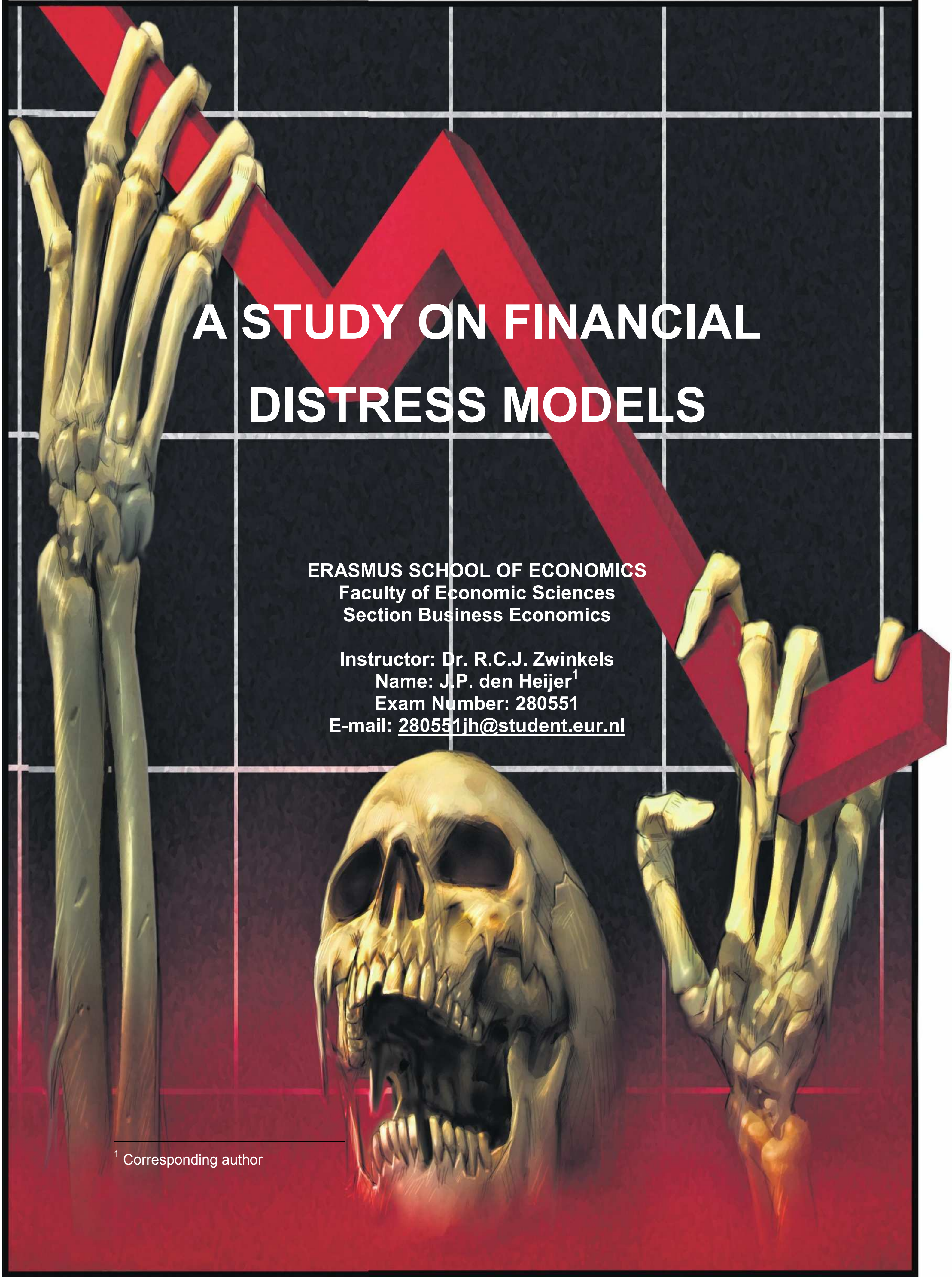
LAP



A STUDY ON FINANCIAL DISTRESS MODELS

ERASMUS SCHOOL OF ECONOMICS
Faculty of Economic Sciences
Section Business Economics

Instructor: Dr. R.C.J. Zwinkels
Name: J.P. den Heijer[1]
Exam Number: 280551
E-mail: 280551jh@student.eur.nl

[1] Corresponding author

Acknowledgements

A sign of gratitude is in order, since my graduation on this research paper would not have been possible without the support of at least a few people. Firstly, I would like to thank all my friends for putting up with my endless discussions on the causes of the financial crisis as related to this paper, also my sister Astrid for being there when it seemed like this paper was never going to finish. I would like to thank my fiancée Lucia for having unrelenting faith in me and putting up with my absence while I was locked away in my study. Lastly but certainly not the least important, I would like to express my gratitude to the Erasmus University Rotterdam for having me as a student, and more specifically my instructor in this research, dr. Remco Zwinkels who was there to give me that last push in the right direction for this paper to be completed.

Summary

This paper investigates the evolution of financial distress likelihood models over the course of the past five decades, as motivated in the academic world by the financial crisis of 2008. We put this in a capital framework, more specifically the existence of theory on information-asymmetry. A firm may always have debt holders as well as equity holders. All stakeholders within a firm will act in such a way that it protects their own investment in the firm as best as possible. This could cause one stakeholder party to protect its own investment at the expense of the other. In our research, these are debt holders versus equity holders.

We attempt to illustrate that financial distress likelihood models have become more advanced over time, with the result that their ability to accurately predict financial distress has improved. We select the Altman model of 2000, the Ohlson model of 1980 and finally the Pindado model of 2008. Replicating all three models on a dataset which samples the period of 1990 – 2010 on all G-7 countries, excluding the USA, we find that the latter model has the best overall performance when it comes to correctly predicting financial distress likelihood. We find similar results in our out-of-sample data of the USA.

It is noteworthy that especially the new, logarithmic models of Ohlson and Pindado have far superior results in accurately predicting financial distress likelihood. We do however not find any significant increases in financial distress likelihood in the years directly following the financial crisis. This could be mitigated due to an inherent bias in the dataset acquired from CompuStat. We conclude that the more advanced models on predicting financial distress, being econometric of nature, appear to be the most reliable models available present day.

Index

1. Introduction

1.1 Relevance

At the end of 2007, the financial crisis struck the United States. Within less than a year, the economies of the European Union had become affected as well, quickly followed by the rest of the world (Pugh, 2009). The world's top economists agree that the financial crisis of 2007 is the worst crisis since the Great Depression of 1929. In the United States alone, the government invested several hundreds of billions of dollars in order to save the banking system and to mitigate the damage caused by the financial crisis to the world economy (Pendery, 2009; Hilsenrath *et al.*, 2008).

Therefore it is only logical to see that politicians, academics and financial analysts alike have taken a steep interest in the development of models which measure financial distress likelihood – i.e. the firm's inability on the short term to fulfill its liabilities and other dues (Andrade and Kaplan, 1998). This partly as there is worldwide disappointment with the rating agencies, since their models are not specifically tailored to estimate the probability of financial distress occurring (Pindado *et al.,* 2008). The models these rating agencies have employed, however, merely take into the account the probability of bankruptcy, which is in effect excluded to an extreme form of financial distress. This while academics have attempted to identify other variables which may have statistical relevance to predicting the probability of financial distress occurring as can be found in Grice and Dugan (2001), Grice and Ingram (2001), as well as Begley *et al.* (1996). The academic literature has brought forth an evolution of various models to estimate the probability of financial distress occurring over the course of the last five decades, such as Altman (1968; 1984), Ohlson (1980), Zmijewski (1984) and Pindado *et al.* (2008). This evolution started with a discriminatory model by Altman (1968), which in essence, took a variety of classic financial ratios[2] which are considered to be statistically related to predicting financial distress likelihood. Over time, adjustments were made to the aforementioned model (Altman, 2000) based on various criticisms from academics and financial analysts as in Dichev (1998); Grice and Dugan (2001); Grice and Ingram (2001) and Bhagat *et al.* (2005). As better computer technology became available, researchers were able to build more sophisticated models, as can be found in Ohlson (1980) and Pindado *et al.* (2008). These advanced models aim to be more reliable in predicting financial distress likelihood than their predecessors. This study is to determine to what degree they have succeeded in achieving this goal.

1.2 Research question

The study of financial distress is part of the domain of finance, and within that, capital structure. As has become widely known by the current financial crisis, financial markets are far from perfect. Therefore the theory stating that a firm's capital structure is of no relevance to the firm's valuation as stated in Modigliani and Miller (1958) – which is a factor affected by financial distress as argued by Pindado *et al.* (2008) – is to be considered inapplicable in practice.

Thus this research paper is to examine the reliability in prediction power of three existing financial distress models[3]. Doing so would allow any findings from employing these three very different models to be placed into an existing theoretical framework – that of capital structure, that is – and therefore yield relevant insights into the use, accuracy and relevance of aforementioned financial distress models.

[2] Those being the working capital ratio, retained earnings ratio, earnings ratio, debt ratio, sales-to-assets ratio.
[3] These models are from Altman (1968), Ohlson (1980) and Pindado *et al.* (2008), respectively.

Now, with the aforementioned, we can state our research question as follows:

Q1: is there any significant difference in the reliability of financial distress likelihood prediction between the three models?

With this research question, backed by a well-structured theoretical framework, one might be able point the academic world in which direction it ought to research further with respect to the subject of financial distress; this by taking exclusively into account the one model which would be most reliable in predicting financial distress likelihood. As the three models employed vary in their manner of calculation and variables - the simplest being a multiple discriminant analysis model based on four financial ratios to the most advanced consisting of a variety of econometric formulae including a dynamic component – one may wonder whether the most complicated one actually yields the best results. Putting these three models to the test would bid outcome to aforementioned. This then yields the following hypothesis:

H1: the financial distress likelihood model of Pindado *et al.* (2008) is superior to the other two models in reliability of predictive power.

Bernanke *et al.* (1988) amongst others, find that during the eighties corporations have relied more and more on debt financing over equity financing, which has resulted in higher leverage ratios. Researchers have found a variety of arguments for the preference of debt financing, most notably Myers (1984), which motivates a pecking order theory. This theory states that a firm will chose debt over equity when debt is considered cheaper[4], taking into account all its advantages, such as an interest tax shield. This is a trend which has continued to present day. Chandra and Nayar (2008) find that debt financing is also subject to information-asymmetry, as debt financing is generally received as a positive effect on the stock price of the firm in question. Lenders tend to have private information on the firm after debt has been issued, prior to a performance decline, thus allowing them to set conditions to protect their investment, usually at the cost of equity holders. Since this is a trend which has persisted until present day, it is possible that debt financing is related to the financial crisis. If so, then we should see an increase in financial distress likelihood due to defaults upon debt in the years following the financial crisis. Thus we arrive at our second and third hypotheses:

H2: the capital structure of a firm is of influence on its likelihood to become financially distressed.

H3: Information-asymmetry within publicly-traded firms is of effect on financial distress likelihood when relevant information becomes public knowledge.

To specify further upon these hypotheses I would like to add that a robustness check would be necessary to investigate said hypothesis. Therefore we will acquire data from G-7 countries, excluding the USA, and perform an out-of-sample test using a dataset for the USA[5].

The remainder of this research paper is structured as follows: firstly we will discuss the theoretical framework in which this research is to take place. Then we will discuss the methodological aspects as well as the data. This

[4] The cost of debt is to be considered the total sum of interest costs, banking fees and the like.
[5] For sake of reference, we will name our dataset “G-7 dataset” from here on, and the out-of-sample set “USA dataset”.

followed by a discussion of the results and its implications with respect to the aforementioned research question and hypotheses, which will result in our conclusion.

2. Theoretical framework

In order to have a sound empirical study, some existing theories are required in order to place one's findings. The capital framework of firms is discussed as this may be of effect on financial distress probability. Underlying to aforementioned, agent theory - or more specifically the theory on information-asymmetry - is taken into consideration as it will elaborate further upon the motivations agents may have with respect to capital structure theory.

2.1 The capital framework

The theory surrounding the capital framework of the firm elaborates upon one of the main motivations for firms to attract debt: it allows for the creation of a tax-shield by means of deducting interest payments on debt from the profit before taxes, effectively decreasing taxable income. As such, firms consider a trade-off between the costs of (the risk of) financial distress and said tax advantage (Kim, 1978). Each of the three financial models in this research considers a (series of) variable(s) which are subject to accounting manipulation due to aforementioned trade-off. Such variables could concern itself with aspects like solvability, liquidity, profitability and the like.

Another phenomenon is that firms tend to keep a certain degree of financial flexibility, or a financial buffer, with respect to their capital structure in order to decrease the probability of financial distress. If the public firm in question happens to be quoted on a stock exchange, another two factors to be taken into account is that said firms tend to time their new issues of debt and equity based on the current stock price. We will elaborate further upon timing of debt issues in the next paragraph. Furthermore, Graham and Harvey (2001) and Brounen *et al.* (2005) find that debt can be held to prevent any corporate takeovers, thus functioning as a strategic defensive measure rather than being financially motivated. Jensen and Meckling (1976) find that debt can also be used as a disciplining device for managers who merely have part ownership or none at all in the firm they operate in.

These findings, irrespective of being strategic or financial in nature, may give rise to gaming behavior (Chandra and Nayar, 2008). Therefore we are dealing with the possibility of principal-agent theory in respect to the probability of financial distress occurring. In other words, one group of investors may try to protect their investment at the expense of another group of investors in times of financial distress.

2.2 Information-asymmetry

In the financial distress model of Pindado *et al.* (2008) a dynamic component is taken into consideration as an individual effect. This makes the Pindado model stand out from the Ohlson and Altman models, which do not take this effect into account. Therefore we can state that the Pindado model specifies the effects of information-asymmetry within its model, thus increasing its accuracy in predicting financial distress likelihood.

Studies from Carey *et al.* (1993); Kwan and Carleton (2004) and Denis and Mihov (2003) find that publicly traded firms are to some degree subject to information-asymmetry. Furthermore, Wittenberg-Moerman (2006) finds that publicly-traded firms suffer from a lesser degree of information-asymmetry than would privately-held firms. This is because the former is legally obligated to make financial statements publicly available where the latter is not so. Since employees within the firm tend to know certain information relevant to the firms financial health before outsiders do, it is typically referred to as *inside information*. Its effect on the interest rate of debt equity is depicted in figure 1 below.

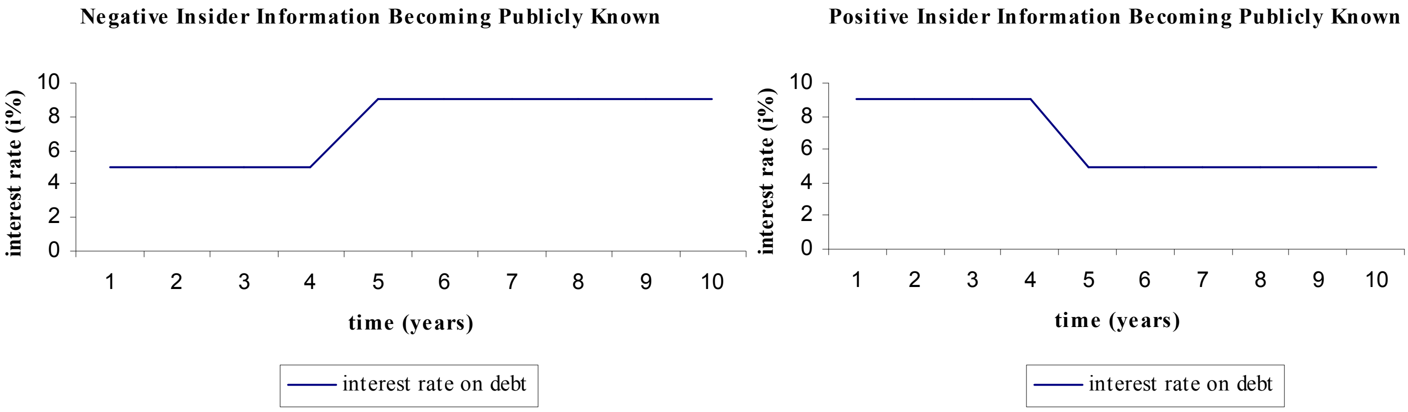


Figure 1: *in the left graph we can see that in the fourth year negative insider information becomes publicly known after the firm in question purposely delayed its publication, causing the interest rate on debt equity for the firm to go up significantly due to increased risk. In the right graph we can see the opposite happening: here, in the fourth year positive insider information is immediately made publicly known, allowing the firm from then on to acquire debt against a more favorable interest rate.*

Because inside information is not known to the public at the same time as it is to employees within the firm, there is the issue of information-asymmetry. Chandra and Nayar (2008) find that insiders tend to exploit this information-asymmetry. For example, if insiders know the firm will perform less in the near future, the firm will attempt to attract as much debt capital as possible – so, before t = 4 in figure 1, left panel. A firm would do this to prevent that it would have to pay a higher interest rate upon said debt if it would have attracted this debt in the future, simply because its financial position has worsened. Intuitively, the opposite is valid as well: firms would be compelled to publish news which positively affects the financial position of the respective firm as soon as possible to gain more favorable interest rates on debt capital – thus it will attract debt if t > 4, see right panel of figure 1 - *ceteris paribus*.

The financial distress models of Ohlson (1980) and Altman (1968; 2000) do not take this dynamic variable of timing debt issuance into consideration. In contrast, the Pindado (2007) model attempts to incorporate the time-effect in one of its two conditions for a firm to be classified as financially distressed:

(1) **EBITDA is lower than financial expenses for two consecutive years.**

Ohlson (1980) employs a dummy variable similar to this condition. It states that if net income was negative for the last two years, this variable takes the value of one. In any other case, it takes the value of zero. Such classification is rather coarse, especially on an international comparative scale as it fails to take into account any differences in national and local tax structures, as well as putting asset-heavy industries at a potential disadvantage in comparison to other industries – the latter due to annual heavy depreciation. In contrast, Pindado (2007) takes these criticisms into account and instead takes the underlying two variables to this dummy variable, replacing EBIT by EBITDA.

(2) **A fall in market value occurs between two consecutive periods.**

This is the dynamic aspect of the Pindado model. The underlying assumption of this second condition is derived from the first one: if a firm has consecutively bad results (read: two consecutive years of net losses), this will have its effect on the market value of the company. The dynamic aspect lies in the fact that when both conditions are met, the following year the firm in question will suffer from the consequences of financial distress. It will continue to do so until it manages to show improved results, provided it does not file for bankruptcy.

These two time-effects allow the modeling of the potential gaming behavior of firms with respect to information-asymmetry as discussed by Wittenberg-Moerman (2006). Note however that this only applies to the Ohlson and Pindado models only, which have the effects of information-asymmetry specified. In contrast, the Altman model does not explicitly specify said effects.

Since this gaming behavior is part of the models specifications, it allows us to indirectly relate our investigation of the effectiveness in the ability to predict financial distress likelihood to said behavior. Thus in years following the financial crisis, we may find a sharp increase in financially distressed observations, as stakeholders attempt to protect their investments at the cost of other stakeholders.

3. Methodology

The following section describes the manner in which this research is executed. Secondly, we will discuss how we acquired the dataset as well as a secondary dataset to perform an out-of-sample test later on. Lastly, it also gives a detailed overview of the actual financial distress probability models employed.

3.1 Research setup

Following Pindado *et al.* (2008), this research will be executed by using a longitudinal, unbalanced panel study. Although the other two models in this research, these of Altman (1968; 2000)[6] and Ohlson (1980) are however static in nature, a longitudinal study is necessary to properly compare the model of Pindado *et al.* (2008) with the other two. Without it, the dynamic component of the latter model would become obsolete, therefore rendering its added value effectively useless.

Following Babbie (2004), in order to meet the requirements of an unbalanced panel study, we acquired data for this research from CompuStat Global and CompuStat North America databanks for G-7 countries. To perform an out-of-sample test to verify our findings based on said dataset for any of the three financial distress probability models, we have also acquired data for the USA from the CompuStat North America databank. The latter dataset will undergo a similar analysis as the former, allowing us to immediately verify the robustness of our findings throughout this paper.

To have a broad range of observations[7] we have acquired data for the years 1990 till the end of 2010 for both datasets. Additional criteria added during the data selection procedure are the following:

- The firms classified under SIC categories 4 and 6, those being utilities and logistic firms in category 4, and financial services in category 6, respectively are eliminated from the dataset. As utilities are usually monitored or controlled by governmental institutions, their business structure and their financial distress environment significantly differs from other firms (Ohlson, 1980). Financial firms are removed due to their different accounting methods, potentially resulting in biased comparisons.

- The firm had or has[8] to be operational in any of the G-7 countries. Those being France, Germany, Italy, Japan, UK, Canada and the USA. Note that the USA dataset is kept as an out-of sample dataset. We will report our findings as G-7 excluding the USA, as well as our findings from the USA dataset.

- The firm had or has to be publicly traded on a global stock exchange or an over-the-counter (OTC) market. This criterion is added due to the fact that we acquired our data from CompuStat, which exclusively records publicly traded firms and thus in effect, excludes small, privately held firms.

[6] In this study, we replicate the Altman (2000) model only. The original, Altman (1968) model is exclusively tailored to manufacturing firms, whereas the Altman (2000) model is an updated version of the original model, taking into account the dominance of service-oriented firms in modern G-7 country economies.
[7] For a logistic regression model such as that of Pindado (2008) to function properly, a large dataset is a prerequisite. Secondly, a large number of observations strengthen the results of our findings, effectively making them more reliable due to increase in predictive power. Secondly, it will allow us to verify whether the older models, being those of Altman (1968; 2000), and Ohlson (1980) can handle large volumes of data while retaining their predictive qualities. The older models were both designed with a more limited amount of observations.
[8] Depending whether the firm is still operational at the end of 2010 which is when our dataset ends, or whether the firm has terminated its operational activities somewhere during the period 1990 – 2010.

- The firm had or has to have at least ten consecutive years of data. The number is arbitrarily chosen. This to allow for a sufficient timespan for financial distress to manifest.

With these criteria, we acquire a total of 3.251 companies (57.113 observations) for the G-7 excluding the USA. 1.592 companies (26.111 observations) are in our USA dataset, which is used for our out-of-sample test, consisting of USA-based companies only. The data has to be processed by applying each of the three models, thus effectively yielding the dataset we are actually interested in. Table 1 summarizes the details of both datasets.

Table 1 - Number of observations per year, per country of the G-7
Time period: 1990 - 2010

	Number of observations per country								
Year	**Canada**	**Germany**	**France**	**United Kingdom**	**Italy**	**Japan**	**G-7 excluding USA**	**USA**	**Total G-7**
1990	76	149	110	452	22	1.992	2.801	1.563	4.364
1991	76	147	111	452	21	2.056	2.863	1.566	4.429
1992	76	147	111	452	23	2.111	2.920	1.571	4.491
1993	76	148	111	463	26	2.117	2.941	1.572	4.513
1994	77	152	120	468	39	2.161	3.017	1.578	4.595
1995	77	155	122	500	40	2.153	3.047	1.582	4.629
1996	77	174	128	603	58	2.152	3.192	1.586	4.778
1997	77	177	130	635	59	2.147	3.225	1.586	4.811
1998	77	181	135	643	60	2.149	3.245	1.589	4.834
1999	63	177	128	595	58	2.131	3.152	1.477	4.629
2000	52	175	125	556	54	2.101	3.063	1.353	4.416
2001	45	160	115	541	54	2.068	2.983	1.259	4.242
2002	44	144	110	522	53	2.010	2.883	1.186	4.069
2003	38	130	104	496	51	1.954	2.773	1.118	3.891
2004	35	123	101	476	50	1.912	2.697	1.061	3.758
2005	34	115	96	451	49	1.861	2.606	985	3.591
2006	30	109	95	419	48	1.828	2.529	919	3.448
2007	24	110	93	389	46	1.778	2.440	848	3.288
2008	21	104	90	364	43	1.721	2.343	799	3.142
2009	21	100	86	341	41	1.675	2.264	767	3.031
2010	0	2	4	48	1	74	129	146	275
Total	**1.096**	**2.879**	**2.225**	**9.866**	**896**	**40.151**	**57.113**	**26.111**	**83.224**

Table 1: *our main dataset is labeled as the G-7 countries, excluding the USA. The G-7 consists of Canada, Germany, France, the UK, Italy and Japan. The USA dataset is kept separate as an out-of-sample dataset. For each country, we report the number of observations per year rather than the number of companies. We are interested in the characteristics of financial distress; henceforth we are interested in a sufficient number of observations. We employ a time span of 20 years for both the G-7 dataset, as well as the USA dataset.*

An overview of the variables selected from the CompuStat and CompuStat North America databanks can be found in Appendix A at the end of this research paper.

Secondly, we will display the results of the replicated three models. Since the models discussed are all based on a separate methodology, we cannot simply compare such statistics as R-square between the three models. Therefore it is prudent to directly compare the outcomes of the models, being their accuracy in correctly predicting financial distress likelihood over the time period investigated. This, too, we will check with both the G-7 dataset as well as the USA dataset.

3.2 The three models

We will elaborate upon the three financial distress probability models in chronological order. This is to depict the actual evolution in financial distress models. Firstly, we have Altman's Z-score model, which is a discriminant analysis model. We have taken the Altman (2000) edition, as it is reviewed to accustom service firms as well, while Altman (1968) was more tailored to production firms only. It is generally denoted as[9]:

$$Z = 0{,}012\ X_1 + 0{,}014\ X_2 + 0{,}033\ X_3 + 0{,}006\ X_4 + 0{,}999\ X_5 \quad (1)$$

Where:

- Z = the Z-score for publicly-traded firms
- X_1 = (current assets – current liabilities) / total assets
- X_2 = retained earnings / total assets
- X_3 = earnings before interest and taxes (EBIT) / total assets
- X_4 = equity / total liabilities
- X_5 = sales / total assets

Deakin (1972) finds that the financial ratios stated above are statistically significant predictors of financial distress. The then eventual Z-score can take a variety of values matching the following criteria:

- $Z > 2{,}99$ is the domain in which a firm is considered safely away from financial distress
- $1{,}81 < Z < 2{,}99$ is when a firm has a chance to become financially distressed
- $Z < 1{,}81$ is the domain in which a firm is considered financially distressed

Pindado *et al.* (2008) argue however that these values are not logical to be interpreted as would an actual probability, which could be expressed as a percentage. Thus in effect, the Altman (2000) model gives no absolute guarantee that if a firm has a Z-score of say, 2,99, that it is free from any risks of financial distress. In contrast, Ohlson (1980) developed a logistic analysis model based on nine variables he derived from literature stating said variables are of influence in a firm's bankruptcy (Grice and Dugan, 2003). Ohlson (1980) logistic analysis model is denoted as follows:

$$Y = 1 / [\ 1 + e^{-1{,}3 - 0{,}4 X_1 + 6{,}0 X_2 - 1{,}4 X_3 + 0{,}1 X_4 - 2{,}4 X_5 - 1{,}8 X_6 + 0{,}3 X_7 - 1{,}7 X_8 - 0{,}5 X_9}\ ] \quad (2)$$

[9] Note that the Z-score model has no intercept, which is uncommon for regression models. Altman (2000) states that the statistics package used did not allow for it. The intercept was however taken into account in the criteria as described above.

Where:

Y = overall index, i.e. probability of bankruptcy
X_1 = log (total assets / GNP price-level index)
X_2 = total liabilities / total assets
X_3 = working capital / total assets
X_4 = current liabilities / current assets
X_5 = 1 (if total liabilities > total assets), otherwise $X_5 = 0$
X_6 = net income / total assets
X_7 = funds provided by operations / total liabilities
X_8 = 1 (if net income < 0 over the last two years), otherwise $X_8 = 0$
X_9 = measure of change in net income

To which the dependent variable is held subject to the following criteria:

- When $Y \geq 0,5$ the firm is classified as bankrupt
- When $Y < 0,5$ the firm is classified as healthy

Ohlson (1980) states no other reason for the cutoff at 0,5 other than implicitly assuming that the function is symmetric across bankrupt and healthy firms.

The Ohlson model is a clear advance compared to Altman's model: it uses a wider variety of variables, although it yet tends to utilize the same theoretical framework with respect to the choice of the above variables. Another perk of the Ohlson model is that it is easy to interpret: it will give a percentage score on a scale of zero to one hundred to assign the probability of bankruptcy occurring. Typically, this is the kind of models used till present day by the rating agencies. The criticism given on the Ohlson model is that it exclusively focuses on the probability of bankruptcy, which is in effect an extreme form of financial distress. It therefore ignores any lesser form of financial distress which in turn could lead to bankruptcy however if not managed properly, which is exactly the cause of the current financial crisis. As an answer to this we finally have Pindado *et al.* (2008) with a logistic regression model. It is generally denoted by the following set for formulae:

(3) $$\text{Log}\ (P(\text{event}) / P(\text{no event})) = \beta_0 + \beta_1\ EBIT_{it} / RTA_{it-1} + \beta_2\ FE_{it} / RTA_{it-1} + \beta_3\ RE_{it-1} / RTA_{it-1} + d_t + \eta_i + u_{it}$$

(4) $$RTA_{it} = RF_{it} + (TA_{it} - BF_{it})$$

(5) $$RF_{it} = RF_{it-1}\ [\ (1 + \Phi_t) / (1 + \delta_{it})\ ] + I_{it}$$

(6) $$\Phi_t = (GCGP_t - GCGP_{t-1}) / GCGP_{t-1}$$

(7) $$\delta_{it} = D_{it} / BF_{it}$$

(8) $$0 = PFD = EBITDA_t + EBITDA_{t-1} > FE_t + FE_t$$
$$1 = PFD = EBITDA_t + EBITDA_{t-1} < FE_t + FE_t$$

Where:

Log (P(event) / P(no event)) = overall index, probability of financial distress

β_0 = constant

β_1, β_2, β_3 = constants obtained from the regression model

$EBIT_{it}$ = earnings before interest and taxes for index *i* at time *t*.

RTA_{it-1} = replacement value of total assets for index *i* at time *t - 1*.

FE_{it} = financial expenses for index *i* at time *t*.

RE_{it-1} = retained earnings for index *i* at time *t - 1*.

RF_{it} = replacement value tangible fixed assets for index *i* at time *t*, following Perfect and Wiles (1994).

TA_{it} = book value total assets for index *i* at time *t*.

BF_{it} = book value tangible fixed assets for index *i* at time *t*.

Φ_t = growth of capital goods prices at time *t*[10].

δ_{it} = real depreciation rate for index *i* at time *t*.

I_{it} = investments done in plant and equipment for index *i* at time *t*.

$GCGP_t$ = growth of capital good prices at time *t*.

D_{it} = book depreciation for index *i* at time *t*.

d_t = time effect

η_i = individual effect

u_{it} = random disturbance

PFD = probability of financial distress

EBITDA = earnings before interest, taxes, depreciation and amortization

This model takes the three classical financial ratios of profitability, retained earnings and substitutes debt for financial expenses, which is a more accurate predictor of financial distress. The logic behind financial expenses is rather straightforward: financial expenses entail all costs related to debt and debt servicing, including interest payments, brokerage fees, issuance costs for bonds and the like. The reasoning behind this is that having debt itself does not affect the probability of financial distress, see Begley *et al.* (1996). It is the aforementioned costs which come with it which may cause an increase in the probability of financial distress. This is reflected in the explanatory power the debt variable has versus the financial expense variable in studies such as that of Andrade and Kaplan (1998), which illustrate that the explanatory power of financial expenses is superior to that of debt. Next to these, it adds d_t and η_i to make it into a dynamic model as random and fixed factors, respectively.

The Altman and Ohlson models described above contain Bèta's with respect to the datasets they were derived from. For the sake of accuracy, we will replicate the aforementioned two models with our own Bèta's, based on the dataset of the G-7 (excluding USA) as well as the out-of-sample dataset of the USA. The dependents used herein allow for benchmarking between the three models as financial distress likelihood is set to be a binary variable. Following Cleary (1999), financial distress likelihood is met under the condition that earnings over the previous two years are smaller than the financial expenses over the same time period, the financial distress likelihood variable is assigned value one. In any other case, it is assigned value zero.

Table 2 below displays some summary statistics based on the discussed variables of the three models.

[10] See Appendix B.

Table 2 - Summary statistics for the G-7 (excluding USA), and the USA
Time period: 1990 - 2010

Altman (2000) model

	WC / TA		RE / TA		EBIT / TA		Equity / TL		Sales / TA	
	G-7	**USA**	**G-7**	**USA**	**G-7**	**USA**	**G-7**	**USA**	**G-7**	**USA**
Observations	59.823	27.566	59.823	27.584	59.801	27.612	59.773	27.616	59.821	27.613
Mean	0,154	-0,346	-0,054	-6,572	0,044	-0,040	1,215	2,367	1,170	1,352
Standard deviation	0,698	29,477	11,418	284,067	0,204	2,858	2,574	20,924	0,727	2,311
Minimum	-98,179	-2.471,000	-1.465,375	-28.169,000	-32,792	-354,500	-45,182	-11,020	0,000	-0,010
Maximum	0,991	1,000	1,686	12,120	5,807	9,710	124,875	2.233,000	39,292	243,670

Ohlson (1980) model

	TA / GCGP		TL / TA		WC / TA		CL / CA		Net / TA		FO / TL	
	G-7	**USA**	**G-7**	**USA**	**G-7**	**USA**	**G-7**	**USA**	**G-7**	**USA**	**G-7**	**USA**
Observations	59.825	27.647	59.775	27.624	59.823	27.566	59.822	27.550	46.410	27.612	59.746	27.535
Mean	2,009	0,342	0,591	1,154	0,154	-0,346	0,899	2,679	0,001	-0,116	0,152	0,180
Standard deviation	1,276	2,596	1,099	29,479	0,698	29,477	8,302	103,054	0,257	6,200	0,422	1,903
Minimum	-3,919	-11,990	0,000	0,000	-98,179	-2.471,000	-6,669	0,000	-44,805	-677,000	-38,755	-154,000
Maximum	5,950	7,350	127,936	2.472,000	0,991	1,000	1.266,875	15.099,000	6,870	319,000	10,843	40,250

Pindado (2008)

	EBIT / RTA		FE / RTA		RE / RTA	
	G-7	**USA**	**G-7**	**USA**	**G-7**	**USA**
Observations	56.480	26.057	56.480	26.057	56.480	26.057
Mean	0,087	0,035	0,017	0,030	0,204	-1,023
Standard deviation	2,941	1,060	0,334	0,200	4,142	18,073
Minimum	-55,601	-32,600	-0,631	-0,001	-412,008	-1.402,000
Maximum	651,448	88,179	54,966	17,214	771,103	38,932

***Table 2:** The above three tables give the summary statistics for the variables employed in the three models. For the Altman (2000) model, the variables are working capital (WC), total assets (TA), retained earnings (RE), earnings before interest and taxes (EBIT), Equity, total liabilities (TL) and net sales (Sales). For the Ohlson (1980) model, these variables are respectively the growth of capital goods prices (GCGP), current liabilities (CL), current assets (CA), net income (Net), funds provided by operations (FO). Lastly, for the Pindado (2008) model, we depict the replacement value of total assets (RTA) and the financial expenses (FE). Any variables not specifically mentioned in the last last models are equally defined as in the first model.*

4. Analysis

We will execute the methodology and present our preliminary findings in this chapter. To keep it orderly, I have divided this chapter into three paragraphs, each paragraph dealing with one of the financial distress models in detail.

4.1 The Altman Z-score model

For the Altman model, we employ a discriminate function, which looks similar to a regression. In Table 3 we see from the P-values for the five variables employed. Their statistical significance indicates that there is strong evidence of differences between means of financially distressed and healthy firms, based on the five variables employed by the replicated Altman (2000) model.

Table 3 - Altman Group Means

Altman Variable	G-7 Dataset F-value	G-7 Dataset P-value	USA dataset F-value	USA dataset P-value
Altman X_1	17,691	0,000	4,686	0,030
Altman X_2	11,722	0,001	3,266	0,710
Altman X_3	9,000	0,003	1,774	0,183
Altman X_4	3,316	0,069	5,823	0,016
Altman X_5	166,301	0,000	4,295	0,038

Table 3: *the five variables of Altman (2000) are displayed for the G-7 dataset, as well as the USA dataset. For the G-7 dataset, all but the fourth variable are statistically significant at the 1%. The fourth is however significant at the 10% level. Our out-of-sample dataset seems less robust: for the USA, the first, fourth and fifth variable are statistically significant at the 5% level, while the second and third variable appear as insignificant.*

Table 4 shows a similar problem found in the original Altman (1968) model: between X_1 and X_2, X_1 and X_3 as well as X_2 and X_3 are quite strongly positively correlated. This applies to both the G-7 excluding USA dataset, as well as findings from the USA dataset. Altman (1968) however states in this regard that it "…*has the advantage of yielding a model with a relatively small number of selected measurements which has the potential of conveying a great deal of information",* thus effectively implying that the positive correlation should not put the model at any disadvantage.

Table 4 - Correlation Between Altman Model Variables

Altman Variable	G-7 Dataset Altman X_1	G-7 Dataset Altman X_2	G-7 Dataset Altman X_3	G-7 Dataset Altman X_4	G-7 Dataset Altman X_5	USA dataset Altman X_1	USA dataset Altman X_2	USA dataset Altman X_3	USA dataset Altman X_4	USA dataset Altman X_5
Altman X_1	1,000	0,626	0,405	0,116	-0,043	1,000	0,706	0,382	0,004	-0,260
Altman X_2	0,626	1,000	0,474	0,027	-0,016	0,706	1,000	0,486	0,004	-0,140
Altman X_3	0,405	0,474	1,000	0,001	-0,080	0,382	0,462	1,000	0,004	-0,531
Altman X_4	0,116	0,027	0,001	1,000	-0,172	0,004	0,004	0,004	1,000	-0,028
Altman X_5	-0,043	-0,016	-0,080	-0,172	1,000	-0,260	-0,140	-0,531	-0,028	1,000

Table 4: *a correlation matrix between the five Altman (2000) variables is depicted for both the G-7 excluding USA dataset as well as the USA dataset. For both our dataset as well as our out-of-sample dataset we find that X_1 and X_2, X_1 and X_3 as well as X_2 and X_3 are quite strongly positively correlated.*

Next, Table 5 contains the unstandardized coefficients based on our G-7 dataset. Thus we can write for the G-7 excluding USA:

$$(9) \qquad Z = -1{,}519 - 0{,}296\ X_1 - 0.009\ X_2 - 0{,}06\ X_3 + 0{,}022\ X_4 + 1{,}316\ X_5$$

And for the USA:

$$(10) \qquad Z = -1{,}446 + 0.02\ X_1 - 0.001\ X_2 + 0.213\ X_3 - 0.023\ X_4 + 0.378\ X_5$$

Whereas the variables are defined the same as in (1).

Table 5 - Altman Unstandardized Coefficients

Altman Variable	G-7 Dataset Beta Value	USA dataset Beta Value
(Constant)	-1,519	-0,446
Altman X1	-0,296	0,020
Altman X2	-0,009	-0,001
Altman X3	-0,060	0,213
Altman X4	0,022	-0,023
Altman X5	1,316	0,378

Table 5: *this displays the unstandardized coefficients of our replicated Altman (2000) for both the G-7 and the USA. What draws our attention immediately is the large difference in the values as well as the sign between the G-7 and the USA datasets. It appears the G-7 dataset, which has far more observations yields different results than does the USA dataset. A reason for this could be that neither dataset has been subject to the strict limits of firm characteristics as does Altman (2000). Secondly, the model might simply not be designed to handle such large number of observations.*

Although our replica of the Altman model yields negative signs for X_1 through X_3, all of them are however, like (1), close to zero, with X_5 having a coefficient larger than one and positive. This could be because the original Altman (1968; 2000) models are derived under more strict assumptions than our dataset. An example could be firm size: the original models are restricted to firms with an stratified asset total of \$1 million to a maximum of \$25 million. Another possibility is the wide variety amongst the firm characteristics as well as the fact that the Altman models employ matched pairing between financially distressed and healthy firms, whereas we obtained a random, large sample without any additional criteria. Similar findings apply to our out-of-sample USA dataset. Lastly, there is sample period to consider. The revisited Altman model originally took the period of 1969 to 1999, while our dataset entails the period of 1990 – 2010.

Following Altman (1968; 2000), we will incorporate the intercept value of -1,519 into the criteria rather than in the model, thus we rewrite the model for the G-7 excluding USA as:

$$(11) \qquad Z = -\ 0{,}296\ X_1 - 0.009\ X_2 - 0{,}06\ X_3 + 0{,}022\ X_4 + 1{,}316\ X_5$$

And for the USA:

$$(12) \qquad Z = 0.02\,X_1 - 0.001\,X_2 + 0.213\,X_3 - 0.023\,X_4 + 0.378\,X_5$$

Where the criteria for our replicated model is defined as:

- $Z > -1{,}487$ is the domain in which a firm is considered safely away from financial distress
- $-1{,}551 < Z < -1{,}487$ is when a firm has a chance to become financially distressed
- $Z < -1{,}551$ is the domain in which a firm is considered financially distressed

The above classification criteria is obtained through Table 6 and following Altman (1968; 2000), employing a cutoff point of 0,5. Taking the average of the group centroids and add it, respectively subtract it, to the intercept value which then becomes the model's criteria as described above.

Table 6 - Altman Group Centroids

Firm Status	G-7 Value	USA Value
Financially Distressed	0,096	0,036
Healthy	-0,032	-0,023

Table 6: *the group centroids are used to determine the cutoff values to determine from what arbitrary value a firm is classified as financially distressed, having a chance to become financially distressed, or healthy.*

Our replica model correctly classifies 75,2% for the G-7 dataset of the firms as financially distressed or healthy, which is considered sufficient to employ the model for our purposes. For the USA dataset, this percentage is 61,2%, being less robust than our original dataset.

4.2 The Ohlson model

In order to get a timeframe analysis we employ a Cox regression model with the purpose of replicating the Ohlson (1980) model. In Table 7 we find the overall statistical significance of our replicated model. It shows the total degrees of freedom (df), being 9, as well as the P-value which is 0,000. The Chi-square value is extremely high, which in turn confirms the significance of the P-value. We observe similar results for our out-of-sample dataset from the USA. Thus we conclude that the model is statistically significant, which allows us to proceed.

Table 7 - Overall Significance Ohlson Model

G-7			USA		
Chi-square	**df**	**P-value**	**Chi-square**	**df**	**P-value**
1.508,642	9	0,000	1.136,101	9	0,000

Table 7: *for both the G-7 dataset as well as our out-of-sample dataset from the USA, we find that both models are statistically significant. Following Ohlson (1980), this check is a prerequisite to proceed with our analysis.*

Table 8 depicts the variable coefficients of the model, together with their respective P-values. The beta coefficients can be found in the second column of the table and represent their respective weights within the model for the G-7 dataset. X_1 has a negative sign, implying that smaller firms are more likely to go bankrupt than do larger ones. A similar relationship exists for our USA dataset. X_2 has a positive sign, which Ohlson (1980) interprets as "almost certain bankruptcy" when total liabilities exceed total assets. This is in contrasts with the USA dataset, although statistically significant at the 10% level. A possible explanation for this difference could be a difference in legal structure between the USA and other G-7 countries. X_3 shows a positive sign for both datasets: a good working capital ratio means a company has a buffer of assets available for economic difficult times. This relation is however not statistically significant for the USA dataset, making this relation spurious. X_4 then is indeterminate and therefore does not appear in our replica model as it is statistically insignificant in both datasets. In contrast, X_5 being the total liabilities to total assets ratio, shows a negative sign for both datasets, with the USA being statistically significant at the 10% level. One could expect this, as when the total debt of a company is larger than its total assets, its solvability is easily threatened. Logically, X_6 shows a positive sign, indicating that generating net income steers a company away from bankruptcy likelihood. For the USA, X_6 is only statistically significant at the 5% level. X_7 has a negative sign for the G-7 dataset, indicating that when earnings fall back, solvability may be at risk. This appears not to be true for our USA dataset. X_8 carries a positive sign, reinforcing the thought that when a company maintains a positive net income, bankruptcy is less likely to occur. For the G-7 excluding the USA, this relation is statistically significant at the 10% level. Finally, X_9 has a negative sign, indicating that downward relative changes in net income may be harmful to the company's financial health. X_9 is not statistically significant for the USA dataset.

Table 8 - Ohlson Coefficients and P-values

	G-7		USA	
Ohlson Variable	**Beta Value**	**P-value**	**Beta Value**	**P-value**
Ohlson X_1	-0,320	0,000	-0,130	0,000
Ohlson X_2	0,066	0,000	-0,031	0,053
Ohlson X_3	0,067	0,000	0,011	0,527
Ohlson X_4	0,000	0,421	0,000	0,352
Ohlson X_5	-0,753	0,000	-0,074	0,096
Ohlson X_6	0,396	0,000	0,004	0,039
Ohlson X_7	-0,264	0,000	0,035	0,000
Ohlson X_8	0,043	0,074	0,109	0,000
Ohlson X_9	-0,126	0,000	-0,002	0,926

Table 8: *the beta values of our replica models of Ohlson (1980) for the G-7 and the USA are displayed here. On overall, our USA dataset yields less robust results than does our G-7 dataset. It does however hold for our purposes in this paper.*

Table 9 then, shows the correlations between the replica model variables. No real surprise here as we find a positive, high correlation (0,837) between X_2 and X_3. Since this is the total liabilities to total assets ratio and the working capital to total assets ratio, respectively, these two are mathematically correlated. We also find a positive, high correlation (0,686) between X_2 and X_6, which is the net income to total assets ratio. This correlation is weaker for the USA dataset. For X_3, we find a similar positive correlation to X_6 (0,562). These variables are all known to be classical variables of bankruptcy, and thus of financial distress.

Table 9 - Correlation Matrix of Regression Coefficients

	G-7 Dataset							
Ohlson Variable	Ohlson X_1	Ohlson X_2	Ohlson X_3	Ohlson X_4	Ohlson X_5	Ohlson X_6	Ohlson X_7	Ohlson X_8
Ohlson X_2	0,080							
Ohlson X_3	0,030	0,837						
Ohlson X_4	0,046	0,033	0,092					
Ohlson X_5	-0,128	0,017	-0,068	0,123				
Ohlson X_6	0,016	0,686	0,562	0,013	-0,124			
Ohlson X_7	-0,196	-0,262	-0,208	-0,014	0,081	-0,363		
Ohlson X_8	-0,061	-0,053	-0,053	0,004	-0,167	-0,079	-0,167	
Ohlson X_9	-0,026	-0,097	-0,062	-0,020	0,013	-0,115	0,009	-0,053

	USA dataset							
Ohlson Variable	Ohlson X_1	Ohlson X_2	Ohlson X_3	Ohlson X_4	Ohlson X_5	Ohlson X_6	Ohlson X_7	Ohlson X_8
Ohlson X_2	0,017							
Ohlson X_3	-0,021	0,881						
Ohlson X_4	0,012	0,034	0,057					
Ohlson X_5	0,159	-0,253	-0,069	-0,041				
Ohlson X_6	0,038	0,038	0,029	0,006	-0,011			
Ohlson X_7	-0,079	0,005	0,003	0,000	-0,019	-0,004		
Ohlson X_8	0,271	0,000	0,005	-0,006	-0,163	0,063	0,255	
Ohlson X_9	0,019	-0,014	-0,012	-0,005	-0,008	-0,045	-0,081	0,025

Table 9: *The top panel displays the correlations between the nine variables employed by our Ohlson (1980) replicated model for the G-7 dataset. The classical variables of financial distress all show strong, positive correlations. Although in weaker form, we find similar results for our USA dataset, as displayed in the bottom panel.*

An important note is that the Ohlson model does not take into account that some firms may experience financial distress, but are able to rebound – or simply restart after bankruptcy. Such observations do occur in the G-7 dataset, while the Ohlson model simply classifies them as bankrupt. This is in line with the findings of Grice and Dugan (2001) and Grice and Ingram (2001) on older models, such as Ohlson's. The Pindado (2008) model we are to discuss next has a dynamic variable which does take into account a firm's possibility to be "temporarily" financially distressed.

4.3 The Pindado model

This model is based on an *ex ante* condition as specified in formula (8). Following Cleary (1999), it considers a firm financially distressed if its EBITDA over the past two years is less than its financial expenses over the same period. If this condition is met, the firm is classified as financially distressed for the current period. If this condition is not met, the firm is classified as not financially distressed for the current period. This is done for each firm in the dataset as shown in table 10.

Table 10 - Number of non-financial distressed and financially distressed observations
Time period: 1990 - 2010

	G-7 sample				US-sample				Total			
	N	FD	Total	%	N	FD	Total	%	N	FD	Total	%
1990	2.396	286	2.682	10,66%	1.225	288	1.513	19,04%	3.621	574	4.195	13,68%
1991	2.401	354	2.755	12,85%	1.213	297	1.510	19,67%	3.614	651	4.265	15,26%
1992	2.345	481	2.826	17,02%	1.245	266	1.511	17,60%	3.590	747	4.337	17,22%
1993	2.292	601	2.893	20,77%	1.255	258	1.513	17,05%	3.547	859	4.406	19,50%
1994	2.321	611	2.932	20,84%	1.255	258	1.513	17,05%	3.576	869	4.445	19,55%
1995	2.560	426	2.986	14,27%	1.274	243	1.517	16,02%	3.834	669	4.503	14,86%
1996	2.714	314	3.028	10,37%	1.268	236	1.504	15,69%	3.982	550	4.532	12,14%
1997	2.816	342	3.158	10,83%	1.250	242	1.492	16,22%	4.066	584	4.650	12,56%
1998	2.765	438	3.203	13,67%	1.211	277	1.488	18,62%	3.976	715	4.691	15,24%
1999	2.748	365	3.113	11,73%	1.136	254	1.390	18,27%	3.884	619	4.503	13,75%
2000	2.758	270	3.028	8,92%	1.028	238	1.266	18,80%	3.786	508	4.294	11,83%
2001	2.696	260	2.956	8,80%	952	225	1.177	19,12%	3.648	485	4.133	11,73%
2002	2.595	261	2.856	9,14%	871	236	1.107	21,32%	3.466	497	3.963	12,54%
2003	2.546	203	2.749	7,38%	851	200	1.051	19,03%	3.397	403	3.800	10,61%
2004	2.508	168	2.676	6,28%	846	150	996	15,06%	3.354	318	3.672	8,66%
2005	2.424	160	2.584	6,19%	797	131	928	14,12%	3.221	291	3.512	8,29%
2006	2.334	178	2.512	7,09%	744	131	875	14,97%	3.078	309	3.387	9,12%
2007	2.233	168	2.401	7,00%	690	118	808	14,60%	2.923	286	3.209	8,91%
2008	2.132	164	2.296	7,14%	638	122	760	16,05%	2.770	286	3.056	9,36%
2009	2.012	206	2.218	9,29%	608	123	731	16,83%	2.620	329	2.949	11,16%
2010	108	10	118	8,47%	113	19	132	14,39%	221	29	250	11,60%
Total	**49.704**	**6.266**	**55.970**	**11,20%**	**20.470**	**4.312**	**24.782**	**17,40%**	**70.174**	**10.578**	**80.752**	**13,10%**

Table 10: *when the condition as specified in formula (8) are applied to the dataset, we are able to the number of observations which are classified as normal and financially distressed. N is to be interpreted as normal, not financially distressed firms. FD stands for financially distressed. The percentages are the total number of observations which are classified as financially distressed for that year. Note that the USA has a relatively larger number of financially distressed observations than does the G-7 dataset. On overall, 13,1% of all observations are classified as distressed.*

Note that this gives rise to an interesting dynamic: it is possible for a firm to be financially distressed somewhere within our sample period, but not permanently. In other words, under this definition of financial distress, it is possible for a firm to “rebound” out of financial distress within the sample period. This is what makes this model stand out from the previous two discussed. Granted, the Ohlson model has a similar dynamic, but employs a Cox regression while the Pindado model employs a binary logistic regression. The latter has far more technical features available which allow for larger datasets to be accurately modeled in comparison to the former. Secondly, it is subject to far less constraints, such as normality of the data. A prerequisite for the Pindado model is that it requires a dataset with a large number of observations to be accurate, this in contrast with the requirements of the previously discussed models, which perform better on smaller, more restricted datasets.

Following Pindado (2006; 2008), we reiterate the binary logistic model for each year for which we have observations. This allows us to view the robustness of the model taking into account the time and individual factors of the model. Important here is to determine whether the signs are of the correct order of each of the three variables.

EBIT / RTA is expected to have a negative sign, since earnings are needed to at least replace assets at replacement values over time. This reduces the probability of financial distress. This matches findings of previous studies, such as Altman (1968), Ohlson (1980) and Zmijewski (1984). Note that creditors also tend to use this ratio in order to estimate the return on capital (Claessens *et al.*, 2003). FE / RTA is expected to be positive, thus increasing the probability of financial distress. As discussed previously, this could be due to the various costs attached to debt. Finally, RE / RTA is expected to have a negative sign. Since retained earnings allow firms to have a financial buffer in lesser times. Altman (1968) argues that younger firms are more susceptible to financial distress as they generally lack such a buffer. Table 11 below displays the beta values, the corresponding standard errors as well as several reliability checks.

For the first ratio, we find a negative sign over the period investigated, being statistically significant for all years. A similar result is found in our USA dataset. The financial theory discussed is thus supported by our model in this respect. The second ratio shows an inconsistent result, however. Up to and including 1994 it is statistically significant and has a positive sign. From 1995 to 1996, the sign turns negative and again becomes statistically significant up to 2009. Our USA dataset shows a statistically significant result with a negative sign over the investigated period. For the third ratio, we find a negative sign in both the G-7 dataset, as well as in our out-of-sample test with the USA dataset. We can therefore conclude that the generation of revenues to be able to replace its assets over time indeed is a statistically significant factor for firms with respect to the probability of financial distress. We do not however find such a consistent statistical relationship for the various financial expenses in relation to the replacement value of assets. As discussed, a financial reserve in the shape of retained earnings reduces the probability of financial distress occurring, as it allows the firm to replace its assets over time using said retained earnings.

Table 11 - Estimation results of the Pindado model

G-7 dataset

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
$EBIT_{it} / RTA_{it-1}$	-69,387*	-62,789*	-70,617*	-59,427*	-72,671*	-60,536*	-54,886*	-50,905*	-47,502*	-36,859*	-35,747*	-24,290*	-34,275*	-52,869*	-63,785*	-51,075*	-70,177*	-71,430*	-42,540*	-43,801*	-67,354*
	(3,919)	(3,435)	(3,557)	(2,840)	(3,535)	(3,239)	(3,050)	(2,770)	(2,435)	(2,081)	(2,089)	(1,502)	(1,903)	(2,931)	(3,421)	(3,094)	(4,070)	(4,494)	(2,751)	(2,703)	(16,835)
Fe_{it} / RTA_{it-1}	50,889*	36,965*	37,369*	22,109*	52,739*	5,845	-22,871*	-45,733*	-44,983*	-52,387*	-101,937*	-126,013*	-138,246*	-109,889*	-116,729*	-123,124*	-99,179*	-174,313*	-164,969*	-238,454*	36,531
	(4,473)	(3,669)	(3,668)	(3,416)	(4,459)	(5,042)	(5,437)	(6,085)	(4,844)	(5,382)	(6,915)	(7,127)	(8,483)	(9,940)	(9,262)	(10,542)	(12,634)	(14,058)	(9,859)	(17,146)	(43,438)
Re_{it-1} / RTA_{it-1}	-6,785*	-4,534*	-3,463*	-2,691*	-1,774*	-3,272*	-4,407*	-5,058*	-3,879*	-4,129*	-5,467*	-6,344*	-3,595*	-1,519*	-1,739*	-2,771*	-1,632*	-2,545*	-5,351*	-3,709*	-1,511
	(0,792)	(0,609)	(0,490)	(0,363)	(0,366)	(0,399)	(0,451)	(0,411)	(0,300)	(0,307)	(0,368)	(0,347)	(0,249)	(0,239)	(0,249)	(0,251)	(0,173)	(0,252)	(0,304)	(0,289)	(1,088)
Psuedo R-squared	0,865	0,842	0,823	0,775	0,795	0,826	0,856	0,862	0,808	0,795	0,848	0,793	0,788	0,815	0,862	0,852	0,881	0,901	0,836	0,816	0,868
Likelihood Ratio	909,106	1.066,461	1.199,920	1.489,002	1.403,671	1.250,399	1.085,213	1.091,446	1.457,772	1.490,803	1.136,667	1.427,289	1.403,689	1.213,065	924,578	949,192	768,061	620,866	915,726	972,487	39,015
Observations	2.673	2.747	2.815	2.883	2.923	2.977	3.019	3.150	3.200	3.107	3.027	2.956	2.853	2.746	2.673	2.581	2.510	2.398	2.292	2.214	117

USA dataset

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
$EBIT_{it} / RTA_{it-1}$	-24,046*	-26,256*	-17,318*	-28,742*	-37,135*	-29,746*	-27,701*	-25,748*	-18,543*	-17,709*	-30,634*	-17,137*	-21,810*	-31,121*	-43,522*				-39,898*		-39,327*
	(1,348)	(1,608)	(1,078)	(1,817)	(2,424)	(1,931)	(1,868)	(1,737)	(1,201)	(1,191)	(2,188)	(1,208)	(1,631)	(2,400)	(3,872)				(3,861)		(9,803)
Fe_{it} / RTA_{it-1}	-6,833*	-11,559*	-12,043*	-28,597*	-19,111*	-26,460*	-25,774*	-27,253*	-14,743*	-25,203*	-11,257*	-19,409*	-13,169*	-3,983	-4,704				-23,565*		-21,483*
	(2,184)	(2,711)	(2,834)	(4,160)	(4,490)	(3,856)	(4,445)	(4,239)	(2,830)	(3,169)	(3,016)	(2,825)	(3,557)	(4,129)	(6,462)				(6,033)		(6,888)
Re_{it-1} / RTA_{it-1}	-0,450*	-0,437*	-0,322*	-0,262**	-0,252***	-0,316*	-0,493*	-0,928*	-0,774*	-0,712*	-0,511*	-0,699*	-1,640*	-1,315*	-0,476**				-0,480*		-1,740**
	(0,099)	(0,083)	(0,071)	(0,112)	(0,141)	(0,062)	(0,142)	(0,135)	(0,128)	(0,105)	(0,160)	(0,124)	(0,171)	(0,186)	(0,204)				(0,112)		(0,696)
Psuedo R-squared	0,752	0,737	0,612	0,789	0,843	0,828	0,829	0,834	0,740	0,741	0,829	0,676	0,734	0,769	0,856				0,865		0,905
Likelihood Ratio	840,777	878,162	1.165,641	741,084	582,562	630,383	623,837	600,872	856,387	799,040	525,195	799,861	648,682	552,881	356,560				258,028		32,974
Observations	1.513	1.510	1.510	1.513	1.513	1.517	1.503	1.490	1.487	1.389	1.266	1.177	1.107	1.051	995	927	872	808	760	729	132

Table 11 *- Standard error in parenthesis. Psuedo R-squared is the Nagelkerke R-square, which is an indication of the goodness of fit of the model based on χ2 statistic. The Likelihood Ratio test the joint significance of the variables employed in the model. It's a rough indicator of which its value is represented as a χ2 statistic.*
** means significant at 1%, ** means significant at 5%, *** means significant at 10%.*
For the USA dataset, 2005, 2006, 2007 and 2009 data unavailable due to complete seperation of observations. See Albert and Anderson (1984): in essence, the relative number of firms classified as financially distressed is too small on such limited population for the years in question, causing the model to fail to converge to maximum likelihood.

Note that for the Psuedo-R Square in table 11 we have chosen for the Nagelkerke R-square. It is a reliable indication of the strength of the relationship between the predicted variable, being the probability of financial distress, and the predictors, being the three variables as previously discussed in the methodology chapter. From table 11 we can conclude that this relationship is moderate to strong.

When we execute the model on both datasets and list the probabilities of financial distress, as well as their corresponding means and standard deviations, we acquire table 12. From here we see that 11,18% of all observations of the G-7 dataset, are classified as financially distressed by the model. For our out-of-sample dataset from the USA, this percentage is 17,37%. See that the probability of financial distress for US firms in the datasets appear to be higher than in other G-7 countries. This is also the case in Pindado (2008). Note that our replica model correctly predicts *ex ante* 93,47% of the total observations for the G-7 excluding USA, and 93,12% for the USA, respectively. The Type I and Type II errors are to be interpreted as follows:

- **Type I:** an observation is classified as financially distressed by the model, while it is actually not financially distressed.
- **Type II:** an observation is classified as normal by the model, while actually it is financially distressed.

Regardless of Type I or Type II error, the model in question misclassifies the observation in question. The observations which are not misclassified are thus correctly classified, either as financially distressed (FD), otherwise as normal (Normal).

Note that for the years 2005, 2006, 2007 and 2008 the model did not iterate, due to a complete separation of observations, see Albert and Anderson (1984).

Table 12 - Estimation results on the probability of financial distress

G-7 dataset

Year	Classification status				Type I error		Type II error		Correct classification		Mean	Standard
	Normal	%	FD	%	Absolute	Relative	Absolute	Relative	Absolute	Relative		deviation
1990	2.388	89,34%	285	10,66%	35	1,31%	131	4,90%	2.507	93,79%	0,110	0,309
1991	2.394	87,15%	353	12,85%	48	1,75%	167	6,08%	2.532	92,17%	0,130	0,335
1992	2.336	82,98%	479	17,02%	77	2,74%	177	6,29%	2.561	90,98%	0,170	0,376
1993	2.285	79,26%	598	20,74%	79	2,74%	228	7,91%	2.576	89,35%	0,210	0,406
1994	2.313	79,13%	610	20,87%	49	1,68%	213	7,29%	2.661	91,04%	0,210	0,406
1995	2.552	85,72%	425	14,28%	51	1,71%	199	6,68%	2.727	91,60%	0,140	0,350
1996	2.706	89,63%	313	10,37%	50	1,66%	171	5,66%	2.798	92,68%	0,100	0,305
1997	2.808	89,14%	342	10,86%	48	1,52%	154	4,89%	2.948	93,59%	0,110	0,311
1998	2.762	86,31%	438	13,69%	85	2,66%	168	5,25%	2.947	92,09%	0,140	0,344
1999	2.743	88,28%	364	11,72%	71	2,29%	170	5,47%	2.866	92,24%	0,120	0,322
2000	2.757	91,08%	270	8,92%	54	1,78%	94	3,11%	2.879	95,11%	0,090	0,285
2001	2.696	91,20%	260	8,80%	68	2,30%	113	3,82%	2.775	93,88%	0,090	0,283
2002	2.594	90,92%	259	9,08%	64	2,24%	115	4,03%	2.674	93,73%	0,090	0,288
2003	2.545	92,68%	201	7,32%	70	2,55%	62	2,26%	2.614	95,19%	0,070	0,262
2004	2.506	93,75%	167	6,25%	56	2,10%	35	1,31%	2.582	96,60%	0,060	0,243
2005	2.423	93,88%	158	6,12%	61	2,36%	46	1,78%	2.474	95,85%	0,060	0,241
2006	2.332	92,91%	178	7,09%	49	1,95%	38	1,51%	2.423	96,53%	0,070	0,257
2007	2.231	93,04%	167	6,96%	48	2,00%	38	1,58%	2.312	96,41%	0,070	0,255
2008	2.130	92,93%	162	7,07%	74	3,23%	57	2,49%	2.161	94,28%	0,070	0,258
2009	2.008	90,70%	206	9,30%	58	2,62%	73	3,30%	2.083	94,08%	0,090	0,290
2010	107	91,45%	10	8,55%	4	3,42%	2	1,71%	111	94,87%	0,090	0,280
Total	**49.616**	**88,82%**	**6.245**	**11,18%**	**1.199**	**2,15%**	**2.451**	**4,39%**	**52.211**	**93,47%**	**0,110**	**0,315**

USA dataset

Year	Classification status				Type I error		Type II error		Correct classification		Mean	Standard
	Normal	%	FD	%	Absolute	Relative	Absolute	Relative	Absolute	Relative		deviation
1990	1.225	80,96%	288	19,04%	89	5,88%	54	3,57%	1.370	90,55%	0,190	0,393
1991	1.213	80,33%	297	19,67%	82	5,43%	71	4,70%	1.357	89,87%	0,200	0,398
1992	1.245	82,45%	265	17,55%	81	5,36%	65	4,30%	1.364	90,33%	0,180	0,381
1993	1.255	82,95%	258	17,05%	75	4,96%	63	4,16%	1.375	90,88%	0,170	0,376
1994	1.255	82,95%	258	17,05%	57	3,77%	48	3,17%	1.408	93,06%	0,170	0,376
1995	1.274	83,98%	243	16,02%	50	3,30%	53	3,49%	1.414	93,21%	0,160	0,367
1996	1.268	84,36%	235	15,64%	54	3,59%	50	3,33%	1.399	93,08%	0,160	0,364
1997	1.249	83,83%	241	16,17%	57	3,83%	35	2,35%	1.398	93,83%	0,160	0,369
1998	1.210	81,37%	277	18,63%	83	5,58%	35	2,35%	1.369	92,06%	0,190	0,389
1999	1.135	81,71%	254	18,29%	54	3,89%	48	3,46%	1.287	92,66%	0,180	0,387
2000	1.028	81,20%	238	18,80%	58	4,58%	25	1,97%	1.183	93,44%	0,190	0,391
2001	952	80,88%	225	19,12%	80	6,80%	44	3,74%	1.053	89,46%	0,190	0,393
2002	871	78,68%	236	21,32%	57	5,15%	33	2,98%	1.017	91,87%	0,210	0,341
2003	851	80,97%	200	19,03%	61	5,80%	20	1,90%	970	92,29%	0,190	0,393
2004	846	85,03%	149	14,97%	43	4,32%	19	1,91%	933	93,77%	0,150	0,358
2005	797	85,98%	130	14,02%		0,00%		0,00%	927	100,00%	0,140	0,347
2006	744	85,32%	128	14,68%		0,00%		0,00%	872	100,00%	0,150	0,357
2007	690	85,40%	118	14,60%		0,00%		0,00%	808	100,00%	0,150	0,353
2008	638	83,95%	122	16,05%	44	5,79%	9	1,18%	707	93,03%	0,160	0,367
2009	608	83,40%	121	16,60%		0,00%		0,00%	729	100,00%	0,170	0,374
2010	113	85,61%	19	14,39%	5	3,79%	2	1,52%	125	94,70%	0,140	0,352
Total	**20.467**	**82,63%**	**4.302**	**17,37%**	**1.030**	**4,16%**	**674**	**2,72%**	**23.065**	**93,12%**	**0,170**	**0,379**

Table 12: *the top panel displays the findings on the G-7 dataset. We display the firms which are predicted to be healthy (Normal) and financially distressed (FD). Secondly, we display the number of Type I and Type II errors for each year investigated as well as the number of observations which are correctly classified. Lastly, we display the mean and standard deviation of the predicted variable. The bottom panel shows similar statistics for the USA dataset. We conclude that we find quite robust results: within the G-7 dataset, 93,47% of the observations are correctly classified. For the USA, this percentage is 93,12%.*

5. Comparing the models

In this chapter we will discuss the accuracy of the models based on their classification results. We will summarize our findings of the previous chapter and discuss the specifics of each of these models with respect to their results in predicting financial distress.

Within statistics it is common to consider the R-square values. There is a catch here, however. The pseudo-R square, such as the Nagelkerke Pseudo R-square reported in the Pindado model, merely illustrates how much the fitted model improves the log-likelihood from the null model. Therefore it is theoretically possible to compare the Ohlson and Pindado replicated models, as they are both logistic regressions of some sort. This would however exclude the Altman model from the comparison, thus effectively missing the goal of this paper.

Since we are dealing with three different types of regression, it is not intuitive on how to compare these models directly. The advantage we have here is that we employed all three models on the same dataset. Even though the models may have different variables and methodologies, the underlying data is the same. For our out-of-sample dataset, this logic also applies.

This allows us to regard each model in its own respect. One would employ a model with predictive power in order to predict something – here, the probability of financial distress. Therefore, the reliability of this predictive power, being the amount of observations correctly predicted would be the most obvious manner of comparison. Woolridge (2009) points out that this is a suitable manner to compare goodness-of-fit between various models. It is noted, however, that this can be very misleading when the probability of the event – here, financial distress likelihood – occurring is very rare. This becomes especially true when the number of observations is very small. Therefore Woolridge (2009) suggests to calculate the number of correctly predicted observations for each year.

When we employ this method, we get table 13 for the replicated Altman model. From the table we can see that the replicated Altman model misclassifies a large number of observations. This is especially true for Type I errors, thus resulting in false positives. This is true for both the G-7 as well as the USA dataset – 33,37% and 41,04% of the observations are Type I errors, respectively. The replicated Altman model finds roughly three times as much financially distressed observations as does the benchmark, being 36,94% versus 11,19% for the G-7 dataset, and 46,59% versus 17,33% for the USA dataset. The replicated Altman model correctly predicts 59,02% of the observations for the G-7 dataset, and 47,19% of the observations for the USA dataset.

Table 13 - number of observations correctly classified, Type I and Type II errors - Altman replicated model

G-7 dataset

	Benchmark classification				Altman						Altman classification			
	Normal		FD		Type I		Type II		Correct		Normal		FD	
Year	Absolute	%	Absolute	%	Absolute	Relative	Absolute	Relative	Absolute	Relative	Absolute	%	Absolute	%
1990	2.396	89,37%	285	10,63%	991	36,96%	168	6,27%	1.522	56,77%	1.573	58,67%	1.108	41,33%
1991	2.401	87,18%	353	12,82%	1.001	36,35%	227	8,24%	1.526	55,41%	1.627	59,08%	1.127	40,92%
1992	2.345	82,98%	481	17,02%	989	35,00%	353	12,49%	1.484	52,51%	1.709	60,47%	1.117	39,53%
1993	2.292	79,23%	601	20,77%	958	33,11%	447	15,45%	1.488	51,43%	1.781	61,56%	1.112	38,44%
1994	2.321	79,16%	611	20,84%	956	32,61%	443	15,11%	1.533	52,29%	1.808	61,66%	1.124	38,34%
1995	2.559	85,73%	426	14,27%	1.002	33,57%	288	9,65%	1.695	56,78%	1.845	61,81%	1.140	38,19%
1996	2.714	89,63%	314	10,37%	1.082	35,73%	194	6,41%	1.752	57,86%	1.826	60,30%	1.202	39,70%
1997	2.815	89,17%	342	10,83%	1.118	35,41%	218	6,91%	1.821	57,68%	1.915	60,66%	1.242	39,34%
1998	2.765	86,33%	438	13,67%	1.060	33,09%	303	9,46%	1.840	57,45%	2.008	62,69%	1.195	37,31%
1999	2.748	88,27%	365	11,73%	966	31,03%	268	8,61%	1.879	60,36%	2.050	65,85%	1.063	34,15%
2000	2.757	91,08%	270	8,92%	937	30,95%	178	5,88%	1.912	63,16%	1.998	66,01%	1.029	33,99%
2001	2.696	91,24%	259	8,76%	915	30,96%	171	5,79%	1.869	63,25%	1.952	66,06%	1.003	33,94%
2002	2.595	90,86%	261	9,14%	919	32,18%	173	6,06%	1.764	61,76%	1.849	64,74%	1.007	35,26%
2003	2.546	92,62%	203	7,38%	873	31,76%	133	4,84%	1.743	63,40%	1.806	65,70%	943	34,30%
2004	2.508	93,72%	168	6,28%	896	33,48%	109	4,07%	1.671	62,44%	1.721	64,31%	955	35,69%
2005	2.424	93,81%	160	6,19%	820	31,73%	109	4,22%	1.655	64,05%	1.713	66,29%	871	33,71%
2006	2.334	92,91%	178	7,09%	846	33,68%	131	5,21%	1.535	61,11%	1.619	64,45%	893	35,55%
2007	2.233	93,04%	167	6,96%	860	35,83%	108	4,50%	1.432	59,67%	1.481	61,71%	919	38,29%
2008	2.130	92,85%	164	7,15%	832	36,27%	95	4,14%	1.367	59,59%	1.393	60,72%	901	39,28%
2009	2.009	90,70%	206	9,30%	612	27,63%	139	6,28%	1.464	66,09%	1.536	69,35%	679	30,65%
2010	106	91,38%	10	8,62%	37	31,90%	7	6,03%	72	62,07%	76	65,52%	40	34,48%
Total	**49.694**	**88,81%**	**6.262**	**11,19%**	**18.670**	**33,37%**	**4.262**	**7,62%**	**33.024**	**59,02%**	**35.286**	**63,06%**	**20.670**	**36,94%**

USA dataset

	Benchmark classification				Altman						Altman classification			
	Normal		FD		Type I		Type II		Correct		Normal		FD	
Year	Absolute	%	Absolute	%	Absolute	Relative	Absolute	Relative	Absolute	Relative	Absolute	%	Absolute	%
1990	1.214	80,83%	288	19,17%	693	46,14%	178	11,85%	631	42,01%	699	46,54%	803	53,46%
1991	1.206	80,56%	291	19,44%	673	44,96%	181	12,09%	643	42,95%	714	47,70%	783	52,30%
1992	1.240	82,56%	262	17,44%	694	46,21%	168	11,19%	640	42,61%	714	47,54%	788	52,46%
1993	1.250	83,00%	256	17,00%	700	46,48%	174	11,55%	632	41,97%	724	48,07%	782	51,93%
1994	1.252	83,02%	256	16,98%	703	46,62%	172	11,41%	633	41,98%	721	47,81%	787	52,19%
1995	1.271	84,12%	240	15,88%	691	45,73%	157	10,39%	663	43,88%	737	48,78%	774	51,22%
1996	1.267	84,41%	234	15,59%	672	44,77%	161	10,73%	668	44,50%	756	50,37%	745	49,63%
1997	1.248	83,81%	241	16,19%	654	43,92%	151	10,14%	684	45,94%	745	50,03%	744	49,97%
1998	1.209	81,52%	274	18,48%	604	40,73%	194	13,08%	685	46,19%	799	53,88%	684	46,12%
1999	1.134	81,70%	254	18,30%	536	38,62%	175	12,61%	677	48,78%	773	55,69%	615	44,31%
2000	1.027	81,31%	236	18,69%	491	38,88%	169	13,38%	603	47,74%	705	55,82%	558	44,18%
2001	951	80,94%	224	19,06%	420	35,74%	152	12,94%	603	51,32%	683	58,13%	492	41,87%
2002	867	78,68%	235	21,32%	373	33,85%	166	15,06%	563	51,09%	660	59,89%	442	40,11%
2003	848	81,07%	198	18,93%	344	32,89%	135	12,91%	567	54,21%	639	61,09%	407	38,91%
2004	843	84,98%	149	15,02%	366	36,90%	104	10,48%	522	52,62%	581	58,57%	411	41,43%
2005	796	86,05%	129	13,95%	349	37,73%	94	10,16%	482	52,11%	541	58,49%	384	41,51%
2006	742	84,99%	131	15,01%	324	37,11%	92	10,54%	457	52,35%	510	58,42%	363	41,58%
2007	690	85,40%	118	14,60%	293	36,26%	91	11,26%	424	52,48%	488	60,40%	320	39,60%
2008	638	83,95%	122	16,05%	293	38,55%	88	11,58%	379	49,87%	433	56,97%	327	43,03%
2009	606	83,24%	122	16,76%	213	29,26%	90	12,36%	425	58,38%	483	66,35%	245	33,65%
2010	112	85,50%	19	14,50%	46	35,11%	16	12,21%	69	52,67%	82	62,60%	49	37,40%
Total	**20.411**	**82,67%**	**4.279**	**17,33%**	**10.132**	**41,04%**	**2.908**	**11,78%**	**11.650**	**47,19%**	**13.187**	**53,41%**	**11.503**	**46,59%**

***Table 13:** the top panel is made up of three windows and concerns the G-7 dataset. The left hand window displays the binary variable of financial distress likelihood. This is considered the benchmark for all three models in this paper. It displays for each year which firm is ex-post classified as normal (Normal) or financially distressed (FD). The middle window displays the Type I and Type II errors from our replicated Altman model. The right hand window displays the number of observations as classified by our replicated Altman model. The bottom panel is to be interpreted in a similar way, but then for the USA dataset.*

We apply the same format to display the correctly classified observations for the Ohlson replicated model. This gives table 14. It appears that the Ohlson replicated model shows better results than the Altman model. It is noteworthy that the Type I and Type II errors are low, except for the Type I errors in the USA dataset, being 30,97%. On overall, the Ohlson model classifies 2,04% of the observations of the G-7 dataset as financially distressed, whereas the benchmark states this is 9,89%. For the USA dataset, these values are 46,07% versus 17,33%. The replicated Ohlson model correctly classifies 89,43% of the G-7 dataset observations and 66,8% of the USA dataset.

This leaves us to conclude that the Ohlson model does far better than the Altman model, whereas both perform less consistent at the out-of-sample dataset of the USA. The Ohlson model seems to be less volatile in this respect than the Altman model. This leaves us to conclude that there is still room for improvement.

When we repeat this exercise for the Pindado replicated model, we obtain table 12. From here we can see that the Type I and Type II errors are relatively the lowest of all three models on both the G-7 as well as the USA dataset. The Pindado replicated model correctly classifies 93,47% of the G-7 dataset, whereas this is 93,12% for the USA dataset.

We employed the binary variable for financial distress as a benchmark for all three models. A slight variation in the total number of observations between the three models comes forth out of partial unavailability of data from CompuStat for the variables employed within the three models.

To make it easier to interpret tables 12, 13 and 14 as to which of the three models has the best overall consistency in correctly predicting financial distress likelihood, see figure 2 and 3 below for a graphical display. From here we see that the Pindado model performs best in both the G-7 dataset as well as in our out-of-sample dataset for the USA. The Ohlson model comes as a close second best to the Pindado model. To a lesser degree, this is also true for the USA dataset. The large number of Type I errors in our out-of-sample test seems to harm its reliability. In both datasets, the Altman model performed the least of the three models. As discussed previously, the technical features of logarithmic models really shine when employed on large datasets, which could be a clarification why the Altman model lags behind the other two on both occasions.

Table 14 - number of observations correctly classified, Type I and Type II errors - Ohlson replicated model

G-7 dataset														
	Benchmark classification				Ohlson						Ohlson classification			
	Normal		FD		Type I		Type II		Correct		Normal		FD	
Year	Absolute	%	Absolute	%	Absolute	Relative	Absolute	Relative	Absolute	Relative	Absolute	%	Absolute	%
1990	611	86,42%	96	13,58%	21	2,97%	83	11,74%	603	85,29%	673	95,19%	34	4,81%
1991	626	84,37%	116	15,63%	19	2,56%	102	13,75%	621	83,69%	709	95,55%	33	4,45%
1992	608	80,96%	143	19,04%	17	2,26%	129	17,18%	605	80,56%	720	95,87%	31	4,13%
1993	591	77,25%	174	22,75%	17	2,22%	160	20,92%	588	76,86%	734	95,95%	31	4,05%
1994	605	77,96%	171	22,04%	19	2,45%	161	20,75%	596	76,80%	747	96,26%	29	3,74%
1995	1.391	82,90%	287	17,10%	21	1,25%	281	16,75%	1.376	82,00%	1.651	98,39%	27	1,61%
1996	1.731	87,51%	247	12,49%	24	1,21%	239	12,08%	1.715	86,70%	1.946	98,38%	32	1,62%
1997	1.839	87,74%	257	12,26%	19	0,91%	243	11,59%	1.834	87,50%	2.063	98,43%	33	1,57%
1998	1.781	84,09%	337	15,91%	17	0,80%	323	15,25%	1.778	83,95%	2.087	98,54%	31	1,46%
1999	1.635	86,69%	251	13,31%	24	1,27%	235	12,46%	1.627	86,27%	1.846	97,88%	40	2,12%
2000	1.714	91,56%	158	8,44%	28	1,50%	141	7,53%	1.703	90,97%	1.827	97,60%	45	2,40%
2001	2.114	92,96%	160	7,04%	57	2,51%	146	6,42%	2.071	91,07%	2.203	96,88%	71	3,12%
2002	2.141	92,68%	169	7,32%	40	1,73%	157	6,80%	2.113	91,47%	2.258	97,75%	52	2,25%
2003	2.107	94,53%	122	5,47%	40	1,79%	107	4,80%	2.082	93,41%	2.174	97,53%	55	2,47%
2004	2.080	95,54%	97	4,46%	35	1,61%	80	3,67%	2.062	94,72%	2.125	97,61%	52	2,39%
2005	2.026	95,75%	90	4,25%	45	2,13%	78	3,69%	1.993	94,19%	2.059	97,31%	57	2,69%
2006	1.774	94,61%	101	5,39%	12	0,64%	97	5,17%	1.766	94,19%	1.859	99,15%	16	0,85%
2007	1.680	94,33%	101	5,67%	1	0,06%	96	5,39%	1.684	94,55%	1.775	99,66%	6	0,34%
2008	1.608	93,93%	104	6,07%	1	0,06%	99	5,78%	1.612	94,16%	1.706	99,65%	6	0,35%
2009	1.531	91,79%	137	8,21%	0	0,00%	132	7,91%	1.536	92,09%	1.663	99,70%	5	0,30%
2010	68	93,15%	5	6,85%	0	0,00%	5	6,85%	68	93,15%	73	100,00%	0	0,00%
Total	**30.261**	**90,11%**	**3.323**	**9,89%**	**457**	**1,36%**	**3.094**	**9,21%**	**30.033**	**89,43%**	**32.898**	**97,96%**	**686**	**2,04%**
USA dataset														
	Benchmark classification				Ohlson						Ohlson classification			
	Normal		FD		Type I		Type II		Correct		Normal		FD	
Year	Absolute	%	Absolute	%	Absolute	Relative	Absolute	Relative	Absolute	Relative	Absolute	%	Absolute	%
1990	1.214	80,83%	288	19,17%	561	37,35%	33	2,20%	908	60,45%	686	45,67%	816	54,33%
1991	1.206	80,56%	291	19,44%	559	37,34%	37	2,47%	901	60,19%	684	45,69%	813	54,31%
1992	1.240	82,56%	262	17,44%	585	38,95%	37	2,46%	880	58,59%	692	46,07%	810	53,93%
1993	1.250	83,00%	256	17,00%	576	38,25%	23	1,53%	907	60,23%	697	46,28%	809	53,72%
1994	1.252	83,02%	256	16,98%	553	36,67%	29	1,92%	926	61,41%	728	48,28%	780	51,72%
1995	1.270	84,11%	240	15,89%	533	35,30%	16	1,06%	961	63,64%	753	49,87%	757	50,13%
1996	1.267	84,41%	234	15,59%	519	34,58%	22	1,47%	960	63,96%	770	51,30%	731	48,70%
1997	1.248	83,81%	241	16,19%	495	33,24%	23	1,54%	971	65,21%	776	52,12%	713	47,88%
1998	1.209	81,47%	275	18,53%	460	31,00%	32	2,16%	992	66,85%	781	52,63%	703	47,37%
1999	1.134	81,70%	254	18,30%	394	28,39%	27	1,95%	967	69,67%	767	55,26%	621	44,74%
2000	1.027	81,31%	236	18,69%	335	26,52%	26	2,06%	902	71,42%	718	56,85%	545	43,15%
2001	950	80,92%	224	19,08%	312	26,58%	28	2,39%	834	71,04%	666	56,73%	508	43,27%
2002	869	78,71%	235	21,29%	289	26,18%	49	4,44%	766	69,38%	629	56,97%	475	43,03%
2003	850	81,11%	198	18,89%	271	25,86%	34	3,24%	743	70,90%	613	58,49%	435	41,51%
2004	845	85,01%	149	14,99%	260	26,16%	23	2,31%	711	71,53%	608	61,17%	386	38,83%
2005	796	86,05%	129	13,95%	220	23,78%	20	2,16%	685	74,05%	596	64,43%	329	35,57%
2006	742	84,99%	131	15,01%	197	22,57%	22	2,52%	654	74,91%	567	64,95%	306	35,05%
2007	690	85,40%	118	14,60%	179	22,15%	19	2,35%	610	75,50%	530	65,59%	278	34,41%
2008	638	83,95%	122	16,05%	162	21,32%	22	2,89%	576	75,79%	498	65,53%	262	34,47%
2009	606	83,24%	122	16,76%	156	21,43%	25	3,43%	547	75,14%	475	65,25%	253	34,75%
2010	112	85,50%	19	14,50%	32	24,43%	3	2,29%	96	73,28%	83	63,36%	48	36,64%
Total	**20.415**	**82,67%**	**4.280**	**17,33%**	**7.648**	**30,97%**	**550**	**2,23%**	**16.497**	**66,80%**	**13.317**	**53,93%**	**11.378**	**46,07%**

Table 14: *the top panel is made up of three windows and concerns the G-7 dataset. The left hand window displays the binary variable of financial distress likelihood. This is considered the benchmark for all three models in this paper. It displays for each year which firm is ex-post classified as normal (Normal) or financially distressed (FD). The middle window displays the Type I and Type II errors from our replicated Ohlson model. The right hand window displays the number of observations as classified by our replicated Altman model. The bottom panel is to be interpreted in a similar way, but then for the USA dataset.*

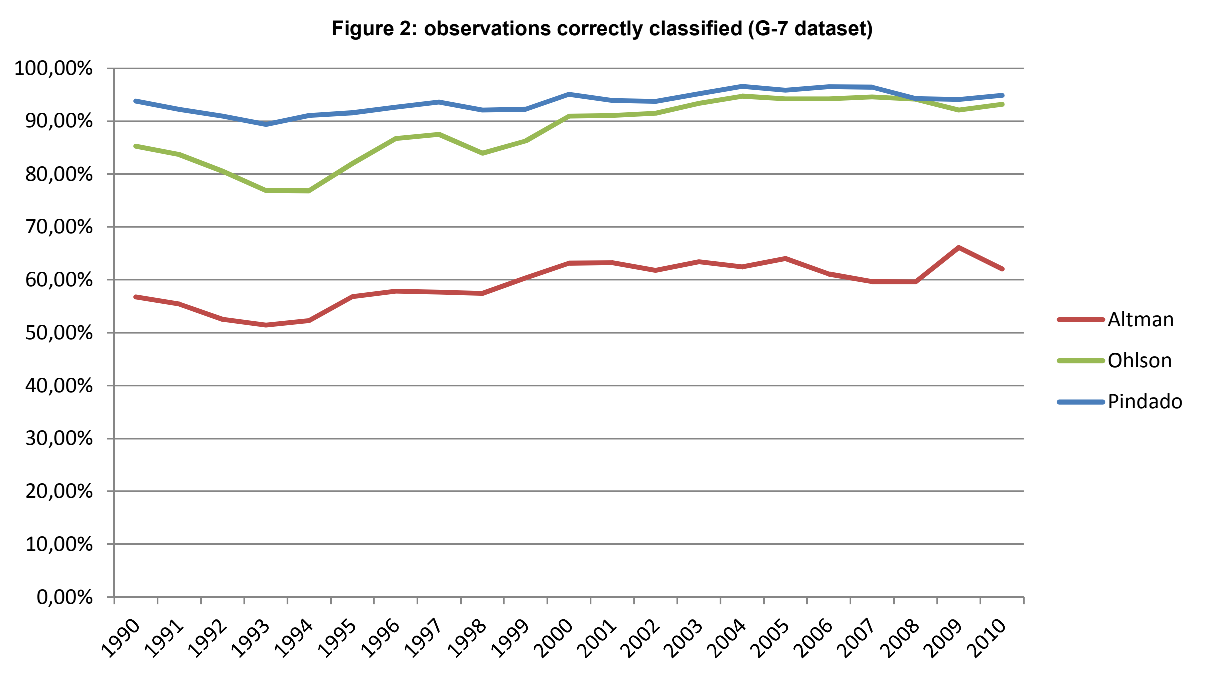


Figure 2: *this is a graphical display of the relative number of observations correctly classified as either financially distressed, either normal by all three models discussed. The above graph concerns itself with the G-7 dataset. As can be seen, the Pindado model consistently yields the best results for the G-7 dataset. The Ohlson model is, especially in the 2000 – 2010 period, a close second. The Altman model performs the worst.*

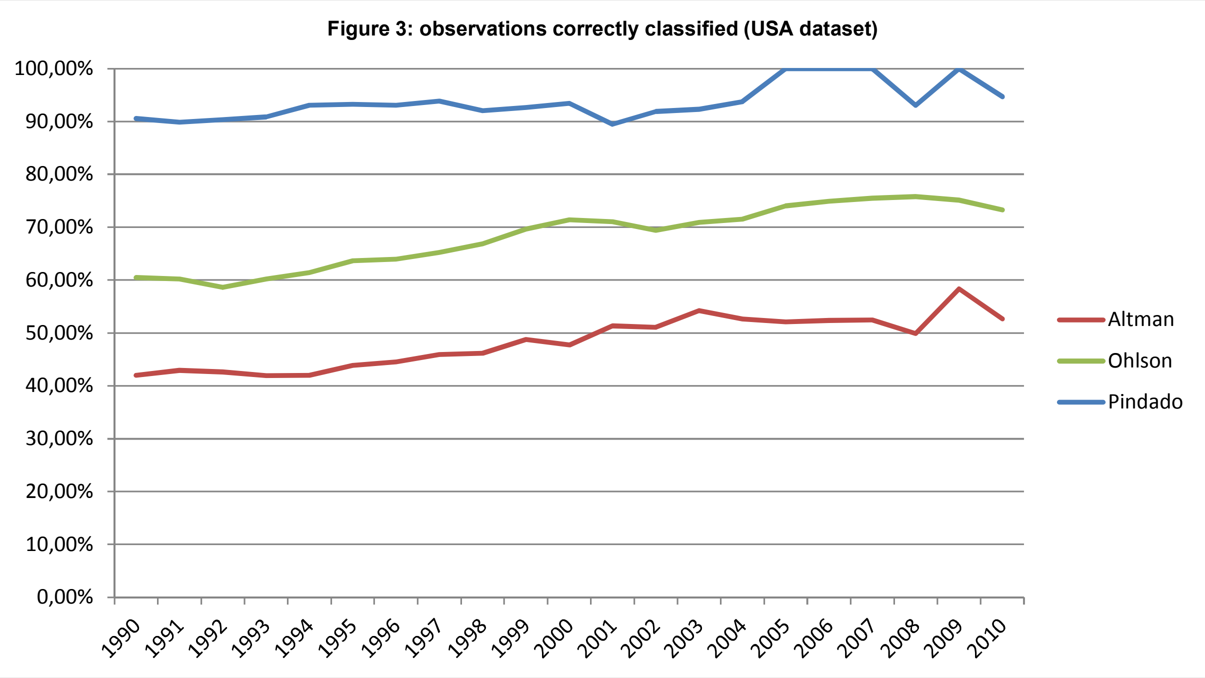


Figure 3: *this is a graphical display of the relative number of observations correctly classified as either financially distressed, either normal by all three models discussed. The above graph concerns itself with the USA dataset. From this out-of-sample test we see that again the Pindado model yields the best results. Note that for the years 2005, 2006, 2007 and 2009, the model did not iterate. The Pindado model performs less than it did in the G-7 dataset due to the large amount of Type I errors, but is still second best to the Pindado model. Again, the Altman model performs worst.*

These figures allow us to answer our first hypothesis, that the Pindado model is indeed superior to the other two models in regard of reliability of predictive power concerning financial distress likelihood.

Another interesting aspect would be to see the practical implication of information-asymmetry as discussed in our theoretical framework. Chandra and Nayar (2008) state that a firm would attract debt prior to announcing a downturn in expected performance. Following Cleary (1999), taking into account our binary definition of the financial distress variable, this should cause an increase in the number of observations classified as financially distressed following the years of the financial crisis. See figure 4 for the G-7 dataset, and figure 5 for the USA dataset, respectively.

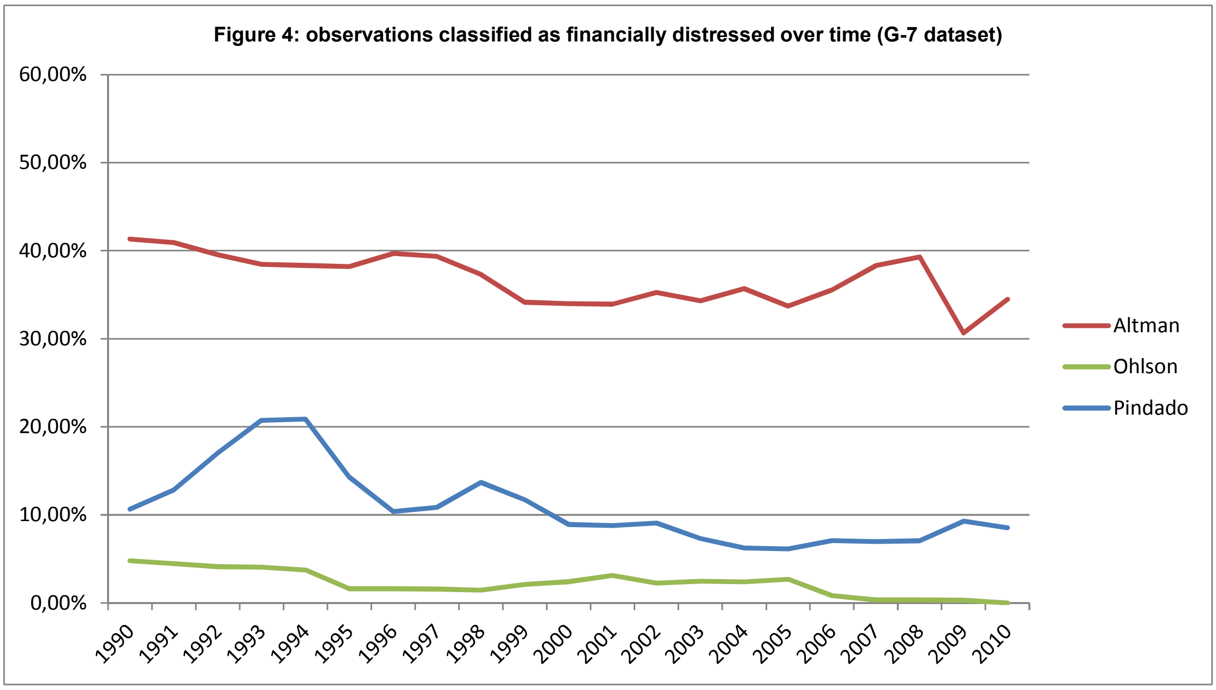


Figure 4: *this figure displays the number of observations classified as financially distressed within the G-7 dataset. As discussed, the Altman model has a large number of Type I errors, thus resulting in many false positives in misclassifying observations as financially distressed. The Pindado model shows a spike in financially distressed firms in the nineties, and declines steadily over time. The Ohlson model seems the most conservative of the three models, in classifying the least observations as financially distressed.*

Neither in figure 4, nor in figure 5 can we establish a significant increase in the number of observations classified as financially distressed. Note however that the financial distress variable is backwards looking over the past two years. It could be that the effects of financially distressed firms have not yet manifested itself within the sample period. Additionally, neither do we find any significant increases in financially distressed observations during the period of the burst of the internet bubble early 2000's.

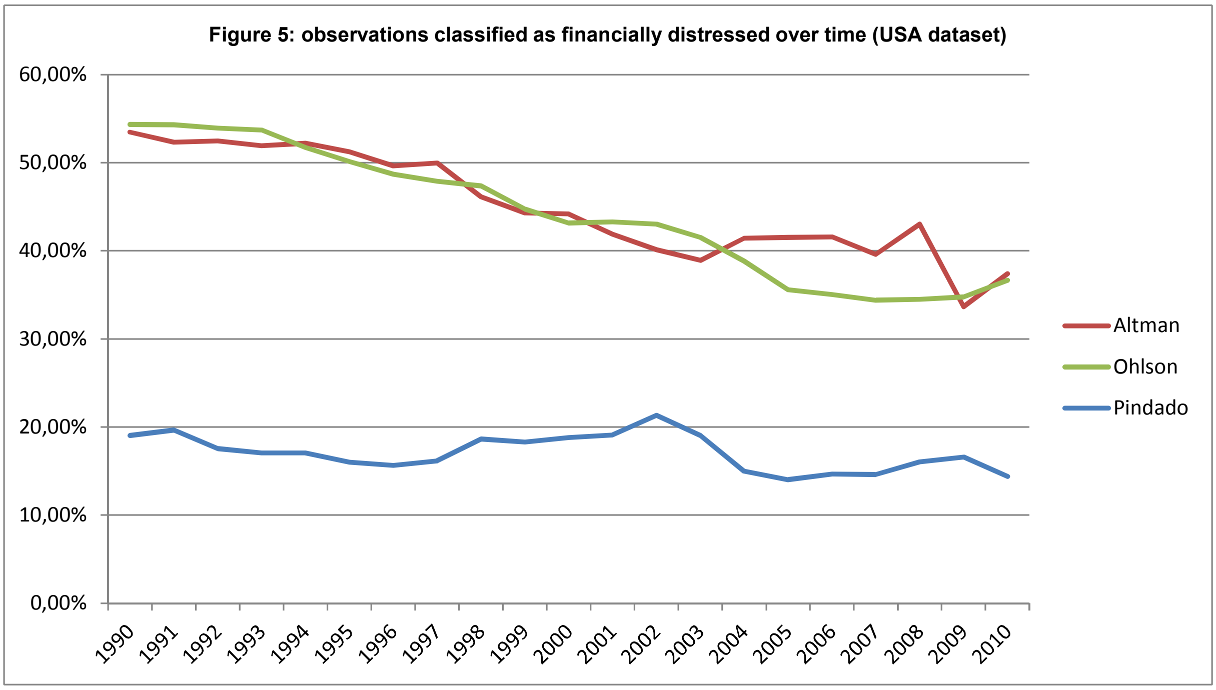


Figure 5: *the figure above displays the number of observations classified as financially distressed within the USA dataset. Here, we see that the Ohlson model shares the same trend in Type I errors as the Altman model, making a strong contrast with the G-7 dataset. The Pindado model remains robust, even in this out-of-sample test.*

This leads us to conclude our second and third hypotheses, that we cannot find any clear establishment of information-asymmetry taking effect in the classification of financial distress likelihood. This contrasts our findings with that of Chandra and Nayar (2008). Bushee *et al.* (2010) argues that the media generally focuses on large firms, which is what the CompuStat database – and thus our dataset as well as our out-of-sample dataset – primarily consists of. Therefore it cannot be stated that information-asymmetry does not occur. We simply fail to establish empirical evidence in this case due to inherently biased data due to firm size.

6. Conclusion

In this research paper we illustrate the evolution of financial distress likelihood models over the past five decades. We selected the Altman model, originally from 1968 and revisited in 2000, the Ohlson model of 1980 and the Pindado model of 2008. In order to let financial distress manifest itself, debt – or its related costs – is a necessary factor. Therefore a theoretical framework in the shape of the capital framework is briefly discussed. Debt holders and equity holders have their own stakes in a firm. When there are multiple parties involved in a firm, information-asymmetry may arise, as one group of investors may try to secure their investment at the expense of the other group. Therefore, we also formulated possible hypotheses illustrating the relation between information-asymmetry and financial distress likelihood.

Upon discussing the theoretical background of the three financial distress likelihood models, we proceeded with the methodology. Since we are dealing with a discriminant model and two logarithmic models, there are methodological implications in comparing the outcomes of the three models. We then proceed by replicating the three models based on our dataset acquired from CompuStat. To verify our findings, we created two datasets: one consisting of the G-7 countries excluding the USA, and an out-of-sample dataset containing data on USA firms. Both concern themselves with the period of 1990 to 2010, allowing us to work with a large number of observations in separate geographical locations. This is especially interesting for the Ohlson (1980) and Pindado (2008) models, as they are logarithmic in nature. The econometric specifications of these models really shine when employed on large datasets. Secondly, it allows us to verify whether the Altman (2000) model, while not being a logarithmic model, is equally able in the sense of predictive reliability to work with large numbers of observations.

Each model has its own approach and manner of interpretation. This causes a conflict to rise in making a rational comparison with respect to the effectiveness of these three models, with effectiveness being defined as the reliability in accurately predicting financial distress likelihood of occurring. This we solve by following Cleary (1999), by comparing the number of observations which are correctly classified – either as normal, otherwise as financially distressed. Since financial distress is a rare occurrence, Woolridge (2009) argues that our chosen approach may yield possible results which may be misleading, effectively overstating the frequency of the event – being financial distress – occurring. To correct for this, we calculate the number of correctly classified observations on a yearly basis.

Based on this method, we find that the Pindado model is indeed superior to the Altman and Ohlson models. It confirms the improving evolution of computer technology (read: the ability to handle large volumes of data) as well as the evolution in predictive reliability with respect to financial distress likelihood. In contrast with Chandra and Nayar (2008), we find no evidence for information-asymmetry manifesting itself in our sample period. This is true for both the internet bubble collapse early 2000's, as well as the financial crisis of 2008. Bushee *et al.* (2010) points out that this could be due to inherent dataset bias. The media tends to coverage events of large firms in a frequent fashion. Since we obtained our data from the CompuStat database, which primarily exists of large firms, this could be a reason for our lack of empirical evidence with respect to information-asymmetry, and as a consequence, the capital structure of the firm.

We conclude that financial distress likelihood models have indeed improved over the past four decades. Whereas the early models merely concerned themselves with the likelihood of bankruptcy (Ohlson; 1980; Altman, 1968), these models were reviewed and quickly developed to entail a broader span, being the probability of financial distress occurring. The practical added value of such models is obvious, as a firm is more likely to rebound from financial distress than it does from filing for bankruptcy, provided that it is aware of its current financial position.

Reference List

Research Papers

Altman, E. 1968. "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy." *Journal of Finance*, Vol. 23, p. 589 – 609.

Altman, E. 1984. "A Further Empirical Investigation of the Bankruptcy Cost Question." *Journal of Finance*, Vol. 39, p. 1067 – 1089.

Altman, E. 2000. "Predicting Financial Distress of Companies: Revisiting the Z-Score and Zeta® Models." *http://pages.stern.nyu.edu/~ealtman/Zscores.pdf*.

Andrade, Gregor, and Steven N. Kaplan. 1998. "How Costly is Financial (Not Economic) Distress? Evidence from Highly Leveraged Transactions That Became Distressed." *The Journal of Finance*, Vol. 53, No. 5.

Begley, J., J. Ming, S. Watts. 1996. "Bankruptcy Classification Errors in the 1980's: An Empirical Analysis of Altman and Ohlson's Models." *Review of Accounting Studies*, Vol. 1, p. 267 – 284.

Bernanke, Ben S., John Y. Campbell, Benjamin M. Friedman and Lawrence H. Summers. 1988. "Is there a corporate debt crisis?" *Brookings Papers on Economic Activity*, Vol. 1988, No. 1, p 83 – 139

Bhagat, S., N. Moyen, S. Inchul. 2005. "Investment and Internal Funds of Distressed Firms." *Journal of Corporate Finance*, Vol. 11, No. 3, p. 449 – 472.

Bushee, B., J. Core, W. Guay, and S. Hamm. 2010. "The role of the business press as an information intermediary." *Journal of Accounting Research*.

Brounen, Dirk, Abe de Jong, Kees Koedijk. 2005. "Capital Structure Policies in Europe: Survey Evidence." *ERIM Report Series Research in Management*.

Carey, M., S. Prowse, J. Rhea, G. Udell. 1993. "The Economics of the Private Placement Markets: A New Look." *Financial Markets, Institutions and Instruments*, Vol. 2, p. 1 – 66.

Chandra, Uday, Nandkumar Nayar. 2008. "The Information Content of Private Debt Placements." *Journal of Business Finance & Accounting*, Vol 35 (9 and 10), p. 1164 – 1195, November / December.

Deakin, E.B. 1972. "A Discriminant Analysis of Predictors of Business Failure." *Journal of Accounting Research*, March.

Denis, David J., Vassil T. Mihov. 2003. "The Choice Among Bank Debt, Non-Bank Private Debt, and Public Debt: Evidence From New Corporate Borrowings." *Journal of Financial Economics*, Vol. 70, Issue 1, p. 3 – 28.

Dichev, I. 1998. "Is the Risk of Bankruptcy a Systematic Risk?" *Journal of Finance*, Vol. 53, p. 1131 – 1147.

Grayham, John R., and Campbell Harvey. 2001. "The Theory and Practice of Corporate Finance: Evidence From the Field." *Journal of Financial Economics*, Vol. 60, p. 187 – 243.

Grice, J., and M. Dugan. 2001. "The Limitations of Bankruptcy Prediction Models: Some Cautions for the Researcher." *Review of Quantitative Finance and Accounting*, Vol. 17, No. 2, September, p. 151 – 166.

Grice, J. and M. Dugan. 2003. "Re-estimations of the Zmijewski and Ohlson Bankruptcy Prediction Models." *Advances in Accounting*, Vol. 20, p. 77 – 93.

Kim, E. 1978. "A Mean-Variance Theory of Optimal Capital Structure and Corporate Debt Capacity." *Journal of Finance*, Vol. 33, p. 45 – 64.

Kwan, Simon, Willard T. Carleton. 2004. "Financial Contracting and the Choice Between Private Placement and Publicly Offered Bonds." *FRB of San Francisco*, Working Paper No. 2004-20.

Modigliani, Franco, and Merton H. Miller. 1958. "The Cost of Capital, Corporation Finance and the Theory of Investment." *The American Economic Review*, Vol. 43, No. 3, June.

Myers, Stewart C. 1984. "Capital Structure Puzzle." *Journal of Finance,* Vol. 39, No. 3, p. 575 – 592.

Ohlson, J. 1980. "Financial Ratios and the Probabilistic Prediction of Bankruptcy." *Journal of Accounting Research*, Vol. 18, p. 109 – 131.

Pindado, Julio. 2006. "Estimating the Probability of Financial Distress: International Evidence." *Journal of Economic Literature,* January 6.

Pindado, Julio, Luis Rodrigues, Chabela de la Torre. 2008. "Estimating Financial Distress Likelihood." *Journal of Business Research*, Vol. 61, p. 995 – 1003.

Wittenberg-Moerman, Regina. 2008. "The Role of Information Asymmetry and Financial Reporting Quality in Debt Trading: Evidence from the Secondary Loan Market." *Journal of Accounting and Economics*, No. 46, p. 240 – 260.

Zmijewski, M. 1984. "Methodological Issues Related to the Estimation of Financial Distress Prediction Models." *Journal of Accounting Research* (Supplement), Vol. 22, p. 59 – 86.

Books

Babbie, Earl. 2004. "The Practice of Social Research", 10th Ed., p. 104 – 114.

Wooldridge, J. 2009. "Introductory Econometrics", 4th Ed., South-Western Cengage Learning.

Media

CompuStat Global, http://wrds-web.wharton.upenn.edu

CompuStat North America, http://wrds-web.wharton.upenn.edu

Hilsenrath, Jon, Serena Ng, Damian Paletta. 2008. "Worst Crisis Since 30's, With No End Yet in Sight." *The Wall Street Journal*, September 18, A1.

Pendery, David. 2009. "Top Three Economists Agree 2009 Worst Financial Crisis Since Great Depression; Risks Increase if Right Steps are Not Taken." *http://www.businesswire.com*, February 13.

OECD Statistics, http://stats.oecd.org

Pugh, Tony. 2009. "Businesses Struggle as Bank Loans Remain Elusive." *The News & The Observer*, May 14, Politics Section.

Appendix A

Variables employed for the three financial distress probability models	
Variable name	**CompuStat code**
total liabilities	LT
total assets	AT
current assets	ACT
current liabilities	LCT
EBITDA	EBITDA
EBIT	EBIT
financial expenses	XINT
retained earnings	RE
capital expenditure	CAPX
common equity	CEQ
depreciation and amortization	DP
intangible assets	INTAN
net sales	SALE
net income	UNNP / NINC*
*Substitution of UNNP in case of unavailability of data	

Appendix B

Capital Goods Price Index.
(source: http://stats.oecd.org)

Capital Goods Price Index (1990 = 100)																						
	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Canada	95.734	100.000	106.233	108.420	110.513	110.667	113.044	114.753	116.593	118.224	119.972	121.984	124.533	127.879	130.968	132.716	134.503	136.205	138.982	140.441	142.834	144.617
France	96.989	100.000	103.442	106.952	109.798	111.896	114.037	116.184	117.273	118.217	118.915	119.528	120.876	123.439	126.047	128.868	130.504	131.713	133.494	135.419	137.785	139.028
Germany	97.559	100.000	103.852	109.789	115.825	119.168	121.651	123.653	126.024	127.628	128.410	129.413	131.026	133.080	134.339	136.645	138.007	138.969	141.574	143.358	145.292	146.338
Italy	94.237	100.000	105.770	111.801	117.756	122.669	128.869	134.285	137.776	141.176	143.658	146.524	150.356	154.341	158.217	161.682	164.751	167.455	169.993	173.724	176.652	179.504
Japan	97.462	100.000	102.699	105.247	106.812	107.767	108.264	108.714	110.564	111.390	111.280	110.675	109.694	108.916	108.602	108.144	107.789	107.350	107.149	107.291	106.665	105.425
United Kingdom	91.156	100.000	105.786	110.138	111.672	114.480	118.292	120.873	123.121	124.933	125.805	125.917	127.270	129.149	130.805	132.236	134.227	135.972	138.276	140.513	142.996	147.112
United States	95.214	100.000	104.909	108.748	112.341	115.546	119.003	122.221	125.137	127.998	130.655	133.836	137.410	140.597	142.645	145.155	148.305	152.015	155.567	159.142	161.847	163.398