achieved. Separately for non-bankrupt firms the classification rates were higher than in case of failed firms. The lowest classification accuracy was in Gilbert's *et al*. (1990) model where for bankrupt firms the highest classification percentage was 62.5% and for non-bankrupt it was 97.9%. Mostly, the overall classification rates 1-5 years before the failure varied between 88-92% for one year before the failure and between 81-83% for 5 years before the failure (Casey and Bartczak (1985), Aziz *et al.* (1988)). Pindado *et al.* (2008) classified 87% of bankruptcies correctly for U.S firms and 83% for G-7 countries, and Platt and Platt's (1990) model's correct classification rate was 90%. Altman *et al.* (2016) had a model with AUC value of 0.743 which is average classification accuracy. But as they used data from 31 countries and from thousands of firms, this AUC score is quite fair (Altman *et al.* 2016). In Zavgren's (1985) model, the classification error rate for a year prior bankruptcy was 18 per cent (see Table 1).

Table 1. Bankruptcy prediction models based on logit method

Authors	Year	Data	Statistically significant variables	Sample size	Model classification accuracy
Ohlson	1980	1970 – 1976	7 variables: firm size, total liabilities/total assets, net income/total assets, fixed assets to total liabilities, working capital/total assets, current assets/current liabilities	105 bankrupt firms and 2058 non-bankrupt firms	1 year before – 96.12 %, 2 years before – 95.55%
Zavgren	1985	1972 - 1978	7 variables: total income/total capital, sales/net plant, inventory/sales, debt/total capital, receivables/inventory, quick assets/current liabilities, cash/total assets	45 failed and 45 non-failed	
Casey and Bartczak	1985	1971-1982	6 variables: cash/total assets, current assets/total assets, current assets/current liabilities, sales/current assets, net income/total assets, and total liabilities/owners' equity and 3 operating cash flow ratios	60 failed and 230 non-failed	Overall: 88%;86%;87%;85%;83% (1-5 years before the failure) with cut off score that maximized the accuracy
Aziz, Emanuel and Lawson	1988	1971-1982	6 variables: cash flow from operations, taxes assessed on the corporation, lender cash flows, net investment in long term investment, cash used for liquidity changes, stockholder cash flows	49 failed and 49 non-failed	Overall: 91.8%, 84.7%, 78.6%, 80.2%, 80.9% (1-5 years before failure)
Gilbert, Menon, Schwartz	1990	1974-1983	6 variables: cash/total assets ,earnings before interest and taxes/total assets, cash flow from operations/total liabilities, cash flow from operations/current liabilities , stockholders' equity/total liabilities, retained earnings/total assets	75 bankrupt and 304 non-bankrupt	Estimation Sample overall errors: 47 (18.1%) Bankrupt/Distressed Holdout errors: 26 (21.7%)
Platt, Platt	1990	1972-1976	7 variables: four financial ratios: cash-flow to sales, net-fixed assets to total assets, total debt to total assets, short-term debt to total debt; one operating ratio (growth in sales) and the percentage change in industry output interacted with two financial ratios	114 companies (57 failed companies, 57 non-failed)	Overall: 90%

Authors	Year	Data	Statistically significant variables	Sample size	Model classification accuracy
Pindado, Rodrigues, De la Torre	2008	1990-2002	3: Profitability, Financial expenses, Retained earnings	For US: 17439 normal and 721 financial distress; For G-7 countries: 14514 normal and 1188 financial distress	For US: mean value of 87% For G-7 countries: mean value of 83%
Altman, Iwanicz-Drozdowska, Laitinen, Suvas	2016	2007-2010	4 variables: WCTA=working capital/total assets, RETA=retained earnings/total assets, EBITTA= EBIT/total assets, BVETD=book value of equity/total liabilities	Estimation sample includes data from 2602563 non-failed and 38215 failed firms	AUC = 0.743

Source: compiled by the author

3. Data and method

In this master thesis the logistic regression method was applied. As mentioned in the second chapter of this thesis Ohlson (1980) was the first who used this method in bankruptcy prediction models. Firstly, Ohlson (1980) did not agree with discriminant analysis because of the requirement of identical covariance-variance matrices for both failed and non-failed groups. The second reason was that while applying MDA there was a requirement of normally distributed predictors. Moreover, logit does not require restrictive statistical assumptions and it offers better empirical discrimination (Klieštik *et al.* 2015).

Logistic regression describes data and explains the relationship between one binary dependent variable and one or more predictive variables (Hosmer and Lemeshow 2000). Binary variable means that the variable can only take two values. In our case the binary dependent variable is bankruptcy. Bankruptcy is marked with value 1 and non-bankruptcy with value 0. Logit model result is a score value between one and zero with what it is possible to classify the firms as bankrupt or non-bankrupt firms based on whether the score is under or over the cut off score. If score is less than a cut-off score (in this paper it is set to 0.5) then the firm is classified as non-bankrupt firm and if the score is above the cut-off score the firm is classified as a bankrupt firm. This ensures that the firms are put in the group they are most similar with as the resemblance principle is used in logit models (Balcaen, Ooghe 2006).

Therefore, logit analysis has two outcomes in the probability prediction: event either occurs or not and therefore the calculated probability is either 1 (for those firms whose calculated score was between 0.5 and 1) which means that event happens or 0 (for those firms the calculated score by model was between 0-0.5) which means that event does not occur (in that case will not go bankrupt) (Klieštik *et al.* 2015). Other variables used for predicting the result are all the financial ratios calculated with the data from annual reports.

This method has also many assumptions. Firstly, the outcome has to be discrete. Secondly, there should not be any outliers in the data. “An outlier is an observation that lies an abnormal distance from other values in a random sample from a population“ (NIST/SEMATECH e-Handbook of Statistical Methods 2013). Thirdly, there cannot be multicollinearity among the predictive variables (Hosmer, Lemeshow 2000). Multicollinearity means that the intercorrelation between the independents is too high and in that case the independents effects cannot be separated (Garson 2012).

Additionally, two types of errors can happen while classifying the firms as bankrupt or non-bankrupt firms. Type I error means that bankrupt firms is classified as non-bankrupt firm and type II error means that non-bankrupt firms is classified as bankrupt firm (Ohlson 1980).

Most of the bankruptcy prediction models in previous researches that use Estonian firms’ data have been industry-specific: Lukason (2006) used commercials firms’ data, Grünberg (2013) used industrial companies and Onno (2015) used road transportation companies in their master theses. This thesis model is multisector model which means that it includes companies from all the industries and to author’s knowledge no one in Estonia has previously composed a universal model.

Bankruptcy prediction models that are used in this master thesis use information from annual reports of Estonian companies. There were 138223 firms that submitted the annual report or reports between 2011 and 2015. According to the Estonian Commercial Code §32 all of the Estonian companies must submit annual reports containing financial information to the Business Register. The annual report must be submitted yearly for the 30th of June (Estonian Business Code, 1995). The data used for bankruptcy modelling is from Centre of Registers and Information System. It includes financial information about bankrupt and non-bankrupt firms presented in annual reports between 2011 and 2015 and bankruptcies that have occurred in firms between 2012 and 2016.

To define all the bankrupt firms the register codes were used in order to match the bankrupt firms with the correct year of their annual report data. Annual reports from one to two years before the bankruptcy year were analysed. When the firm went bankrupt in the first half of the year (from January to June) then the annual report data from the year before the last year was used. For example, if the firm went bankrupt in March 2015 then the financial variables from 2013 annual report were applied.

Annual report from previous year before the bankruptcy was used if firm went bankrupt in the second half of the year (from July to December). This approach is necessary because many companies do not submit their annual report information after they have gone bankrupt. During the period (2011-2015) under review 970 Estonian companies went bankrupt. With the method described above it was not possible to get the correct annual report information for almost half of those firms and therefore the initial sample of bankrupt firms decreased to 522 firms. Afterwards, the firms that did not have all the needed data for the financial ratios calculation were erased. The final sample included 325 bankrupt firms.

All the firms that were not defined as bankrupt firms were considered as non-bankrupt firms. Firstly, those firms that had not submitted all the annual reports between 2011 and 2015 were not included in the non-bankrupt firms' sample. Secondly, to get the correct final sample the firms that had missing values in their annual reports data that was needed for financial ratios calculations were deleted. Otherwise, firms with all kind of sizes were included: micro, small, medium and big size companies. Sample size of non-bankrupt firms was 1712 firms and their data of 2011-2015 so together there were 8560 data series in that sample. As there were significantly more non-bankrupt firms than bankrupt firms then there are three randomly generated samples for bankruptcy prediction model. Each includes 325 bankrupt firms and 325 non-bankrupt firms that were chosen randomly from all 8560 data series. Paired sample technique is widely used in bankruptcy prediction models (Beaver 1966, Altman 1968, Zavgren 1985). With paired sample technique it is easy to interpret the logit model results with cut-off score of 0.5. When the number of bankrupt firms and non-bankrupt firms in the sample is the same then the correct classification accuracy for bankrupt firms' increases and decreases for non-bankrupt firms. It is the case when comparing the paired sample to the sample where non-bankrupt firms outnumber the bankrupt firms (Platt, Platt 2002).

Last step was to replace the outliers before getting the final sample. Financial ratios often include outliers that may be the result from errors in the data or they can be just the extreme values as well. For this bankrupt model a method called "Winsorizing" is used. It means that outliers' values were changed to the closest non-outlier value (Barnes 1987). Most of the financial ratios had outliers and the "Winsorizing" method was applied.

Financial ratios were chosen based on how often they have been used previously and how well they have performed in previous bankruptcy prediction models that were described in chapter 2.3. Financial ratios that were used in the logit model were divided between 4 dimensions that were also used by Lukason *et al*. (2016). Those dimensions were liquidity (reflects the ability to pay short-term liabilities with assets), leverage (reflects the ability to pay the liabilities in the long run), efficiency (shows how the firm is operationally functioning), profitability (ratios reflect some profit account or cash flow to assets or turnover) and also control variables like firm's age, sector and size (see Table 2).

Table 2. Financial ratios dimensions

Dimension	Variable name	Variable formula
Liquidity	CS/TA	Cash stock/total assets
	CS/CL	Cash stock/current liabilities
	CA/TA	Current assets/total assets
	CS/CL	Current assets/current liabilities
	WC/TA	Working capital/total assets
Leverage	TL/TA	Total liabilities/total assets
	E/TA	Equity/total assets
	RE/TA	Retained earnings/total assets
	TL/E	Total liabilities/equity
Efficiency	TR/TA	Total revenues/total assets
Profitability	BP/TR	Business profit/total revenues
	NI/TR	Net income/total revenues
	NI/TA	Net income/total assets
	BP/TA	Business profit/total assets
	EBT/TA	EBT/total assets
Control variables	size	Ln(total assets)
	age	Firm's age
	sector	Firm's sector

Source: compiled by the author

For each dimension many variables were calculated as seen in Table 2. In total there were 15 financial variables. The profitability ratios dimension is the biggest and includes 5 variables: net income to total assets, net income to total revenues, business profit to total assets, business profit to total revenues and earnings before taxes to total assets. Liquidity dimension included ratios like cash stock to total assets, cash stock to current liabilities, current assets to total assets and current assets to current liabilities.

Descriptive statistics table (see Table 3) includes values for medians and standard deviation. It shows how values for bankrupt firms differ from non-bankrupt firms. As seen from the table then all the profitability ratios for bankrupt firms are negative which is also logical because bankrupt firms are not usually very profitable a year before they go bankrupt**.** For most of the variables standard deviation values between non-bankrupt and bankrupt firms are quite big.

Table 3. Descriptive statistics for financial ratios (random 1 sample)

Variables names	**Median**		**Standard deviation**	
	Bankrupt	Non-bankrupt	Bankrupt	Non-bankrupt
CS/TA	0.03	0.17	0.22	0.32
CS/CL	0.03	0.97	1.08	2.31
CA/TA	0.88	0.79	0.31	0.33
CA/CL	1	2.82	2.48	4.66
TL/TA	0.92	0.25	2.15	0.28
E/TA	0.08	0.75	2.15	0.28
TR/TA	2.59	1.14	6.04	2.11
BP/TR	-0.03	0.02	1.83	0.89
NI/TR	-0.04	0.02	1.63	0.81
NI/TA	-0.07	0.03	2.97	0.25
EBT/TA	-0.07	0.03	2.97	0.25
BP/TA	-0.07	0.50	2.19	0.53
RE/TA	0.00	0.54	4.16	0.58
TL/E	0.88	0.31	10.87	4.53
WC/TA	0.00	0.43	2.11	0.38
Size	11.29	11.18	2.03	2.14
Age	7	20	5.44	2.86

Source: compiled by the author

4. Results and discussion

The goal was to build a bankruptcy prediction model that would have high classification accuracy to predict whether it is likely that the firm will go bankrupt in following year. Also, the model had to include only statistically significant explanatory variables (statistically significant at confidence level 95% means that when p-value is less than 0.05, then the parameter is statistically significant at significance level 5%) and the coefficient signs had to be theoretically correct.

The first stage of building a bankruptcy model is to check whether there is multicollinearity between the variables. Variables that belong to the same dimension for example as profitability ratios: net income to total assets, net income to total revenues, business profit to total assets and business profit to total revenues and earnings before taxes to total assets tend to have multicollinearity (see Appendix 1). From the collinearity matrix it is seen which variables have high correlation coefficient value (over 0.4) and cannot be used together at the same time in the model. From each of the dimensions it was tested which variable of the correlated variables distinguish the bankrupt and non-bankrupt groups better in the model with other variables.

So from every dimension only the variables that had the highest prediction probability were chosen to the model and those were: equity to total assets, current assets to current liabilities, net income to total revenues and total revenues to total assets. However, when backward stepwise method was applied to random sample 1, current assets to current liabilities ratio was not included to the final model (hereinafter referred to as K-Model) and, as a result, net income to total revenues significance in the model also improved compared to the model where all of the 4 initial variables were included. Therefore, the K-Model had three explanatory variables: equity to total assets, net income to total revenues and total revenues to total assets (see Table 4).

Table 4. K-Model results

Variables names	B	S.E.	Wald	df	Sig.	Exp(B)
equity/total assets	-4.132	0.351	138.573	1	0.000	0.016
net income/total revenues	-0.211	0.105	4.029	1	0.045	0.810
total revenues/total assets	0.152	0.044	11.967	1	0.001	1.164
Constant	1.236	0.225	30.258	1	0.000	3.441

Source: compiled by the author

In this model all three variables are statistically significant at 95% confidence level (significance level is 5%), which means that the probability that null hypothesis will not be rejected given that it is true is 0.95.

Overall classification accuracy for K-Model is 81.5%, which means that 81.5% of the firms were classified correctly as being bankrupt or non-bankrupt firms. For non-bankrupt firms the classification 82.5% of correctly classified firms was slightly higher than for bankrupt firms where it was 80.6% of correct results (see Table 5).

Table 5. Bankruptcy prediction model classification accuracy with random 1 sample

		Predicted		
		Non-bankrupt	Bankrupt	Correct Percentage
Observed	Non-bankrupt	268	57	82.5
	Bankrupt	63	262	80.6
	Overall Percentage			81.5

Source: compiled by the author

ROC (receiver operating characteristic) curve is a graphical plot and the area under curve (AUC) measures models accuracy. The receiver operating characteristic curve is constructed by varying the cut-off probability. It means that for every cut-off probability, the ROC curve defines "true positive rate" (percentage of bankruptcies that the model correctly classifies as bankruptcies) on the y-axis that is a function of corresponding "false positive rate" which is a percentage of non-bankrupt firms that are mistakenly classified as bankrupt firms on the x-axis (see Chart 1). The perfect model (with zero false negatives and zero false positives) would have an AUC value of 1. For sample number 1 the AUC value is 0.89. For the second and third sample the AUC value is 0.87 for both. Altman *et al.* (2016) model had AUC value 0.74 which means that this bankruptcy prediction model has higher classification accuracy than Altman *et al.* (2016) model (see Appendices 2 and 3).

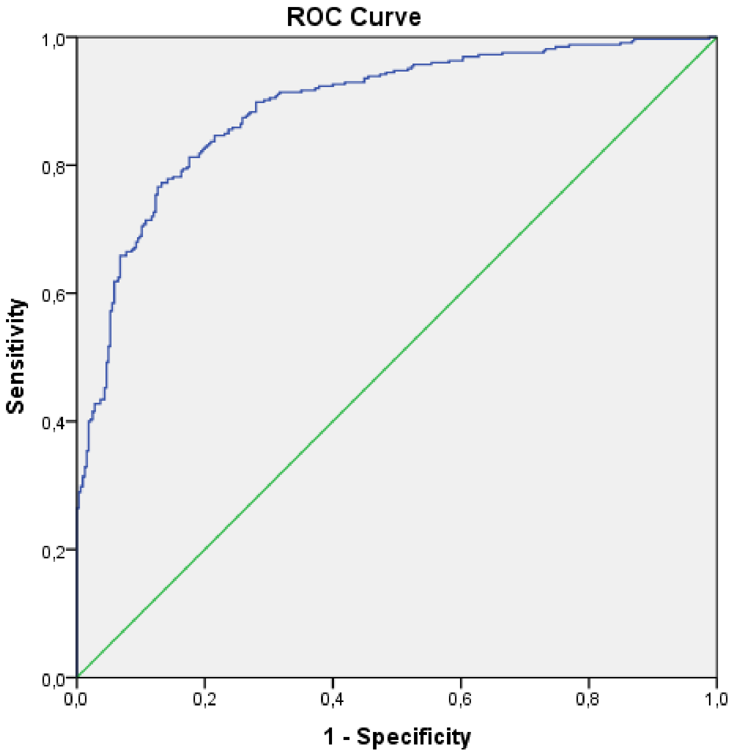


Chart 1. Random sample 1 ROC curve

As described before logit model calculates a predictive probability between 0 and 1 for each of the firms and if the probability is lower than a cut-off score, it is predicted to be in a non-bankrupt group of firms and, if the predicted probability is higher than cut-off score value 0.5, it is predicted to be in a group of bankrupt firms. For random samples 1, 2 and 3 the calculated predictions for non-bankrupt firms are for most of the firms quite low (around zero) and same apply for bankrupt firms where most of the bankrupt firms have score calculated near 1. Although there are few exceptions, where non-bankrupt firms have classified as bankrupt firms and the calculated score is lower than 0.3 or bankrupt firms as classified as non-bankrupt firms where score is higher than 0.7. But most of the misclassified firms are between 0.3 and 0.7 which means that they are not that far away from the 0.5 cut-off score (see Appendices 4, 5 and 6). For further implications the model can be improved with machine learning technique for classifying the values between 0.3 and 0.7 after the logit model has classified the firms that have a score value less than 0.3 or over 0.7.

To compare the statistically significant variables used in K-Model with variables that have been used in logit models before (see Table 1), none of those models included exactly those 3 variables. They included many variables that were also first inserted to K-Model but turned out to have multicollinearity or not to be significant, such as total liabilities to total assets, net income to total assets, current assets to current liabilities, working capital to total assets, current assets to total assets, but with Estonian firms data those variables were not included to the final bankruptcy prediction model.

The same method was applied to random sample 2 and random sample 3. The results stayed stable. In random samples number 2 and 3 there were the same three ratios: equity to total assets, net income to total revenues and total revenues to total assets that were included to the model and all were statistically significant with 5% significance level (see Table 6). Variables had also logical coefficient signs. The higher the equity to total assets and net income to total revenues values, the smaller is the probability that firm will go bankrupt. The situation is different with total revenues to total assets value where the coefficient sign was positive, which means that the higher the total revenues to total assets ratio value, the higher the probability that the firm will go bankrupt. It may be assumed to have negative coefficient sign for that ratio as well. For bankrupt firms' total assets' amount usually becomes quite small, but at the same time, total revenues do not drop that much and this causes the anomaly for this ratio. Also, as this is a multisector model, firms belong to different industries like construction, service, wholesale and industrial industry, thus this ratio behaves differently among the sectors. Due to these two reasons the total revenues to total assets ratio is positive in this model.

In all of the random samples' models equity to total assets had p-value 0.000. For other predictive variables p-value varied from 0.001 to 0.045. Overall classification accuracy for sample 2 was 78.6% and for sample 3 it was 79.7%. So classification accuracy stayed between 78.6% – 81.5% for the random samples.

Overall classification accuracies have been quite high for logit models. Ohlson (1980) managed to classify correctly 96.12 % of the firms one year before and 2 years before the failure the correct classification was 95.55%. Casey and Bartczak (1985) reached model classification accuracy as high as 88% one year before the failure but their sample also included 290 firms in total. Similar accuracy (90%) was achieved by Platt and Platt's (1990) model but they also had a small sample with 114 firms in total (see Table 1). K-Model classification accuracy of 81.5% is quite a good result taking into account that the firms are not industry specific and compared to the Bellovary *et al.* (2007) results.

Table 6. Results with backward logit analysis for random 2 (R2) and random sample 3 (R3)

Variables names	B		S.E.		Wald		df	Sig.		Exp(B)	
	R2	R3	R2	R3	R2	R3	R2/3	R2	R3	R2	R3
equity/total assets	-2.53	-2.92	0.27	0.29	86.66	104.18	1	0.000	0.000	0.080	0.054
net income/total revenues	-0.41	-0.22	0.15	0.10	7.26	4.65	1	0.007	0.031	0.665	0.802
total revenues/ total assets	0.12	0.08	0.03	0.03	13.10	5.82	1	0.000	0.016	1.127	1.082
Constant	0.49	0.80	0.17	0.18	8.24	18.85	1	0.004	0.000	1.626	2.229

Source: compiled by the author

As mentioned at the beginning of chapter 3, logistic regression is very sensitive to multicollinearity and there cannot be multicollinearity among the predictive variables. For testing the multicollinearity among the predictive variables VIF (variance inflation factor) values were calculated to see whether the values belong in the accepted interval or not (see Table 7). If the model predictive variables have VIF value over 4, there is said to be multicollinearity (O'Brien 2007).

Table 7. Multicollinearity diagnostics

Variables names	VIF	Tolerance
equity/total assets	1.88	0.5329
net income/total revenues	1.22	0.8183
total revenues/ total assets	1.71	0.5862

Source: compiled by the author

Multicollinearity is not an issue in the model because for all of the predictive variables the VIF value is smaller than 4 and the tolerance indicator is greater than 0.1.

Additionally, marginal effects are of use while interpreting the models results. Marginal effects measure the rate of change which means that in bankruptcy prediction model it shows how much the probability of bankruptcy changes when one of the predictive variables values change by one unit. When equity to total assets value increases by one unit (other variables stay the same), the probability of bankruptcy decreases by 53%. With one unit increase in net income to total revenues ratios (other variables stay the same) the probability of bankruptcy decreases by 2.7% and a unit increase in total revenues to total assets (other variables stay the same) the probability of bankruptcy increases by 1.97%. Also, all the marginal effects are statistically significant at 5% significance level (see Table 8).

Table 8. Marginal effects of model variables for random sample number 1

Variables names	**dy/dx**	**Std. Err.**	**z**	**P>|z|**	**[95% C.I.]**		**X**
equity/total assets	-0.534	0.04332	-12.34	0.000	-0.6194	-0.4496	0.0174
net income/total revenues	-0.027	0 .0137	-1.99	0.046	-0.0542	-0.0004	-0.3186
total revenues/ total assets	0.0197	0.0059	3.33	0.001	0.0081	0.0312	3.1777

Source: compiled by the author

In order to check if the model is stable among the control variables such as firm's age, size and sector, the model is tested with these control variables. It is analyzed whether the model yields the same results among the firms of different age, sectors and size (see Table 9). Firm size was measured as the natural logarithm of total assets, firm's age was calculated as of how old the firm was when annual report was submitted and sectors were fixed based on their field of activity. When firm size was added to the K- Model, this variable was not statistically significant at significance level 5% (p=0.767) but other variables (equity to total assets, net income to total revenues and total revenues to total assets) stayed statistically significant. On the other hand, Altman *et al.* (2016) found that model with size variables performed better than the benchmark model but it has to be mentioned that the size effect was very small for micro firms.

In case only firm sectors were added, the result was same as for firm size that this control variable was not statistically significant at 5% significance level (among the sectors p-value > 0.2 and p-value < 0.45) but other three variables stayed significant. Opposite results were attained by Altman *et al.* (2016) where model with industry dummies outperformed the benchmark model. Although the model performed notably better with industry variables for France, Portugal, Latvia and Spain and for other countries the model improvement was not that much higher.

However, firm's age was the only statistically significant control variable. It means that financial ratios' values vary accordingly to company's age. When only firm's age was added to the K-Model, total revenues to total assets and net income to total revenues turned out not to be statistically significant anymore and firm's age and equity to total assets were statistically significant at 5% significance level. This result is in accordance with Altman *et al.* (2016) result where it was pointed out that younger firms have lower profitability ratio values because they have not had time to build up their cumulative profits and therefore for younger firms the likelihood of being classified as a bankrupt firm is higher than it is for older firms. As Altman *et al.* (2016) found in their model for firms 6 years or younger, the risk of going bankrupt is very high. Moreover, the model classification accuracy with firm's age dummy was as high as 93.7% for random sample number 1 while applying backward stepwise method and in the final step it used only two variables: firm's age and equity to total assets.

Predictive variables had also logical coefficient signs: both equity to total assets and firm's age had negative signs in front of the coefficient, which means that if the predictive values increase, the probability of going bankrupt decreases. This means that older companies have smaller probability of going bankrupt than the younger ones. As mentioned in the second chapter of the paper, Argenti (1976) had 3 failure trajectories where two of those were: young firms that cannot achieve sustainable profit and old firms that at some point of time lose their place in the market. Comparing the model variables with Argenti (1976) trajectories, it seems that in these samples of Estonian firms the younger companies that cannot find their place in the market and therefore go bankrupt in the few first years are overruling the old ones that lose their place in the market and then go bankrupt.

Firm age variable was also added to model with random sample number 2 and 3 and the model was able to classify correctly accordingly 85% (benchmark model had 78.6%) and 82.1% (benchmark model had 79.7%) which means that both were higher results than with the benchmark model that used only three variables.

Table 9. Results with backward logit analysis for random 1 sample including control variable firm age.

Variables names	B	S.E.	Wald	Df	Sig.	Exp(B)
equity/total assets	-4.39	0.63	49.15	1	0.000	0.01
net income/total revenues	-0.15	0.19	0.61	1	**0.435**	0.86
total revenues/total assets	0.03	0.06	0.28	1	**0.598**	1.03
firm age	-0.48	0.05	88.81	1	0.000	0.62
Constant	8.92	1.01	78.55	1	0.000	7484.74

Source: compiled by the author

When all the control variables were added together with the three of interest variables which were equity to total assets, net income to total revenues and total revenues to total assets and backward stepwise method was applied, in the final step as expected only equity to total assets and firm's age were included to the model with both statistically significant with p-value 0.000.

One of the recent bankruptcy prediction models is Altman *et al.* (2016) research where ratios used in their model were:

1) WCTA=working capital to total assets;

2) RETA=retained earnings to total assets;

3) EBITTA= EBIT to total assets;

4) BVETD=book value of equity to total liabilities.

Model that used the same variables as it was applied in Altman *et al.* (2016) model was built and it was tested with all 3 random samples. The results were unexpected. For all of the three samples at least one variable in each model had statistical significance p-values over 0.1, which means that those variables were not statistically significant at significance level 5% and even 10%. When backward stepwise method was applied for random sample number 2, working capital to total assets had p=0.234 and the final model did not include this ratio. Also, the classification accuracy was lower (75.8%) than the bankruptcy model built above that had classification accuracy 78.6%. Similarly as with random sample number 2 the same happened with random sample number 3 when backward stepwise method was applied, but this time retained earnings to total assets was not statistically significant with p-value of 0.494 and 74.3% of firms were correctly classified. Moreover, there was high multicollinearity among Altman *et al*. (2016) 4 predictive variables. VIF value for retained earnings to total assets was 6.88 and the mean value for all the ratios was 4.33 and, as mentioned before, when the value is greater than 4, there is a strong multicollinearity among the variables.

This bankruptcy prediction model can be also compared with other models that have used Estonian firms but where firms in the data have been industry specific. For example, Lukason (2006) used commercial firms' data, Grünberg (2013) has used companies from industry area and Onno (2015) has used Estonian road transportation companies in their master theses. Lukason's (2006) logit model correctly classified 98.9% of the firms, Grünberg's (2013) model classification accuracy was 72% for bankrupt firms and 88% for non-bankrupt firms a year before the failure. Onno (2015) had 82.3% in training sample and 79.3% correctly predicted results in test sample. However, all those models were industry or firms specific and it is not certain how well they would perform with firms among all sectors.

This thesis can be improved in several ways. Firstly, it would be possible to test this model on a test sample as well. Also, the model could be tested with sample weights. As there was information about 1712 non-bankrupt firms and 325 bankrupt firms, it could be tested if model weighted sample would give different results or not. Moreover, as mentioned in the second chapter of this paper, it is difficult to predict failure based on annual reports in the event of acute failure processes. In that case it is hard to predict the bankruptcy with a model based on financial ratios because they may not show any signs of problems in the firm even one year before it goes bankrupt. The solution for more accurate prediction could be achieved by using some additional variables that would help to see the problems in the firms. For example, Back (2005) used payment delays in the bankruptcy prediction model. In case of Estonia, a similar variable could be tax arrears data. It may help to distinguish better the firms with acute failure processes from "healthy" firms, so bankruptcy prediction would be more accurate.

5. Conclusion and acknowledgements

Since 1966 when the first corporate failure prediction model was built by Beaver, the development of models has grown into a wide research domain. Many models have been built using different methods and data that have been, for example, country or industry specific. Due to the arbitrary definition of failure many of the authors have used the phrase "failure prediction" by what they have actually meant "bankruptcy prediction".

Ohlson (1980) was the first to apply the logit model in bankruptcy prediction models. In this paper the logistic regression method is used as well. The model built in this thesis is unique because it is a multisector model and to author's knowledge no one in Estonia has previously compiled such a model. Multisector model is useful for investors and creditors who have to decide to whom they lend their money or which stocks to invest in. Bankruptcies that occurred between 2012 and 2016 have used and the sample includes 325 bankrupt and 1712 non-bankrupt firms. Whereas the previous models with Estonian data have focused on one industry at a time and models' variables have been also industry specific, for this paper's final model the variables that could yield high prediction results with all the sectors were used. Finally, equity to total assets, total revenue to total assets and net income to total revenues are used in the model.

The model is tested with three randomly generated samples. Each random sample includes all available data for 325 bankrupt firms and then randomly selected 325 non-bankrupt firms' data as well. The overall classification accuracy for the first sample is 81.5%, in which the results for non-bankrupt firms the result is 82.5% and for bankrupt firms 80.6%, which shows that the results among samples are stable. Control variable firm's age is the only control variable that is statistically significant at 5% significance level. By adding firm age to the model two other variables net income to total revenues and total revenues to total assets, were not statistically significant anymore. Also, the prediction accuracy rises for sample 1 to 93.5% and for the other two samples to 85% and 82.1%.

For further implications the usage of training and test samples could be useful. Also, instead of paired sample technique the weighted sample could be tested as well to see whether it would yield different results. Moreover, this thesis can be improved for example by adding the tax arrears data to the model and, as a result, the model prediction could be more accurate. Only with the usage of financial ratios it is difficult to correctly predict the result for the firms that go through the acute failure process.

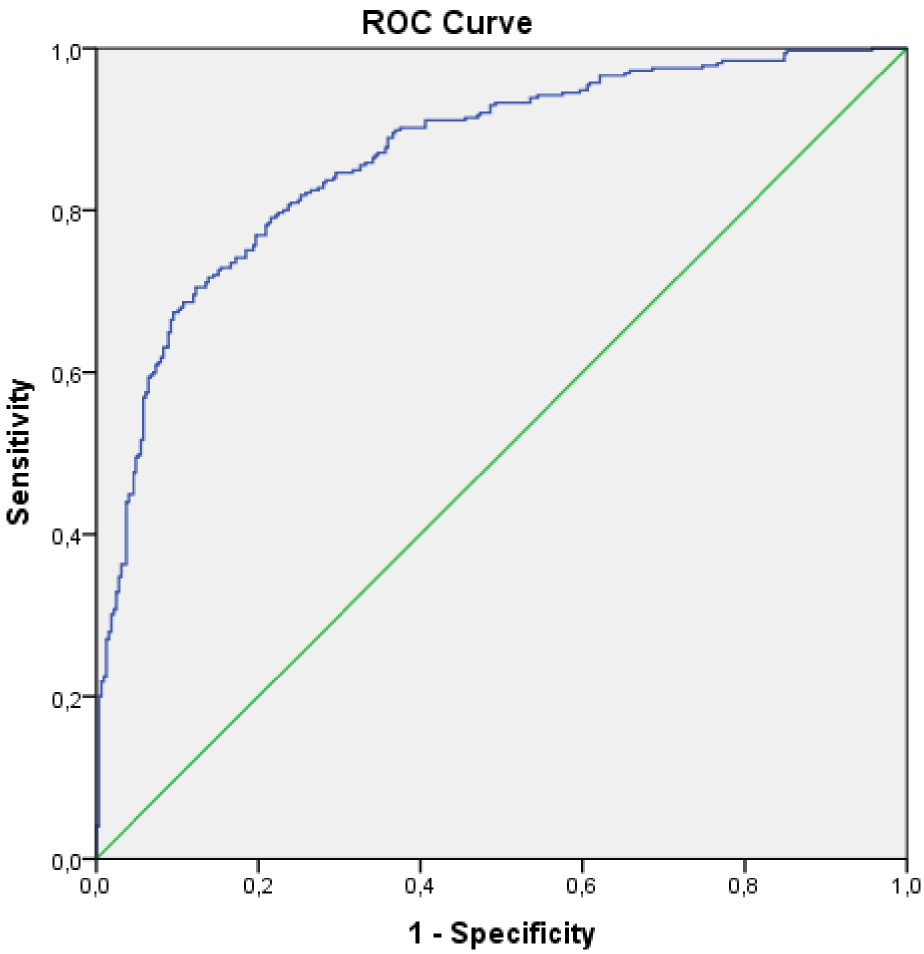
The author would like to thank her supervisor Oliver Lukason for guidance and useful comments during the entire writing process and former collegues from Bigbank for the support through the master studies.

Appendices

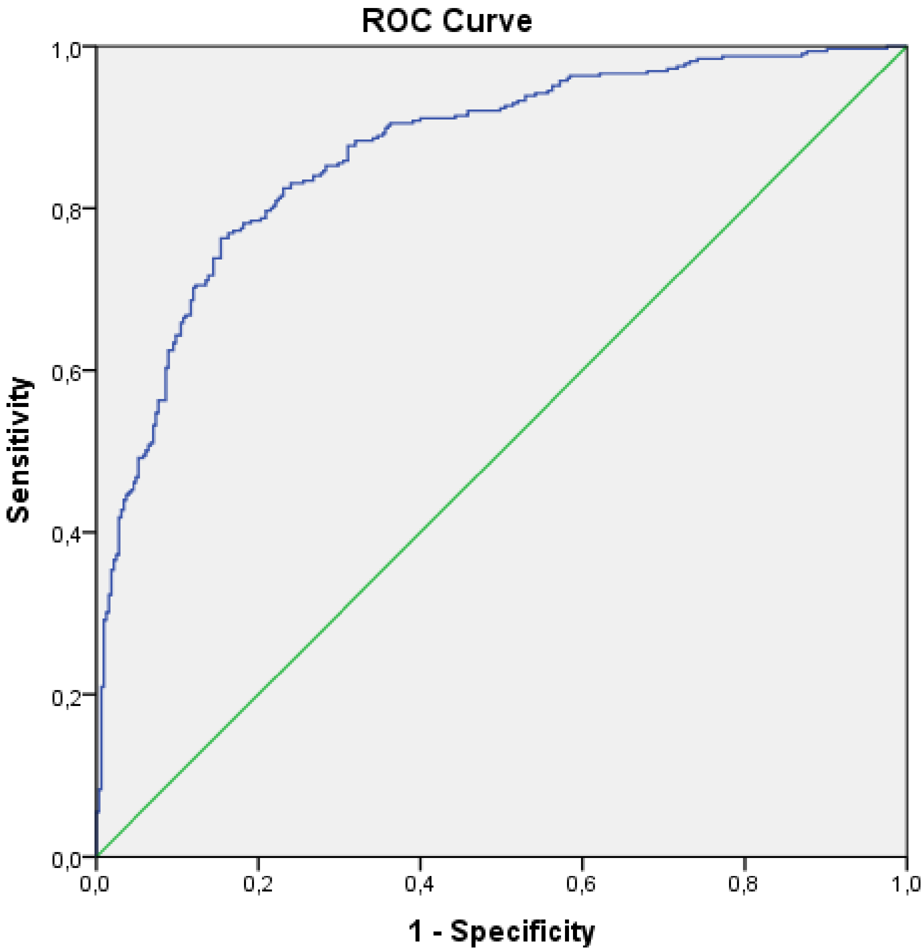
Appendix 1. Correlation matrix

	CS/TA	CS/CL	CA/TA	CA/CL	TL/TA	E/TA	TR/TA	BP/TR	NI/TR	NI/TA	EBT/TA	BP/TA	RE/TA	TL/E	WC/TA	size
CS/TA	1.00															
CS/CL	**0.67**	1.00														
CA/TA	**0.49**	0.28	1.00													
CA/CL	**0.42**	**0.84**	0.34	1.00												
TL/TA	-0.17	-0.36	-0.03	-0.38	1.00											
E/TA	0.17	0.36	0.03	0.38	**-1.00**	1.00										
TR/TA	0.05	-0.16	0.20	-0.20	0.39	-0.39	1.00									
BP/TR	0.08	0.08	0.05	0.06	-0.15	0.15	0.05	1.00								
NI/TR	0.09	0.09	0.05	0.08	-0.18	0.18	0.04	**0.89**	1.00							
NI/TA	0.06	0.12	0.00	0.12	**-0.63**	**0.63**	-0.34	0.25	0.29	1.00						
EBT/TA	0.07	0.12	0.01	0.12	**-0.63**	**0.63**	-0.34	0.25	0.29	**1.00**	1.00					
BP/TA	0.00	0.16	-0.04	0.18	**-0.69**	**-0.69**	-0.37	0.12	0.13	**0.43**	**0.43**	1.00				
RE/TA	0.03	0.17	-0.02	0.19	**-0.80**	**0.80**	**-0.40**	0.16	0.19	**0.70**	**0.70**	**0.82**	1.00			
TL/E	-0.14	-0.20	-0.06	-0.20	0.14	-0.14	0.01	-0.01	-0.02	0.00	0.00	-0.05	-0.04	1.00		
WC/TA	0.32	**0.43**	**0.40**	**0.50**	**-0.87**	**0.88**	-0.30	0.16	0.19	**0.58**	**0.58**	**0.58**	**0.69**	-0.12	1.00	
size	-0.39	-0.24	-0.31	-0.16	-0.06	-0.06	-0.25	0.00	0.02	0.14	0.14	0.21	0.23	0.09	-0.03	1.00

Appendix 2. Random sample 2 ROC curve



Appendix 3. Random sample 3 ROC curve



Appendix 4. Random sample 1 classification plots

```
             Step number: 1

             Observed Groups and Predicted Probabilities

      80 +                                                                                                    +
         I                                                                                                    I
         I                                                                                                   1I
F        I                                                                                                   1I
R     60 +                                                                                                   1+
E        I                                                                                                   1I
Q        I                                                                                                   1I
U        I                                                                                                   1I
E     40 +                                                                                                   1+
N        I                                                                                                   1I
C        I      1                                                                                            1I
Y        I      0                                                                                            1I
      20 +     10  1                                                                                         1+
         I     000000010                                                                                     1I
         I     000000000 1      00   1                    1                         1  111   1    1       1 11I
         I    000000000000000000000000000111101 1110   01 011 110  011 11101101001101111111 10 11111111 1 1111I
Predicted ---------+---------+---------+---------+---------+---------+---------+---------+---------+----------
  Prob:   0        ,1        ,2        ,3        ,4        ,5        ,6        ,7        ,8        ,9         1
  Group:  0000000000000000000000000000000000000000000000000011111111111111111111111111111111111111111111111111

          Predicted Probability is of Membership for 1
          The Cut Value is ,50
          Symbols: 0 - 0
                   1 - 1
          Each Symbol Represents 5 Cases.
```

Appendix 5. Random sample 2 classification plots

```
             Step number: 1

             Observed Groups and Predicted Probabilities

      80 +                                                                                                    +
         I                                                                                                    I
         I                                                                                                    I
F        I                                                                                                    I
R     60 +                                                                                                    +
E        I                                                                                                   1I
Q        I                                                                                                   1I
U        I                                                                                                   1I
E     40 +                                                                                                   1+
N        I                                                                                                   1I
C        I                                                                                                   1I
Y        I                                                                                                   1I
      20 +                                                                                                   1+
         I           0  0 1   1                                                                              1I
         I          00010000000100  01 0 1  1 1      1 1 1  1 1 111  1111111        1                        1I
         I        0000000000000000000000000 01011100000010100 0000000110011011111 111 1 1 1 1 1  1 1  11111111I
Predicted ---------+---------+---------+---------+---------+---------+---------+---------+---------+----------
  Prob:   0        ,1        ,2        ,3        ,4        ,5        ,6        ,7        ,8        ,9         1
  Group:  0000000000000000000000000000000000000000000000000011111111111111111111111111111111111111111111111111

          Predicted Probability is of Membership for 1
          The Cut Value is ,50
          Symbols: 0 - 0
                   1 - 1
          Each Symbol Represents 5 Cases.
```

Appendix 6. Random sample 3 classification plots

```
             Step number: 1

             Observed Groups and Predicted Probabilities

      80 +                                                                                                    +
         I                                                                                                    I
         I                                                                                                    I
F        I                                                                                                    I
R     60 +                                                                                                   1+
E        I                                                                                                   1I
Q        I                                                                                                   1I
U        I                                                                                                   1I
E     40 +                                                                                                   1+
N        I                                                                                                   1I
C        I                                                                                                   1I
Y        I                                                                                                   1I
      20 +           01 0                                                                                    1+
         I           00 001 0                                                      1                         1I
         I         0 0000000001  10    01   101    1   1 1       1  1  11  111111                            1I
         I         000000000000000000000001001  00000101000 0010100 1110011010000 1111 111 11  1  1  1 111111I
Predicted ---------+---------+---------+---------+---------+---------+---------+---------+---------+----------
  Prob:   0       ,1        ,2        ,3        ,4        ,5        ,6        ,7        ,8        ,9         1
  Group:  0000000000000000000000000000000000000000000000000011111111111111111111111111111111111111111111111111

          Predicted Probability is of Membership for 1
          The Cut Value is ,50
          Symbols: 0 - 0
                   1 - 1
          Each Symbol Represents 5 Cases.
```

References

1) Altman, E. I. Financial Ratios, Discriminant Analysis And The Prediction Of Corporate Bankruptcy. – Journal of Finance, 1968, Vol. 23, No. 4, pp. 589–609.

2) Altman, E.I., Iwanicz-Drozdowska, M., Laitinen, E.K., Suvas, A. Financial distress prediction in an international context: a review and empirical analysis of Altman's Z-score model. J. Int. Financial Manage. Account , 2016

3) Argenti, J. Corporate Collapse: the Causes and Symptoms. Berkshire: McGraw– Hill Book Company, 1976, 193 p.

4) Aziz, A., Emanuel D., Lawson G. Bankruptcy prediction - An investigation of cash flow based models. Journal of Management Studies, 1988, Vol. 25, No. 5, pp. 419–437.

5) Back, P. Explaining financial difficulties based on previous payment behavior, management background variables and financial ratios, European Accounting Review, 2005, Vol. 14, No. 4, pp. 839–868.

6) Balcaen, S., Ooghe, H. 35 years of studies on business failure: an overview of classic statistical methodologies and their related problems. British Accounting Review, 2006, Vol. 38, No. 1, pp. 63–93.

7) Bankruptcy Act. Riigi Teataja. 2003. [https://www.riigiteataja.ee/en/eli/ee/Riigikogu/act/504072016002/consolide]. 07.04.2017

8) Bariatti, S., van Galen, R. Study on a New Approach to Business Failure and Insolvency- Comparative Legal Analysis of the Member States‘ Relevant Provisions and Practices. INSOL Europe, 2014, 44 p.

9) Barnes, P. The Analysis and Use of Financial Ratios: A Review Article. – Journal of Business Finance and Accounting, 1987, Vol. 14, No. 4, pp. 449–461.

10) Beaver, W. H. Financial Ratios as Predictors of Failure, Journal of Accounting: Selected Studies, 1966, Vol. 4, pp. 71–111.

11) Bellovary, J., Giacomino, D., Akers, M. A review of bankruptcy prediction studies: 1930 to present, 2007, Journal of Financial Education 33 (Winter), pp. 1–42.

12) Bernard, V. L., Stober, J. L. The Nature and Amount of Information in Cash Flows and Accruals, 1989, The Accounting Review 64 (October), pp. 624–651.

13) Bhandari, S., Iyer, R. Predicting Business Failure Using Cash Flow Statement Based Measures. Managerial Finance, 2013, pp. 667–676.

14) Black, F., Scholes, M. The Pricing of Options and Corporate Liabilities. Journal of Political Economy 7, 1973, pp. 637–654.

15) Carslaw, C. A., Mills, J. R. Developing Ratios for Effective Cash Flow Statement Analysis, Journal of Accountancy, 1991, Vol. 172 (November), pp. 63–70.

16) Casey C., Bartczak, N. Using operating cash flow data to predict financial distress: Some extensions. Journal of Accounting Research, 1985, Vol. 23, No. 1, pp. 384–401.

17) Charitou, A., Neophytou, E., Charalambous, C. Predicting corporate failure: empirical evidence for the UK. European Accounting Review, 2004, Vol. 13, No. 3, pp. 465–497.

18) Cochran, A. B. Small business mortality rates: A review of the literature. Journal of Small Business Management, 1981, Vol. 19, No. 4, pp. 50–59

19) D'Aveni, R. A. The Aftermath of Organizational Decline: A Longitudinal Study of the Strategic and Managerial Characteristics of Declining Firms. – The Academy of Management Journal, 1989, Vol. 32, No. 3, pp. 577–605.

20) Daubie M., Meskens N. Business failure prediction : a review and analysis of the literature. Working Paper, Department of Productions and Operations Management, Catholic University of Mons, Belgium, 2002, pp. 1–15.

21) Davydenko, S. A., Franks, J. R. Do bankruptcy codes matter? A study of defaults in France, Germany, and the UK. Journal of Finance, 2008, Vol. 63, No. 2, pp. 565–608.

22) Dimitras, A. I., Zankis, S. H., Zopounidis, C. A Survey of Business Failures With Emphasis on Prediction Methods and Industrial Applications. – European Journal of Operational Research, 1996, Vol. 90, No. 3, pp. 487–513.

23) Estonian Business Code. Riigi Teataja. 1995. [https://www.riigiteataja.ee/akt/131122010019]. 07.04.2017.

24) Garson, G. D. Testing Statistical Assumptions, Statistical Associates Publishing: Asheboro, NC, USA, 2012

25) Gentry, J. Bankruptcy, Working Capital and Funds Flow Components. Urbana, Ill.: Bureau of Economics and Business Research, 1984

26) Giacomino, D.E., Mielke, D.E. Using the statement of Cash flows to analyze corporate performance. Management of Accounting, 1988, pp. 54–57.

27) Gilbert, L. R., Menon, K., Schwartz, K. B. Predicting bankruptcy for firms in financial distress. Journal of Business Finance and Accounting, 1990, Vol. 17, No. 1, pp. 161–171.

28) Grünberg, M. Pankroti prognoosimise mudeli koostamine Eesti tööstussektori ettevõtete näitel, TÜ Ettevõttemajanduse instituut, 2013 (master thesis).

29) Hambrick, D. C., D'Aveni, R. A. Large Corporate Failures as Downward Spirals. – Administratice Science Quarterly, 1988, Vol. 33, No. 1, pp. 1–23.

30) Hillegeist, S.A., Keating, E.K., Cram, D.P. and Lundstedt, K.G. Assessing the probability of bankruptcy. Review of Accounting Studies, 2004, Vol. 9, No. 1, pp. 5–34.

31) Hosmer, D., Lemeshow W. Applied Logistic Regression, Stanley, 2000, 2nd ed.

32) Keener, M. H. Predicting the financial failure of retail companies in the United States. Journal of Business & Economics Research, 2013, Vol. 11, No. 8, pp. 373–380.

33) Klieštik, T., Kočišová, K., Mišanková, M. Logit and probit model used for prediction of financial health of company, Procedia Econ. Finance 23, 2015, pp. 850–855.

34) Kumar, P. R., Ravi, V. Bankruptcy prediction in banks and firms via statistical and intelligent techniques–A review. European journal of operational research, 2007, Vol. 180, No. 1, pp. 1–28.

35) Laitinen, E. Financial Ratios and Different Failure Processes. – Journal of Business Finance & Accounting, 1991, Vol. 18, No. 5, pp. 649–673.

36) Laitinen, E. K. Financial Predictors for Different Phases of the Failure process. – Omega, 1993, Vol. 21, No. 2, pp. 215–228.

37) Lawrence, J. R., Pongsatat, S., Lawrence, H. The Use Of Ohlson's O-Score For Bankruptcy Prediction In Thailand. The Journal of Applied Business Research, 2015, Vol. 31, No. 6, pp. 2069–2078.

38) Lukason, O. Charactersitics of firm failure processes in an international context, TÜ, 2016 (Ph.D. Thesis).

39) Lukason, O. Pankrotistumiste modelleerimine Eesti kaubandusettevõtete näitel. TÜ ärirahanduse ja investeeringute õppetool, 2006 (master thesis).

40) Lukason, O., Laitinen, E. K., Suvas, A. Failure processes of young manufacturing micro firms in Europe, Management Decision, 2016, Vol. 54, No. 8, pp. 1966–1985.

41) Merton, R. On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. Journal of Finance 29, 1974, pp. 449–470.

42) NIST/SEMATECH e-Handbook of Statistical Methods, http://www.itl.nist.gov/div898/handbook//prc/section1/prc16.htm , 08.04.2017.

43) O'Brien, R. M. A Caution Regarding Rules of Thumb for Variance Inflation Factors. Quality & Quantity, 2007, Vol. 41, pp. 673–690.

44) Ohlson, J. A. Financial ratios and the probabilistic prediction of bankruptcy. Journal of Accounting Research, 1980, Vol. 18 , No. 1, pp. 109–131.

45) Onno, A. Pankroti prognoosimise mudelite koostamine Eesti maanteetranspordi ettevõtete näitel, TÜ, 2015 (master thesis).

46) Pervan, I., Kuvek, T. The relative importance of financial ratios and nonfinancial variables in predicting of insolvency. Croatian Operational Research Review, 2013, Vol. 4, No. 1, pp. 187–197.

47) Pindado, J., Rodrigues, L., De la Torre, C. Estimating financial distress likelihood. Journal of Business Research, 2008, Vol. 61, No. 9, pp. 995–1003.

48) Platt, H. D., Platt, M. B. Development of a class of stable predictive variables: The case of bankruptcy predic- tion, 1990, Journal of Business Finance and Accounting, Vol. 17, No. 1, pp. 31–51.

49) Platt, H. D., Platt, M. B. Predicting Corporate Financial Distress: Reflections on Choice–Based Sample Bias. – Journal of Economics and Finance, 2002, Vol. 26, No. 2, pp. 184–199.

50) Pretorius, M. Defining business decline, failure and turnaround: a content analysis. Southern African Journal of Entrepreneurship and Small Business Management, 2009, Vol. 2, No. 1, pp. 1–16.

51) Probst, G., Raisch, S. Organizational crisis: The logic of failure. Academy of Management Executive, 2005, Vol. 19, No. 1, pp. 90–105.

52) Scott, J. The probability of bankruptcy: A comparison of empirical predictions and theoretical models. Journal of Banking and Finance, 1981, Vol. 5, No. 3, pp. 317– 344.

53) Tamari, M. Financial Ratios as a Means of Forecasting Bankruptcy, Management International Review, 1966, Vol. 4, pp. 15–21.

54) Ulmer, M. J., Nielsen, A. Business turn-over and causes of failure, Survey of Current Business (April), 1947, pp. 10–16.

55) Von Stein, J. H., Ziegler, W. The Prognosis and Surveillance of Risks from Commercial Credit Borrowers, Journal of Banking and Finance, 1984, Vol. 8, pp. 249 – 268.

56) Watson, J., Everett, J. Small Business Failure Rates: Choice of Definition and the Size Effect, Journal of Entrepreneurial and Small Business Finance, 1996, Vol. 5, No. 3, pp. 271–285.

57) Zavgren, C. V. Assessing the vulnerability to failure of American industrial firms. A logistic analysis, Journal of Business Finance and Accounting, 1985, Vol. 12, No. 1, pp. 19–45.

58) Zeller, T. L., Figlewicz, R. E. New investing and financing activity ratios from the statement of cash flows enhance traditional ratio analysis in assessing future cash flows. The Mid-Atlantic Journal of Business, 1990, Vol. 26, No. 2, pp. 49–59.

Non-exclusive licence to reproduce thesis and make thesis public

I,

Kerttu Lääne,

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to:

1.1. reproduce, for the purpose of preservation and making available to the public, including for addition to the DSpace digital archives until expiry of the term of validity of the copyright, and

1.2. make available to the public via the web environment of the University of Tartu, including via the DSpace digital archives until expiry of the term of validity of the copyright,

“Corporate bankruptcy prediction of Estonian firms”,
supervised by Oliver Lukason.

2. I am aware of the fact that the author retains these rights.

3. I certify that granting the non-exclusive licence does not infringe the intellectual property rights or rights arising from the Personal Data Protection Act.

Tartu, 25.05.2017