UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Scienze Economiche e Aziendali "M. Fanno"

CORSO DI LAUREA MAGISTRALE IN
BUSINESS ADMINISTRATION

DO THE ACCOUNTING-BASED MODELS FOR BANKRUPTCY PREDICTION STILL WORK? A TEST ON THE FIRMS FROM PADOVA AND VICENZA

Relatore:
Ch.mo prof. Antonio Parbonetti

Laureando: Matteo Belfiori

Matricola n. 1179412

Anno Accademico 2018 – 2019

Il candidato dichiara che il presente lavoro è originale e non è già stato sottoposto, in tutto o in parte, per il conseguimento di un titolo accademico in altre Università italiane o straniere.
Il candidato dichiara altresì che tutti i materiali utilizzati durante la preparazione dell'elaborato sono stati indicati nel testo e nella sezione "Riferimenti bibliografici" e che le eventuali citazioni testuali sono individuabili attraverso l'esplicito richiamo alla pubblicazione originale.

The candidate declares that the present work is original and has not already been submitted, totally or in part, for the purposes of attaining an academic degree in other Italian or foreign universities. The candidate also declares that all the materials used during the preparation of the thesis have been explicitly indicated in the text and in the section "Bibliographical references" and that any textual citations can be identified through an explicit reference to the original publication.

Firma dello studente
____

Ringraziamenti

Per il costante aiuto che ho ricevuto nella scrittura di questa tesi, ringrazio il Professor Antonio Parbonetti, di grande supporto nella parte teorica e nella costruzione del dataset.

Ringrazio i miei genitori, i quali, nonostante non fossero sempre d'accordo con le mie scelte, mi hanno sempre supportato e mi hanno sostenuto nel percorrere la mia strada.

Ringrazio la mia nonna materna, Marilena, per essere stata il mio hotel in questi cinque anni di università.

Ringrazio tutti i miei familiari, per essere stati sempre delle persone di grande supporto morale.

Ringrazio Padova, la mia città, che mi ha permesso di passare dei meravigliosi anni e di vivere al meglio l'esperienza universitaria.

Ringrazio i miei amici di sempre, per esserci sempre stati nonostante le distanze e gli impegni.

Ringrazio l'università di Padova e il progetto Erasmus per avermi aiutato ad aprire gli occhi, per avermi fatto conoscere persone di mille culture diverse, per avermi arricchito da tantissimi punti di vista e per avermi permesso di conoscere la mia ragazza, Beatriz.

Ringrazio quest'ultima per avermi sempre supportato e aver creduto in me anche quando nemmeno io a volte ci credevo.

Infine, ringrazio AEP ESN Padova per avermi fatto conoscere un bel gruppo di amici e per avermi fatto sentire in Erasmus nonostante fossi a casa.

Summary

Introduction

The following thesis deals with the prediction of the failure of a company through the analysis of the income statements and the balance sheets in the years before the bankruptcy. In particular, the models used in this thesis are based on the data taken from the firms' annual report.

The so named accounting-based models will predict whether a certain company, taken from a sample of small and medium firms from Padova and Vicenza, will fail in the following years.

In the first chapter it is provided a definition of bankruptcy and the consequences of laggard firms in the market; in the second chapter before are presented all the most used kinds of bankruptcy prediction models (the Altman Z', Z'' and ZETA models, Ohlson model and the Beaver, as well as the ratios proposed by the italian order of Chartered Accountants and Accounting Expert. Going forward, in the third chapter all the above mentioned literature models are tested. In the fourth chapter a new accounting-based model will be created and will be tested as well and then compared with the literature models, to understand which one is more accurate in the bankruptcy prediction of this thesis sample. Finally, in the fifth chapter, also the model proposed by the by the italian order of Chartered Accountants and Accounting Expert will be tested.

1. CHAPTER ONE

Why should we predict firm's bankruptcy?

1.1 Definition of bankruptcy

According to Business dictionary, bankruptcy is the "Legal procedure for liquidating a business (or property owned by an individual) which cannot fully pay its debts out of its current assets. Bankruptcy can be brought upon itself by an insolvent debtor (called 'voluntary bankruptcy') or it can be forced on court orders issued on creditors' petition (called 'involuntary bankruptcy'). Two major objectives of a bankruptcy are (1) fair settlement of the legal claims of the creditors through an equitable distribution of debtor's assets, and (2) to provide the debtor an opportunity for fresh start."

As it is mentioned above, understanding whether a firm will fail, is important from the point of view of:

- The current creditors because they have to take into account that they can lose a part of the money borrowed to the lender;
- The potential lenders because they have to consider if they should lend money to the borrowers, which collaterals they should require, if they should increase the interest rate, etc.
- The shareholders, because they are the ones that are going to lose money first in the case of company failure and they are those that can appoint new managers if the actual ones are not working correctly;
- The managers/entrepreneurs, because they can create a recovery plan to go out from a negative situation. Indeed, according with Jensen (1989) and Whitaker (1999), a financial distress situation forces management to implement a series of actions aimed at improving the firm's overall performance. This is true when the poor management performance is the cause of the difficult firm situation and, in the sample of Whitaker, this is the only factor that leads to the crisis in the 39.3% of the times and it is accompanied by the economic distress in the 37.5 of the times. On the other side, when the economic distress or other reasons are the only reason of the financial distress, the considerations written above are not valid.

As intuitively the reader can get, when the firm faces bankruptcy, it is too late. The managers should detect the signals of a potential bankruptcy before, just when the company starts being in a financial distress situation. This last three words represent "[…] a condition in which a company or individual cannot generate revenue or income because it is unable to meet or cannot pay its financial obligations. This is generally due to high fixed costs, illiquid assets, or revenues sensitive to economic downturns." (Kenton, 2019). From an accounting point of view, it is the inability of the company to repay the creditors (supplier or lenders, for example) with its operating performance. If this happens just one year, the company will face just a small crisis moment, if it persists, it can lead to the bankruptcy.

1.2 Zombie firms

Not only the above-mentioned categories but also the regulators have advantages in predicting the failure of a company. Lagging companies can damage the market and the regulators must detect them in order to improve the efficiency of the competitive environment.

Lagging and close to bankruptcy firms will be also called henceforth "zombie firms". They can be classified as zombies "whether they are receiving subsidized credit. […] If instead we were to define zombies based on their operating characteristics, then almost by definition industries dominated by zombie firms would have low profitability, and likely also have low growth." (Caballero, et al., 2008) Similar definition were provided in other researches in this same fields, like "In economic terms, a zombie is a firm that is not viable and therefore, when competitive forces are at play, should be compelled to exit the market or, where feasible, restructure." (Fontoura Gouveia & Osterhold, 2018). The state of Korea classifies a firm as zombie if the operating income is lower than the interest expenses. According to Storz, et al. (2017), a firm is classified as a zombie if, for two years in a row, the firm performs a negative ROA, a negative net investment and EBITDA to total financial debt lower than 5%.

An interesting finding in Portugal was that "Zombie firms are on average larger companies and significantly less productive than their healthy counterparts, pushing labor productivity down." (Fontoura Gouveia & Osterhold, 2018) Moreover, in that country 1 worker out of 5 is employed in a zombie firm.

"The zombies' distortions came in many ways, including depressing market prices for their products, raising market wages by hanging on to the workers whose productivity at the current firms declined, and, more generally, congesting the markets where they participated. Effectively, the growing government liability that came from guaranteeing the deposits of banks that supported the zombies served as a very inefficient program to sustain employment. Thus, the normal competitive outcome whereby the zombies would shed workers and lose market share was thwarted. More importantly, the low prices and high wages reduce the profits and collateral that new and more productive firms could generate, thereby discouraging their entry and investment. Therefore, even solvent banks saw no particularly good lending opportunities in Japan." (Caballero, et al., 2008) Moreover, the average zombie firm, according with Adalet McGowan, et al. (2017) "inflates wages relative to productivity and depresses market prices and (non-zombie) market shares".

Their negative influence damages the market especially in the difficult moments, when a recovery is needed but it is depressed by the zombie firms. Indeed, when a shock hits the economy or a part of it, the sectors that have zombie firms are the ones fall more and from the point of view of the granting of credit, the collateral values are likely to be lower (also for the firms considered "healthy").

Moreover, "increases in percentages of zombie firms operating in an industry significantly reduce both investment and employment growth for the healthy firms in the industry. Second, […], the productivity gap between zombies and non-zombies rises significantly as the percentage of zombies in an industry rises." (Caballero, et al., 2008) This thesis is supported by the figure 1, provided by Andrews, et al. (2016), that shows huge increase in the gap between the best and worst performing firms.

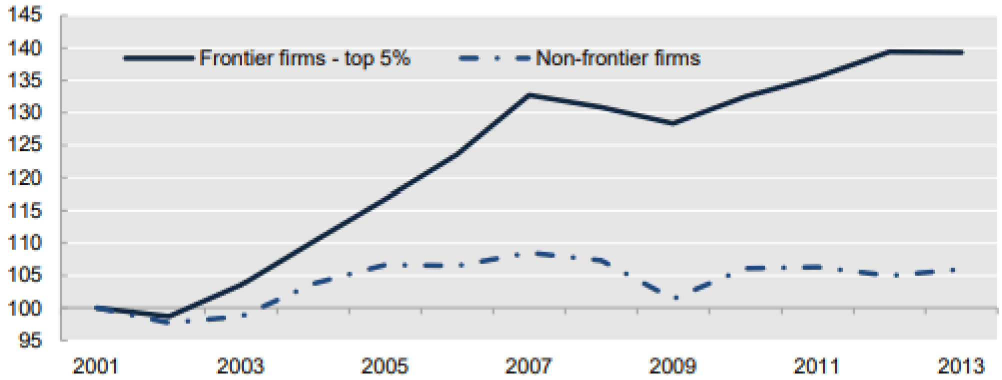

Figure 1 - Widening productivity gap between frontier and laggard firms; based on 24 OECD countries

The presence of zombie firms in an industry decreases the average productivity of a sector, both because of the low performance of the zombie firms but also because of the entrance barriers imposed on those efficient firms that would like to join the market but whose entrance is blocked by the unhealthy firms. At the same time, "a reduction in exit and restructuring barriers promotes a more effective exit channel" (Fontoura Gouveia & Osterhold, 2018).

Furthermore, in those industries where the capital allocated in the zombie firms sinks, a decline in the ability of attracting capital by the productive firms is observed.

Adalet McGowan, et al. (2017) sustained that in some countries, firms that should exit a market, do not and this is due to the fact that "In some countries, these problems are likely symptomatic of structural policy weaknesses, particularly with respect to insolvency regimes. But there are reasons to suspect that nonviable firms may also be increasingly kept alive by the legacy of the financial crisis, with bank forbearance, prolonged monetary stimulus and the persistence of crisis-induced SME support policy initiatives emerging as possible culprits". Furthermore, according to Andrews, et al. (2016) the slowdown in the technology innovation is worsened by weak reforms that don't induce the laggard firms to abandon the market. This creates an innovation gap with the frontier firms and, so, the average performance worsens.

A worrying finding of Adalet McGowan, et al. (2017) is that the number of zombie firms has increased since the mid 2000-s and this is a datum that also the Italy should take into consideration, because in the same paper, they sustained that, if a causal relationship is assumed, the business investment of the average non-zombie firm would have been 2% greater in 2013, whenever the number of zombie firms remained at the same level of 2007, with great benefits for Italy, Finland and Spain for example. At the same time, if the number of zombie firms remained at the pre-crisis level, "the contribution of capital reallocation to aggregate MFP in 2013 would have been around 0.7% to 1% higher in Italy and Spain, respectively. In other countries, reducing zombie congestion to the lowest level observed within each industry could yield gains to MFP of up to 0.5%." (Adalet McGowan, et al., 2017)

The low productivity of the zombie firms should represent an incentive for the well performing firms to increase their share and kick them off the market. Instead in Italy and Spain the difference in capital growth between a healthy and a non-healthy firm declined by around 2% in the period since 2004 to 2013, as the figure 2, taken from Adalet McGowan & Andrews (2017), shows.

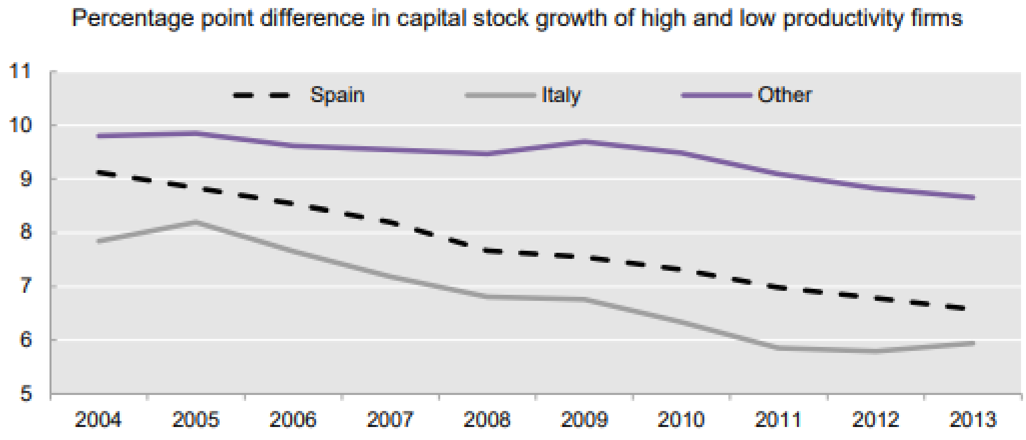

Figure 2 - Micro-level dimensions to the productivity slowdown: capital allocation

Another interesting datum in the comprehension of the consequences of the presence of non-performing firms is that "NPLs in Italy have reached high levels, hindering the recovery. Cleaning up banks' balance sheets is crucial to encourage credit growth, especially to SMEs that are more reliant on bank financing. Resolving impaired loans would also help facilitate restructuring or resolution of distressed SMEs." (Garrido, et al., 2016). The researchers also highlighted the need of an improvement of the weak banking Italian system.

The regulators could focus in the presence of zombie firms, whose presence in the market in positively correlated with the number of NPLs, as the figure 3, got through OECD calculations based on IMF, Financial Soundness Indicators and ORBIS, shows:

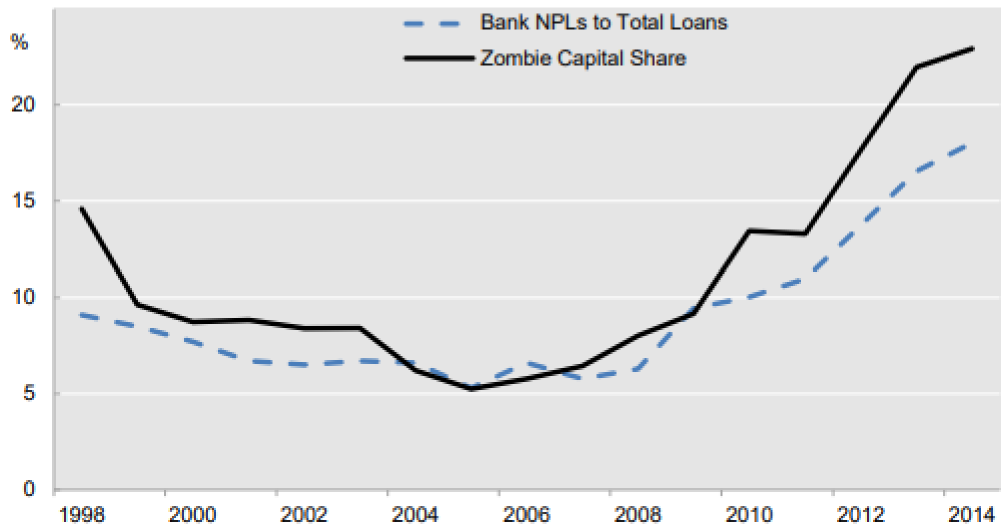

Figure 3 - Zombie firms and NPLs: the case of Italy (Source: OECD calculations)

This kind of policy should be accompanied by other political maneuvers that facilitate the job turnover and the labor mobility. They should be implemented together with other policies to manage the costs related to the move and reallocation of the workers. This is fundamental because there is proof that many high skilled workers is employed in the zombie laggard firms. Zombie firms are expected to meet more difficulties in borrowing money from financial entities and banks; instead the following perverse mechanism was documented: "Firms are more likely to receive additional bank credit if they are in poor financial condition, because troubled Japanese banks have an incentive to allocate credit to severely impaired borrowers in order to avoid the realization of losses on their own balance sheets". (Peek & Rosengren, 2005). Moreover, the governments, with the aim of avoiding the bankruptcy of the weakest firms, make pressure on the banks to lend money to them. This way of work allows banks to increase their income through money lending to unworthy firms. From the point of view of troubled banks, this mechanism allows them to not fail and to achieve this aim without much effort and without reform the entity. Furthermore, the Outright Monetary Transactions (OTM), i.e. a ECB program under which the Euro system buys or sells eligible assets outright on the market, helps this practice. Many banks benefit of the inflow of liquidity brought by these operations. Some undercapitalized banks, since 2012, started lending to zombie firms with that amount of money. An improvement in the capital allocation, favoring complementary benefits of bank health and insolvency regimes, can reduce the practices of bank forbearance and survival of environment damaging firms. Fontoura Gouveia & Osterhold (2018) published that “The evidence shows that there are additional policy complementarities, beyond the ones related to bank health, that need to be promptly addressed. For instance, ensuring a fit for purpose regulatory environment is an important challenge for policy makers, as product market distortions and administrative barriers to entry are also positively associated with higher zombie congestion and lower exit”. Also they found out that in Portugal, even if two third of the zombie firms remained as such for two following years, the market managed in doing partially a positive selection, with the most productive zombie firms restructuring.

All these researches prove that an analysis and a detection of these zombie firms can lead to a better market environment if the regulators will be able to create an effective policy, able to improve the performance or force the exit of those firms from the market.

2. CHAPTER TWO

Bankruptcy prediction models

2.1 A comparison of the existent forecast bankruptcy models

The economists tried in the past to provide some models in order to forecast whether a firm will fail in the future according with the present information.

The main methods that are going to be analysed in this chapter are:

- The market-based bankruptcy prediction model, that it is based on the application of the Black and Scholes and the Merton models on the future trend of the company;
- The accounting-based bankruptcy model, which uses a big number of accounting ratios to forecast the future of the company. Clearly, since the balance sheets and the income statements change over time, the analysis must be done on several years in order to identify a trend;
- The macroeconomic-based bankruptcy model, which rely on the use interest rate of short-term bills (for example the one-year treasury bills) and the inflation, related to a weak macroeconomic environment and hence connected positively to a financial distress condition.

The last model is the less effective one, according with Pham, et al. (2018) and Hernandez T. & Wilson (2013) but its variables can be combined with the other 2 models to improve their efficiency.

In the doctrine, many researchers have discussed about which model could be deemed the best among the market-based and the accounting-based models but the authors that tried to provide an answer to this question came out with different answers. Up to now, a final judgement hasn't been found. This is because these 2 models analysed in this paper have warts and all.

Agarwal & Taffler (2008) tried to go into deep of "the very nature of the accounting statements on which these models are based casts doubt on their validity: (i) accounting statements present past performance of a firm and may or may not be informative in predicting the future, (ii) conservatism and historical cost accounting mean that the true asset values may be very different from the recorded book values, (iii) accounting numbers are subject to manipulation by management, and in addition, (iv) Hillegeist, et al. (2004) argue that since the accounting statements are prepared on a going-concern basis, they are, by design, of limited utility in predicting bankruptcy."

"The market-based model is appealing on several grounds. First, the timeliness of corporate bankruptcy predictions may be increased exponentially by combining market-based variables. Second, the volatility of market-based variables is calculated directly using a market index to enhance the power of indicators of default risk. The fluctuation plays a key role in default prediction. Third, information from financial and other statements are not part of accounting statements, which generally reflect the market price. Fourth, the market price is likely to be more suitable for default prediction because it reflects forward-looking information or future expectations of cash flow, whereas the accounting-based model reveals only backward-looking or past performance." (Pham, et al., 2018). Hillegeist, et al. (2004) resumes all of that writing that on average information coming from the stock markets has a better quality.

On the other side, the market-based model relies on a series of assumption about the forecast that not always are realized in the reality, like the normal distribution of the stock returns or the ownership of only zero-coupon bonds by the examined companies. Furthermore, it requires some measures of the volatility and asset value that are hardly achievable.

According with Pham, et al. (2018), the accounting-based model has a better forecast accuracy than the market-based approach, even when powered by the variables of the macroeconomic approach. Hernandez T. & Wilson (2013) Instead, following the trend of Hillegeist, et al. (2004), found out that in their sample that the market variables have a greater explanatory power than the accounting ones (even if we combine these ones with the macroeconomic variables), because the first model add information not contained in the financial reporting.

Instead, Agarwal & Taffler (2008) reported that the accounting-based and the market-based models have the same predictive power. Reisz & Perlich (2004) found that the ratios-based model is slightly better in the years before the bankruptcy while the model based on the models of Merton (1974) and Black & Scholes (1973) in the years previous the bankruptcy, while, if the 5-10 years before the bankruptcy are considered, the second model provides better results.

Even if a part of the doctrine criticizes the first model, the results don't show that a model is more reliable than the other. Depending on the sample, the results differ and there is not a univocal answer. Indeed, the ratios-based approach has, according to (Agarwal & Taffler, 2008), 3 main foundations on which to build:

- The failure is not something that happens suddenly, but it is that can be noted in the annual reports of the previous years and so the model should report it;
- The update in the accounting policy never changed dramatically the values of the ratios and this guarantees a continuity in the time;

- Loan covenants, elements very interesting from the point of view of the creditors (that are among the main beneficiaries of these models) are generally present in the annual report and, so, they influence the accounting-based ratios.

Provided that the past literature was not able to provide a model with a better performance compared with the others and that neither those models that mixed variables coming from both the accounting-based and the market-based approaches over performed the previous 2 models, we can feel free to choose the one we prefer in order to forecast the failure or not of the firms in the sample. Since the motivation is still not strong enough, it must be anticipated that the sample considered includes only firms that are not listed. The consequence is that the approach based on the works of Merton and Black and Scholes cannot work, since a fair stock price of our test cases cannot be provided.

2.2 Accounting based-model

2.2.1 Introduction to the accounting based-model

In order to forecast a firm bankruptcy using the accounting-based approach, the scholar should collect a number of ratios and items coming from the income statement and the balance sheet of the companies in the sample. The user is interested in detecting whether the firm will fail. The year before bankruptcy is the one where there is the strongest deterioration of the ratios and so it is easier to detect unhealthy firms but limiting the observations at just one year it is not useful because in that case the firm often is already irretrievably destined to go bankruptcy. The firm is supposed to be already in a dramatic situation both from the point of view of the managers, that can't improve the performance of the firm, and of the banks, that will see their loan impaired soon. Because of this, it is more interesting to perform the same analysis in a certain amount of years before the bankruptcy event. Usually, the chosen number of years applied by the doctrine is 5. When also further years have been considered, the models didn't forecast with a satisfying accuracy if the firm would have failed in the following years.

2.2.2 Altman Z score

The pioneer and the most known researcher in the construction of an accounting-based model aimed at predicting the failure of a company is Edward J. Altman, who, at the end of the '60s, with its study that was published under the name of "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy", started a line of research that led to many further models and reviews.

He considered a sample of 66 firms. Half of them failed, the other half didn't. In order to forecast the future condition of that company, he applied a model that was based on a MDA (multiple discriminant analysis). This is a method, used mainly to classify or make prediction when the output is a qualitative variable (i.e. male or female) that classified a sample is groups according to the data points individual characteristics. The scholar should create groups and divide all the data points according to them. MDA derives the best linear combination of the qualities of each group in order to get some discriminant criteria. If a certain entity (like a company) analysed has a series of characteristics (a series of ratios) present in all the objects of the group, MDA creates a series of coefficients according with the characteristics influence the output and, on the base of the values attributed to each data point, it will end in a certain group. A peculiarity of this analysis is that it considers also the interactions of the data points qualities in the allocation to a certain category. This kind of analysis has as response, in the case of Altman, a dummy where the alternatives were "Bankrupt" and "Non-Bankrupt". A firm is considered destined to the bankruptcy if the likelihood of being sound doesn't exceed a certain threshold, otherwise not.

To reach the final result, he started using a huge number of ratios, related to the measurement of the liquidity, profitability, leverage, solvency and activity situation of the subjects of the sample. The final model, in order to catch all the aspects of the firm, was supposed to include those ratios which could catch every aspect of the company.

The criteria used by Altman to reach his final model have been:

- Weighting of the significance of many created functions, as well as the contribution of each single variable;
- Analysis of possible correlations between the independent variables;
- Evaluation of the success matrix of each model;
- Personal judgement of the author.

The best function found by Altman was:

$$Z = 0.012X_1 + 0.014X_2 + 0.033X_3 + 0.006X_4 + 0.999X_5$$

where the variables were the following:

- Working capital/Total assets (X_1) is a popular measure of liquidity because it compares the net liquid assets with the capitalization. If the firm expects operating losses, the net working capital probably will decrease.

- Retained Earnings/Total assets (X_2) is a measure of cumulative profitability since it reflects what was gained in the previous years. Clearly, the age of the firm is implicitly included in this ratio. The consequence is that a discrimination is performed between the young and the older firms. Despite this, also in the real world there is a discrimination among firms that have a different age and the younger ones are more likely to fail;
- Earnings before interest and taxes/Total assets (X_3) is a measure of the operating profitability of the firm (caught by the nominator) since the interests and the taxes are excluded and, thanks to the denominator, the productivity is associated to the company amount of assets.
- Market value equity/Book value of total debt (X_4) compares the equity and the total debt and it helps in recognizing which value the equity can reach before the firm becomes insolvent.
- Sales/Total assets (X_5) is a financial ratio that show the sales generating ability of the company depending on the assets. Its importance is due to the relationship with the other variables. This ratio strongly depends on the industry where the firms operates.

X_3, that measures the profitability of the companies, is the most important one since the earnings are the most important source of financing in a firm. Notwithstanding this, even if a firm is having negative results, it can repay its debt with the liquidity left from the previous years. Surprisingly, as it was mentioned before, also X_5, due to the relation with the other variables, and especially with X_3, had a great impact (table 1).

Table 1 - Relative contribution of the variables in the Altman model (Altman, 1968)

Variable	Scaled Vector	Ranking
X_1	3.29	5
X_2	6.04	4
X_3	9.89	1
X_4	7.42	3
X_5	8.41	2

In the first year before the bankruptcy, the prediction accuracy of the model corresponds to 95%, while it decreases to 72% 2 years before the failure. In the third, fourth and fifth years the ratios are not significant since their accuracy is under 50% (respectively 48%, 29% and 36% - table 2). An explanation to these results is that many firms were not already in a dramatic situation some years before the bankruptcy and, at the same time, some companies hide voluntarily the negative items in order to show outside a sound image of the firm until they fail. Moreover, it must be noted that Altman didn't focus in the recognition of non-bankruptcy firms.

Table 2 - Five years predictive accuracy of the Altman's model (Altman, 1968)

Year Prior to Bankruptcy	Hits	Misses	Per cent Correct
1st n = 33	31	2	95
2nd n = 32	23	9	72
3rd n = 29	14	15	48
4th n = 28	8	20	29
5th n = 25	9	16	36

Since one of the main aims of the Altman's study was to provide a model that could have been used by the rating agencies or by the banks in order to judge the credit worthiness of an institution, a seed was planted but, given that the most interesting years in the decision of granting a debt are the 3rd, the 4th and the 5th years, the model still needed some improvements. This model still can be useful to those investors that don't keep the ownership of a certain share for a long period, but they try to gain in the short run. They can invest in the firms that are not supposed to go bankruptcy or they can short sell the ones that the stocks relative to those companies that are likely to fail.

2.2.3 Beaver model

Even if the works of Altman are probably the most well known in the prediction of failure through the study of the ratios, another scholar, Beaver, must be cited. In the 1966 he published a research that can be considered, together with the Altman's Z score, at the basis of the further studies that will be developed in this field.

Beaver wanted to extend the use of the accounting-based model, declaring that "The emphasis upon financial ratios does not imply that ratios are the only predictors of failure. The primary concern is not with predictors of failure per se but rather with financial ratios as predictors of important events - one of which is failure of the firm." (Beaver, 1966)

Moreover, he emphasized that a certain value of a ratio leads to different conclusions according with the sector (ex. a liquidity ratio of 2 can be positive in certain fields, while it is not in others). It follows that the firms in the sample must come from a heterogeneous number of sectors and the failed and non-failed firms belong to comparable businesses. Another factor that influences the way the results are read is the size of the firm. "If firms are viewed as aggregates of assets and if asset returns are less than perfectly correlated with one another, statistical formulae suggest that the variability of total return to the firm will increase less than proportionately to the size of the firm. The rate of return to the firm will become more stable as asset size increases. Empirical evidence indicates that the variability of rate of return does behave in this manner.

The implication is that larger firms are more solvent, even if the value of their ratios is the same as that of smaller firms." (Beaver, 1966)

The solution that was applied by Beaver and by many following scholars, was to create a sample with matching couples: for each failed company, a non-failed company with similar characteristics and coming from a similar sector were found. The results, contrarily to the expectations, showed that the residuals deriving from the use of test subjects coming from different industries or sectors were not great. The reduction in the accuracy of the predictive model was not overwhelming, even if a small reduction was noted. It follows that, it is suggested to use subjects from the same industry, even if this is not compulsory. Instead, choosing firms with similar assets dimensions is important because Beaver found out that the asset size of healthy firms was higher than the failed firms. The analysis indicated that the mean asset size of the non-failed firms was greater than that of the failed firms. Whether the ratios depend on the assets and if those assets already help in explain the output, the risk is that the predictive power of ratios would be overestimated.

Like Altman, from a big group of ratios (30), he selected a small number on the basis of the popularity in the literature and of their effectiveness in the previous studies. Moreover, the ratios were analysed considering the ability that they intrinsically have to generate cash flow, i.e. the best measure to create liquidity to pay back the debt and avoid the bankruptcy condition. As well as with Altman, the ratios were computed from the 1st to the 5th year before bankruptcy.

"Four concepts are important in drawing the relationship between the liquid-asset-flow model and the ratios. The first is the size of the reservoir itself. The second is the net liquid-asset flow from operations, which measures the net amount of liquid assets supplied to (or drained from) the reservoir by current operations. The third is the debt held by the firm and is one measure of the potential drain upon the reservoir. The fourth is the fund expenditures for operations and is the amount of liquid assets drained from the reservoir by operating expenditures." (Beaver, 1966). The consequence is that:

- The reservoir is negatively correlated with the bankruptcy likelihood;
- A relatively high number of liquid-assets from operations, i.e. cash flow is negatively correlated with the probability of being a laggard firm;
- A high debt is positively correlated with the bankruptcy likelihood;
- A high usage of liquidity for operative purposes is positively associated with failure probability.

Failed firms showed lower cash flows and liquid assets compared with the non-failed firms. This led to more difficulties in paying the debt, even if, on the other side, the amount of debt due by the bankruptcy firms is greater. The chosen ratios were:

1. Cash flow/total debt;
2. Net income/total assets;
3. Total debt/total assets;
4. Working capital/total assets;
5. Current ratio (current assets/current liabilities)
6. No-credit interval (defensive assets minus current liabilities to fund expenditures for operations, where "Defensive assets mean assets for which the effects of inflation and currency devaluation are minimized. This term refers, in particular, to real assets such as real estate and precious metals such as gold." (Martins, s.d.))

The 4th ratio is the same used by Altman, while the 2nd one is very close to X_3 (earnings before interest and taxes/Total assets), even if, in theory the one used by Altman, given that it focuses more on the operating activities, should be more faithful to the real situation of the company. Also, the 3rd ratio shows the relation between the total debt and the total assets and it is quite similar to the ratio the ratio used by Altman, named as X_4 (Market value equity/Book value of total debt). Beaver, instead, focused particularly on the ratios that derive from the cash flows. Indeed, the 1st ratio directly includes them, while the 4th one includes the working capital, which, on the basis of the change between one year and the previous one, influences the cash flow and the short-term liquidity. Also, the 5th ratio, which compares the current assets with the current liabilities it is a commonly used measure of the ability of the firm to repay the current debt using the liquid assets. Finally, the 6th ratio deals with the credit worthiness of paying of the firm and the safety of the creditors. Altman focuses more on the operative performance and this should be a good way, provided that this is the basis to measure the ability of a company of repaying the debt. Beaver, instead, taking 4 out of 6 ratios from only from the balance sheet, focuses much more on the reliability of the borrower at a certain moment in time, without looking at its history and focusing less on its future performance.

Beaver observed that the most useful ratio was the cash flow to total debt ratio, followed by the net income to total asset, total debt to total assets ratio, while the size of the reservoir is the less important one. His model predicted correctly the 87% of the sample in the first year, 79% in the second year, 77% in the third year, 76% in the fourth year and 78% in the fifth year before the failure. Its power is stronger than the model developed by Altman in the 3rd, 4th and 5th years before the failure. After the 5th year, the author admits that the error increases and that the model is not reliable anymore.
An alternative analysis of the bankruptcy probability is provided by the use of the odds. "Odds of an event happening is defined as the likelihood that an event will occur, expressed as a

proportion of the likelihood that the event will not occur. Therefore, if A is the probability of subjects affected and B is probability of subjects not affected, then odds = A /B." (Anon., s.d.) This method is used, for example, in the calculation of the probability in the horse races bets. If the odds are lower than one, it is more likely that the firm won't fail. If the odds are higher than one, the company is not likely to fail.

2.2.4 Ohlson model

After the studies published by Beaver and Altman, a huge number of scholars started focusing on the prediction bankruptcy models. Among these papers, surely the one that can be considered the main competitor of the Altman's Z score, is the O-Score developed by Ohlson (1980). His model was composed by the following ratios:

1. "SIZE = log(total assets/GNP price-level index). The index assumes a base value of 100 for 1968 (i.e. the beginning year of the research). Total assets are as reported in dollars. The index year is as of the year prior to the year of the balance sheet date. The procedure assures a real-time implementation of the model. The log transform has an important implication. Suppose two firms, A and B, have a balance sheet date in the same year, then the sign of $P_A – P_B$ is independent of the price-level index. (This will not follow unless the log transform is applied.) The latter is, of course, a desirable property.
2. TLTA = Total liabilities divided by total assets.
3. WCTA = Working capital divided by total assets.
4. CLCA = Current liabilities divided by current assets.
5. OENEG = One if total liabilities exceed total assets, zero otherwise.
6. NITA= Net income divided by total assets.
7. FUTL = Funds provided by operations divided by total liabilities.
8. INTWO = One if net income was negative for the last two years, zero otherwise.
9. CHIN = $(NI_t – Ni_{t-1})/(|NI_t| + |NI_{t-1}|)$, where NI_t is net income for the most recent period. The denominator acts as a level indicator. The variable is thus intended to measure change in net income." (Ohlson, 1980)

Ohlson used a logit model in order to create its prediction function. He created three different models using the above-mentioned ratios. “Model 1 predicts bankruptcy within one year; Model 2 predicts bankruptcy within two years, given that the company did not fail within the subsequent year; Model 3 predicts bankruptcy within one or two years”. (Ohlson, 1980)
The logit model uses a cut-off point equal to 0.038 in order to decide whether a firm will fail. With this model with the above-mentioned variables, the percent correctly predicted in the 1st year before was 96.12%, 95.55% and 92.84% for the Models 1, 2 and 3 respectively, as can be seen in table 3.

Table 3 - Ohlson models prediction results

	Variable									
	SIZE	*TLTA*	*WCTA*	*CLCA*	*NITA*	*FUTL*	*INTWO*	*OENEG*	*CHIN*	*CONST*
Model 1										
Estimates	−.407	6.03	−1.43	.0757	−2.37	−1.83	0.285	−1.72	−.521	−1.32
t-statistics	−3.78	6.61	−1.89	.761	−1.85	−2.36	.812	−2.450	−2.21	−.970
Model 2										
Estimates	−.519	4.76	−1.71	−.297	−2.74	−2.18	−.780	−1.98	.4218	1.84
t-statistics	−5.34	5.46	−1.78	−.733	−1.80	−2.73	−1.92	−2.42	2.10	1.38
Model 3										
Estimates	−.478	5.29	−.990	0.062	−4.62	−2.25	−.521	−1.91	.212	1.13
t-statistics	−6.23	7.72	−1.74	.738	−3.60	−3.42	−1.73	−3.11	1.30	1.15

	Likelihood Ratio Index	Percent Correctly Predicted
Model 1	0.8388	96.12
Model 2	0.7970	95.55
Model 3	0.719	92.84

The non-significant variables are INTWO, CLCA and WCTA. Instead, it is surprising is that the variable SIZE is significant in all the models. This is consistent with the work of (Horrigan, 1966), that found that in his research that both accounting data and financial ratios are important in the determination of ratings to be assigned to corporates. The amount of assets alone is enough to predict more than half of the bond ratings of the sample considered.

“[…] the four factors derived from financial statements which are statistically significant for purposes of assessing the probability of bankruptcy are: (i) size (SIZE); (ii) the financial structure as reflected by a measure of leverage (TLTA); (iii) some performance measure or combination of performance measures (NITA and/or FUTL); (iv) some measure(s) of current liquidity (WCTA or WCTA and CLCA jointly).” (Ohlson, 1980)

The different performance of the model applied by Ohlson, according with the author himself, is due to the fact that its sample is one of the few examples in literature coming from the ’70 and this particular historic period is also associate to a scenario not similar to the one of the ’50s or of the ’60s (and, even less, to the present time). Moreover, he deemed that the kind of ratios chosen didn’t influence too much the final result. Notwithstanding this, the presence of further variables, like the size of the test subjects, can increase the accuracy of the model. Finally, the use of the logit model against the MDA can lead to different results. In particular, "Strictly speaking a linear approach is only appropriate when the dispersion matrices of the groups are identical. When this is not the case, quadratic discriminant analysis may be more appropriate" (Moyer, 1977)

Comparing the ratios of the models of Altman and Ohlson other than the size, the ratios used by them are quite close. The factor OENEG, is quite interesting. It is obvious that it influences the bankruptcy probability, but it was never used in the previous models. The past earnings are considered in both the models, but in the case of Altman, they are considered simply as a cumulative measure, Ohlson, instead looks at the deterioration of the net income compared with the previous years. Since the distress condition can show up just in the last years, the second ratio is more interesting. Another interesting ratio is FUTL. This last ratio looks at the liquidity brought by the operating management of the company, excluding the investment and the financial part, and it is an important frame for the constant collection of new funds. The two model are much closer than how much they are with the model developed by Beaver.

2.2.5 Begley criticism

Further reviews and improvements will be mentioned in this thesis. Begley, et al. (1996) applied both the Altman and the Ohlson models to a group of firms operating in the '80s, finding out that the results achieved changed, led mostly by the leverage ratios. Indeed, in the '80s the economic conditions changed and an increase in the quantity of debt compared with equity was performed by many firms. Since this was common at the time, a certain leverage, that could have seen as a negative signal in the '60-'70, it was deemed normal in the '80 and it wasn't seen as a worrying signal. Begley, et al. (1996) reported that in his research a high level of debt didn't necessarily have effect on the default probability while the liquidity variables increased their importance in the re-estimated models, consistently with the higher focus of the firms in their cash flow level. These reasons confirm the fact that the original models had a great prediction accuracy in the period they were created but their performance worsened a lot in the recent periods, even when the coefficients are re-calculated.

In all the previous works, the misclassification errors can be conducted mainly to 2 groups. Since the names assigned to these mistakes of the model were very wide, henceforth called in the statistic way, namely:

- Type I error: misclassifying a bankrupt firm as non-bankrupt;

- Type II error: misclassifying a non-bankrupt firm as bankrupt.

The first kind of error is the most important from the point of view of the creditors/banks because it means that these entities, due to the error of the model, grant a loan to a borrower that is not likely to pay back the money received.

The second kind of error, instead, it is the one that mostly affects the community. Due to this error, borrowers worth of receiving a credit are not going to benefit of a loan. From the point of view of the creditors, instead, the damage is not significant as the one undergone by the type I errors because they only lose the interests that they could have collected if a loan was granted.

"Where the prior probability of membership in group i is q_i, the cost of misclassifying an entity as belonging to group j when it actually belongs to group i is c_{ij}. And where the objective is to minimize the expected total cost of misclassification, a critical value of z,

$$z_c = log \frac{q_2 c_{21}}{q_1 c_{12}}$$

should be employed." (Joy & Tollefson, 1975)

While the probabilities can be proxied but models like the logit, the cost of error I and, notably, the cost of error II are very difficult to estimate. Because of this, it is important to define if a research is made to reduce the losses incurred by the creditors or by the community.

An empirical study was performed to assess the costs of these lending errors with the following specification for the equivalent type I (C_1) and type II (C_{11}) error costs.

$$C_1 = 1 - \frac{LLR}{GLL}, \quad C_{11} = r - i$$

where: LLR = amount of loan losses recovered,

GLL = gross loan losses (charged-off),

r = effective interest rate on the loan,

i = effective opportunity cost for the bank.

This formula is interesting but at the same time there is the need to have data, that only banks have. In this thesis the error costs could be calculated in a different way but a bank, that has a greater data availability and is willing to do some assumption on the opportunity cost, can calculate C_1 and C_{11} in order to have a more precise and consistent model.

After this small parenthesis on the Type I and II errors, let's come back to the Begley's analysis, where a worsening of the Ohlson model can be seen in the figure 4, where the original cut-off point was 3.8%.

Table 4 - Comparison of the errors I and II – Ohlson's model 1 (Begley, et al., 1996)

Estimated Probability of Bankruptcy used as Cutoff Points	Ohlson's Sample		1980s Sample[a]	
	Type-I	Type-II	Type-I	Type-II
0.00	0.0%	100.0%	0.0%	100.0%
0.02	7.6	28.7	9.2	38.0
0.038	**12.4**	**17.4**	**10.8**	**26.6**
0.04	14.3	16.7	10.8	25.5
0.06	20.6	11.6	20.0	19.1
0.08	25.7	9.3	26.1	15.7
0.10	26.7	7.2	30.8	13.1
0.20	44.8	3.3	53.8	7.5
0.30	48.6	1.7	58.5	5.1
0.40	57.1	1.1	66.1	3.5
0.50	67.6	0.6	70.8	2.5
0.60	71.4	0.3	73.8	1.8
0.70	76.2	0.2	80.0	1.5
0.80	81.9	0.1	90.8	0.9
0.90	88.6	0.0	95.4	0.5
1.00	100.0%	0.0%	100.0%	0.0%

The change of conditions led the type I error decrease slightly (1.6%), while type 2 error increased by 9.2%. The total percentage error amounts to 14.9%.

Altman's model, instead changed from a 95.5% correct and 4.5% wrong predictions, to 78.2% correct and 21.8% wrong predictions (table 5).

Table 5 - Comparison of the errors I and II - Altman's model (Begley, et al., 1996)

Panel A: Altman's Original Sample[a]

Predictive Ability One Year Prior to Bankruptcy

	N	Number Correct	Percent Correct	Percent Error
Type-I[b]	33	31	93.9	6.1
Type-II[c]	33	32	97.0	3.0
Overall Error Rate			95.5%	4.5%

Panel B: 1980s Holdout Sample[e]

Predictive Ability One Year Prior to Bankruptcy

	N	Number Correct	Percent Correct	Percent Error
Type-I[b]	65	53	81.5	18.5
Type-II[c]	1300	974	74.6	25.1
Overall Error Rate			78.2%	21.8%

In order to try to fix the predictive ability of both the models, (Begley, et al., 1996) re-estimated them.

The re-estimated Altman's model, became:

$$Z = 0.012X_1 + 0.014X_2 + 0.033X_3 + 0.006X_4 + 0.999X_5$$

Despite the change, the forecast accuracy didn't variate. The type I error increased slightly, while the type II error decreased slightly, compensating each other (table 6).

Table 6 - predictive ability of the re-estimated Altman model (Begley, et al., 1996)

One Year Prior to Bankruptcy				
	N	Number Correct	Percent Correct	Percent Error
Type-I	65	51	78.5	21.5
Type-II	1300	1019	78.4	21.6
Overall Error Rate			78.4%	21.6%

Ohlson's model was re-calculated and the model can be appreciated in table 7.

WCTA (even if with a lower coefficient), CLCA and INTWO became significant with the new sample, while OENEG and CHIN lost their predictive power. It must be reminded that in the model CLCA and OENEG reflect the assets between the assets and the liabilities, while CHIN and INTWO reflect the relationship between the insolvency and the previous net incomes. Probably the increase of importance of one ratio of these couple, was compensated by the lower importance of the second one.

Table 7 - comparison between Ohlson old original and re-estimated models (Begley, et al., 1996)

Variable[a]	Ohson's Model	Re-estimated Model	Wald Test[b]
	Estimate	Estimate	(χ^2)
INTERCEPT	−1.320	−1.249	0.00
SIZE	−0.407	−0.211	5.13*
TLTA	6.030	2.262	41.21**
WCTA	−1.430	−3.451	7.57**
CLCA	0.076	−0.293	5.96*
OENEG	−1.720	−0.907	2.24
NITA	−2.370	1.080	21.69**
FUTL	−1.830	−0.838	11.29**
INTWO	0.285	1.266	12.52**
CHIN	−0.521	−0.960	3.62

The new model 1 (that measures the failure probability one year before bankruptcy) led to the error percentage reported in the table 8. The cut-off is chosen according with the point that registers the lower Type I and II errors, is 0.061 now. The combined error is 22.1%, still not satisfying.

Table 8 - errors I and II – re-estimated Ohlson's model (Begley, et al., 1996)

Cutoff Points	Type-I	Type-II
0.00	0.0%	100.0%
0.02	16.9	40.7
0.04	24.6	21.8
0.06	29.2	15.1
0.061	**29.2**	**14.9**
0.08	33.8	10.9
0.10	46.1	8.7
0.20	67.7	4.1
0.30	80.0	1.8
0.40	89.2	1.1
0.50	93.8	0.5
0.60	98.5	0.3
0.70	100.0	0.1
0.80	100.0	0.1
0.90	100.0	0.0
1.00	100.0%	0.0%

Begley, et al., (1996) Concluded that a review of the models didn't lead to any improvements and that the most reliable model is still the original model 1 developed by Ohlson. The reason we have gone so much into deep in the work of this last researcher is due to the fact that the first part of the empirical analysis of this paper is very close to the one of Begley, et al. (1996).

2.2.6 Altman reviews

Altman himself reviewed his model many times. For the purpose of this thesis, the most interesting model is the one published in 2000. He published other works where he tried to adjust his original model to face the change in the composition of the firms, in particular the size and the leverage.

The original model that was described above, was

$$Z = 0.012X_1 + 0.014X_2 + 0.033X_3 + 0.006X_4 + 0.999X_5$$

Where:

X1 = working capital/total assets

X2 = retained earnings/total assets

X3 = earnings before interest and taxes/total assets

X4 = MV of equity/BV of total liabilities

X5 = sales/total assets, and

Z = overall index.

The representation of the original formula was changed in the following years and it became: $Z=1.2X_1+1.4X_2+3.3X_3+0.6X_4+1.0X_5$. Simply the author did a stylistic choice, preferring the use of percentages in the writing of the formula and rounding the 0.999 to 1.

The new model was:

$$Z' = 0.717X_1 + 0.847X_2 + 3.107X_3 + 0.420X_4 + 0.998X_5$$

Where he kept the same variables of the original Z score (even if the coefficients changed), expect for X_4, whose market value of equity was substituted by the book value of equity. The coefficient for X_1 decreased from 1.2 to 0.7, but still the model looks quite similar to the original Z score. The actual variable that was modified, X_4, showed a coefficient change to 0.42 from 0.6001 and now it now has less of an impact on the Z-Score. X_3 and X_5 are virtually unchanged." (Altman, 2000)

The model was tested on a sample offered by a model published by Moody's in 2000, which used middle market firms, mainly private and the original sample amounted to 1600 test units, while the sample used by Altman amounted to only 66 firms. The results are reported in the table 9.

Table 9 - Revised Z score model: classification results (Altman, 2000)

Actual	**Classified**		
	Bankrupt	Nonbankrupt	Total
Bankrupt	30 (90.9%)	3 (9.1%)	33
Nonbankrupt	1 (3.0%)	32 (97.0%)	33

Note: Bankrupt group mean = 0.15; nonbankrupt group mean = 4.14.
Z'<1.21 = Zone I (no errors in bankruptcy classification):
Z'>2.90 = Zone II (no errors in nonbankruptcy classification):
gray area = 1.23 to 2.90.

The sensitivity (the percentage of true defaulters that are identified) is 91%, while the specificity (the percentage of non-defaulters that are correctly identified) is 97%. The results are satisfying from an absolute point of view but are discouraging if compared with the performance of the original Altman's model, where the sensitivity was 94% (table 10).

Table 10 - Classification and prediction accuracy of the Z score (1968) failure model (Altman, 2000)

Year Prior To Failure	**Original Sample (33)**	**Holdout Sample (25)**	**1969-1975 Predictive Sample (86)**	**1976-1995 Predictive Sample (110)**	**1997-1999 Predictive Sample (120)**
1	**94% (88%)**	**96% (92%)**	**82% (75%)**	**85% (78%)**	**94% (84%)**
2	**72%**	**80%**	**68%**	**75%**	**74%**
3	**48%**	-	-	-	-
4	**29%**	-	-	-	-
5	**36%**	-	-	-	-

*** Using 2.67 as cutoff score (1.81 cutoff accuracy in parenthesis)**

In this model, Altman tried to solve the critical points underlined by Ohlson (1980), Moyer (1977), Joy & Tollefson (1975). They criticized his model because it lacked an important assumption, which in this last model is considered. We remind that the critics dealt with the fact that a linear model is appropriate only when the dispersion matrices of the groups formed by the MDA division should be identical. Instead, if the variance-covariance matrices differ, a quadratic structure will perform better, given the advantage that the qualities of each group and the relationships with the other groups can be appreciated independently.

Because of this problem connected to the MDA, this assumption must be proved before applying the model. Actually, it is difficult to distinguish the group of failed to the non-failed firms. This is why the linear model can be appropriate.

In order to increase the efficiency of the new model, Altman created also an adaptation of the model designed to the non-manufacturers, where X_5 was eliminated. The resulting model was

$$Z" = 6.56X_1 + 3.26X_2 + 6.72X_3 + 1.05X_4$$

All the coefficients changed, because of the different kind of firms considered and the omission of a variable. “All of the coefficients for variables X_1 to X_4 are changed as are the group means and cut off scores. This particular model is also useful within an industry where the type of financing of assets differs greatly among firms and important adjustments, like lease capitalization, are not made.” (Altman, 2000)

An alternative model, named the 7-variable model, that was proved to be the most reliable in many conditions. The variables are described below:

“X_1: Return on assets, measured by the earnings before interest and taxes/total assets. […]

X_2: Stability of earnings, measured by a normalized measure of the standard error of estimate around a five to ten-year trend in X_1. Business risk is often expressed in terms of earnings fluctuations and this measure proved to be particularly effective. […]

X_3: Debt service, measured by the familiar interest coverage ratio, i.e., earnings before interest and taxes/total interest payments (including that amount imputed from the capitalized lease liability). We have transposed this measure by taking the log 10 in order to improve the normality and homoscedasticity of this measure.

X_4: Cumulative profitability, measured by the firm's retained earnings (balance sheet)/total assets. This ratio, which imputes such factors as the age of the firm, debt and dividend policy as well as its profitability record over time, […].

X_5: Liquidity, measured by the familiar current ratio. […] we now find it slightly more informative than others, such as the working capital/total assets ratio.

X_6: Capitalization, measured by common equity/total capital. […] the common equity is measured by a five-year average of the total market value, rather than book value. The denominator also includes preferred stock at liquidating value, long-term debt and capitalized leases. We have utilized a 5-year average to smooth out possible severe, temporary market fluctuations and to add a trend component (along with X_2 above) to the study.

X_7: Size, measured by the firms' total assets. This variable, as is the case with the others, was adjusted for financial reporting changes.” (Altman, 2000)

This model focuses a lot in the continuity and in the ability to generate liquidity, through the indices X_1, X_2 and X_4, while the linkage between its ability to repay and its net income is provided by X_3. The solvability of the firm is resumed by the index X_5, while the leverage is indirectly represented by X_6. Finally, following the model of (Ohlson, 1980), Altman added the size of the firm (X_7). The most significant ratio is X_4, followed by X_2. This is a confirmation that the ability of the firm to generate constantly liquidity is the main feature that determines the ability of the company to avoid the bankruptcy condition.

Table 11 - comparison of the prediction accuracy between the Z and the ZETA Altman's models (Altman, 2000)

Years prior to bankruptcy (1)	ZETA Model		Altman's 1968 Model		1968 Model, ZETA Sample		1968 Variables, ZETA parameters	
	Bankrupt (2)	Non-bankrupt (3)	Bankrupt (4)	Non-bankrupt (5)	Bankrupt (6)	Non-bankrupt (7)	Bankrupt (8)	Non-bankrupt (9)
1	96.2%	89.7%	93.9%	97.0%	86.8%	82.4%	92.5%	84.5%
2	84.9	93.1	71.9	93.9	83.0	89.3	83.0	86.2
3	74.5	91.4	48.3	n.a.	70.6	91.4	72.7	89.7
4	68.1	89.5	28.6	n.a.	61.7	86.0	57.5	83.0
5	69.8	82.1	36.0	n.a.	55.8	86.2	44.2	82.1

The 7-variable model, in the table named "ZETA model", is represented in the columns (2) and (3), while the original model is represented in the columns (4) and (5). The following columns apply a certain model on the sample of the other model – namely applies the original model to the ZETA sample and vice versa. If in the first year the original model performs better, in the models applied to the preceding years, the new formula worsens its performance but just slightly, while the old one after the second year, it is not reliable anymore. At the 5th year, the accuracy rate of the new model is 69.8%, against the 36% of the other one.

2.2.7 Further insights

The works of Altman, Ohlson and Beaver are at the basis of the bankruptcy prediction model discipline. Some of the criticises/improvements/observations to the model are already presented above, others are going to be showed below. A paper that could provide the biggest help to the development of future models is the one provided by Deakin (1972). He observed that "the failed firms tended to expand rapidly in the third and fourth years prior to failure. If we look at the capital structure, it seems that the expansion was financed by increased debt and preferred stock rather than common stock or retained earnings. Therefore, funds raised were invested in plant and equipment rather than in liquid assets. […] These firms were unable later to generate the sales and net income to support their heavier debt, and so they lost their assets rather rapidly after the third year prior to failure. At that point, their asset ratios and debt ratios tended to fall back in line with the ratios shown in the earlier study."

The consequence is that the leverage, used in all the models above mentioned, usually increased (when the expansion was not financed by preferred stocks) and, so, the ratio measuring the relative indebtedness, could become even more important. This effect is counterbalanced by the higher amount of debt held by firms in last decades.

Moyer (1977) also argued that the Altman's model parameters depend on the firms size of the sample and the time span considered. In his study, indeed, the results are worse compared to the ones collected by Altman (1968) and, in particular, the attribution to "Sales/total assets" as the 2nd most important variable is questioned the use of linear variables as well, while the use of quadratic variables could be more appropriate. Despite this, in practice, it doesn't provide better results.

Blum (1974), examining the accounting based ratios issued previously, found out that cash flow/total debt (used by Beaver) was one of the most important ratios, as well as net quick assets/inventory and trend inversion of the income amount. A further finding was using non-ratio variables, as Ohlson did, to improve the models. Moreover, he used in his analysis many ratios related to the inventory and what he found out was that inventories used to decrease when bankruptcy was close. Firms didn't fail because of the excessive accumulation of inventory.

Pongsatat, et al. (2004) tested both the Z-score and the O-score to small and large Thailand firms. From an overall point of view, the first model showed to be less powerful but none of the 2 models was deemed good enough to predict the bankruptcy of large firms.

El-Ansary & Bassam (2019), applying similar models to the ones of the literature to predict the bankruptcy failure of listed firms in Middle East and North Africa, concluded that distress firms showed low or negative earnings, actual and cumulated, and cash flow, as well as a WC deficit. Also, a smaller size is associated to fall in a distress situation and this fact is probably also due to the greater instability of the area considered.

A work very similar to the one done in this thesis is the one of Altman, et al. (2013), that tested the model proposed by Altman to Italian firms with at least 200 million income. The best model found out by the authors for their sample, made of different size firms (whose the majority was manufacturer) was the Z'', deemed better because Italian firms were manufacturer, quite connected to the local banking system, with a low capitalization and budget policies not transparent. Moreover, they based their work knowing that for non-US companies the Z'' score model was better.

2.2.8 Italian law situation

Before starting analysing the literature model, it should be underlined that in Italy a regulation in interception of financial distress situations has been developed in these years.

In Italy also from a legislative point of view, the use of indices to predict the bankruptcy of a company is being recognized as a valid method of health self-assessment of the firms.

In the recent meeting focused on the firm crisis that took place in Florence on the 25th and 26th October, organized by the Order of Chartered Accountants and Auditors (in Italy named "CNDCEC"), it was provided a new law enactment on the bankruptcy recognition and prevention.

In particular, for the purpose of this thesis, the most important part is the use of ratios as a way to understand better the firm difficulties. The art. 13 c. 2 of the crisis and insolvency code assigned to the CNDCEC the role of elaborate the necessary indices to complete the alarm system, which was introduced in the law system with the law n. 155/2017. This law, named the "reform of the firm crisis and insolvency regulation", it is located in a national framework aimed at restructuring firms in the early stage of their financial distress and it was implemented to align Italy to the recommendation of the European Commission dated 12 march 2014. The biggest change provided recently by the CNDCEC is that the distress situation can be caught in advance thanks in particular to the 3rd point of the following list.

The method below is aimed at recognizing a financial distress situation, not the bankruptcy probability of a firm, even if they are connected.

The firm health assessment is divided in the following steps:

1. Equity monitoring, defined as the item A) in the passive side of the balance sheet, net of the deduction of credits from shareholders), own shares and dividends approved but not accounted for;
2. Even if the equity is greater than 0, also the DSCR (Debt Service Coverage Ratio), that measures the ability of a firm to create, in the following months, enough income to cover the debt that must be paid during that period, must be calculated. The value of this index must be greater than 1. In this case, it means that the firm is able to sustain the debt at least for the following 6 months, otherwise not. There are 2 approaches to calculate it, according with the CDNCEC:
    a. At the denominator the next 6 months contractually agreed payments for the financial debt reimbursement and at the numerator the initial liquidity, from which must be summed up the liquidity input and subtracted the liquidity output;

   b. At the numerator the sum of next six months FCFO, of the initial liquidity and of the credit, while at the denominator the sum of the next 6 months debt interests, the social security and tax debt and the non-current debt against suppliers and other creditors.
3. Whenever the DSCR can't be calculated and if the calculation of equity and the presence of continuous failure in debt payments are not registered, five indices are used:
   a. Ratio between financial expenses and income (sustainability of financial expenses index);
   b. Ratio between equity and total debt (balance sheet adequacy index);
   c. Ratio between cash flow and total assets (conversion of assets into liquidity index);
   d. Ratio between the short-term assets and the short-term liabilities (liquidity index);
   e. Ratio between the social security and tax debt and total assets (social security and tax indebtedness ratio).

If all the 5 ratios show negative results, the firm is passing a period of financial distress. If less than 5 ratios are below the threshold, the situation is more or less worrying and it must be analyzed case by case. The threshold are showed in the table 12:

Table 12 - cut-offs by sector

	Financial expenses/revenues	Equity/total debt	ST liabilities/ST assets	Cash flow/total assets	Social security and tax debt/total assets
Agricolture, forestry and fishing	2,8%	9,4%	92,1%	0,3%	5,6%
Extraction; manifacture; energy and water production	3,0%	7,6%	93,7%	0,5%	4,9%
Energy and water trasmission; water supply; sewerage and waste management	2,6%	6,7%	84,2%	1,9%	6,5%
Buildings construction	3,8%	4,9%	108,0%	0,4%	3,8%
Civil engineering; specialized construction	2,8%	5,3%	101,1%	1,4%	5,3%
Aumotive trade; wholesale trade; energy/gas distribution	2,1%	6,3%	101,4%	0,6%	2,9%
Retail trade; bars and restaurants	1,5%	4,2%	89,8%	1,0%	7,8%
Transportation and storage; hotel	1,5%	4,1%	86,0%	1,4%	10,2%
Firms sevices	1,8%	5,2%	95,4%	1,7%	11,9%
People services	2,7%	2,3%	69,8%	0,5%	14,6%

Firms whose performance cannot be calculated through the use of indices can write the reasons in the notes to the financial statement. Special indices are addressed to innovative start-ups.

The choice method of these indices has been very similar to the one used for the model created by the thesis author, where the ratios were chosen as the best between many indices.

3. CHAPTER THREE

Literature models analysis

3.1 Introduction

The empirical part is aimed at recognizing which bankruptcy prediction literature model is the best to be applied to a sample of firms from Padova and Vicenza and finally, comparing the results from these models with a new model.

3.2 Data

A sample of 1101 non-listed firms were taken from the provinces of Padova and Vicenza, cities in the region of Veneto, in Italy.

In this sample, the ratio between failed and non-failed firms is 1 to 5. For each firm, 5 years of activity were considered. The years considered change firm by firm but the range of years considered is 2010-2017. For the non-failed firms the years are simply 5 years of activity, for the failed firms the 5 years considered are the ones before the bankruptcy declaration year. The firms have different sizes and come from different sector, in order to create a heterogeneous sample.

A con of the choice of a sample like this is that there are not matching couples (namely pairs of firms, 1 failed and 1 not, of similar size and belonging to the same sector), like in the samples used by Altman (1968 and 2000) or Ohlson (1980). The advantage of having matching couples is that it is easier to isolate the differences between healthy and laggard firms.

3.3 Ratios analysis

In this paragraph, we explain the ratios used. Only the ones that involved a certain discretion in the choice of the elements that constitute them will be mentioned. The ratios that involve simple elements from the balance sheet won't be deepened.

The first group of indices will be the one used in the Z, Z' and Z'' models in (Altman, 1968) and (Altman, 2000):

- Net working capital/Total assets (X_1). Altman named this ratio Working capital/Total assets but this definition can be misleading according with the definition provided by most of the books and websites. The author in this thesis will apply what the majority of the doctrine believes, i.e. "Working capital is sometimes used to refer only to current assets, while net working capital is defined to be the difference between current assets and current liabilities." (Anon., s.d.)
- Retained earnings/Total assets (X_2). The item that is closer to the retained earnings is "utile/perdita a nuovo (section AVIII of the passive section of the balance sheet)". The item retained earnings is present in the annual reports drawn up according with the international accounting principle. With that representation method is easier to analyze the cumulative profitability of the firm because the items of the equity are designed to be relevant, while the Italian accounting principles follow the criterion of reliability. Some reserves, especially the statutory and the legal ones, often contain a part of the cumulated earnings. Notwithstanding this, a part of the amount included in those reserves can come from other sources, like the shareholders. Because of the above mentioned reasons and because the small/medium firms are not used to keep a great amount of earnings with them, the importance of this ratio could decrease.
- Earnings before interests and taxes/Total assets (X_3). The first is a flow measure and the second is a stock measure. Because of this, since the flow measured depends on the assets of the firm, an average of the stock measure of t-1 and t is used. This method will be applied with all the following ratios created by the use of a flow and a stock measure.
- Book value of equity/Book value of total Debt (X_4). This ratio is not the original one used by (Altman, 1968), but the one used in the Z' model by (Altman, 2000). The fact that the firms considered are not listed, leads to the impossibility of achieving a market value of equity. For this reason, book value will be used in this thesis.
- Sales/Total assets (X_5). The sales considers only the item A1 of the income statement, namely the revenues from the sales of good and services.

The following group of indices is constituted by the one used by (Altman, 2000) to create the ZETA model.

- Debt service measured by the interest coverage ratio (EBIT/total interest payments, X_3). In the total interest payments also the financial income was included, because it compensates the need for liquid money to pay the financial expenses.

- The size (X_7) is only measured by the firms' total assets.
- The capitalization (X_6) and the stability of earnings (X_2) won't be used because Altman applied a 5 years average of the past values of the items considered. It means that, for example, for the 1st year considered, namely the one that is located 5 years before the bankruptcy, would need also the data coming from the 6th, 7th, 8th and 9th years before the bankruptcy.
- X_1 and X_4 have already been described, while X_5 is a normal current ratio (current assets/current liabilities).

My calculation of (Ohlson, 1980) indices are as follows:

- Log(total assets/GNP price-level index), named by the author "SIZE". It was used the natural logarithm in the calculation of the ratio, while the inflation considered for the denominator was taken from (Anon., s.d.). The starting point, whose value of the assets was not influenced by the inflation was 2009, while for the following years, up to 2017, the value of the assets was divided by the inflation of their year, multiplied by the inflation of the previous years. The inflation values are the following:
    1. Average inflation in 2010: 1.5%;
    2. Average inflation in 2011: 2.8%;
    3. Average inflation in 2012: 3.0%;
    4. Average inflation in 2013: 1.2%;
    5. Average inflation in 2014: 0.2%;
    6. Average inflation in 2015: 0.1%;
    7. Average inflation in 2016: -0.1%;
    8. Average inflation in 2017: 1.2%;
    9. Average inflation in 2018:1.1%.
- Fund provided by operations/total liabilities (FUTL). The nominator, also named FFO (funds from operations), "is the cash flows generated by the operations of a business. The term is most commonly used in relation to the cash flows from real estate investment trusts (REITs)." (Bragg, 2018)

  FFO, following the formula of (Bragg, 2018), is got starting from the net income, adding the amortization, the depreciation, the capital losses and the interest expense and subtracting the capital gains and the interest income. The formulas that can be found in the web are not equal. Some websites didn't include the interest income/expense, but the author preferred this one since in his opinion it isolates better the operating income of the firms.

- TLTA, CLCA, OENEG and NITA have already been explained and their formulation didn't need a high grade of discretion in their construction, while WCTA has been already seen in the Z model.

The further model is the one developed by (Beaver, 1966).

The ratios that need a little bit of more attention are the following:

- No credit interval: (defensive assets-current liabilities)/funds expenditure for operations. The defensive assets are defined as quick assets, namely the sum between the receivables and cash. The denominator is approximated by the COGS.
- Cash flow/total assets. Cash flow is defined by Beaver as cash flus marketable securities. The ratio will be calculated like the original one.

The data are taken from 2009 to 2017. As anticipated, since the firms were analysed only for 5 years, some firms where analysed since 2010 to 2014, others since 2012 to 2016 for example. The following table shows a preliminary observation on the net income of the firms analysed:

Table 13 – data points net income

Net income in last year considered	Number of firms
Loss	400
n.a.	43
€0	9
Profit – less than €8k	165
Profit – more thank €8k	484

The number of bankrupt firms is 189. The cut-off of €8k has been chosen arbitrarily because it is the minimum amount for an entrepreneur to live in cities like Padova or Vicenza.

In the analysis, we will start using a logistic model. Basically, the model is made by a regression of the independent variable "treated" on a series of dependant variables. The Z', Z'', Zeta models of Altman, as well as the ones of Ohlson and Beaver will be analysed. "Treated" is a dummy variable, equal to 0 didn't fail, 1 otherwise. All the firms with a variable "Treated" equal to 1 will fail one year after the 5 years considered in our sample. For example, if the years taken into consideration are 2013, 2014, 2015, 2016 and 2017, the firm will fail in 2018. A great number of researchers from the economic and statistical fields supports the logistic regression against the discriminant analysis (used by Altman). The main critics come from (Ohlson, 1980), (Moyer, 1977), (Joy & Tollefson, 1975).

The model has already been explained in the previous chapter. Basically, a cut-off point is selected to decide under which score a firm is considered by the model as non-failed. If the likelihood of failure is greater than the cut-off point, the model predicts that the firm will fail, otherwise not.

The first model considered is the Z' Altman score. The Z Altman score, as above mentioned, won't be analysed because it is a model designed for the big companies, whose market value of equity can be easily calculated. The firms of the following sample, being not-listed, have only the book value.

Each year will be analysed separately. The first one investigated will be the first year before bankruptcy, which is supposed to be the one with the most meaningful results.

3.4 Altman Z' score

3.4.1 Variables analysis

An initial analysis of the data points is provided to understand a little bit the sample. Among the variables, the one named "lag", that measures the year before bankruptcy considered, it is not included when the single variables will be considered but will be included when the full model is observed. The purpose is to understand the correlation with the other variables but then, not complicate too much the graphs.

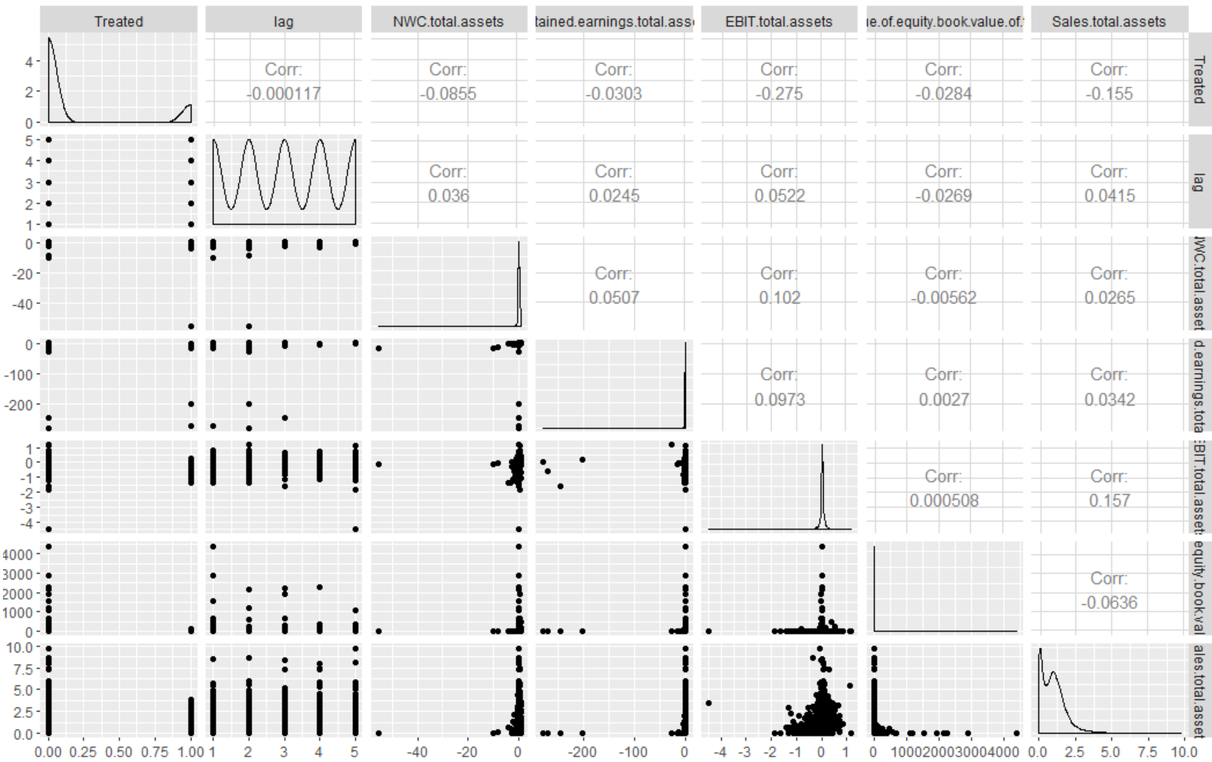

Figure 4 - Altman Z' score correlation matrix

As shown in the figure 4, the correlation between the variable “Treated” and the variables of the Z score of Altman is not really high, except for the variables “EBIT/total assets” and “Book value of equity/book value of debt”. The other interesting relation is the one between lag and the variables of the model.

Table 14 - multicollinearity Z' score of Altman

Variable	**VIF**
Lag	1.000
NWC/total assets.	1.01
Retained earnings/ total assets	1.01
EBIT/total assets	1.05
BV of equity/ BV of total debt	1.00
Sales/total assets	1.03

We can see that there is a small but not insignificant relation between all lag and all of the variables. The correlation between the ratios is very low. Also the table 14, that shows the multicollinearity between the variables, reports a value lower than 5 for each variable and this means that the multicollinearity is low.

The figures 5, 6 and 7 have been created in order to check if there are differences between the subgroups of the healthy and the non-healthy firms.

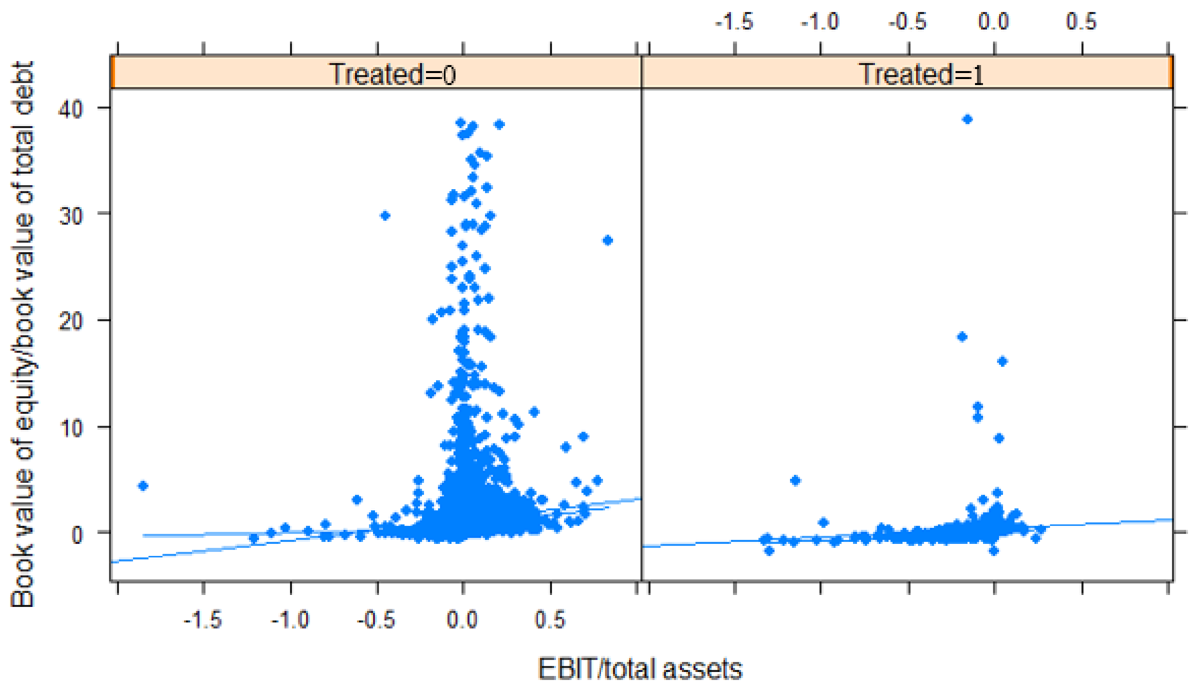

Figure 5 - Z' model variables: X_3 and X_4

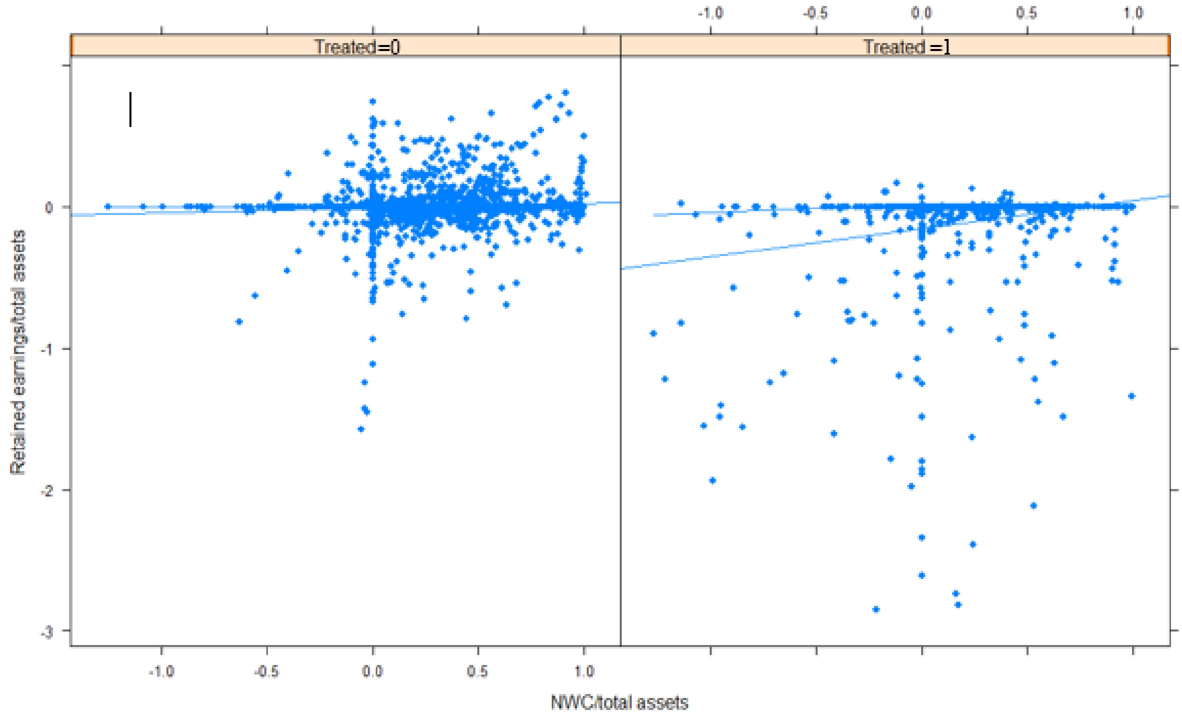

Figure 6 - Z' model variables: X_1 and X_2

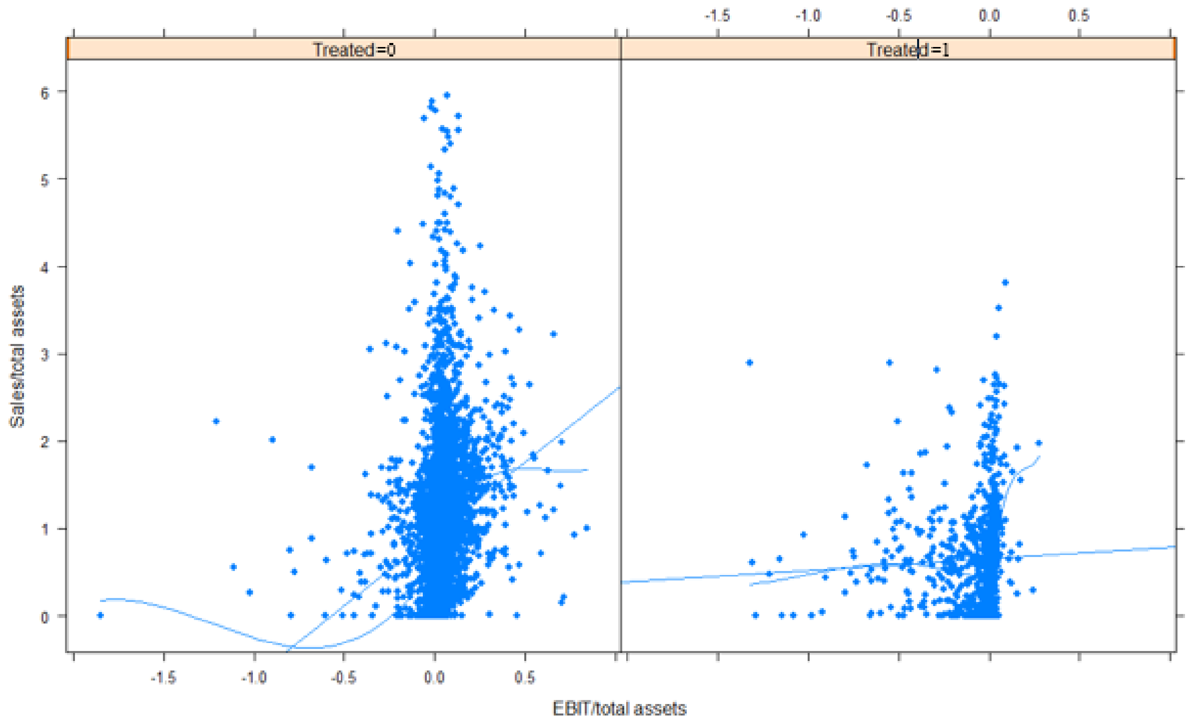

Figure 7 - Z' model variables: X_3 and X_5

The graphs above show that on average, non-bankrupt firms have a higher equity/debt ratio, sales/total assets ratio, higher retained earnings/total assets, a slightly higher EBIT/total assets ratio and NWC/total assets. Only for the 3 previous graphs 105 outliers were removed to improve their readability.

3.4.2 Missing values

With the purpose of understanding how to treat the missing values, their distribution and their weight are going to be estimated in table 15.

Table 15 - Z' model. Number of missing values

Ratio	Number of firms with n.a. values
NWC/total assets	78
Ret. earnings/total assets	78
EBIT/total assets	23
BV of equity/BV of total debt	88
Sales/total assets	23

The relative amount of missing data is quite small but not enough to be ignored. To better understand the dispersion of the errors, 105 outliers, responsible of decrease the readability of the error's dispersion, have been eliminated from figure 8.

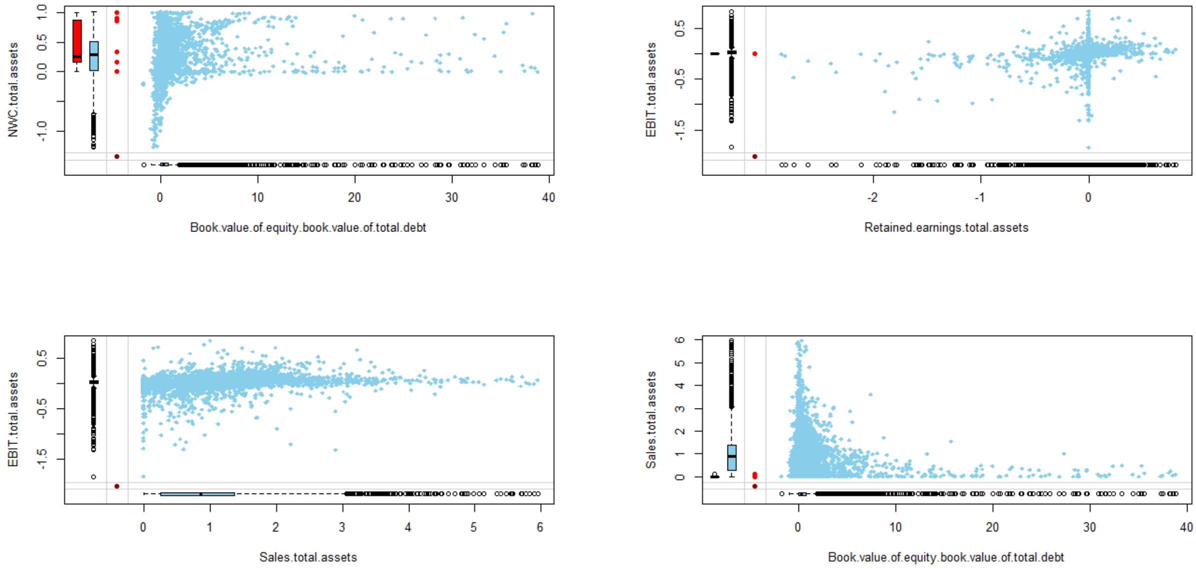

Figure 8 - Z' model. Location of missing values

These graphs allow us know where data points, whose variables are missing, could be located in the plots according with the other variable whose value doesn't miss. Where one of the 2 variables is missing, some red points can be observed at the margins of the plot at the height/length where the other variable is present. In particular, the most meaningful graphs are the 2 on the right. They show that almost all those firms with missing data are almost not active. We get it from the fact that "EBIT/total assets" and "Sales/total assets", but the 2nd variable of the graph doesn't, they are close to 0. Despite this, we can't consider them empty boxes since the NWC/total assets is quite distributed.

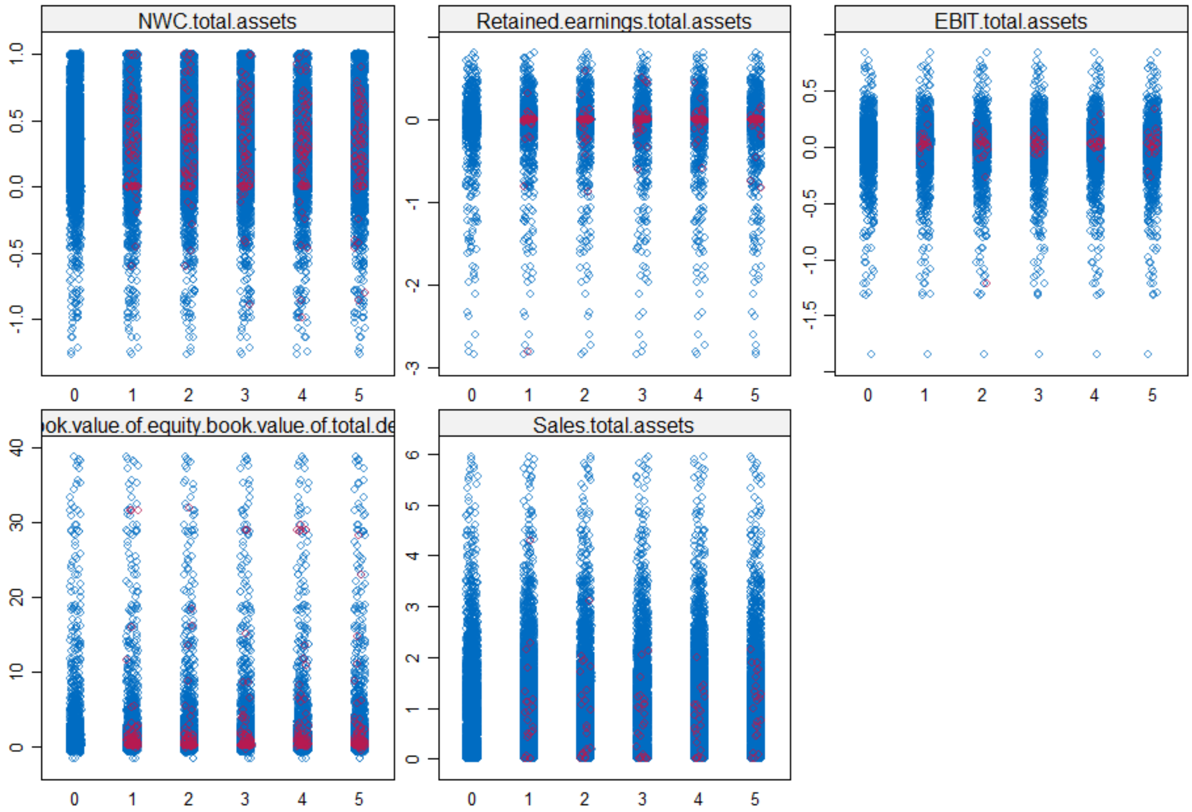

Figure 9 - Z' model. Estimated value of missing ratios

The figure 9 shows were the missing data point are likely to be according with the estimation of R based on the results achieved from the other variables. The hunch that the missing values are related to non-active firms is correct but not enough to assume that they are 0. The empty values represent just a small percentage of the sample. Assigning median or mean values to these empty spaces could create a bias, since, on average, the firms with empty spaces are expected to perform worse than the average firm. At the same time, assigning a value equal to 0 could fit for "Retained earnings/total assets" and for "EBIT/total assets" but for the others could be misleading because the range of values is wider. Instead of changing artificially the results and the variance, the empty values could be eliminated. Even if the risk is to achieve results that are slightly higher than the real ones, they don't influence the considerations at the end of the thesis.

3.4.3 Model test

The first step in the determination of the reliability of the Altman models is the check of the success rate to understand if it is still proper. That model is based on the following equation:

$$Z' = 0.717X_1 + 0.847X_2 + 3.107X_3 + 0.420X_4 + 0.998X_5,$$

where the variables X_1, X_2, X_3, X_4 and X_5 have been explained above. The variable Z' will be named Z' in the following graphs.

The analysis will start with the prediction ability of this model in the first year.

The Initial sample is equal to 1101. After having eliminated all the rows with empty space, a sample of 1058 firms remained.

The author tried to follow the original formula with the multivariate discriminant analysis method, but the results are not positive.

Following the setting of the Altman's formula, this model has a difference compared with the following models: if a firm is going to be bankrupt, the variable Treated will be equal to one, otherwise. In table 16, the application of the original Z score is shown.

Table 16 - confusion mattrix - 1 year before bankruptcy dataset

	0 real	**1 real**
0 predicted	0	0
1 predicted	155	903

It was also tried to apply this model with the sample of firms 5 years before the bankruptcy in table 17. The results are the same.

Table 17 - confusion matrix - 5 years before bankruptcy dataset

	0 real	**1 real**
0 predicted	0	0
1 predicted	189	902

Since the analysis of the sample with the pre-defined coefficients of Z' was a failure, the author will try to analyse them, still through the discriminant analysis method, but letting R calculating the most fitting coefficient values of the variables.

The following analysis splits up the sample in 2 groups, the train and the test groups. The first is composed by the first 4 years, while the second one by the 5th year. The coefficients are shown in the table 18.

Table 18 - Z' model. Coefficients with the 4 years before bankruptcy train dataset

Coefficients of linear discriminants	**LD1**
NWC/total assets	-0.27
Retained earnings/total assets	-0.007
EBIT/total assets	-5.205
BV of equity/BV of total debt	-0.002
Sales/total assets	-0.56

The equation that comes from the overall analysis is:

$$Z' = -0.27 * X_1 - 0.007 * X_2 - 5.205 * X_3 - 0.002 * X_4 - 0.56 * X_5$$

The success matrix of the train dataset is represented in table 19, while the success matrix of the test dataset in table 20.

Table 19 - Z' model confusion matrix, 4 years before bankruptcy train dataset. Train.

	0 real	**1 real**
0 predicted	3609	700
1 predicted	19	32

Table 20 - Z' model confusion matrix, 4 years before bankruptcy train dataset. Test.

	0 real	**1 real**
0 predicted	900	126
1 predicted	3	29

And the percentage of correct predictions is 87.7% in the train sample and 83.5% in the test sample. The results are not satisfying since the model fails in classifying the laggard firms. Table 21 represents the percentage success with that equation. The % of correctly classified data points is quite high but only because the proportion between non-bankrupt firms and bankrupt firms is very high. The most worrying column is the one related to the sensitivity, too low to make the model acceptable.

Table 21 - Z' score, 4 years before bankruptcy train dataset. Success table

	% correctly classified	**Sensitivity**	**Specificity**	**I type error %**	**II type error %**
1 year before bankruptcy	87.8%	18.7%	99.7%	81.3%	0.3%
2 years before bankruptcy	85.5%	13.3%	99.9%	86.7%	0.1%
3 years before bankruptcy	83.2%	3.3%	99.3%	96.7%	0.7%
4 years before bankruptcy	82.8%	1.1%	99.4%	98.9%	0.6%
5 years before bankruptcy	82.6%	0%	99.2%	100%	0.8%

An alternative to the previous equation is to get a formula from the sample with data coming from 1 year before bankruptcy and then apply that equation to the other years. The intuition is that it is more evident in that year if the firm is going to fail or not and so, the Z scores can be more meaningful.

The coefficients are represented in the table 22.

Table 22 - Z' model. Coefficients with the 1st years before bankruptcy train dataset

Coefficients of linear discriminants	LD1	Group 1 mean	Group 0 mean
NWC/total assets	-0.262	-0.013	-0.262
Retained earnings/total assets	-0.005	-2.173	-0.005
EBIT/total assets	-6.862	-0.245	-6.862
BV of equity/BV of total debt	-0.0004	-0.134	-0.0004
Sales/total assets	-0.178	-0.544	-0.178

The equation of the new Z' score is:

$$Z' = -0.262 * X_1 - 0.005 * X_2 - 6.862 * X_3 - 0.0004 * X_4 - 0.178 * X_5$$

The success rate is represented in table 23.

Table 23 - Z' score, 1 year before bankruptcy train dataset. Success table

	% correctly classified	Sensitivity	Specificity	I type error %	II type error %
1 year before bankruptcy	91.5%	47.7%	99.0%	52.3%	1.0%
2 years before bankruptcy	85.9%	21.1%	98.8%	78.9%	1.2%
3 years before bankruptcy	82.6%	5.4%	98.2%	94.6%	1.8%
4 years before bankruptcy	82.7%	2.7%	99.0%	97.3%	1.0%
5 years before bankruptcy	82.0%	2.2%	98.1%	97.8%	1.9%

Still, the sensitivity is too low. From the point of view of a bank, these analysis are useless. As last test, the previous sample was changed in the following way:

- If one or more ratios of a certain firm is/are not present, the raw relative to that year or years is/are deleted, as well as all the other raws with the name of that firm;

- The number of firms that face the bankruptcy and the ones that do not is the same.

A 146 failed firms and 146 non failed firms is made. This analysis is made to check if the sample matters and if the missing data and the the proportion between bankruptcy and non-bankruptcy firms determines a difference in the final result. The doubt comes from the fact that (Altman, 1968) used a sample where the number of the bankruptcy and the non-bankruptcy firms was the same. Coefficients are represented in Table 24.

Table 24 - Z' score model. 1 year before bankruptcy test dataset. Equal sample.

Coefficients of linear discriminants	**LD1**	**Group 1 mean**	**Group 0 mean**
NWC/total assets	-0.56	-0.023	0.328
Retained earnings/total assets	-0.001	2.283	-0.008
EBIT/total assets	-4.199	-0.253	-0.034
BV of equity/BV of total debt	-0.0005	-0.135	33.465
Sales/total assets	-0.61	0.536	-0.928

The % of correctly predited output (82.5%) is lower than in the previous analysises but in this case the % of the type I error smaller.

$$Z' = -0.5605 * X_1 - 0.0009 * X_2 - 4.1991 * X_3 - 0.0005 * X_4 - 0.6097 * X_5$$

The table 25, applying this equation, shows the success rate:

Table 25 - Z' score, 1 year before bankruptcy train dataset. Equal sample. Success table

	% correctly classified	**Sensitivity**	**Specificity**	**I type error %**	**II type error %**
1 year before bankruptcy	82.5%	77.4%	87.7%	22.6%	12.3%
2 years before bankruptcy	66.4%	43.2%	89.7%	56.8%	10.3%
3 years before bankruptcy	55.1%	17.8%	92.5%	82.2%	7.5%
4 years before bankruptcy	54.8%	15.8%	93.8%	84.2%	6.2%
5 years before bankruptcy	56.2%	17.1%	95.2%	82.9%	4.8%

With this model the confusion matrix gave results a little bit more interesting. As it was written above, the proportion of the dataset of (Altman, 1968) was reproduced and, indeed, the was an improve in the sensitivity.

Still, the problem hasn't been resolved yet because this dataset doesn't represent the reality, where the number of non-defaulted firms is much bigger than the defaulted ones. If we apply this model to a group of random firms, where the proportion of non-bankruptcy and bankrutpcy firms is high, we could get biased results and a very low success rate.

A last analysis will be done. Since the sample is unbalanced, the model tends to care more about reducing the II type errors, but the author is more interested in reducing the I type error. Because of that, still keeping the variables of the Altman model, a logit model will be applied, adding a further column that measures the burdersome of the errors. A weight equal to 1 is given to the type II errors, while a weight equal to 1/(number of defaulted firms/number of non-defaulted firms)) is given to the type I errors. Coefficients in table 26 are got.

Table 26 - Z' score model. 1 year before bankruptcy test dataset. Application of weights. Coefficients

Variable	**Estimate**	**Std. Error**	**Z value**	**Pr(>│z│)**
(Intercept)	0.0155	0.1208	0.128	0.898
NWC/Total assets	-0.1339	0.2100	-0.638	0.523
Retained earnings/Total assets	-0.9070	0.2006	-4.522	6-13e-06
EBIT/total assets	-15.2044	0.9210	-16.509	<2e-16
BV of equity/BV of total debt	-0.2597	0.5012	-5.185	2.16e-07
Sales/total assets	-1.2121	0.1298	-9.335	<2e-16

The equation is:

$$Z' = 0.155 - 0.13398 * X_1 - 0.90704 * X_2 - 15.2044 * X_3 - 0.25986 * X_4 - 1.2121 * X_5$$

And the success rate is shown in table 27.

Table 27 - Z' score, 1 year before bankruptcy train dataset. Weights applied. Success table.

	% correctly classified	**Sensitivity**	**Specificity**	**I type error %**	**II type error %**
1 year before bankruptcy	91.4%	85.2%	92.5%	14.8%	7.5%
2 years before bankruptcy	84.2%	49.4%	91.2%	50.6%	8.8%
3 years before bankruptcy	72.8%	19.1%	84.3%	80.9%	15.7%

4 years before bankruptcy	73.3%	17.3%	84.8%	82.7%	15.2%
5 years before bankruptcy	78.9%	15.3%	91.7%	84.7%	8.3%

The results are not satysfing yet because the sensitivity is still quite low in the 3rd, 4th and 5th years. Despite this, like in the analysis of the equal sample, the results improved. The model can be applied by somebody that is worried in the II type errors but not by an entity interested in I type error.

Using weights as explained before is a way to have a more successful model. In this case the weight of the first type errors was more or less 5, namely the proportion between non-bankrupt firms and bankrupt firms. If somebody else wants to use this model, an idea could be use the ratio in his/her territory of non-bankrupt firms to bankrupt firms in the time interval considered. For banks that rank the firms they lend money to in categories, it should be relatively easy to assign a weight based on the probability of failure of these firms. The con of this method is that if the weights are assigned in presence of a shock, the proportions could change and so the reliability of this method could decrease.

Partially satysfied by the previous model, in order to decrease the I type error, a new regression will be made starting from the 3rd year and not anymore from the 1st. The logic behind is that in the 3rd year the characteristics of a bankruptcy firm are fuzzier and, because of this, when a model that has like data train the 1st year before bankruptcy is used, it is difficult recognize a failed firm some years before because its situation was not that bad. Table 28 shows the new coefficients got and table 29 shows the success rate.

Table 28 - Z' score model. 3rd year before bankruptcy test dataset. Application of weights

Variable	**Estimate**	**Std. Error**	**Z value**	**Pr(>│z│)**
(Intercept)	1.2793	0.1119	11.440	<2e-16
NWC/Total assets	-0.6648	0.1848	-3.598	0.0003
Retained earnings/Total assets	0.0540	0.0208	2.596	0.009
EBIT/total assets	-4.7470	0.6454	-7.356	1.90e-13
BV of equity/BV of total debt	-1.1047	0.1039	-10.635	<2e-16
Sales/total assets	-0.6083	0.0772	-7.873	3.45e-15

$$Z' = 1.2799 - 0.6649 * X_1 + 0.05405 * X_2 - 4.74704 * X_3 - 1.1047 * X_4 - 0.60826 * X_5$$

Table 29 - Z' score, 3rd year before bankruptcy train dataset. Weights applied. Success table

	% correctly classified	Sensitivity	Specificity	I type error %	II type error %
1 year before bankruptcy	68.3%	95.5%	63.7%	4.5%	36.3%
2 years before bankruptcy	67.7%	83.9%	64.5%	16.1%	35.5%
3 years before bankruptcy	66.5%	75.0%	64.8%	25.0%	35.2%
4 years before bankruptcy	54.2%	54.2%	55.4%	51.9%	44.6%
5 years before bankruptcy	53.1%	53.1%	54.7%	55.2%	45.3%

This model has a higher sensitivity but still it is not enough to be used by an entity that wants to decide if to grant a credit with a reimbursement period of 4 years or more for example. On the other side, also the specificity is quite low and this makes the previous model not really interesting.

A limitation of this model is the presence of the weight: the assumption relative to that column is that the ratio between non-failed and failed firms is equal to 5. We know that this is true just in the artificial example that was created for this thesis. If somebody else wants to use the weights in his/her model, the proportion must be completely changed. Let's pretend that in a certain period this ratio is true: wheter there is a negative or a positive shock in the economy, the weights could change. If for example, the region considered is suffering a severe crisis and the ratio between non-failed and failed firms decreases to 2.5, the sensitivity decreases, up to reach almost 0 in the 5th year. It can be considered an approximation of the average of the results between the model with a weight equal to 5 and one with out the weight.

If, instead, the economy was living a period of strong expansion and, because of that, the ratio between non-failed and failed firms would increase to 10, the specificity would reach very low levels and would create another useless model. This test is going to be done. It can also be seen from another point of view as a "stress test" made by the banks to be more sure of the ability to pay of some potential borrowers.

Two tries were made by the author to check which model was better in order to increase the sensitivity, namely the ones that use respectively 1 and 3 years before bankruptcy as train set using a weight equal to 10. The 2nd model, whose coefficients are in table 30 and the success rate in table 31, showed to be the best.

Table 30 - Z' score model. 3rd year before bankruptcy test dataset. Application of a weight equal to 10. Coefficients.

Variable	Estimate	Std. Error	Z value	Pr(>│z│)
(Intercept)	2.0395	0.1020	19.987	<2e-16
NWC/Total assets	-0.7538	0.1684	-4.478	7.55e-06
Retained earnings/Total assets	0.0601	0.0200	3.000	0.0027
EBIT/total assets	-5.2372	0.6328	-8.277	<2e-16
BV of equity/BV of total debt	-1.1171	0.0860	-12.997	<2e-16
Sales/total assets	-0.6315	0.0665	-9.498	<2e-16

The equation used is:

$$Z' = 2.03925 - 0.7538 * X_1 - 0.06013 * X_2 - 5.2372 * X_3 - 1.11711 * X_4 - 0.6315 * X_5$$

The statistics are as follows:

Table 31 - Z' score, 3rd year before bankruptcy train dataset. Application of weights equal to 10. Success table

	% correctly classified	Sensitivity	Specificity	I type error %	II type error %
1 year before bankruptcy	50.9%	98.7%	42.6%	1.3%	57.4%
2 years before bankruptcy	49.1%	97.8%	39.4%	2.2%	60.6%
3 years before bankruptcy	49.6%	95.1%	40.4%	4.9%	59.6%
4 years before bankruptcy	48.4%	93.5%	39.2%	6.5%	60.8%
5 years before bankruptcy	44.5%	91.8%	35.0%	8.2%	65.0%

The model is still not good. If a coin were tossed, the probability to guess the right exit was the same. On the other side, this model can be used by those banks that don't want to grant money to insolvent firms. The sensitivity is always more than 90%. A possible idea that comes from this model is that different models could be used by different entities with different needs. A model that can satisfy everybody doesn't exit, according with this thesis.

The model of Altman can work only if the average bankruptcy firms percentage in a certain area is known and will consequently be considered. Moreover, the author believes that also the influence of the variables is not the same worldwide. Because of that, if some entities want to use the Altman model to predict the future of a firm, it should use some reviews of the model considering their characteristics. For example, italian firms are on average small, managed by the entrepreneur and with a limitied use of external capital.

Finally, a test on the difference of each ratio between a certain year and the previous one could be an interesting base to determine a model. Basically, the variables used won't be the ratios themselves but the difference between the ratio of the year t and t-1. The deterioration of the ratios could be a clue of the failure of a firm.

The variables "Delta book value of equity/book value of total debt" and "Delta sales/total assets" are not significant, so they have been removed. Coefficients are shown in table 32, while the success rate in table 33.

Table 32 – Delta Z' score, 1st year before bankruptcy train dataset. Weights applied. Coefficients.

Variable	**Estimate**	**Std. Error**	**Z value**	**Pr(>│z│)**
(Intercept)	-1.0493	0.0706	-14.860	<2e-16
Delta NWC/Total assets	-0.7871	0.3063	-2.569	0.0102
Delta Retained earnings/Total assets	-11.405	0.8317	-13.713	<2e-16
Delta EBIT/total assets	-8.5878	0.5865	-14.642	<2e-16

Where,

$$Y = -1.04929 - 0.78709 * X_1 - 11.40550 * X_2 - 8.58784 * X_3$$

The result got is:

Table 33 – Delta Z' score, 1st year before bankruptcy train dataset. Weights applied. Success table

	% correctly classified	**Sensitivity**	**Specificity**	**I type error %**	**II type error %**
1 year before bankruptcy	91.8%	72.4%	95.1%	27.6%	4.9%
2 years before bankruptcy	84%	35.4%	93.6%	64.6%	6.4%
3 years before bankruptcy	80.6%	13.7%	94.1%	86.3%	5.9%
4 years before bankruptcy	81.1%	10.6%	95.2%	89.4%	4.8%

Only the first year predicts correctly the output of the sample.
The author tried to do a further model, got with the difference between the values of the variables in the years t and t-1 and other 5 variables got with the interaction between the time and the delta (variable) but the try was unsuccessful.
The fact that these models are not reliable can be because 3 or 4 year before the bankruptcy the difficulties of the firms can be partially hidden by the accountants and only in the last year, when a recovery is difficult, the accounting numbers plummet.
Finally, some final observation will be written about the role of the ratios in the application of the Altman models.
It has been observed that the less important ratio is "Retained earnings/total assets". This ratio is not significative and its coefficient is very low. It can be connected to the fact that the majority of small and medium firms in Veneto is represented by firms where the manager is also the entrepreneur and that has as a unique income the earnings of the company. It follows that they withdraw constantly money, letting a small amount of retained earnings in the firm.
The other ratio that shows a small coefficient is "Book value of equity/book value of total debt". This can be due to the fact that the way of financing in a country like the USA, where the model was originally applied, compared with Italy is different. The USA are known to be more market dependant, while Italy is a hybrid. Moreover, the small firms in Veneto avoid as more as possible increase in the equity value, while they prefer using debt. At the same time, the book value of equity is less informative than the market value of debt.

3.5 Altman Z'' score

3.5.1 Model test

After having analyzed the Z' Atman score, the Z'' model will be taken into consideration. The hope that this model can bring more information comes from the positive results collected by (Altman, et al., 2013). The model is like the original Z score, but the last ratio (X_5 - Sales/total assets) is eliminated, in order to remove the differences among firms of different industries, with different assets turnover.
Since Z' was already analyzed, it is pointless to make again all the models. The analysis will be limited to a model that uses the first year before bankruptcy as data train and the previous 4 years as data test and one that has uses the 3rd year before bankruptcy as data train and the other

years as data test. Since the use of the weights showed in Z' an improvements of the accuracy, the weights directly applied also in this model. The author deems useless applying the original models or also the models without the weights. The model that uses the change in the values of the variables, i.e. the last model used in the Z' score analysis, to predict the failure of the model won't be used because also in Z' the variable "delta Sales/total assets" was not significant and, because of this, was neither applied in the other try. The consequence is that identical model to the previous model would be got.

The weighted model shows a big improvement compared with the previous one even if the success rate is not much higher than 50%.

It has characteristics very similar to the ones of the Z' score and the variable "NWC/total assets" is not statistically significant, as the following figure shows, even if it influences the model a little bit more than before. Weights will be applied to ponder the weight of the different kinds of error.

Table 34 – Z'' score, 1st year before bankruptcy train dataset. Weights applied. Coefficients.

Variable	**Estimate**	**Std. Error**	**Z value**	**Pr(>│z│)**
(Intercept)	-0.7334	0.0956	-7.671	1.70e-14
NWC/Total assets	-0.1455	0.2069	-0.703	0.482
Retained earnings/Total assets	-1.1574	0.2246	-5.154	2.55e-07
EBIT/total assets	-15.4400	0.9377	-16.465	<2e-16
BV of equity/BV of total debt	-0.2302	0.0494	-4.656	3.22e-06

The equation applied will be:

$$Z'' = -0.73338 - 0.14547 * X1 - 1.15739 * X2 - 15.43991 * X3 - 0.23018 * X4$$

Table 35 contains the success rate:

Table 35 - Delta Z'' score, 3rd year before bankruptcy train dataset. Weights applied. Success table

	% correctly classified	**Sensitivity**	**Specificity**	**I type error %**	**II type error %**
1 year before bankruptcy	92.3%	80.6%	94.4%	19.4%	5.6%
2 years before bankruptcy	83.6%	42.8%	92.1%	57.2%	7.9%
3 years before bankruptcy	80.7%	20.1%	93.1%	79.9%	6.9%

4 years before bankruptcy	80.4%	15.1%	93.8%	84.9%	6.2%
5 years before bankruptcy	79.4%	13.1%	92.9%	86.9%	7.1%

The model has a lower predicted power compared with the Z' score so, the results of this thesis contradict the ones of (Altman, et al., 2013), that found out that the Z'' score was the best for their sample.

The 2nd model, as was already anticipated, will be equal to the previous one but it will be made using as train dataset the 3rd year before bankruptcy.

Table 36 - Z'' score, 3rd year before bankruptcy train dataset. Weights applied. Coefficients.

Variable	**Estimate**	**Std. Error**	**Z value**	**Pr(>│z│)**
(Intercept)	0.6999	0.0844	8.288	<2e-16
NWC/Total assets	-0.6070	0.1797	-3.377	0.0007
Retained earnings/Total assets	0.0583	0.0185	3.156	0.0016
EBIT/total assets	-5.9458	0.7102	-8.372	<2e-16
BV of equity/BV of total debt	-1.0637	0.1087	-9.791	<2e-16

All the variables, contrarily to the model with the 1st year before bankruptcy as train data, are significant. The equation is the following:

$$Z'' = 0.69984 - 0.60704 * X_1 + 0.05829 * X_2 - 5.94577 * X_3 - 1.06368 * X_4$$

Table 37 - Delta Z'' score, 3rd year before bankruptcy train dataset. Weights applied. Success table

	% correctly classified	**Sensitivity**	**Specificity**	**I type error %**	**II type error %**
1 year before bankruptcy	64.4%	91.6%	59.7%	8.4%	40.8%
2 years before bankruptcy	64.0%	82.8%	60.2%	17.2%	39.8%
3 years before bankruptcy	63.5%	76.1%	60.9%	23.9%	39.1%
4 years before bankruptcy	63.2%	81.0%	52.9%	19.0%	47.1%
5 years before bankruptcy	66.1%	88.0%	54.8%	12.0%	45.2%

The results of this model are, surprisingly, much better than the version of this model got using the Z'. Probably the influence of the ratios NWC/total assets and Retained earnings/total assets, both non-significant in the previous model but significant here, influenced the output more than the variables Sales/total assets in the Z' model. The sensitivity is very satisfying, but the specificity is close to 50%. The risk is that the costs of not granting debts to worthy people could be higher than the costs of non-paying of the debt, especially considering the fact that the non-failed firms are 5 times the failed firms.

3.6 ZETA model

3.6.1 Variables analysis

(Altman, 2000) showed a new model, with 7 new variables, explained before. As was previously mentioned, the capitalization and the stability of earnings variables will be removed, due to the impossibility of the author to calculate them.

The analysis of the model will start checking the correlations of the variables, with the output and the other input in Figure 10.

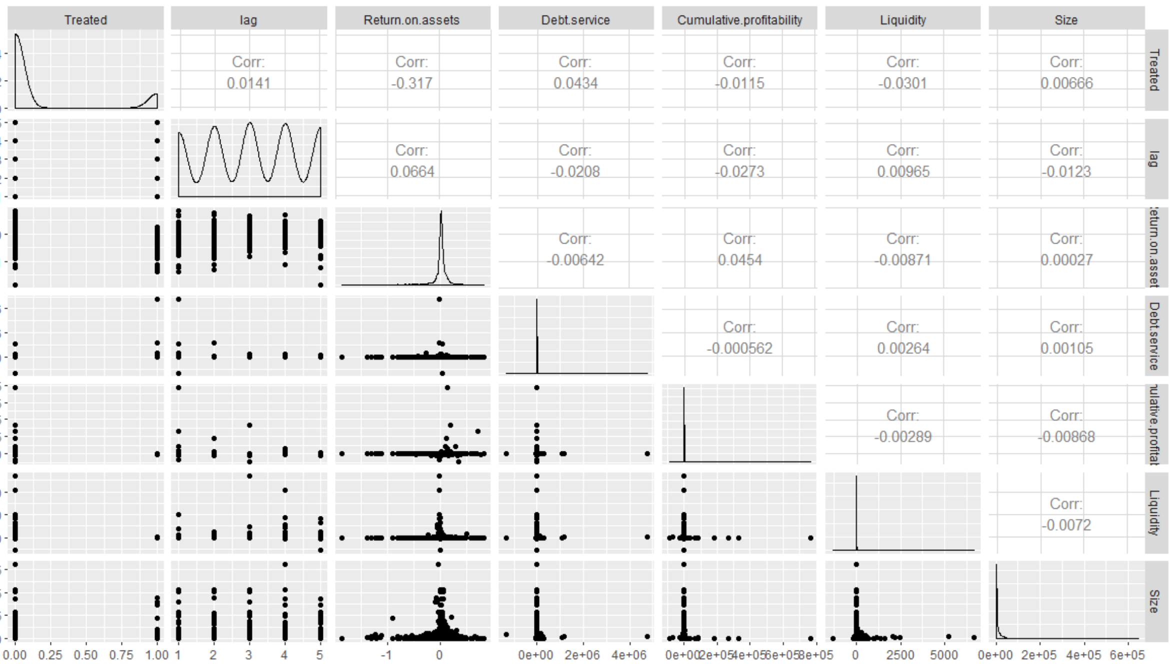

Figure 10 - Altman ZETA model correlation matrix

The variables with a greater correlation with the output is the Return on Assets. This result is very similar to the Altman Z' score, where the indicator measuring the performance of the firm in a certain year is the most explicative.

Table 38 - multicollinearity Altman ZETA model

Variable	VIF
Lag	1.02
Debt service	1.00
Liquidity	1.00
Return on assets	1.03
Cumulative profitability	1.00
Size	1.00

The VIF function is close to 1. It means that there is no multicollinearity.

The next graphs, instead, are made to check if a difference between the there is a difference in the sample between those rows that have "Treated=0" and "Treated=1". 57 outliers were removed to make the following graphs.

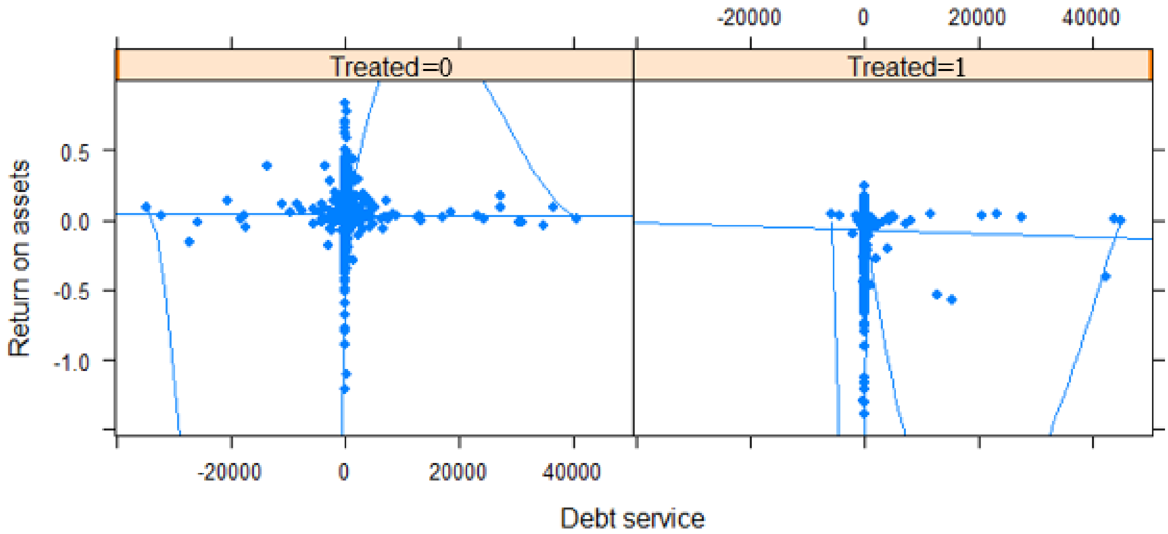

Figure 11 - ZETA model variables: Return on assets and debt service

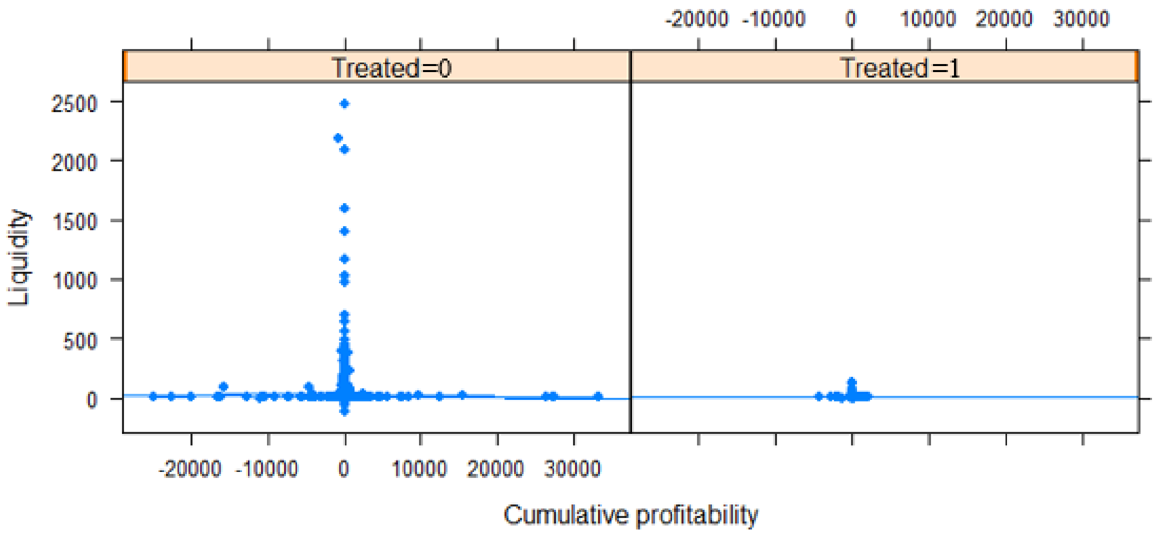

Figure 12 – ZETA model variables: Liquidity and debt service

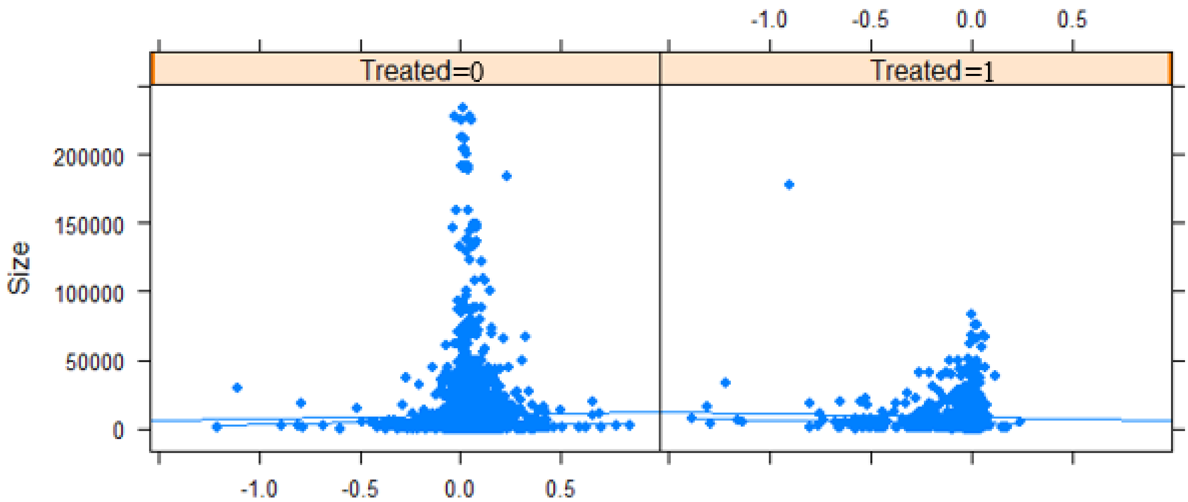

Figure 13- ZETA model variables: Size and Return on assets

The variables whose value is clearly different between failed and non-failed firms are the size, the return on assets and the liquidity. The debt service doesn't look to influence too much the output, while the cumulative profitability has more variance in the non-failed than in the failed firms. This is not expected because some non-failed firms with a cumulated profitability that is much lower than the failed firms, won't fail, while the failed firms show an accumulated profitability close to 0.

3.6.2 Missing values

The 2nd analysis performed is the one of the missing values, to understand if a solution like Altman can be developed or other alternatives should be preferred. First of all, the number of missing values is checked through table 39.

Table 39 - ZETA model. Number of missing values

Ratio	Number of firms with n.a. ratio
Return on assets	23
Debt service	293
Cumulative profitability	130
Liquidity	869
Size	0

The empty values are many more than in the case of the Z' model. Indeed, the missing liquidity ratios are 869, but also the missing debt service empty values are not negligible.

A series a graphs are made to better understand where the missing points are located. In order to improve the comprehension of the graphs, 57 outliers have been removed.

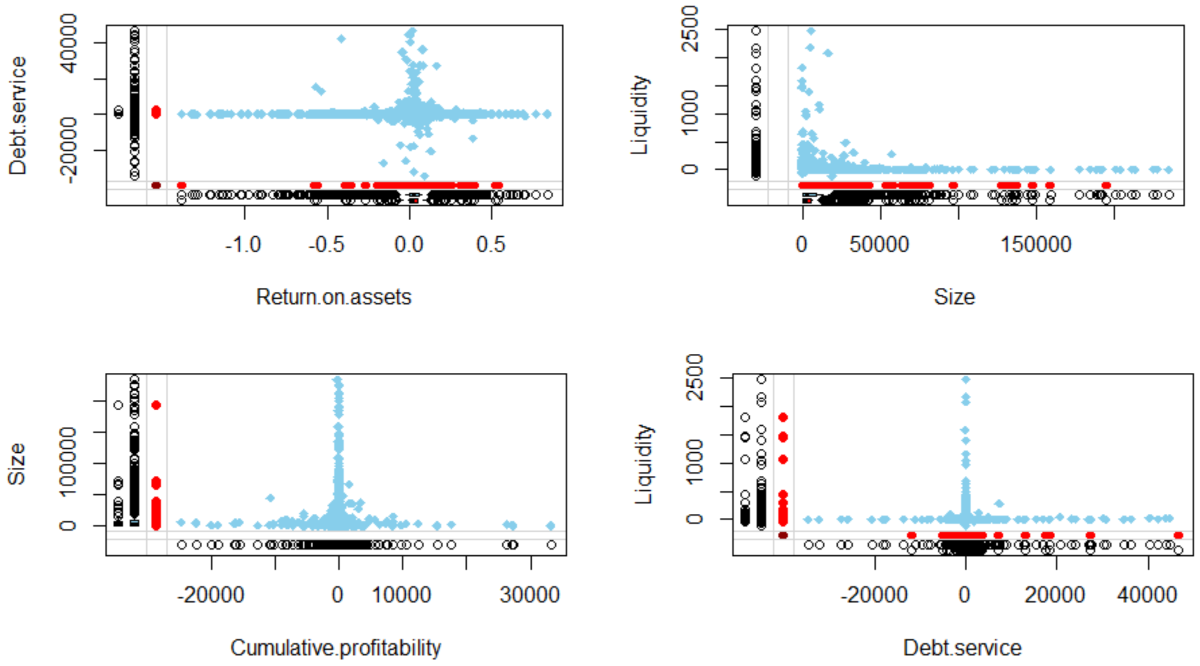

Figure 14 - Altman ZETA model. Location of missing values

The red points appear only when a variable out of 2 is present and they show the place where that data point would have stayed if the other variables hadn't been empty. As with the Altman model, the red points are distributed around the graph. It means that these firms are active and not empty boxes. Also the graph below predicts that, even if the ratios are quite low, they are not 0 and, because of this, the author will delete le rows with empty cells.

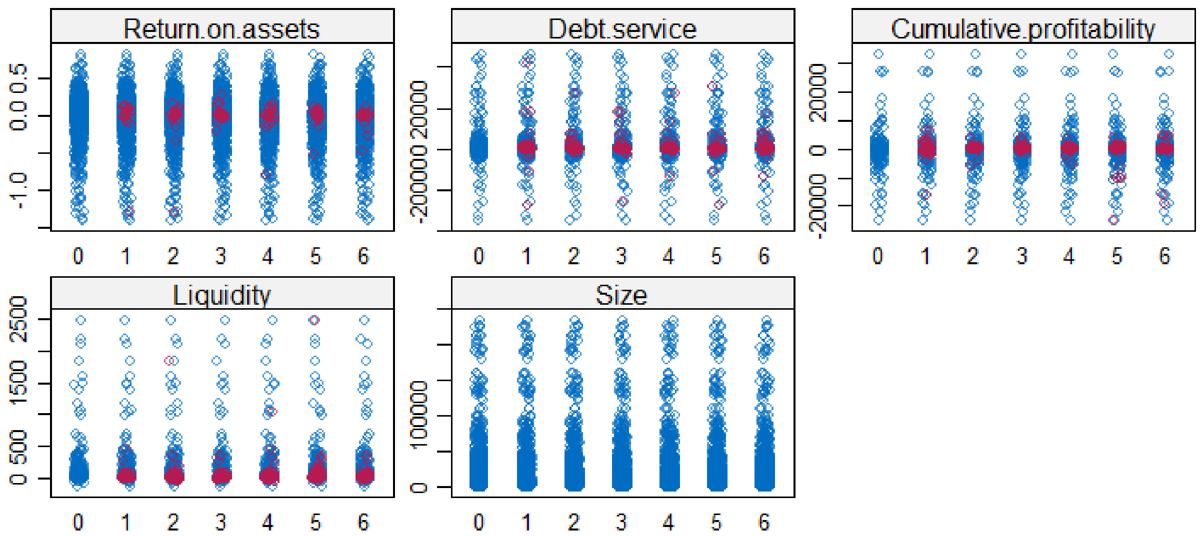

Figure 15 - ZETA model. Estimated value of missing ratios

3.6.3 Model test

No coefficients were showed by Altman, since the model is subject to copyright. The consequence is that the part where the original model is analysed will be skipped.
A quick analysis of the model without the application of the weights was tried but the results are similar to the ones achieved with the application of the Z score model without the weights. The original sample used by Altman was composed of 53 failed firms and 58 non-failed firms. To reproduce artificially the proportion of Altman, the weight applied to the I type error is equal to (53/58)/(number of failed firms in my sample/number of non-failed firms in my sample). The method used will be the logit model to try to have better results than the criticized MDA but the weights (instead of changing the cut-off) will be kept in order to hold a connection with the MDA model. The statistic methodology behind is preferred by the author since the MDA is, in his opinion not a good method to get good results.
The coefficients got from a weighted model is the following:

Table 40 - ZETA model, 1st year before bankruptcy train dataset. Weights applied. Coefficients.

Variable	**Estimate**	**Std. Error**	**Z value**	**Pr(>│z│)**
(Intercept)	-7.699e-01	1.124e-01	-6.848	7.49e-12
Return on assets	-1.883e+01	1.195e+00	-15.761	<2e-16
Debt service	1.611e-06	7.393e-07	2.179	0.0294
Liquidity	-7.764e-02	2.445e-02	-3.175	0.0015
Size	5.019e-06	2.202e-06	-2.280	0.0226

The cumulative profitability is not significant, and it was removed. The equation below is reached:

$$ZETA = -7.669 * e^{-1} - 1.883 * e^{1} * X_1 + 1.611 * e^{-6} * X_3 - 7.764 * e^{-2} * X_5 + 5.019 * e^{-6} * X_7$$

Table 41 - Altman ZETA model, 1st year before bankruptcy train dataset. Weights applied. Success table

	% correctly classified	**Sensitivity**	**Specificity**	**I type error %**	**II type error %**
1 year before bankruptcy	92.1%	80.8%	94.1%	19.2%	5.9%
2 years before bankruptcy	84.5%	41.7%	92.9%	58.3%	7.1%

3 years before bankruptcy	80.8%	19.1%	92.9%	80.9%	7.1%
4 years before bankruptcy	79.8%	14.5%	92.9%	85.5%	7.1%
5 years before bankruptcy	77.5%	15.2%	90.1%	84.8%	9.9%

More interesting is the model that uses the 3rd year as data train, which excludes the cumulative profitability and the size. It is interesting notice that this last variable, that was deemed important by (Altman, 2000), but also by (Ohlson, 1980), here is not useful. Using the coefficients of the table 42, the equation below is got:

Table 42 - ZETA model, 3rd year before bankruptcy train dataset. Weights applied. Coefficients

Variable	**Estimate**	**Std. Error**	**Z value**	**Pr(>│z│)**
(Intercept)	1.514e-01	6.292e-02	2.407	0.0161
Return on assets	-8.826e+00	9.470e-01	-9.319	<2e-16
Debt service	1.274e-05	7.470e-06	1.705	0.0882
Liquidity	-2.561e-02	6.556e-03	-3.906	9.39e-05

The equation got is the following:

$$ZETA = 1.514 - 8.826 * X_1 + 1.274 * e^{-5} * X_3 - 2.561 * e^{-2} * X_5$$

The table representing the results achieved is the following:

Table 43 - Altman ZETA model, 3rd year before bankruptcy train dataset. Weights applied. Success table

	% correctly classified	**Sensitivity**	**Specificity**	**I type error %**	**II type error %**
1 year before bankruptcy	77.5%	90.8%	75.1%	9.2%	24.9%
2 years before bankruptcy	74.0%	71.5%	74.5%	28.5%	25.5%
3 years before bankruptcy	70.2%	47.4%	74.7%	52.6%	25.3%
4 years before bankruptcy	70.6%	46.7%	75.3%	53.3%	24.7%
5 years before bankruptcy	69.8%	47.7%	74.4%	52.3%	25.6%

A prediction accuracy of 70% is quite high, especially considering that the sensitivity is around 50%. A further test, using a weight equal to 10, is tested in order to check if this solution can fit for conservative lenders who are willing to renounce to grant some debts but, in exchange, they don't want to take the risk to lose their money due to insolvent borrower.

2 tests have been made: the first one was based on the data coming from one year before the bankruptcy and another one that used a data train coming from 3 years before the bankruptcy. In the first case, the sensitivity was very low, especially in the 4th and in the 5th year, making the results not usable. In the second case, the sensitivity was very high but at the same time the specificity was very low. The result was that, if from one side almost all the bankrupt firms were identified, on the other side too many non-bankrupt firms were identified as failed. The consequence is that this model can lead to a huge cost in terms of missed chance, due to the non-grant of debt to worthy firms.

As with Z' score, a model using the deltas was developed. Also here, 2 tests were made, using the dataset coming from respectively 1 and 3 years before bankruptcy.

Both the models didn't show the awaited results. In the 2nd try, only "Delta return on assets" and "Delta debt service" (this last one only with a confidence interval equal to 90%) were significant and both the sensitivity and the specificity were close to 50%. In the first test, the results are better but still the problem of almost all the previous tests made using as data train the 1st years before bankruptcy remains, i.e. the sensitivity is very low.

The coefficients are represented in table 44:

Table 44 - ZETA model, 3rd year before bankruptcy train dataset. Weights applied. Coefficients

Variable	**Estimate**	**Std. Error**	**Z value**	**Pr(>│z│)**
(Intercept)	-1.675e-01	6.181e-02	-2.710	0.0067
Delta Return on assets	1.862e+00	2.714e-01	-6.862	6.79e-12
Delta Liquidity	7.984e-03	4.337e-03	1.841	0.0656
Delta Size	-8.536e-05	1.900e-05	-4.492	7.04e-06

And the equation got is

$$Y = -1.675 * e^{-1} - 1.862 * X_1 + 7.984 * e^{-3} * X_5 - 8.536 * e^{-5} * X_7$$

And the results are presented in table 45.

Table 45 - Delta Altman ZETA model, 3rd year before bankruptcy train dataset. Weights applied. Success table

	% correctly classified	**Sensitivity**	**Specificity**	**I type error %**	**II type error %**

1 year before bankruptcy	83.6%	60.2%	87.7%	39.8%	12.3%
2 years before bankruptcy	78.0%	35.3%	86.2%	64.7%	13.8%
3 years before bankruptcy	74.9%	14.0%	86.5%	86.0%	13.5%
4 years before bankruptcy	62.9%	23.4%	70.6%	76.6%	29.4%

The first conclusion that comes from the view of the results is that the ZETA model doesn't predict successfully the value of the output. This is true in our case, but the author reminds the reader that the model applied here is not the full model, because the variables X_2 and X_6 have been omitted due to the fact that with the dataset used it has been impossible include them, since also some previous years were required. Maybe, if the full ZETA model were used, the accuracy would have increased but this test should be made in a further research.
What is common in all the models is that the cumulative profitability, as was observed also in the analysis of the Z' and Z'' scores, is not significant in the failure prediction.
Another observation that comes from the analysis of the results of this ZETA, the Z' and the Z'' models come from the choice of the use of the train dataset. The selection of the most fitting train dataset was quite difficult. The observation that I can do is the following:

- If the train dataset comes from the 1st or the 2nd year, the specificity is always greater than 50%, but the sensitivity is quite low and in the 3rd, 4th and 5th years is very low.
- If the train dataset is represented by the 3rd year, the sensitivity can increase. The problem of this choice is that usually nor the sensitivity nor the specificity are much higher than 50% and sometimes they are even a little bit lower. This is due to the fact that 3 years before bankruptcy the values of the firm are still quite fuzzy and it is not so difficult to recognize the healthy and the unhealthy firms. The consequence is that the model has a better equilibrium in terms of sensitivity and specificity and its predictive ability is greater than 50% but such a percentage is enough to be considered a good model from the point of view of a banker? Having, let's say 60% of the outputs correctly predicted, according with the author, it could not be enough or many lenders.
- If the train dataset is represented by the 4th and the 5th years, the results are very bad both from the point of view of the sensitivity and of the specificity.

- If the train dataset is made of all the years, the sensitivity is very low, especially in the first years (the ones farther from the bankruptcy event) and the specificity is not too high.

What the author concluded is that it doesn't exist a data train that is better than the others. Each one has some pros and some cons. The best 2 are the ones that use as data train 1 and 3 years before bankruptcy. Still no of them provided satisfying results.

3.7 Ohlson O-score

3.7.1 Introduction

The following model analysed is the one proposed by (Ohlson, 1980). Undoubtedly, after the Altman models, it is the most famous accounting-based model to predict bankruptcy.
The author reminds the reader that Ohlson created 3 models:

- Model 1 predicts bankruptcy within one year;
- Model 2 predicts bankruptcy within 2 years. The underlying assumption is that the firm won't fail in the subsequent year;
- Model 3 predicts bankruptcy within 2 years.

To increase the comparability of the Ohlson O-score with the previous models, the author will predict the failure of a firm up to 5 years before bankruptcy. The advantage of this method is that all the 3 models will be included in a bigger model. In this way, the effectiveness of the Ohlson model will be checked and, furthermore, also the 3rd, the 4th and the 5th years before bankruptcy.
In this model the analysis of the data missing won't be done. The results would be the same of the previous 2 cases. Both the analysis showed that the firms whose data miss are firms with a bad performance but that still have an activity that prevents us from substituting the empty values with "0". At the same time, using an average or a median of the other values, would create a reduction of the variability and a bias that could create misleading results.
Instead, the study of the model will be performed again also in this case.

3.7.2 Variables analysis

Only the 4 years before bankruptcy have been included in our model aimed to the analysis of the overall performance. The reason behind is that in the 5th year the variables INTWO and CHIN can't be calculated because they need, for their calculation, both the current and the previous year but the 6th year before bankruptcy is not present.

The analysis will start with the correlation between the variables in the model:

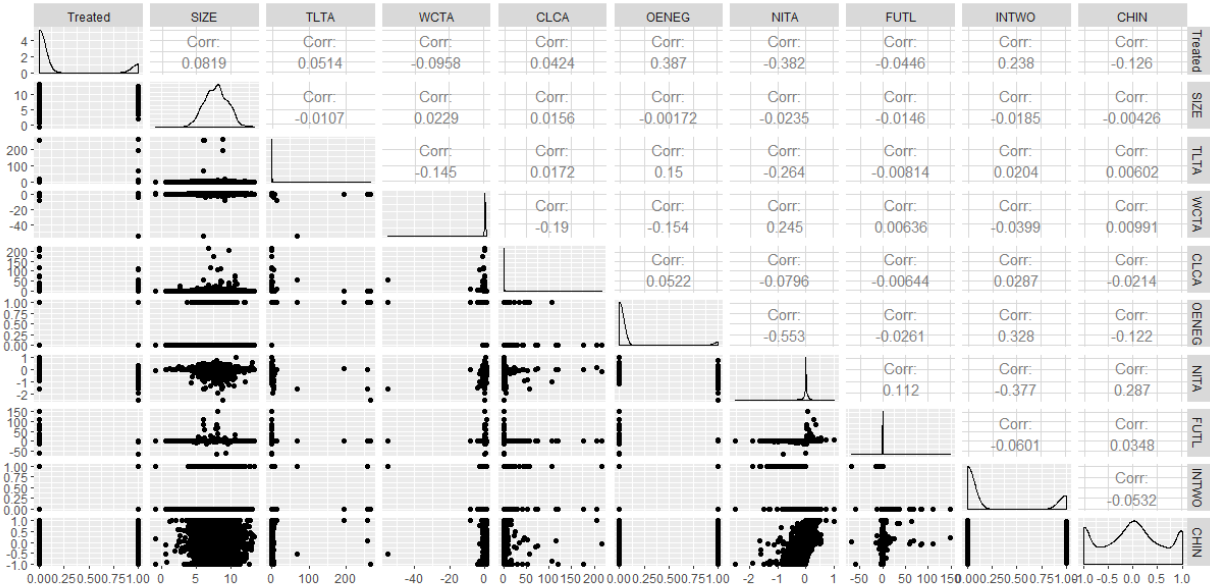

Figure 16 - Ohlson model correlation matrix

The correlation between the single variables is more worrying compared with the Altman's models.

The highest correlation between the variables come from the relationship between:

- WCTA and NITA. These 2 elements are the less expected even if they have some elements in common, for example the presence of the inventories in both the working capital and the net income.
- INTWO and OENEG. They are not directly correlated but it is very likely that a firm that has a net income<0, it has also total assets<total liabilities.
- TLTA and NITA. The reason of this correlation is the same of the one of the previous 2 variables.
- INTWO and NITA. The 2nd variable measures the ability of the firm to create income, while the 1st one checks whether a firm had a negative income in the previous 2 days.

To make a further test, also the presence of multicollinearity is tested:

Table 46 - multicollnearity Ohlson model

Var.	lag	SIZE	TLTA	WCTA	CLCA	OENEG	NITA	FUTL	INTWO	CHIN
VIF	1.07	1.01	1.05	1.02	1.03	1.4	1.74	1.02	1.26	1.16

There is no multicollinearity, provided that all the VIF values are lower than 5.

Finally, a series of graph, to test the differences among those firms that are going to fail and those that are not. In order to increase the readability of the following graphs, 66 outliers were removed.

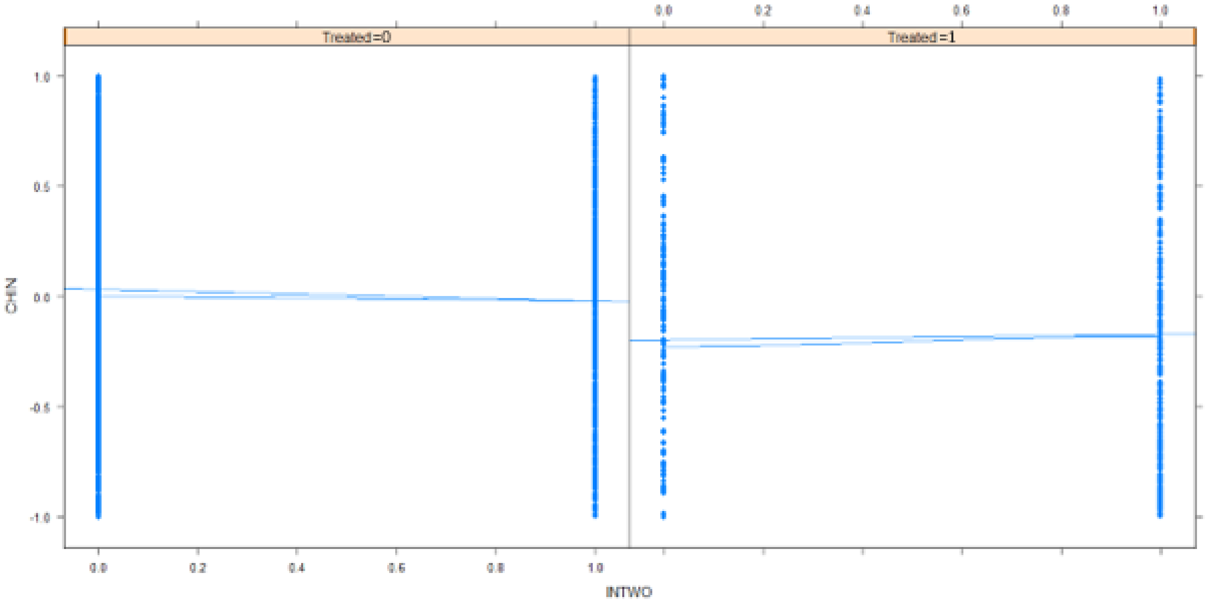

Figure 17 - Ohlson model variables: CHIN and INTWO

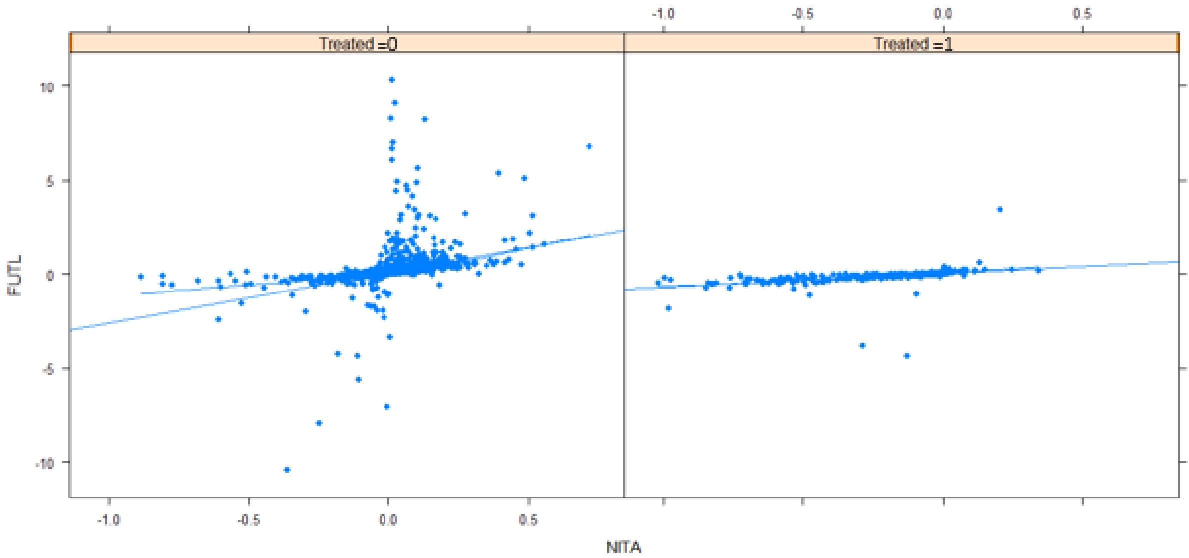

Figure 18 - Ohlson model variables: FUTL and NITA

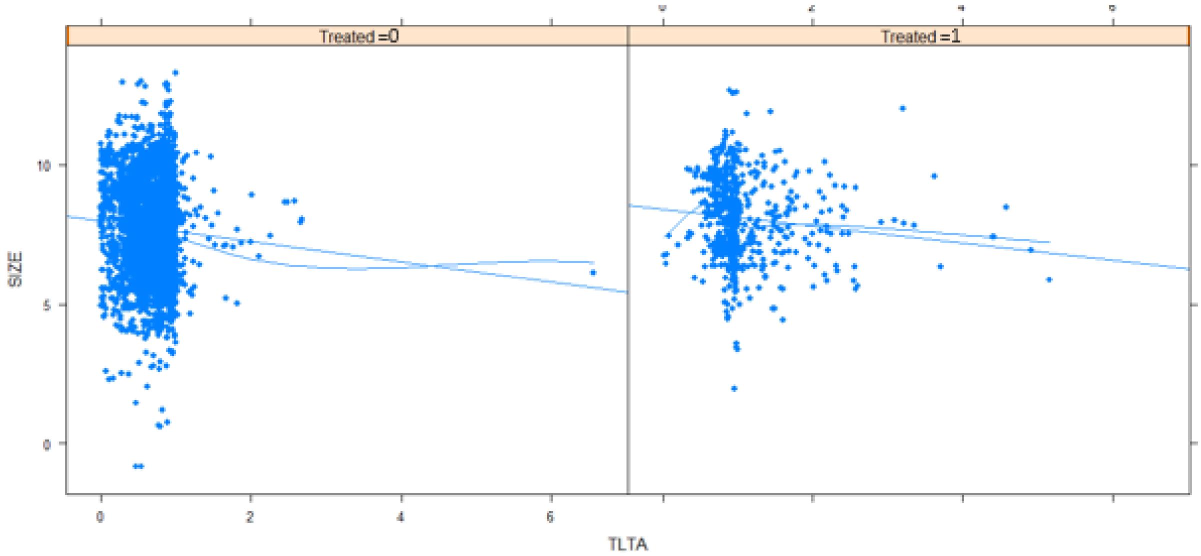

Figure 19 - Ohlson model variables: SIZE and TLTA

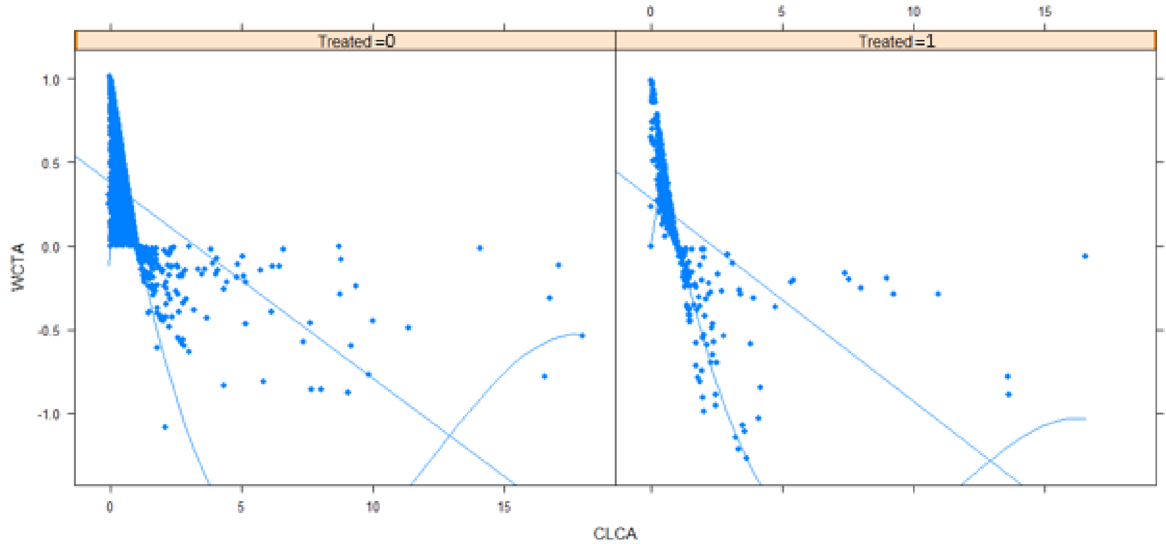

Figure 20 - Ohlson model variables: WCTA and CLCA

From these graphs we can se that on average the firms that go bankrupt are more likely to have a INTWO equal to 1, i.e. they have total liabilities greater than total assets compared with the non-failed ones. This last kind of firms also have a lower value associated to the value of CHIN, where,

$$CHIN = \frac{NI_t - NI_{t-1}}{(|NI_{t|} + |NI_{t-1}|)}$$

This ratio shows a deterioration of the net income, when it is present. Also, on average, the income is higher for non-failed firms. The FUTL, ratio aimed at measuring the ability of firm to repay the debts only with the liquidity coming from the operating performance, is slightly higher for non-failed firms. On average, also the liabilities of the failed firms are greater than the assets, as it is demonstrated by the higher value of TLTA and of CLCA. At the same time, a difference in SIZE and WCTA are not observed in the graphs. In particular, the WCTA could be non-significant, as it happened in the study of the Z' score.

3.7.3 Testing the model

The first test was made on the model 1, namely the model that predicts the failure of the firms one year before bankruptcy. The regression is based on the logit model. In order to test the original O-score, a cut-off equal to 3.8% and no-weights will be performed.

For the 2 versions of the first model, the variables WCTA, CLCA, FUTL and CHIN have been removed. In the 5th year before bankruptcy the variable OENEG can't be calculated so the model will be made by the other 4 variables.

The significant coefficients for the 1st are represented in table 47.

Table 47 - Ohlson model. Model 1. Coefficients.

Variable	Estimate	Std. Error	Z value	Pr(>│z│)
(Intercept)	-7.5607	0.9200	-8.218	<2e-16
SIZE	0.3465	0.0937	3.700	0.0002
TLTA	1.7520	0.4520	3.876	0.0001
OENEG	1.5455	0.4499	3.435	0.0006
NITA	-8.7486	1.3219	-6.618	3.64e-11
INTWO	0.6521	0.3156	2.066	0.0388

What is surprising is the coefficient of the size. According with the literature, firms with a greater size are less likely to fail, while in the sample of this thesis, the coefficient associated with the size shows that a bigger company is more likely to fail.

The equation got is the following:

$$O - score = -7.56075 + 0.34652 * SIZE + 1.75197 * TLTA + 1.54555 * OENEG - 8.74858 * NITA + 0.65209 * INTWO$$

The table 48 represents the success probability of the model.

Table 48 - Ohlson model 1. Cut-off 0.038. Success table.

	% correctly classified	Sensitivity	Specificity	I type error %	II type error %
1 year before bankruptcy	72.3%	95.4%	68.3%	4.6%	31.7%
2 years before bankruptcy	66.8%	79.8%	64.2%	20.2%	35.8%

3 years before bankruptcy	66.0%	65.4%	66.1%	34.6%	33.9%
4 years before bankruptcy	64.5%	63.6%	64.7%	36.3%	35.3%
5 years before bankruptcy	58.9%	59.7%	58.8%	40.3%	41.3%

Undoubtedly, this is the best model so far because both the sensitivity and the specificity are higher than 50%. The cut-off will be kept as 0.038. The presence of such a low cut-off makes the presence of a weight unnecessary. Anyway, a test with a logit model, with a weight applied on the data points with "Treated=1" equal to 1/(number of bankrupt firms/number of non-bankrupt firms) has been made and the result was worse compared with the one with the model with cut-off equal to 0.038. The use of a logit model together with a low cut-off in our case works much better than a multivariate analysis or than a logit model with the weights.

The use of model 2 will be tried as well. This model uses as data train 2 years before the bankruptcy and it has the strict condition that the firms can't flop in the year after the one used as train dataset. The non-significant variables are WCTA, CLCA and FUTL and, consequently, they have been removed.

Table 49 - Ohlson model 2. Coefficients.

Variable	**Estimate**	**Std. Error**	**Z value**	**Pr(>│z│)**
(Intercept)	-3.4469	0.5084	-6.780	1.20e-11
SIZE	0.1583	0.0598	2.647	0.0081
TLTA	0.94202	0.3551	2.653	0.0080
OENEG	-8.0338	1.3774	-5.833	5.45e-09
NITA	0.3964	0.2289	1.732	0.0834
INTWO	-0.3189	0.1690	-1.887	0.05910

The equation got is:

$$O - score = -3.44688 + 0.15826 * SIZE + 0.94202 * OENEG - 8.03380 * NITA + 0.39640 * INTWO - 0.31890 * CHIN$$

The table obtained with the success percentage of model 2 with the original cut-off is the following:

Table 50 - Ohlson model 2. Cut-off 0.038. Success table.

	% correctly classified	Sensitivity	Specificity	I type error %	II type error %
1 year before bankruptcy	22.5%	98.0%	9.4%	2.0%	90.6%
2 years before bankruptcy	22.8%	100.0%	7.4%	7.4%	92.6%
3 years before bankruptcy	23.9%	98.4%	8.7%	8.7%	91.3%
4 years before bankruptcy	23.2%	98.4%	7.8%	7.8%	92.2%
5 years before bankruptcy	22.3%	98.9%	7.0%	7.0%	93.0%

The cut-off, for the model 2, is too low. The consequence is that the sensitivity is very high, but the specificity is too low. Even if somebody were more interested in the sensitivity side, such a high percentage of misclassified non-bankrupt firms could create such a high amount of costs to make a reader prefer a lower sensitivity if compensated by a greater specificity.

After some tries, the best cut-off tested is 0.12. The results got with that are represented in table 51.

Table 51 - Ohlson model 2. Cut-off 0.12. Success table.

	% correctly classified	Sensitivity	Specificity	I type error %	II type error %
1 year before bankruptcy	72.4%	92.1%	69.0%	7.9%	31.0%
2 years before bankruptcy	68.5%	80.3%	66.2%	19.7%	33.8%
3 years before bankruptcy	66.6%	57.7%	68.5%	42.3%	31.5%
4 years before bankruptcy	64.5%	54.9%	66.5%	45.1%	33.5%
5 years before bankruptcy	61.3%	52.8%	63.0%	47.2%	37.0%

The results aren't bad but they are not as good as in the case of the first model. The last model to be analysed is model 3. Its dataset is made of firms that are going to fail within one or 2 years. The solution applied here is to create a dataset merging 516 firms from the dataset containing the ratios calculated on the firms one year before bankruptcy and 516 firms from the dataset containing the ratios calculated on the firms 2 years before bankruptcy. All the ratios, except OENEG and NITA, being non-significant, have being removed. The coefficients got are showed in table 52.

Table 52 - Ohlson model 3. Coefficients.

Variable	**Estimate**	**Std. Error**	**Z value**	**Pr(>│z│)**
(Intercept)	-2.6116	0.1356	-19.263	<2e-16
OENEG	2.1221	0.3217	6.596	4.22e-11
NITA	-9.7289	1.2377	-7.860	3.83e-15

When the datasets are mixed, the number of significant variables decreases a lot and the equation is:

$$O - score = -2.6116 + 2.1221 * OENEG - 9.7289 * NITA$$

Table 53 contains the success rate.

Table 53 - Ohlson model 3. Cut-off 0.17. Success table.

	% correctly classified	**Sensitivity**	**Specificity**	**I type error %**	**II type error %**
1 year before bankruptcy	27.6%	98.0%	15.5%	2.0%	84.5%
2 years before bankruptcy	28.2%	100.0%	13.9%	0.0%	86.1%
3 years before bankruptcy	29.0%	97.8%	15.0%	2.2%	85.0%
4 years before bankruptcy	28.3%	97.8%	14.0%	2.2%	86.0%
5 years before bankruptcy	27.0%	98.9%	12.6%	1.1%	87.4%

Also, in this case the specificity is too low. Some other cut-offs were tested on the 4^{th} and 5^{th} years before bankruptcy, that are the years where the sensitivity is lower among all the years and with all the cut-offs it was impossible to reach a balance between the specificity and the sensitivity.

The next step is to create a model, as was done also with the analysis of the Z' and ZETA scores, that uses as variables the difference in the values between one year and the previous one, to better appreciate the worsening of the ratios for those firms that are going through the bankruptcy condition.

Some tries were made using each year as data train and also, following the approach of (Ohlson, 1980), many cut-offs. The best result collected is definitely the one reached using the cut-off equal to 0.15 and as data train the 2nd year before bankruptcy. Instead, with the other years it was more difficult to find an equilibrium between specificity and sensitivity. Also, the cut-off equal to 0.038 produced a specificity too low.

The significant ratios can be appreciated in table 54.

Table 54 - Delta Ohlson model 2. Coefficients.

Variable	**Estimate**	**Std. Error**	**Z value**	**Pr(>│z│)**
(Intercept)	-1.8617	0.0977	-19.061	<2e-16
Delta SIZE	0.1209	0.0421	2.869	0.0041
Delta OENEG	-0.7656	0.2837	2.698	0.0070
Delta NITA	-4.2263	0.6396	-6.608	3.9e-11

The only significant coefficients are the ones above, while a change in the working capital or a deterioration in the ratio both between current liabilities to current assets as well as between total liabilities to total assets doesn't influence the output. Contrarily with what was expected, also the delta FULT doesn't influence the failure of a firm. A change in the variables INTWO and CHIN was more expected to not influence the model: the first one because it is less likely to change compared with the other ratios from one year to another and, if a firm has a net income around zero, a trend change can be considered a positive/negative signal when in reality it doesn't mean much, while the second one, CHIN, is a relative measure that, due to the signs of the of the net incomes, can be misinterpreted.

So, the equation that will be used is the following:

$$Y = -1.86177 + 0.12089 * delta\ SIZE - 0.76557 * delta\ OENEG - 4.22630 * delta\ NITA$$

The table 55 shows the success rate.

Table 55 - Delta model 2. Cut-off 0.15. Success table.

	% correctly classified	**Sensitivity**	**Specificity**	**I type error %**	**II type error %**
1 year before bankruptcy	67.1%	67.1%	67.2%	32.9%	32.8%
2 years before bankruptcy	70.2%	85.9%	67.5%	14.1%	32.5%
3 years before bankruptcy	63.5%	57.4%	64.7%	42.6%	35.3%
4 years before bankruptcy	62.7%	53.1%	64.7%	46.9%	35.3%

Both sensitivity and specificity are above 50% and this makes it, according with the author's opinion, a good model, even if the overall success rate, compared with other models, is not that high but at least it is more balanced.

To conclude the analysis of the Ohlson's model, it can be noted that some models both an interesting sensitivity and specificity, higher, on average, compared with the ones provided by Altman.

Then, as was mentioned before, the variables CLCA, WCTA and FUTL don't influence the final model. The first variable was expected to be more significant. Indeed, it influence the ZETA score. Probably its influence is decreased by the presence of the variable TLTA but also this last variable was significant only in one model. The variable WCTA, represented by X_1 in the Altman's Z score, is again useless. Finally, also FUTL has never been significant. This last result is quite surprising because at the base of the ability to pay the debts, there is the generate cash through the operating performance. Both the WCTA and the FUTL show that the cash generated by the firm, which can be used to pay the debt, is not significant, while NITA, that measures the firm overall performance, is again the most important ratio to determine the probability to be bankrupt.

3.8 Beaver model

3.8.1 Introduction

The last model that comes from the literature that is going to be analysed in this thesis is the one proposed by (Beaver, 1966). Also that scholar used the logit model like Ohlson and looked for, with an ex-post comparison, the best cut-off to check which one produces the best results. Because of that, the approach followed in this thesis, will be apply many cut-offs to the ratios of Beaver in order to find the most fitting model. Instead, contrary to the solution applied in the Altman's models, the so-called weights are not going to be used. The first step, like in the previous case, will be a general overview of the model, to better understand the relations between the variables and the influence that they exercise on the output. Like in the case of (Ohlson, 1980), the analysis of the missing values will be skipped, since the results of it is obvious (i.e. the missing data will simply be removed).

3.8.2 Variables analysis

A map of the correlations between the variables shows that some of them have a high correlation with the others.

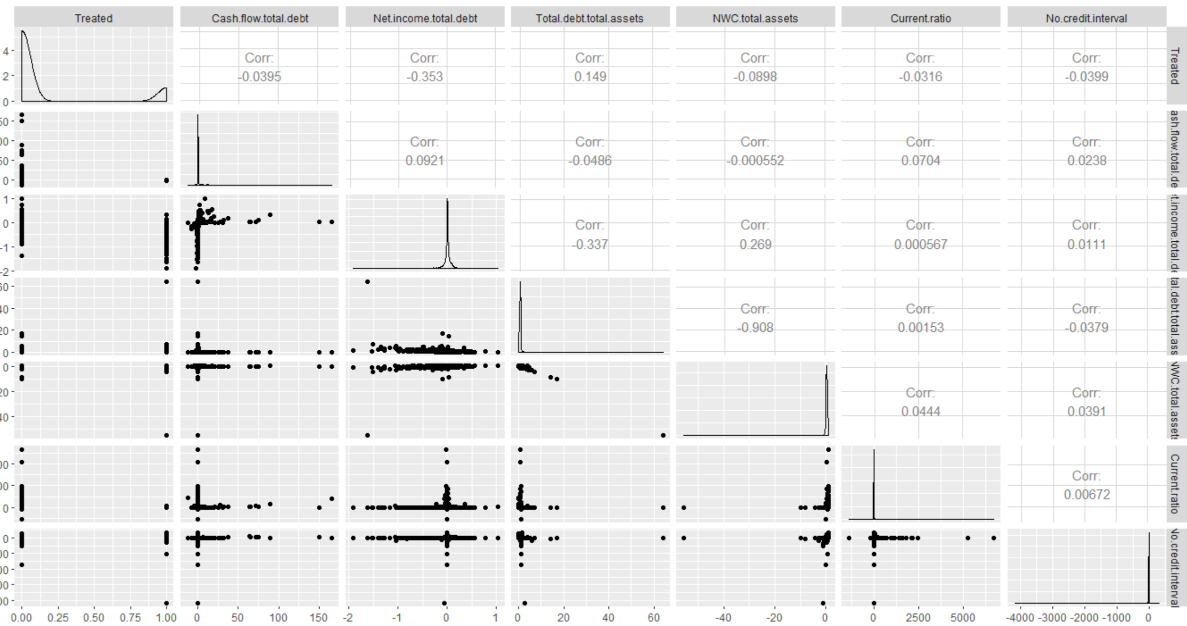

Figure 21 - Beaver model correlation matrix

It must be underlined that a high correlation is seen among:

- NWC/total assets and total debt/total assets (0.908). Clearly, the fact that the denominator is the same increases the correlation but the fact that the numerators are very similar is a clue to believe that the debt has a great influence on the NWC.
- Net income/total assets and total debt/total assets (-0.337). This is due to the fact that probably the interest rates, originated by a high amount of that, influence negatively the

net income. At the same time, high costs, that hit negatively the NI, are accompanied by an increase of the trade accounts payables.

- Net income/total assets and NWC/total assets (0.269). The correlation, probably, is still because the NWC is influenced by trade payables and this debt, through the above-mentioned relation between costs and trade payables, hits the net income. The correlation is positive because both the NWC and the NI are decreased by the presence of the trade payables.

A further check is made, in order to investigate the presence of multicollinearity in table 56.

Table 56 - multicollnearity Beaver model

Variable	**VIF**
Lag	1.037
Cash flow/total debt	1.18
Net income/total debt	1.32
Total debt/total assets	1.141
NWC/total assets	1.163
Current ratio	1.091
No credit interval	1.012

As the table 56 shows, there is no multicollinearity.

The influence of the single variables on the output will be examined with the following graphs. For a better comprehension of the graphs, 54 outliers have been removed.

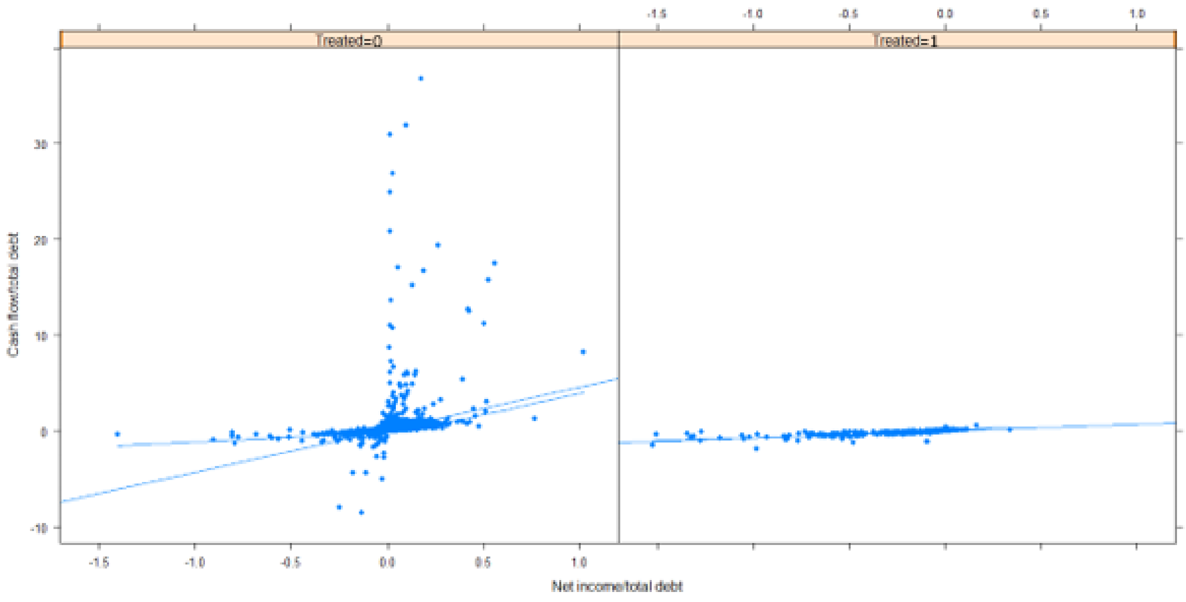

Figure 22 - Beaver model variables: Net income/total debt and Cash flow/total debt

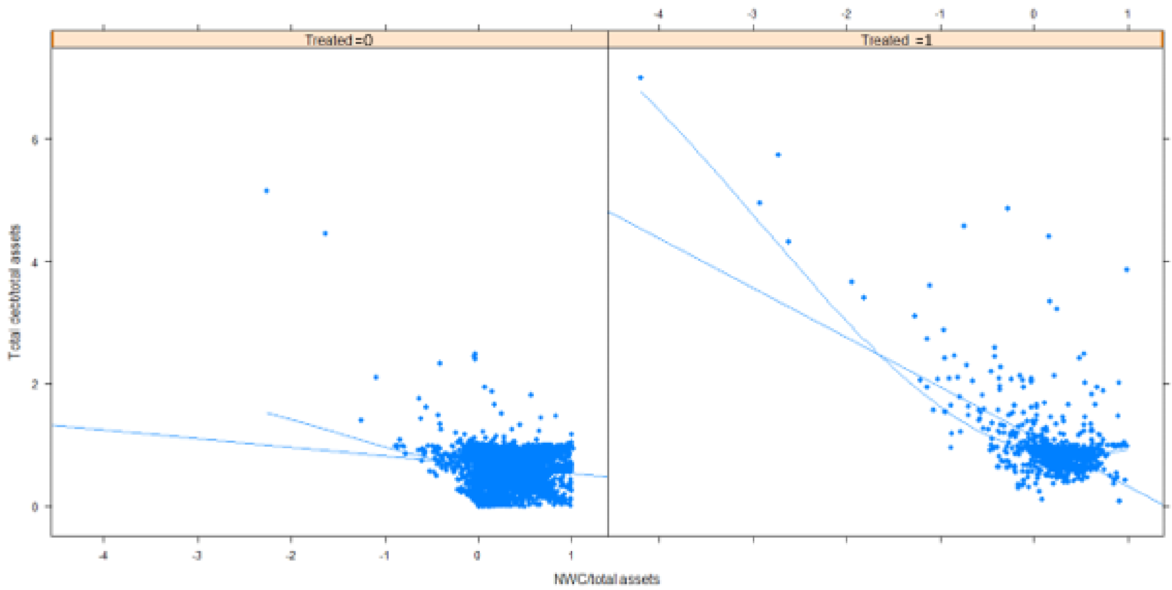

Figure 23 - Beaver model variables: NWC/total assets and Total debt/total assets

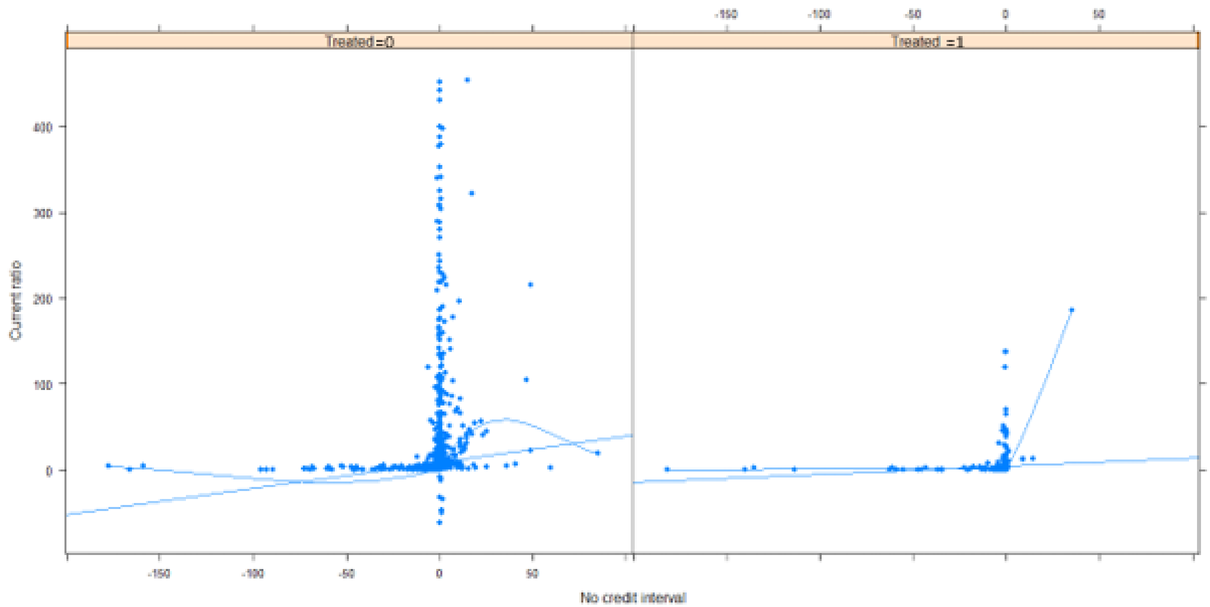

Figure 24 - Beaver model variables: No credit interval and Current ratio

In these graphs, an observer can see that on average the non-bankrupt firms have a higher ratio "cash flow/total assets" and a greater current ratio (current assets/current liabilities). Both these ratios measure the ability of a firm to pay back the debt. The 1st one does it from the point of view of the cash, while the 2nd one from the proportion of current assets to current liabilities, where this is a proxy of the solvability of a firm in the short term.

The net income/total assets is, as expected, on average, higher for the healthy firms while the no credit interval, that measures the ability of a firm to pay the operating expenditures with its quick and safe assets, it is just slightly higher for the firms marked by the output "Treated=1". Finally, the ratio "NWC/total assets" doesn't look like to be higher in a certain category.

3.8.3 Model test

Beaver in his model used a sample where half of the firms failed and half didn't. He did many tries in order to choose the cut-off the minimized the I and the II type errors, also according to the importance that he assigned to those errors. In his paper, he didn't specify the position of cut-offs as well as the coefficients so some tests will be done.

The first model used is one that choose the first year before bankruptcy as data train. The coefficients are represented in table 57.

Table 57 - Beaver model, 1st year before bankruptcy train dataset. Coefficients.

Variable	Estimate	Std. Error	Z value	Pr(>│z│)
(Intercept)	-3.4466	0.2989	-11.531	<2e-16
Net income/total debt	-13.9688	1.4101	-9.906	<2e-16
Total debt/total assets	0.7878	0.3200	2.462	0.0138

The significant variables are just 2. Cash flow/total debt and NWC/total assets are not significant. Again, the variables related the cash flows don't influence the model. Also, the current ratio and the no credit interval, both related to the current assets (even if the first denominator underlines the relation with the current assets, while the second with the operating expenses) do not influence the model. As usual, the most important variable is the one related to the net income and, as happened sometimes in the past, another important variable is the one related to the total debt to total assets. As expected, greater this last ratio, more a firm is likely to fail.

And the equation is the following:

$$Y = -3.4466 - 13.9688 * \frac{Net\ income}{Total\ debt} + 0.7878 * \frac{Total\ debt}{Total\ assets}$$

Table 58 shows the results obtained using the equation above.

Table 58 - Beaver model, 1st year before bankruptcy train dataset. Cut-off 0.055. Success table

	% correctly classified	Sensitivity	Specificity	I type error %	II type error %
1 year before bankruptcy	73.5%	92.9%	70.1%	29.9%	29.9%

2 years before bankruptcy	72.7%	83.8%	70.6%	29.4%	29.4%
3 years before bankruptcy	67.4%	67.1%	67.5%	32.5%	32.5%
4 years before bankruptcy	64.6%	57.7%	66.0%	42.5%	34.0%
5 years before bankruptcy	64.0%	60.3%	64.7%	39.7%	35.3%

With the cut-off equal to 0.05, instead, table 59 is got.

Table 59 - Beaver model, 1st year before bankruptcy train dataset. Cut-off 0.05. Success table

	% correctly classified	**Sensitivity**	**Specificity**	**I type error %**	**II type error %**
1 year before bankruptcy	69.6%	94.4%	65.3%	5.6%	34.7%
2 years before bankruptcy	68.6%	86.5%	65.3%	13.5%	34.7%
3 years before bankruptcy	64.3%	75.9%	62.1%	24.1%	37.9%
4 years before bankruptcy	62.5%	71.1%	60.8%	28.9%	39.2%
5 years before bankruptcy	59.7%	71.2%	57.4%	28.8%	42.6%

The percentage of correctly classified data points is lower in the 2nd case but the sensitivity is greater, so it can be more interesting for those people who focus in the prediction of firms' failure, renouncing also partially to the prediction of the non-failure.

Then, as it was made in the analysis of the Ohlson model, also the 2nd year before bankruptcy as used as data train. In this case, only the No credit interval and the cash flow to total debt ratios were not significant.

Table 60 - Beaver model, 2nd year before bankruptcy train dataset. Coefficients.

Variable	**Estimate**	**Std. Error**	**Z value**	**Pr(>│z│)**
(Intercept)	-2.9395	0.3069	-9.579	<2e-16
Net income/total debt	-11.4451	1.4049	-8.147	3.74e-16
Total debt/total assets	0.9837	0.2592	3.795	0.0001

NWC/total assets	1.0867	0.3576	3.039	0.0024
Current ratio	-0.0362	0.0166	-2.182	0.0291

The equation created is equal to

$$Y = -2.93950 - 11.44515 * \frac{Net\ income}{Total\ debt} + 0.98367 * \frac{Total\ debt}{Total\ assets} + 1.08672 * \frac{NWC}{Total\ assets} - 0.03621 * Current\ ratio$$

The result achieved, using a cut-off equal to 0.115 is as follows:

Table 61 - Beaver model, 2nd year before bankruptcy train dataset. Cut-off 0.115. Success table

	% correctly classified	**Sensitivity**	**Specificity**	**I type error %**	**II type error %**
1 year before bankruptcy	68.9%	95.2%	64.9%	4.8%	35.7%
2 years before bankruptcy	68.3%	87.2%	64.7%	12.8%	35.3%
3 years before bankruptcy	63.9%	76.6%	61.5%	23.4%	38.5%
4 years before bankruptcy	62.6%	69.8%	61.2%	30.2%	38.8%
5 years before bankruptcy	59.9%	71.2%	59.9%	28.8%	42.3%

The results are a little bit worse but still, they are above 50%.

Also, a model using the 3rd, the 4th and the 5th years before bankruptcy was tested but the only positive result was obtained using the 3rd year before bankruptcy as data train and 0.17 as cut-off and the significant variables are represented in table 62.

Table 62 - Beaver model, 3rd year before bankruptcy train dataset. Coefficients.

Variable	**Estimate**	**Std. Error**	**Z value**	**Pr(>│z│)**
(Intercept)	-2.7155	0.2886	-9.409	<2e-16
Net income/total debt	-4.1182	1.0789	-3.817	0.0001
Total debt/total assets	1.5831	0.3707	4.270	1.95e-05
Current ratio	-0.0170	0.0094	-1.806	0.0709

The equation, using the coefficients coming from the table above is the following:

$$Y = -2.715536 - 4.118237 * \frac{Net\ income}{Total\ debt} + 1.583061 * \frac{Total\ debt}{Total\ debt} - 0.016962 * Current\ ratio$$

The table achieved is the following:

Table 63 - Beaver model, 3rd year before bankruptcy train dataset. Cut-off 0.17. Success table

	% correctly classified	**Sensitivity**	**Specificity**	**I type error %**	**II type error %**
1 year before bankruptcy	69.2%	94.4%	64.9%	5.6%	35.1%
2 years before bankruptcy	68.7%	93.8%	65.9%	16.2%	34.1%
3 years before bankruptcy	64.8%	74.1%	63.1%	25.9%	36.9%
4 years before bankruptcy	62.8%	68.6%	61.7%	31.4%	38.3%
5 years before bankruptcy	62.1%	64.1%	61.7%	35.9%	38.3%

Also this model is quite balanced and it achieved nice success percentage both from the point of view of the sensitivity and the specificity.

A problem that comes to the mind of the author, comparing the big number of cut-offs used in this thesis, is that a small variation of the cut-off leads to very different results. For example: a unique cut-off, good for all the models was impossible to find. Moreover, as it happened with the Ohlson model, further is the year of the data train from the bankruptcy year, higher is the cut-off chosen. In the case of the analysis of the model proposed by (Ohlson, 1980), for example, when the data train was set one year before bankruptcy, the cut-off was 0.038, while, when the data train was set 2 years before bankruptcy, the cut-off was 0.15. The results collected with Beaver are the same: if the data train is one year before bankruptcy, the cut-off is 0.05, if it is 2 years before bankruptcy, the cut-off is 0.12. The difference in the cut-off selection is that farther the firm is from bankruptcy moment, fuzzier is the determination of the output "Treated" and, therefore, the cut-off must be increased from values that are below 0.1 to values that are much higher. As usual, also a model using the variations of the variables will be used to predict the bankruptcy of a certain firm.

The first model uses 1 year before bankruptcy as train data and the cut-off is equal to 0.08. The model got is the following, where just two variables are significant. Coefficients are represented in table 64.

Table 64 - Delta Beaver model, 1st year before bankruptcy train dataset. Coefficients

Variable	Estimate	Std. Error	Z value	Pr(>│z│)
(Intercept)	-2.4562	0.1681	-14.608	<2e-16
Delta Net income/total debt	-5.1708	0.7810	-6.621	3.58e-11
Delta Total debt/total assets	1.6209	0.2994	5.414	6.15e-08

The equation is as follows:

$$Y = -2.4562 - 5.1708 * \frac{delta\ Net\ income}{delta\ total\ debt} + 1.6209 * \frac{delta\ Total\ debt}{delta\ total\ assets}$$

And the success table is here below (table 65):

Table 65 - Delta Beaver model, 1st year before bankruptcy train dataset. Cut-off 0.08. Success table

	% correctly classified	Sensitivity	Specificity	I type error %	II type error %
1 year before bankruptcy	69.2%	94.4%	64.9%	5.6%	35.1%
2 years before bankruptcy	68.7%	93.8%	65.9%	16.2%	34.1%
3 years before bankruptcy	64.8%	74.1%	63.1%	25.9%	36.9%
4 years before bankruptcy	62.8%	68.6%	61.7%	31.4%	38.3%

The last model built is used 2 years before bankruptcy, 0.15 cut-off and, again, the difference of the values of the variables. The model is as follows and it has only one significant variable:

Table 66 - Delta Beaver model, 2nd year before bankruptcy train dataset. Cut-off 0.15. Coefficients.

Variable	Estimate	Std. Error	Z value	Pr(>│z│)
(Intercept)	-1.7932	0.1091	-16.430	<2e-16
Net income/total debt	-5.4530	0.7381	-7.388	1.49e-13

The equation is:

$$Y = -1.7932 - 5.4530 * \frac{delta\ Net\ income}{delta\ total\ debt}$$

And, finally, the table containing the success rate is the following:

Table 67 - Delta Beaver model, 2nd year before bankruptcy train dataset. Cut-off 0.15. Success table

	% correctly classified	Sensitivity	Specificity	I type error %	II type error %
1 year before bankruptcy	72.6%	88.2%	69.8%	11.8%	30.2%
2 years before bankruptcy	67.8%	67.9%	67.8%	32.1%	32.2%
3 years before bankruptcy	65.0%	50.3%	68.0%	49.7%	32.0%
4 years before bankruptcy	64.0%	49.0%	67.0%	50.4%	33.0%

The 3rd year and the following ones didn't show good results so no models were developed using them.

What is surprising of this last model is that the success percentage is high even if just one variable is used.

The no-credit interval, that is supposed to measure the ability of a firm to pay its operating expenses with the difference between current assets and liabilities, was never significant. Also, the ratio connected to the cash flow was never useful, while the current ratios and NWC ratios were a little bit more useful but still, again, as in the cases of Altman and Ohlson, the ratio connected to the net income is the most important, followed by the ratio between debt and the assets. The important of the ratio connected to the performance of the firm is confirmed by the fact that also using only that ratio, the predictive ability of the model is more or less the same that the models with other variables.

3.9 Comparison of the literature models

Finally, a comparison of the model will be performed. Many equations were created during this thesis, but not all of them performed at the same manner. An initial filter will be used to understand which model is better with the consequent exclusion of all those models that didn't perform both a specificity and a sensitivity at least equal to 40%.

The only models that pass this selection are the following ones:

- Z2 weights - 3 years before bankruptcy

- ZETA weighted - data train year 1
- Ohlson model 1
- Ohlson model 2 – cut-off 0.12
- Beaver - 1 year before bankruptcy - cut-off 0.055
- Beaver - 2 years before bankruptcy – cut-off 0.115
- Beaver – 3 years before bankruptcy
- Ohlson delta - cut-off 0.15 - 2 year before bankruptcy
- Beaver delta - 1 year before bankruptcy - 0.08 cut-off
- Beaver delta - 2 years before bankruptcy, 0.15 cut-off

What the selection approach adopted implies? It implies that the excluded models focus on a high prediction of the failure 1 or 2 years before bankruptcy, not caring about the deterioration of the results in the 3rd, 4th and 5th years before bankruptcy. A lower percentage of correctly classified data points can be accepted in the 1st and 2nd years before bankruptcy if this is compensated by a better forecast ability in the previous years. The reason behind is that a lender as well as a regulator can be more interesting in the bankruptcy forecast some years before the moment itself and not just 2 or, even worse, 1 year when it is too late for a corrective action. This modus operandi doesn't lead to a high percentage of correctly classified outputs each year, but to a more homogeneous correctly classified number of firms and at the same time.

What it was expected is that the models applied by Altman were not going to be successful. This is not due to the fact that the variables applied by Altman were wrong, but it is because of the analysis methodology chosen by the scholar, namely the discriminant analysis. Also if we collected the results achieved by Altman himself in 1968, they wouldn't have passed the initial selection because they were quite low in the 3rd year before bankruptcy and in the previous ones.

Less models based on the O-score, instead, compared with the one based on Beaver (1966), passed the selection.

Coming back to the choice of the best model, it is very difficult to determine which one forecast better the output. A compromise decision is to determine which one has the highest average % of correctly predicted output, which one has the greatest average sensitivity, and which one has the greatest average sensitivity. The imposition of a value of both sensitivity and specificity higher than 40%, anyway, lead to quite balanced models. Because of this, even if we take the model with the highest specificity, we are ensured that also the sensitivity has an acceptable level and vice versa.

The model with the highest sum between specificity and sensitivity is: Beaver - 1 year before bankruptcy - cut-off 0.05. This model has also a very high sensitivity, overtaken only by slightly by Z'' weights – 3 years before bankruptcy.

The model with the highest specificity is: ZETA weighted - data train year 1

The best model among the ones got using the difference in the variables value between one year and another is: Beaver delta - 1 year before bankruptcy - 0.08 cut-off. Notwithstanding this, the ability prediction of the models that use the delta is much lower compared with the traditional models.

The 2 best models ("ZETA weighted - data train year 1" and "Beaver - 1 year before bankruptcy - cut-off 0.05") will be analysed better in order to provide more information about their use.

The "Beaver model - 1 year before bankruptcy - cut-off 0.05" was applied to our sample, in order to associate a certain value of our Y, coming from the equation got previously, i.e.

$$Y = -3.4466 - 13.9688 * \frac{Net\ income}{Total\ debt} + 0.7878 * \frac{Total\ debt}{Total\ assets}$$

to the likelihood of correctly predict the output of a data point. In table 68 the sample has been divided in groups, in order to catch the likelihood associated to a single group to face the bankruptcy. Before take some conclusions from the table, the reader should be aware that in the Beaver sample, the failed firms represent the 16% of the total number of firms. Y is the output of the model, got using the formula, not the variable "Treated". The sample used a mix of data points coming from the all 5 years.

Table 68 - Beaver – 1 year before bankruptcy. Cut-off 0.05. Success table by Y value.

Y value	**Bankruptcy probability**
Y>0	69.2%
-2.5<Y<0	28.75%
-3<Y<-2.5	20.0%
-3.3<Y<-3	8.0%
Y<-3.3	3.5%

This can be helpful for a person that is trying to apply the model. He/she knows that if Y>0, the probability that the firm is going to fail is pretty high (4,3 times the average failure probability), while if Y<-3.3, the probability of bankruptcy is almost null (0.22 times the average default probability).

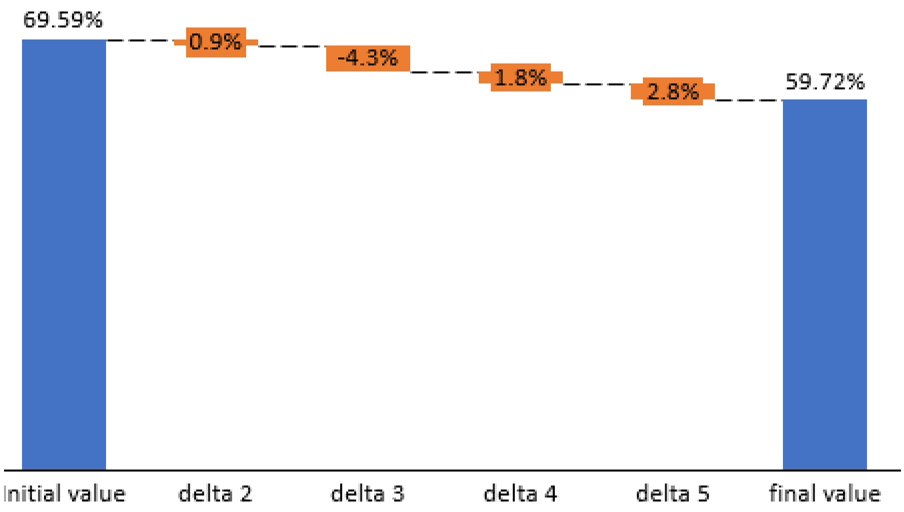

Figure 25 - Beaver – 1 year before bankruptcy. Cut-off 0.05. Prediction accuracy bridge

The following graphs help the reader in better understanding the prediction ability of the Beaver's model. Figure 25 allows us to better understand the decrease of the model performance.

The deltas show the decrease in the accuracy of each year. As could be observed also in the previous models, in the 3rd, 4th and 5th years before bankruptcy the prediction accuracy decreases a lot compared with the first 2 years. The final accuracy is not much higher than 50% but it must be pointed that the sensitivity is quite high in this model and the bankruptcy firms represent less than 20% of the total number of firms.

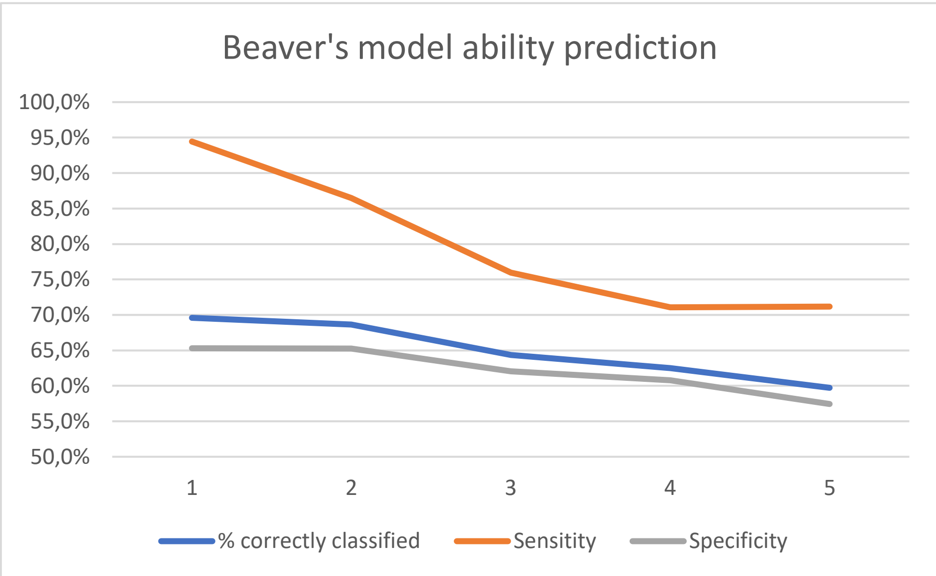

Figure 26 - Beaver model – 1 year before bankruptcy. Cut-off 0.05. Ability prediction by year.

The advantage of figure 26 is that it helps us in observing the percentage of correctly predicted data points and in particular if the decrease in the prediction accuracy is due to the specificity or the sensitivity. In this case, the decrease of the precision of the model is due both to the sensitivity and the specificity.

The other model, "ZETA weighted - data train year 1", has a lower ability prediction. Even if its values are associated to a lower prediction ability. The high specificity is damaged by a low sensitivity. The model is a little bit unbalanced in favour of detecting the non-failed firms, as the following table, that shows which ZETA values are associated with a certain default probability.

Table 69 - ZETA weighted – 1 year before bankruptcy. Success table by ZETA value

ZETA value	**Bankruptcy probability**
ZETA>0	39.2%
-1<ZETA<0	23.0%
-0.5<ZETA<0	16.0%
-1.5<ZETA<-0.5	10.8%
ZETA<-1.5	5.0%

As it can be observed in this table, the ZETA score surely provides some signals about the likelihood of a firm if it is laggard or not, even if it is not accurate like in the case of the Beaver's model.

The right graph, instead, shows the decrease of the prediction ability of the model. In particular, a decrease in the 2nd and 3rd years is registered. While in the 4th and in the 5th years remain more or less constant.

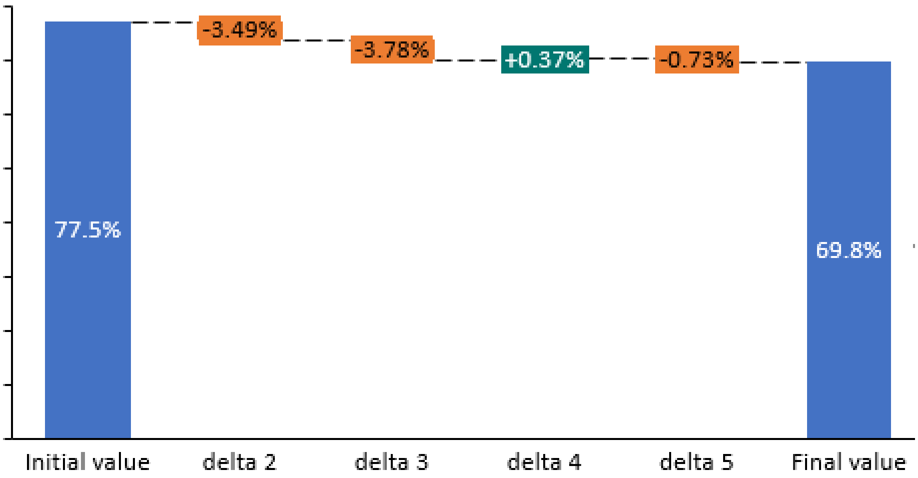

Figure 27 - ZETA weighted – 1 year before bankruptcy. Prediction accuracy bridge

The decrease in the prediction accuracy is mainly due to a drop in the 2nd and 3rd years before bankruptcy of the sensitivity.

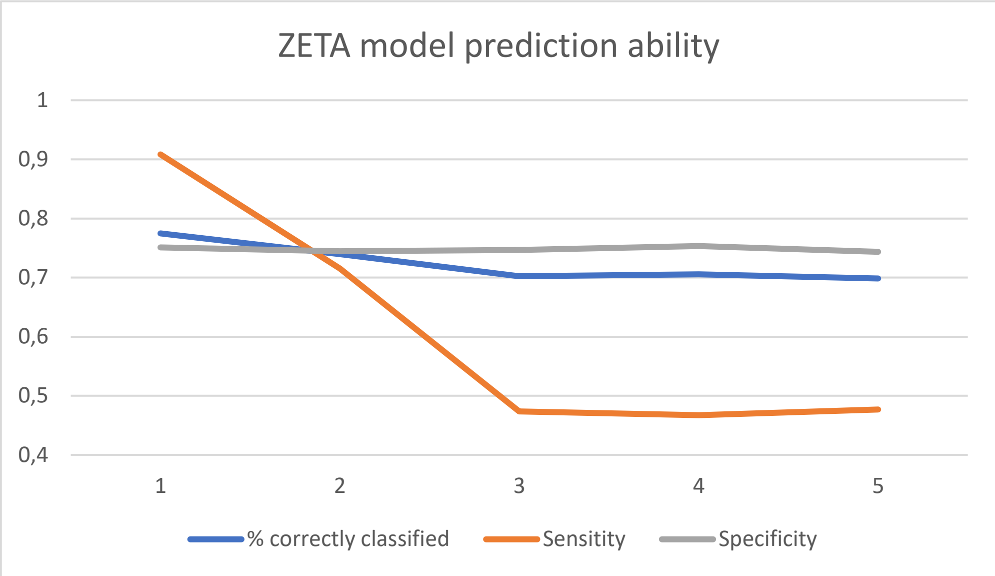

Figure 28 - ZETA weighted – 1 year before bankruptcy. Ability prediction by year.

Even if these are the 2 models chose, it should be noted that also the "Z2 weights - 3 years before bankruptcy" had a very high sensitivity, supported by an acceptable specificity. Since the Beaver model – 1 year before bankruptcy had quite similar results but an overall prediction accuracy higher, it was preferred. Even if the author didn't analyse it carefully, an entity particularly interested in sensitivity, willing to lose something in terms of specificity, could use the above mentioned Z'' model.

4. CHAPTER FOUR

Analysis of a new model

4.1 Additional tests

4.1.1 Variables selection

Last model presented will be based on the author analysis of the best ratios that can be applied to the sample of this thesis. A group of 23 ratios where chosen as potential variables of the model. These indices have been chosen according with their popularity in the literature, their relevance in the prediction accuracy and the lack of multicollinearity between the variables. They were tested on the full sample of firms, composed by all the 5 years each firm was analyzed.

The initial ratios were the following ones:

1. Current ratio;
2. Acid ratio;
3. Cash ratio;
4. Current assets/total assets;
5. Current liabilities/total liabilities;
6. Total debt/total assets;
7. Total debt/total equity;
8. Days payables outstanding;
9. Days receivables outstanding;
10. Days inventories outstanding;
11. CFO/total debt;
12. CFO/EBIT;
13. FCF/total assets;
14. NWC/operating assets;
15. Sales/Net operating assets;
16. EBIT/net operating assets;
17. EBIT/total assets;
18. Net income/EBIT;

19. Net income/total assets;
20. Retained earnings/total assets;
21. Dummy EBIT-(financial income and expenses)>0;
22. Firm size;
23. Profitability deviation.

According with the criteria explained below, only the following variables were selected:

Table 70 - X model, coefficients chosen

Variable	**Z value**	**Pr(>│z│)**
(Intercept)	-22.157	<2e-16
Current ratio	-1.4	0.1510
Net income/total assets	-3.458	0.0005
Total debt/total assets	9.012	<2e-16
Dummy EBIT-(financial income and expenses)	5.736	9.68e-09
Profitability deviation	6.846	7.62e-12
NWC/Operating assets	0.406	0.6845

These variables are going to form the model henceforth named "Model X". The first 1st and the 3rd ratios focus on the balance sheet composition of the firm and they look for a difference in the composition of the assets and liabilities a certain proportion that can differentiate the healthy and the sick firms. The 2nd ratio measures the profitability ability of the firm, while the 4th dummy investigates the ability of paying the financial debt through the operating activity. Finally, the 5th variable measures the difference between the net income of a certain company and the average net income of the sample. The author chose a compromise in the creation of this model. A variable, namely "CFO/EBIT", was eliminated from this model because it was too correlated with the variables "NWC/Operating assets". After the elimination of the first variable, also the second was shown to be inconsistent, but since in the past literature all the investigated models used this variables and since in many cases in the past regressions it showed to be significant, the author decided to keep it. By the way, since a selection of the variables, as it was done in the past, will be done in each model, if NWC/operating assets won't show to be useful in the bankruptcy prediction, it will be removed. This 6th variable represents the ability of the firm to repay its current financial obligations using its current assets. For what concerns the Current ratio, even if the variable is not really significant per sé is not significant, its removal would lower lead to an increase in AIC, so it was kept.

Looking at the coefficients, having a profitability above the average, surprisingly, increases the probability to be bankrupt, as well as having an EBIT higher than the interest expenses.

4.1.2 Variables analysis

Before apply the model, we make some checks, to be sure that we won't get biased results, starting from the multicollinearity.

Table 71 - Multicollinearity, X model

Variable	VIF
Current ratio	1.00
Total debt/total assets	1.13
Profitability deviation	1.33
Net income/total assets	1.62
Dummy EBIT-(financial income and expenses)	1.33
NWC/Operating assets	1.00

No multicollinearity is highlighted. The correlations between the single variables are as follows:

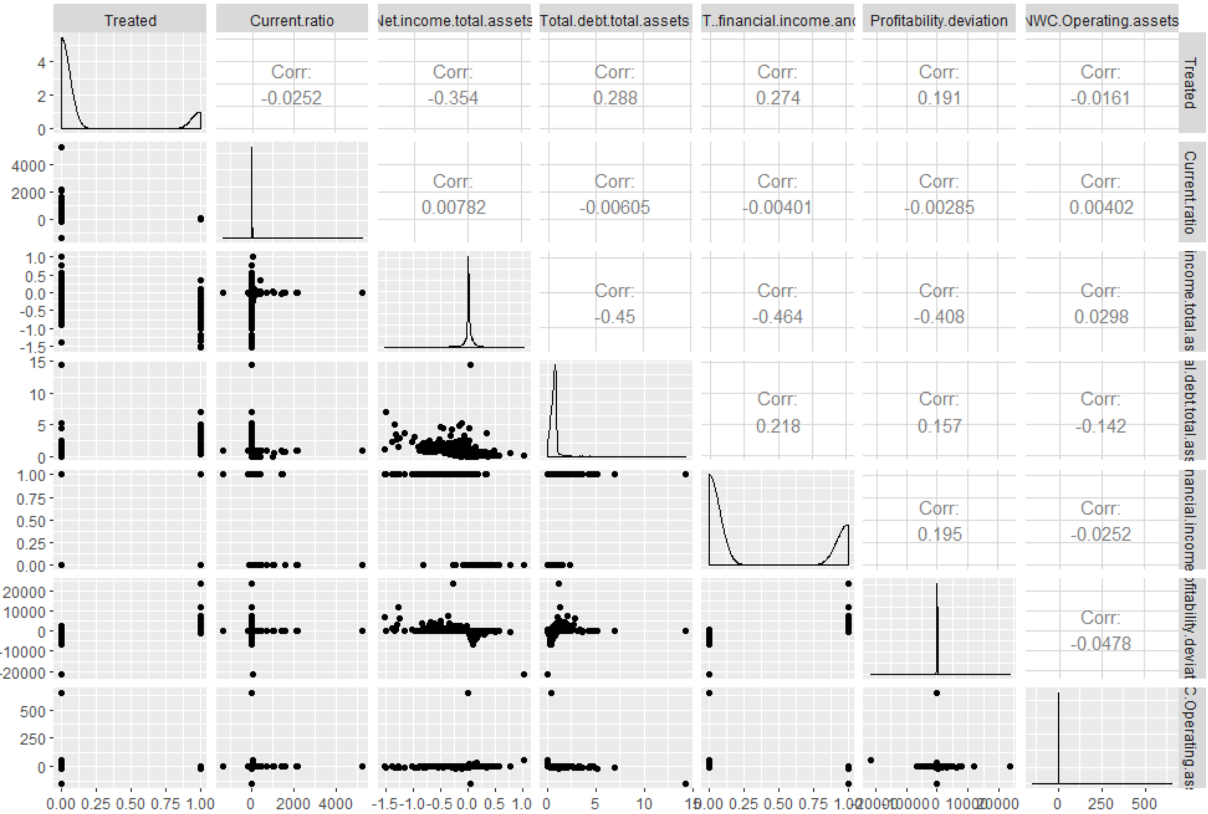

Figure 29 - Beaver model correlation matrix

The variable that shows a greater correlation with the others is net income/total assets. This variable is strongly correlated with the dummy independent variable and the profitability deviation variable. This can be expected because all the 3 variables measure the profitability performance, while the negative correlation between the net income/total assets and total debt/total assets shows the often firms with a higher debt perform worse than the own with a relatively low debt.

For the following graphs, 92 outliers were eliminated.

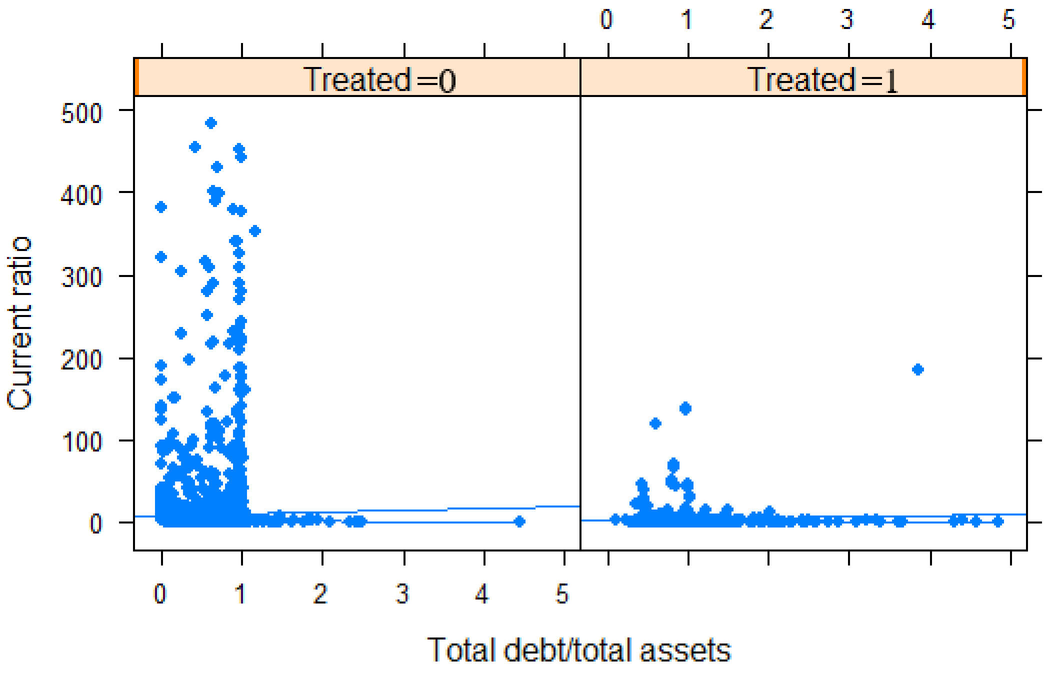

Figure 30 - X model variables: Current ratio and Total debt/total assets

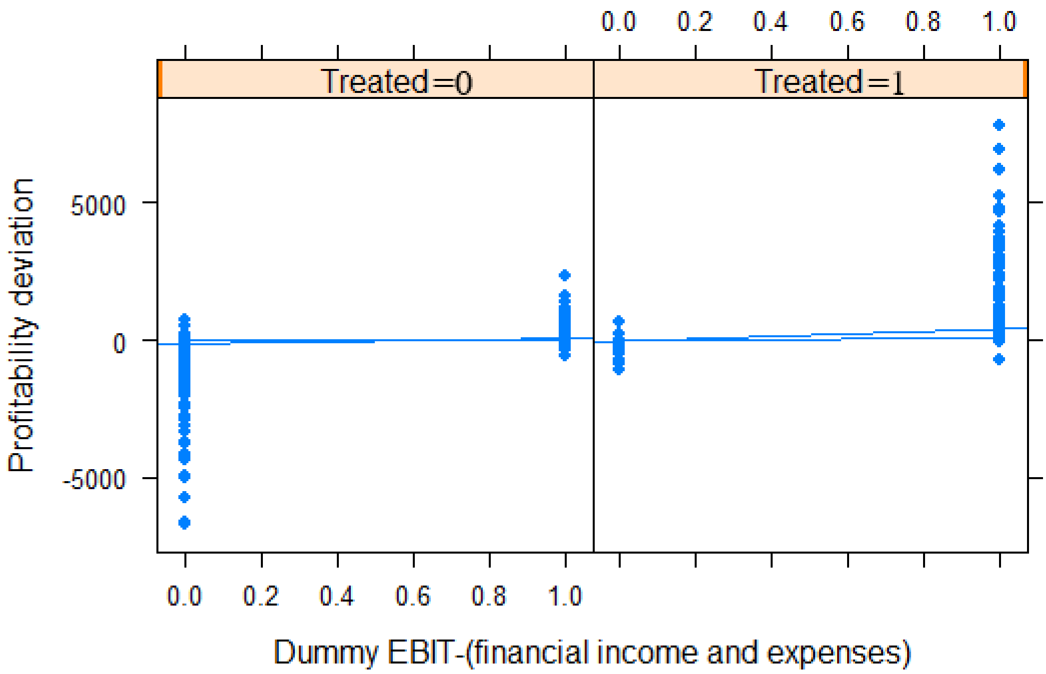

Figure 31 - X model variables: Profitability deviation and Dummy EBIT-(financial income and expenses)

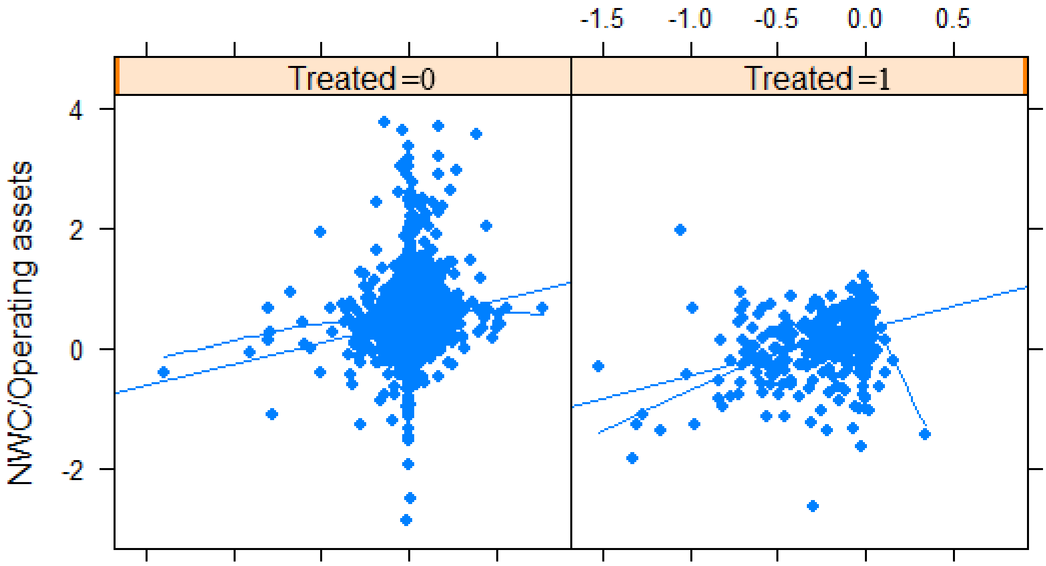

Figure 32 -. X model variables: NWC/Operating assets and Net income/total assets

Net income/total assets

Laggard firms show graphically on average a higher total debt/total assets ratio, while they have a lower current ratio and the ratio between net income/total assets. As it happened in the analysis of the literature models, the NWC/Operating assets ratio, instead, didn't show a different trend in the comparison between Treated=1 and Treated=0. This is coherent with the non-significance of this ratio. Finally, unexpectedly, the profit deviation reports a relatively higher number of non-bankrupt firms with a profitability below the average. Instead, on the other side, from the same graph it looks like that the non-bankrupt firms have usually the difference between EBIT and financial costs higher than 0.

4.1.3 Model test

The model will be tested to observe its prediction ability. The first test will be done using as train dataset the first year before bankruptcy and 0.02 as cut-off of the logit model.

The model got is the following:

Table 72 - X model, 1st year before bankruptcy train dataset. Coefficients.

Variable	**Estimate**	**Std. Error**	**Z value**	**Pr(>|z|)**
(Intercept)	-5.9192	0.5226	-11.326	<2e-16
Current ratio	-0.0298	0.0161	-1.850	0.0643
Net income/total assets	-5.1882	1.6138	-3.215	0.0013
Total debt/total assets	2.5095	0.3955	6.345	<2.22e-10

Dummy EBIT- (financial income and expense)	1.7654	0.4487	3.935	8.33e-0.5
Profitability deviation	0.0035	0.0009	3.733	0.0001
NWC/Operating assets	0.2486	0.0407	6.110	9.97e-10

It is represented by the equation below:

$$B = -5.9192 - 0.0299 * Current\ ratio - 5.1882 * \frac{Net\ income}{Total\ assets} + 2.5095 * \frac{Total\ debt}{Total\ assets} + 1.7654 * [Dummy\ EBIT - (financial\ income\ and\ expenses)] + 0.0035 * Profitability\ deviation + 0.2486 * \frac{NWC}{Operating\ assets}$$

And the table containing the prediction accuracy is as follows:

Table 73 - X model, 1st year before bankruptcy train dataset. Cut-off 0.02. Success table

	% correctly classified	Sensitivity	Specificity	I type error %	II type error %
1 year before bankruptcy	70.2%	96.0%	65.7%	4.0%	34.3%
2 years before bankruptcy	65.9%	84.5%	62.4%	15.5%	37.6%
3 years before bankruptcy	64.5%	74.5%	62.5%	25.5%	37.5%
4 years before bankruptcy	61.1%	67.9%	59.8%	32.1%	40.2%
5 years before bankruptcy	58.3%	67.3%	56.3%	32.7%	43.5%

The first equation already showed satisfying results.

The second model, instead, will be created using as a base the train dataset on the 2nd year before bankruptcy and as cut-off of the logit model 0.11.

Table 74 - X model, 2nd year before bankruptcy train dataset. Coeffcients.

Variable	Estimate	Std. Error	Z value	Pr(>│z│)
(Intercept)	-3.4333	0.2815	-12.195	<2e-16
Current ratio	-0.0480	0.0172	-2.787	0.0053
Net income/total assets	-3.6392	1.4726	-2.471	0.0134

Total debt/total assets	1.6145	0.3241	4.981	6.31e-07
Dummy EBIT- (financial income and expense)	0.7538	0.2651	2.843	0.0045
Profitability deviation	0.0038	0.0009	4.183	2.87e-05
NWC/Operating assets	0.1385	0.0303	4.572	4.83e-06

The equation got is the following:

$$B = -3.4333 - 0.0488 * Current\ ratio - 3.6392 * \frac{Net\ income}{Total\ assets} + 1.6145 * \frac{Total\ debt}{Total\ assets} + 0.7538 * [Dummy\ EBIT - (financial\ income\ and\ expenses)] + 0.0038 * Profitability\ deviation + 0.1385 * \frac{NWC}{Operating}assets$$

And the success percentage table is the following:

Table 75 – X model, 2nd year before bankruptcy train dataset. Cut-off 0.11. Success table

	% correctly classified	**Sensitivity**	**Specificity**	**I type error %**	**II type error %**
1 year before bankruptcy	71.6%	94.4%	67.6%	5.6%	32.4%
2 years before bankruptcy	69.5%	85.1%	66.5%	14.9%	33.5%
3 years before bankruptcy	67.0%	71.4%	66.2%	28.6%	33.8%
4 years before bankruptcy	65.4%	68.6%	64.8%	31.4%	35.2%
5 years before bankruptcy	62.0%	68.6%	60.6%	31.4%	39.4%

Also a try to develop a further will be done using as train dataset the 3rd year before bankruptcy and as cut-off 0.18.

The coefficients for the equation are the following:

Table 76 - X model, 3rd year before bankruptcy train dataset. Coeffcients.

Variable	**Estimate**	**Std. Error**	**Z value**	**Pr(>│z│)**
(Intercept)	-3.2165	0.3312	-9.712	<2e-16
Current ratio	-0.0211	0.0122	-1.735	0.0827

Net income/total assets	-1.7811	1.1941	-1.492	0.1358
Total debt/total assets	2.2956	0.4206	5.457	4.83e-08
Profitability deviation	0.0018	0.0006	3.023	0.0025

The the equation got is here below:

$$B = -3.2165 - 0.0211 * Current\ ratio - 1.7811 * \frac{Net\ income}{Total\ assets} + 2.2956 * \frac{Total\ debt}{Total\ assets} + 0.0018 * Profitability\ deviation$$

And the success table as below:

Table 77 – X model, 3rd year before bankruptcy train dataset. Cut-off 0.18. Success table

	% correctly classified	**Sensitivity**	**Specificity**	**I type error %**	**II type error %**
1 year before bankruptcy	71.7%	92.9%	68.0%	7.1%	32.0%
2 years before bankruptcy	70.1%	81.1%	68.0%	18.9%	32.0%
3 years before bankruptcy	66.8%	69.6%	66.3%	30.4%	33.7%
4 years before bankruptcy	64.3%	65.6%	64.1%	34.4%	35.9%
5 years before bankruptcy	63.9%	61.5%	64.4%	38.5%	35.6%

All the 3 models applied had a specificity and a sensitivity slightly above the models coming from the literature.

Also, a couple of models using as variables the change of the model X variable values, as it was done in the case of the Altman, Ohlson and Beaver models.

The first equation is got using as train dataset the change of the values between 1 and 2 years before bankruptcy. The cut-off used is 0.08.

Table 78 – Delta X model, 1st year before bankruptcy train dataset. Coeffcients.

Variable	**Estimate**	**Std. Error**	**Z value**	**Pr(>│z│)**
(Intercept)	-2.5103	0.1445	-17.372	<2e-16
Delta dummy EBIT- (financial income and expense)	-0.7389	0.2808	2.632	0.0085

Delta total debt/total assets	7.6457	0.8139	9.393	<2e-16

$$B = -2.5103 + 0.7389 * [Delta\ dummy\ EBIT - (financial\ income\ \ and\ expenses)] + 7.6457 * Delta\frac{Total\ debt}{Total\ asets}$$

The success probability table is the following one:

Table 79 – Delta X model, 1st year before bankruptcy train dataset. Cut-off 0.08. Success table

	% correctly classified	**Sensitivity**	**Specificity**	**I type error %**	**II type error %**
1 year before bankruptcy	69%	86.1%	66.2%	13.9%	33.8%
2 years before bankruptcy	67.2%	70.7%	66.5%	29.3%	33.5%
3 years before bankruptcy	63.8%	52.1%	66.1%	47.9%	33.9%
4 years before bankruptcy	60.7%	48.3%	63.1%	51.7%	36.8%

The success percentage is worse compared with the ones got in the original model. Last model, that uses as train dataset, with a cut-off equal to 0.12. The coefficients are as follows:

Table 80 - Delta X model, 2nd year before bankruptcy train dataset. Coeffcients.

Variable	**Estimate**	**Std. Error**	**Z value**	**Pr(>│z│)**
(Intercept)	-1.9814	0.1100	-18.018	<2e-16
Delta NWC/Operating assets	-0.3973	0.0685	5.798	6.7e-09
Delta Profitability deviation	0.0006	0.0003	1.897	0.0578
Delta dummy EBIT- (financial income and expense)	0.6131	0.2108	2.908	0.0036
Delta total debt/total assets	5.7942	0.9444	6.135	<8.5e-10

The equation is here below:

$$B = -1.9814 - 0.3973 * Delta\frac{NWC}{Operating\ assets} + 0.0006 * Delta\ profitability\ deviation + 0.6131 * [Delta\ dummy\ EBIT - (financial\ income\ and\ expenses) + 5.7942 * Delta\frac{Total\ debt}{Total\ assets}$$

And the success table is the following:

Table 81 - Delta X model, 2nd year before bankruptcy train dataset. Cut-off 0.12. Success table

	% correctly classified	**Sensitivity**	**Specificity**	**I type error %**	**II type error %**
1 year before bankruptcy	58.7%	80.0%	55.2%	20.0%	44.8%
2 years before bankruptcy	59.1%	72.1%	56.6%	27.9%	43.4%
3 years before bankruptcy	57.9%	62.5%	57.0%	37.5%	43.0%
4 years before bankruptcy	55.3%	59.4%	54.5%	40.6%	45.5%

This model is from an overall point of view a little bit worse compared with the previous model but the sensitivity is higher. Also a model that uses as data train the 3rd year before bankruptcy was tested but the results were quite unsatisfactory so its representation will be omitted.

4.2 Best model choice

At the end, analyzing and comparing all the models seen so far together with the models based on the variables chosen by the author, the best model is the model X – data train 2 years before bankruptcy, which showed to be slightly better than the model proposed by Beaver – data train 2 years before bankruptcy.

It must be reminded to the reader that the equation is as follows, where B is the equation output:

$$B = -3.4333 - 0.0488 * Current\ ratio - 3.6392 * \frac{Net\ income}{Total\ assets} + 1.6145 * \frac{Total\ debt}{Total\ assets} + 0.7538 * [Dummy\ EBIT - (financial\ income\ and\ expenses)] + 0.0038 * Profitability\ deviation + 0.1385 * \frac{NWC}{Operating\ assets}$$

The reader should remember that a bankrupt firm has a dependant variable equal to 1, while if it is a non-bankrupt firm is 0. Using this equation, in table 82 on the left side the value B (the dependant variable), was calculated and on the right side the bankruptcy.

Before read the results, it must be underlined that the probability of being bankrupt in our overall sample is 17.2%, while to create the table below, the empty rows were eliminated and the average probability became 16.1%.

Table 82 – X model. 2nd year before bankruptcy. Cut-off 0.11. Success table by output value.

B value	**Bankruptcy probability**	**Category prob./average prob.**
B>5	91.7%	5.69
1.5<B<5	71.7%	4.45
-1.5<B<1.5	31.6%	1.96
-3<B<-1.5	12.8%	0.8
-5<B<-3	3.2%	0.2
B<-5	2.0%	0.12

While B>1.5, the probability of being bankruptcy is very high, while if B<-3 is very low. If -3<B<1.5, assigning an output to the firm gets more difficult, even if, when -1.5<B<1.5 the probability of going bankruptcy is twice the average. The following graph shows the worsening of the model performance across the years, even if probably a reader is more interested in the trend of both sensitivity and specificity. By the way, it can already observe that there is a constant decrease of the % of correctly classified data points, with a bigger drop between the prediction ability of the 4th and the 5th year.

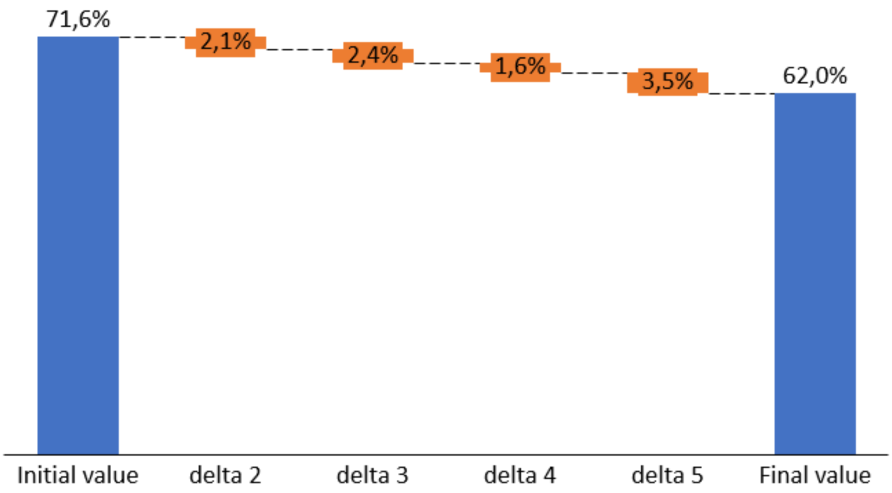

Figure 33 – X model – 2nd year before bankruptcy. Cut-off 0.11. Prediction accuracy bridge

The sensitivity and specificity performance can be more appreciated in the following graph:

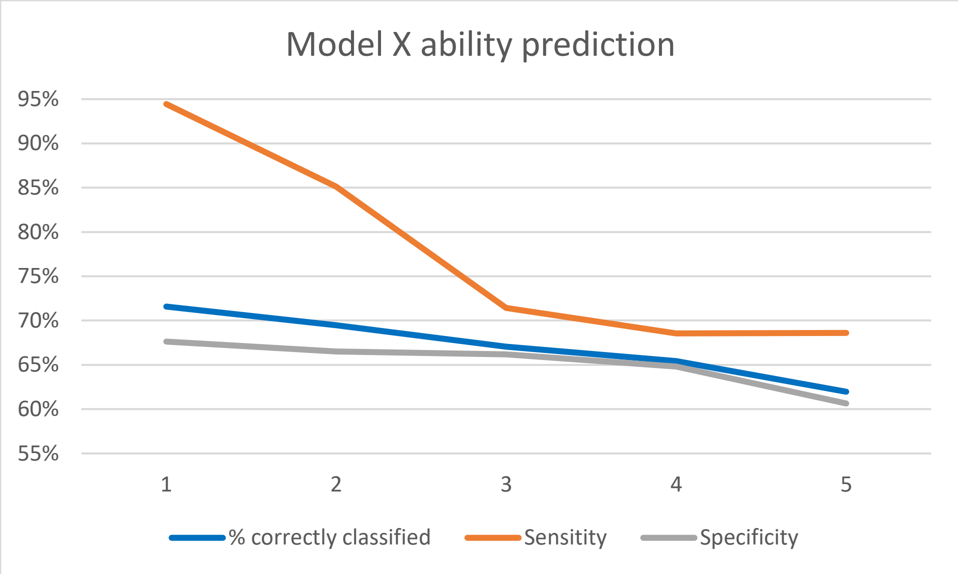

Figure 34 – X model, 2 years before bankruptcy. Cut-off 0.05. Ability prediction by year

The sensitivity has a big fall in since the 1st to the 3rd year (decreasing from 94.4% to 71.4%) but its magnitude is so small compared to the specificity that it doesn't influence too much the % of correctly classified line (that moves only from 71.6% to 67.0%). Instead, during last year the decrease of specificity from 64.8% to 60.6% made the precision of the overall model drop from 64.8% to 60.6%.

The X model is better due to some reasons:

- Models coming from the literature were developed using the annual reports of firms coming from the 60's and the 70's, while the firms analyzed in this thesis come from the 2010s. The consequence is that actual firms have a different balance sheet and income statement composition. The biggest change derives from the relatively higher use of debt compared to equity than in the past.
- The other models were applied on middle-size American firms while this analysis is made on Italian medium and small firms, not listed in a stock exchange, with a limited access to capital and with a different management and organizational culture.
- Since the variables are tested on the sample before being used to predict the bankruptcy, the variables chosen will be obviously better than the ones in thevmodels created from different datasets.
- Also, a higher focus in this thesis was from the point of view of the sensitivity. Notwithstanding the fact that the number of the failed firms was relatively low, we cared more about having a high sensitivity more than the specificity and at the same time caring about not have them decrease under certain thresholds.

5. CHAPTER FIVE

Analysis of the CNDCEC ratios

5.1 Introduction

A quick analysis also of the ratios proposed by the CNDCEC in October 2019 will be done. In this analysis there are 2 limitations:

- 6 months DSCR cannot be calculated since six months values are not available;
- A model similar to the previous ones must be calculated. We know the thresholds that cannot be exceeded by the ratios but, since the limit is different by sector, all the sectors of the firms considered should be known and this is not the case.

Table 83 – accuntants model: number of firms with negative equity by year

Years before bankruptcy	Number of firms with negative equity
5	49
4	50
3	59
2	98
1	186

As a first step, even if the DSCR cannot be calculated, only the negative equity value will be investigated in table 83. Approaching the year of bankruptcy, the number of firms with a negative equity skyrockets. These firms have already a high risk to be in a financial distress situation but legally this cannot be declared as such without the analysis of the DSCR.

5.2 Model test

The best model among all the ones possible was the one with train dataset the 2nd year before bankruptcy and as cut-off 0.18, slightly best than the ones that use the dataset coming respectively from the 1st and the 3rd years before bankruptcy. The coefficients are represented in the table below.

Table 84 - accountants model,, 2nd year before bankruptcy train dataset. Coefficients

Variable	Estimate	Std. Error	Z value	Pr(>│z│)
(Intercept)	-0.7895	0.1272	-6.204	5.51e-10
Financial interests substainability ratio	0.0689	0.0456	1.510	0.131

Capital adequacy ratio	-2.8439	0.3553	-8.004	1.21e-15
Assets liquid turnover ratio	-0.0285	0.0208	-1.368	0.171
Liquidity ratio	-0.0272	0.0221	-1.227	0.220
Social security and tax debt ratio	-0.0216	0.0182	-1.188	0.235

Only the capital adequacy ratio is really significant. Despite this, removing more ratios would worsen the quality of the model. It is unexpected the fact that a higher social security and tax ratio leads to a lower probability to fail.

The equation got is the following one:

$$Y = -0.7895 + 0.0689 * Financial\ interests\ substainability\ ratio - 2.8439 * Capital\ adequacy\ ratio - 0.0285 * Asset\ liquid\ turnover\ ratio - 0.0272 * Liquidity\ ratio - 0.0216 * Social\ security\ and\ tax\ debt\ ratio$$

The table below contains the accuracy percentage:

Table 85 – accountants model,, 2nd year before bankruptcy train dataset. Cut-off 0.18. Success table

	% correctly classified	**Sensitivity**	**Specificity**	**I type error %**	**II type error %**
1 year before bankruptcy	68.0%	88.7%	64.5%	11.3%	35.5%
2 years before bankruptcy	66.4%	75.7%	64.7%	24.3%	35.3%
3 years before bankruptcy	63.2%	62.4%	63.4%	37.6%	36.6%
4 years before bankruptcy	61.1%	59.6%	61.4%	40.4%	38.6%
5 years before bankruptcy	61.1%	56.3%	62.0%	43.8%	38.0%

Still, sensitivity and specificity are much lower compared with the best models. This is probably due to the fact that the indicators don't contain any ratio that deals with the operating performance and this kind of ratios showed to be the most helpful ones in bankruptcy prediction. Despite this, the ratios proposed are not aimed at being put together to create a model but each of them underlines a particular aspect that provides a hint about the financial situation of the firm. All together they can help a stakeholder if the firm is passing a difficult moment.

Conclusions

Finally, making a generalized brief thought about the use of accounting based model in the prediction of firms' bankruptcy, basing the observations on this work, it can be observed that the accounting based models are still an effective way to predict the future of a company.

What was found is that the logit model is more effective compared to the multivariate discriminant analysis. The ineffectiveness of the second method was partially solved by using a weight on the bankrupt firms in order to increase the severity associated to the wrong classification of a bankrupt firms but still was not efficient like the logit model. Moreover, the original ZETA, Z' and Z'' Altman models without any adjustments didn't work. An example about the effectiveness of the logit model comes from the fact that the cut-off to be applied for the Ohlson model was provided and it showed to be effective.

Finally, the model proposed by the CNDCEC didn't have a great prediction accuracy but anyway they are useful in the comprehension of the financial situation of a firm and, in particular, it can help the management understand the criticalities that hit the firm.

Bibliography

Adalet McGowan, A. & Andrews, D., 2017. "Declining resource allocation in Spain: Implications for productivity. *OECD economic department working papers*.

Adalet McGowan, M., Andrews, D. & Millot, V., 2017. "The walking dead? Zombie firms and productivity performance in OECD countries". *Economics department working papers*, Issue 1372.

Agarwal, V. & Taffler, R. J., 2008. "Comparing the performance of market-based and accounting-based bankruptcy prediction models". *Journal of Banking & Finance*, Issue 32, pp. 1541-1551.

Altman, E., 1968. "Financial ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy". *Journal of Finance*, September, Volume 23, pp. 589-609.

Altman, E., 2000. *Predicting financial distress of companies: revisiting the z-score and zeta model,* s.l.: s.n.

Altman, E., Danovi, A. & Fallini, A., 2013. "Z-score models' application to italian companies subject to extraordinary administration". *Journal of applied finance*, Issue 1.

Andrews, D., Criscuolo, C. & Gal, P., 2016. "The global productivity slowdown, technology divergence and public policy: a firm level perspective". *OECD Productivity working papers*, Issue 5.

Beaver, W., 1966. Financial ratios as predictors of failure. *Journal of accounting financce*, Volume 4, pp. 71-111.

Begley, J., Ming, J. & Watts, S., 1996. Bankruptcy classification errors in the 1980s: an empirical analysis of Altman's and Ohlson's Models. *Review of accounting studies*, Issue 1, pp. 267-284.

Black, F. & Scholes, M., 1973. "The pricing of options and corporate liabilities". *Journal of Political economy*, Issue 7, pp. 637-654.

Blum, M., 1974. "Failing discriminant analysis". *Journal of accounting research*, Volume 12, pp. 1-25.

Caballero, R. J., Hoshi, T. & Kashyap, A. K., 2008. "Zombie lending and depressend restructuring in Japan". *The amerrican economic review*, Volume 84 (5), pp. 1350-1368.

Deakin, E., 1972. "A discriminant analysis of predictors of business failure". *Journal of accounting research*, 10(1), pp. 167-179.

El-Ansary, O. & Bassam, L., 2019. "Predicting financial distress for listed MENA firms". *International journal of accounting and financial reporting*, 9(2).

Fontoura Gouveia, A. & Osterhold, C., 2018. "Fear the walking dead: zombie firms, spillovers and exit barriers". *OECD productivity working papers*, June, Issue 13.

Garrido, . J., Kopp, E. & Weber, A., 2016. "Cleaning-up bank balance sheets: economic, legal and supervisory measures for Italy". *IMF working paper*, Issue 135.

Hernandez T., M. & Wilson, N., 2013. "Financial distress and bankruptcy prediction among listed companies using acconting, market and macroeconomic variables". *International review of Financial Analysis*, Issue 30, pp. 394-419.

Hillegeist, S. A., Keating, E. K., Cram, D. P. & Lundstedt, K. G., 2004. "Assessing the probability of bankruptcy". *Review of accouting studies*, March, pp. 5-34.

Horrigan, J., 1966. The determination of long-term credit standing with financial ratios. *Journal of accounting research*, Volume 4, pp. 44-62.

James, G., Witten, D., Hastie, T. & Tibshirani, R., 2013. *"An introduction to statistical learning".* New York: Springer.

Jensen, M. C., 1989. "The Eclipse of the Public Corporation.". *Harvard Business Review*, September-October, pp. 61-74.

Joy, M. & Tollefson, J., 1975. "On the financial applications of discriminant analysis". *The journal of fnancial and quantitative analysis*, December, 10(5), pp. 723-739.

Kaludjerovic, N., Stanojevic, S. & Ljubic, M., 2016. "Hidden losses in financial reporting and the manner of hiding case serbia - part two. *International review*, Issue 1-2.

Kassambara, A., 2018. *"Machine learning essentials - practical guide in R".* 1 ed. s.l.:s.n.

Korea, B. o., 2013. *"Financial stability report"*, s.l.: s.n.

Merton, R. C., 1974. "On the pricing of corporate debt: the risk structure of interest rates". *Journal of Financce*, Issue 29, pp. 449-470.

Moyer, C., 1977. *Forecasting financial failure: a re-examination*, s.l.: s.n.

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

Peek, J. & Rosengren, E. S., 2005. "Unnatural selection: perverse incentives and the misallocation". *The American Economic Review*, 95(4), pp. 1144-1166.

Pham, B., Do, T. & Vo, D., 2018. "Financial distress and bankruptcy prediction: An appropriate model for listed firms in Vietnam". *Economic System*, Issue 42, pp. 616-642.

Pongsatat, S., Ramage, J. & Lawrence, H., 2004. "Bankruptcy prediction for large and small firms in Asia: a comparison of Ohlson and Altman". *Journal of accounting and corporate governance*, Volume 1, pp. 1-13.

Reisz, A. S. & Perlich, C., 2004. *"A market based framework for bankruptcy prediction".* New York: Baruch College, City university of New York.

Storz, M., Koetter, M. & Westphal, A., 2017. "Do we want these two to tango? On zombie firms and stressed banks in Europe". *ECB working papers series*.

Whitaker, R. B., 1999. "The early stage of financial distress". *Journal of Economics and Finance*, Issue 23, pp. 123-133.

Sitography

Anon., 2013. *Stackoverflow.* [Online]
Available at: https://stackoverflow.com/questions/17200114/how-to-split-data-into-training-testing-sets-using-sample-function
[Accessed 2019 August 2019].

Anon., 2019. *"Crisi d'impresa - Indici di allerta dei commercialisti"*, s.l.: Il sole 24 ore documenti.

Anon., n.d. *Business dictionary*. [Online]
Available at: http://www.businessdictionary.com/definition/bankruptcy.html
[Accessed 16 July 2019].

Anon., n.d. *Discussion issues and derivations.* [Online]
Available at: http://people.stern.nyu.edu/adamodar/New_Home_Page/AppldCF/derivn/ch5deriv.html
[Accessed 19 August 2019].

Anon., n.d. *Investopedia.* [Online]
Available at: https://www.investopedia.com/terms/f/fundsfromoperation.asp
[Accessed 16 August 2019].

Anon., n.d. *Piacentini & associati.* [Online]
Available at: http://www.piacentinieassociati.it/crisi-dimpresa/
[Accessed 2019 novembre 9].

Anon., n.d. *Psyche Scene Hub.* [Online]
Available at: https://psychscenehub.com/psychpedia/odds-ratio-2/
[Accessed 23 July 2019].

Anon., n.d. *RIVALUTA.it.* [Online]
Available at: https://www.rivaluta.it/serie-inflazione-media.asp
[Accessed 19 August 2019].

CNDCEC, 2019. *IPSOA.* [Online]
Available at: https://www.ipsoa.it/documents/impresa/fallimento-e-procedure-concorsuali/quotidiano/2019/10/28/crisi-impresa-pubblicato-documento-indici-allerta
[Accessed 28 October 2019].

Anon., n.d. *Rpubs.* [Online]
Available at: https://rpubs.com/Nolan/298913
[Accessed 2019 August 24].

Bragg, S., 2018. *Accounting tools.* [Online]
Available at: https://www.accountingtools.com/articles/what-is-funds-from-operations.html
[Accessed 16 August 2019].

Kenton, W., 2019. *Investopedia.* [Online]

Available at: https://www.investopedia.com/terms/f/financial_distress.asp

[Accessed 16 July 2019].

Lozzi, F., 2019. *IPSOA.* [Online]

Available at: https://www.ipsoa.it/documents/impresa/fallimento-e-procedure-concorsuali/quotidiano/2019/01/02/crisi-impresa-indicatori-presunta-capacita-predittiva

[Accessed 3 November 2019].

Martins, C., n.d. *knoow.* [Online]

Available at: http://knoow.net/en/economics-business/finance/defensive-assets/

[Accessed 22 July 2019].

Petruzzellis, G., 2019. *IPSOA.* [Online]

Available at: https://www.ipsoa.it/documents/impresa/fallimento-e-procedure-concorsuali/quotidiano/2019/10/26/crisi-impresa-indici-allerta-approccio-gerarchico

[Accessed 02 November 2019].