Bachelor's Thesis in Retail Management
Stockholm School of Economics
Spring 2021

Who Saw It Coming?

The Efficacy of Bankruptcy Prediction Models on Private Retail Firms

Erik Larsson (50608)

Edvin Lindhout (50592)

Supervisor: Mariya Ivanova
Date: 5 June 2021

Abstract

Bankruptcy is never considered a desired outcome, and methods of predicting it have received extensive attention during the past decades. However, as a consequence of weak academic interest and peculiar financial profiles, retail bankruptcy prediction remains rather unexplored. Likewise, the majority of existing bankruptcy prediction models have been developed on publicly listed firms—restricting their applicability to privately held companies. Consequently, there is a gap in bankruptcy prediction for privately held retailers—worsening their ability to signal, capture, and divert financial distress. Our study bridges that gap by analysing the efficacy of a selection of bankruptcy prediction models, both general and retail-specific, to establish which one reigns supreme in predicting bankruptcies for private retail firms. We find that Ohlson's O-score performs best in our sample of over 14,000 Swedish private retail firms. Additionally, we find evidence in the literature and our empirical study to support the use of different sensitivity levels in the models, depending on the use case. Finally, we discuss limitations of our research design and give suggestions for further research.

Keywords: Bankruptcy, Private Firms, Retail, Financial Distress, Prediction, Sensitivity, Capitalization of Leases

A special thank you to our competent and pedagogic supervisor Mariya Ivanova, for repeatedly steering this thesis into the right direction. Without your valuable insights, this thesis would not have been as rewarding to write nor as interesting to read.

Table of contents

Contents

Introduction

Bankruptcy prediction has been and continues to be a relevant topic across all major industries considering the negative impacts brought by bankruptcies for associated stakeholders. As bankruptcies are expensive for everyone involved, any efforts to identify and potentially reduce their impact should be pursued (Bhargava, 1998). Because of this, one role of bankruptcy prediction is to foresee bankruptcies and thereby allow managers and other stakeholders to mitigate the damages caused by bankruptcy before the event occurs. However, predicting bankruptcies also serves a broader, more commercial purpose in helping creditors and financial analysts evaluate companies—allowing for greater efficiency in capital markets and more accurate financial evaluations.

The practice of predicting bankruptcies stretch across all industries and creating bankruptcy prediction models tailored for specific sectors has been common practice since the middle of the 20th century. However, while the majority of bankruptcy research is industry agnostic, some industries such as manufacturing and banking have been overrepresented in sector-specific bankruptcy prediction research, while other industries such as the retail sector have received little attention despite its central role to the economy (Bellovary et al., 2007; Fejér-Király, 2015; Bhargava, 1998). Several explanations have been offered to account for the disparity between the retail sector's importance and the attention it has received in bankruptcy literature. Altman & Levallee's (1981) proposed that the noncapitalization of leases can cause an overinflation of assets for retailers owning their stores, leading to a higher degree of misclassification when determining bankruptcy. Sharma & Mahajan (1980) suggested that retail bankruptcies often are regarded as atypical and therefore have not caught the attention of researchers. Regardless, the retail sector's sheer size warrants further investigation to determine if bankruptcy prediction is possible despite the challenges posed.

Another underexplored frontier in bankruptcy research is that of private companies. Despite private firms constituting a vast majority of businesses and business failures (Altman, 1993), the focus of research has been on listed firms. The most likely explanation is that most bankruptcy research publications come from the U.S. and are based on U.S. firms (Bellovary et al., 2007), where access to financial data on private firms is limited. Adding to the body of research on private firm bankruptcy prediction is valuable for several reasons. Firstly, public firms often carry a much wider toolkit to avoid bankruptcy compared to private firms, meaning that it is all the more important to foresee an impending private firm bankruptcy in order to avoid excessive losses. Secondly, most existing models are specifically designed for public firms, such as by using a firm's market value of equity, and are therefore treacherous to use on private firms without careful consideration. Lastly, bankruptcy prediction on Swedish private firms is of particular interest. This because, contrary to filing for Chapter 11 in the U.S., a manager filing for bankruptcy in Sweden surrenders control over firm assets. This creates an incentive to keep the firm in going concern instead of alarming creditors. Paired with the informational advantage managers tend to have over creditors concerning the firm's financial status, the playing field is not level (Thorburn, 2000). Therefore, expanding the field to offer stakeholders of private retailers more tools to predict and avoid bankruptcy-incurred losses, is something to pursue.

An additional research gap was found in that private retailers warrant attention from researchers due to the dynamic and quickly changing business climate they are operating in, often referred to as the retail apocalypse (Helm et al, 2020). A study by Oliver Wyman, discussing the future of retail, identify the emergence of several new trends as potentially devastating for retailers unable to keep up (Harrison & Thomas-Dupuis, 2018). The rise of e-commerce business models and omnichannel retailing pressures retailers to differentiate themselves and redefine the function of the brick-and-mortar store (Shankar et al, 2021). Furthermore, the traditional value chain that has been at the core of retailer operations is changing, with able suppliers establishing direct-to-consumer relationships, relegating retailers from being gatekeepers to the customers (Harrison & Thomas-Dupuis, 2018). Lastly, with only 22 Swedish retailers being publicly traded, privately held companies account for the vast majority of retail firms (Aktiespararna, 2021). Therefore, this study could add valuable insights into the outlook of these pressured retailers and provide tools for a retail manager to assess the financial health of their firm.

Purpose & Research Question

The purpose of this study is to evaluate the accuracy of established bankruptcy prediction models on privately held retailers. To do this, a brief background on bankruptcy prediction research and bankruptcy research within the retailing industry is provided before presenting the models selected for evaluation. Furthermore, with this thesis, we also aim to present which level of sensitivity is favourable in a model attempting to indicate financial distress for private retailers – the benefits of which are twofold. Firstly, by presenting which level holds the highest predictive ability, we aim to contribute to bankruptcy prediction research wherein future model developments may benefit from our findings. Secondly, we aim to offer practical insights and advice on how to treat a prediction, and why a false positive may be a signal for further investigation rather than a wrong prediction.

The above purpose thereby results in the following research question:

- *What is the accuracy, efficacy, and applicability of bankruptcy prediction models on private retail firms?*

With the research question, we aspire to establish the applicability of foreign bankruptcy prediction models on privately held retailers, leading the way for further research and use in practice.

Review of literature

History of Bankruptcy Prediction

Due to the devastating nature of bankruptcies, several researchers have taken it upon themselves to develop a way for which to predict business failure. Since bankruptcy prediction research first emerged roughly 90 years ago, this has been done in several different ways, often based on quantitative financial ratios and accounting measures (Bellovary et al., 2007). The first models presented in the 1930's were based on univariate analysis and often targeted mid-sized manufacturing firms, with the most famous univariate model developed by Beaver (1966). While the use of univariate analysis eventually grew obsolete, the models generally performed quite well with Beaver (1966) presenting variables with a successful predictive ability of more than 90% within a year from bankruptcy. Furthermore, the univariate models also laid the groundwork for multivariate bankruptcy prediction models (Bellovary et al., 2007).

Univariate analysis remained the most popular statistical method of predicting bankruptcy until 1968 when Altman (1968) presented the first multivariate discriminant analysis (MDA) model—summarizing the weights of five individual financial measures to get a cumulative score, referred to as a Z-score. The research presented by Altman (1968) laid the foundation of modern bankruptcy prediction and is often used as a benchmark for modern bankruptcy prediction models despite the model being over 50 years old. After being presented in 1968, Altman's bankruptcy prediction model spurred a steep growth in the number of bankruptcy prediction models being introduced, with the number growing from less than ten during the 1960's to over 70 new models in the 1990's (Bellovary et al., 2007). Beyond aiming to achieve higher accuracies, subsequent bankruptcy prediction models and research in the field has mainly centred around extending bankruptcy prediction models to also include contexts, primarily industry adaptation (Bellovary et al., 2007; Fejér-Király, 2015), but also qualitative settings such as competition (Mokrišová & Horváthová, 2018) and audit reports (Fejér-Király, 2015).

While MDA models still remain relevant, a shift towards more progressive bankruptcy prediction models based on logit analysis, probit analysis, and neural networks occurred in the 1980's-1990's after Ohlson (1980) introduced the first logit-based bankruptcy prediction model and achieved accuracies as high as 96%. The main difference for practitioners between univariate/multivariate analysis models and the aforementioned modern models is that the former is binary while the latter accounts for the probability of the firm going bankrupt (Ohlson, 1980).

Moving into the 2000's, bankruptcy prediction models being developed became evermore data-driven as it became easier to leverage computers in handling larger datasets, allowing for wider analysis using more variables. Therefore, the 1990's growth in Neural Network models continued throughout the 2000's, while the growth of artificial intelligence (AI) based models took off in popularity after 2005. This led to models becoming less transparent due to the neural networks, and most computer-driven models' "black box" nature—giving a practitioner the percentage risk of bankruptcy but being unable to explain the drivers of the risk (Fejér-Király, 2015).

Despite their high predictive accuracies, bankruptcy prediction models remained rather under-utilized by practitioners prior to the 2008 great financial crisis (Bellovary et al., 2007). However, following the crisis, the models grew in popularity, primarily among creditors who were forced to comply with the BASEL agreements, but also among analysts as bankruptcy prediction models were deemed a great tool within corporate finance (Fejér-Király, 2015).

Following the crisis, the growth of computer-driven models continued, which had another implication for bankruptcy research as it allowed for more sophisticated industry-specific models. This led to models, other than those claiming industry agnosticism, going from being primarily developed within manufacturing and banking, to being developed within highly specific industries (Fejér-Király, 2015). However, despite the enhanced ability to predict bankruptcy in specific industries, roughly 80% of the models developed between 2015-2020 were developed using industry agnostic firm data (Clement, 2020) and few to none of the industry-specific models were developed for the retailing sector (Fejér-Király, 2015).

Industry-specific models may have achieved increased adoption among practitioners, but there is little to show that researchers are becoming better at predicting bankruptcy. While neural networks historically perform somewhat better on average than other methods, the difference is slim. Likewise, both MDA models and neural networks have been able to report accuracies over 95% out of sample. Furthermore, the advantage of being able to analyse a superior number of variables through a computer-based model is contradictory to the fact that the number of variables in a bankruptcy prediction model does not affect the predictiveness of the model. The claim that more variables increase predictive ability has been debunked as models analysing two variables have proven equally predictive as models analysing over 50 variables (Fejér-Király, 2015; Altman, 2000).

Bankruptcy prediction in the retailing industry

Despite its central role to the global economy, the retail industry has been widely overlooked in bankruptcy research—with only a handful of models being developed for the sector. Likewise, much of the research conducted within the retail bankruptcy field has been the modification of existing models, such as the model presented by Altman et al (1977). In their modification, Altman et al (1977) extended Altman's previous work, i.e., the original Z-score, to improve classification rates for the retail sector by capitalizing their leases to remove any structural differences between retailers owning and leasing their stores. In their 1977 article, Altman et al further introduced the fact that their developed model seemed to be equally accurate for both retailing and manufacturing firms. Therefore, the many extensions of Altman's original Z-score (developed for manufacturers) into the retail sector may come because of the mentioned similarity of manufacturing and retail firms in predicting bankruptcy. However, through a model introduced merely four years later in 1981, Altman & Levallee (1981) found that their misclassified firms were predominantly from the retailing sector and not manufacturing—casting doubt on the predictive resemblance between the retailing and manufacturing industry.

Similar results to that of Altman & Levallee (1981) were found by McGurr & DeVaney (1998) who applied five different industry salient bankruptcy prediction models to a sample consisting solely of retail firms. Three of the applied models were based on logistic regression while the other two predicted through MDA. The five models were further differentiated by the independent variables taken into account, with models considering both financial and nonfinancial variables being used. The results displayed the models' successfully predicting bankruptcy in 82%-96% for industry salient firms as compared to average accuracies between 57%-78% when the same models were applied to the retail sample. These results further reiterate the findings of Altman & Levallee wherein caution should be taken before predicting bankruptcy for retail firms through industry salient models. Moreover, the need for caution is important when the leases of the sampled retail firms are not capitalized as this will make retailers owning their stores less comparable to retailers leasing them (Altman & Levallee, 1981; Bhargava et al., 1998).

Further extensions of Altman's original work in 1968, can be found in Bengley et al (1997) who reestimated the model and improved its accuracy through their adjusted Z-score. Similar extensions into the retail industry were made by Pang & Kogel (2013), who reestimated Bengley et al's reestimation of Altman's original 1968 model to improve its accuracy within the retail sector. Likewise, Pang & Kogel also reestimated Altman's 2000 reestimation of his 1968 model, to better adapt the model to the retailing sector. Pang & Kogel's reestimations for both Bengley et al and Altman's model proved superior at predicting bankruptcies within the retail sector—reaching an average accuracy of 93% compared to Altman's 2000 reestimation model accuracy of 87% and Bengley et al's accuracy of 80% on the same sample data.

While many of the documented retail-oriented bankruptcy prediction models are reestimations of older, nonretail targeted models, some original models have been developed specifically for the sector. Considered as the first was Sharma & Mahajan (1980) whose MDA model analysed a set of two independent variables (return on assets and current ratio) to predict bankruptcy up to five years before it occurs. The model became rather successful as it was able to predict bankruptcy with an accuracy of 92% up to a year before the event. In their study, Sharma & Mahajan further proposed that retail bankruptcy is not necessarily a long, exhaustive process but rather swift and hard to predict more than one year ahead of time—contradicting Altman's general bankruptcy prediction findings as he believed bankruptcy to be a long and foreseeable occurrence (Altman & Levallee, 1981).

Another MDA based retail bankruptcy prediction model was published by McGurr & DeVaney (1998) who went beyond accounting-based independent variables to introduce variables such as sales per employee—a variable which can be deemed highly retail tailored. However, in spite of McGurr's higher number of independent variables (seven vs Sharma & Mahajan's two), the model's predictive ability was inferior to that of Sharma & Mahajan as its predictive accuracy averaged between 70%-75% on various samples. This reiterates the fact that the number of independent variables used in a model does not necessarily correlate with its predictive accuracy (Bellovary et al., 2007).

While almost all bankruptcy prediction models within the retail space use MDA to foresee bankruptcy, Klepáč & Hampe, (2016) were one of the first to use SVM classifiers to create a retail targeted model. Klepáč & Hampe's model outperformed Sharma & Mahajan's 92% accuracy with the SVM model managing to reach an accuracy of more than 95% one year ahead of bankruptcy—hinting toward the potential relevance of using computer-driven models to predict bankruptcies within the retail sector. However, one central limitation with Klepáč & Hampe's SVM model is that analysts are unable to detect the drivers of the bankruptcy risk.

Looking beyond bankruptcy prediction through modelling, several articles have been published on other variables' correlation to retail bankruptcy. One such publication is that of Chaganti et al (1985) who found that bankrupt retail firms tend to have smaller board sizes as compared to their solvent counterparts. While the predictive nature of nonfinancial variables such as board sizes is hard to establish, it is worth noting that bankruptcy research within the retail sector stretches beyond merely predicting firm failure.

Models Selected for Evaluation

The models selected to be evaluated on our dataset include a range of models, with two being retail-specific (Sharma & Mahajan, 1980; McGurr DeVaney, 1998) and the other three being largely industry agnostic (Altman, 2000; Ohlson, 1980). Since our study is conducted on private firms, we followed the literature and adopted the private-firm version of Altman's Z-model, where the market value of equity is substituted with the book value (See, for example, Altman and Saunders, 1997). See all chosen models and their independent variables in Table 1.

The models were chosen to reflect a diverse set of independent variables as models with overlapping independent variables are more likely to yield similar results. Furthermore, despite the possibility to enhance accuracy through nonfinancial measures (Fejér-Király, 2015), we chose to limit this study to quantifiable variables. This because we see the models as being early indicators of financial distress developed to guide managers who are better suited to evaluate their qualitative surroundings.

While aiming to avoid overlapping variables we chose to include three levels of model complexity in an effort to investigate whether complexity and intricacy correlate to higher accuracy for private retail firms, even though previous model comparison studies have failed to find such a correlation (See, for example, Bellovary, 2007). More specifically, we included Sharma & Mahajan's model (1980), which only uses two variables to capture profitability and debt structure. Altman's two models contain four and five independent variables and attempt to capture the two components with slightly more nuance. The final and most intricate category of models contains Ohlson's and McGurr & DeVaney's models, which use nine variables and seven variables respectively to capture the most nuances of the profitability and debt structure components, including past data to measure changes. By comparing the three categories we aimed to gain insights into what level of complexity is necessary in order to accurately predict bankruptcy. Furthermore, considering that there is a limited number of retail-specific models which are not revised versions of older models (such as Pang & Kogel (2013)), and the selection is reduced even further when excluding public firm models, we aimed to establish how the retail models hold up against Altman's and Ohlson's general models.

Altman (2000)

Altman's 2000 model is, as previously described, a modification of Altman's original Z-score model developed in 1968. The main adjustment Altman did in 2000 was to adjust the model to better fit private firms by adjusting his five independent variables for book value, rather than market value, of equity. Furthermore, in his 2000 model, Altman also adjusts his original 1968 model to better fit nonmanufacturing firms. This is done through the removal of the Sales/TA variable to "*minimize the potential industry effect which is more likely to take place when such an industry-sensitive variable as asset turnover is included*" (Altman, 2000). Due to the adjustment for nonmanufacturing firms, Altman (2000) presented two models in his article, where one is industry agnostic and the other adjusted to better fit nonmanufacturers—with both being adjusted for private firms.

Ohlson (1980)

Ohlson's model broke new ground in 1980 as he introduced logit analysis to reach an accuracy of 96%. While it was initially developed on industrial firms, the model is now regarded to be a generalist model, much like Altman's original 1968 model. Furthermore, Ohlson also introduced independent variables that differentiated Ohlson's model from its peers as he chose to partly focus on the size of the firms and finding it to be one of the main drivers of bankruptcy, while simultaneously elaborating on existing profit measures such as net income by looking at how profitability develops over time (Ohlson, 1980). Therefore, due to Ohlson's differentiated model—both in terms of independent variables and method of analysis—it became a natural model to include in this thesis.

Beyond its intrinsic ability to generate high accuracy rates, Ohlson's model was further selected as it was created in part due to Ohlson's disapproval of Altman's (1968) earlier works within the bankruptcy prediction space. Therefore, we believe that any study with the aim of comparing accuracy rates between bankruptcy prediction models which include Altman's models must also include Ohlson's work.

Sharma & Mahajan (1980)

Analysing the model developed by Sharma & Mahajan (1980) was considered crucial to evaluate the accuracy of bankruptcy prediction models on privately held retailers as Sharma & Mahajan developed the most accurate retail-specific bankruptcy prediction model with an accuracy of 92% one year prior to bankruptcy (Sharma & Mahajan, 1980). Furthermore, the work of Sharma & Mahajan (1980) became increasingly interesting to analyse as the model analyses a mere two independent variables, return on assets and current ratio—displaying a much lower level of intricacy compared to Ohlson and McGurr & DeVaney.

McGurr & DeVaney (1998)

McGurr & DeVaney's model was chosen to be included in this thesis despite their model resulting in rather moderate accuracy rates ranging between 70%-75% one year ahead of bankruptcy in their original study. We mainly chose to include their model due to its sole focus on the retail industry—making it easy to compare against Sharma & Mahajan. Moreover, the model was attractive for us to include as its number of independent variables analysed, as previously mentioned, are far greater than that of Sharma & Mahajan—giving us the possibility to establish a relationship, whether it be positive or negative, between predictive accuracy and model complexity. Moreover, the model's in-depth focus on retailers' liabilities made it increasingly attractive for our study. Lastly, the model's inclusion of measures found outside of a company's three financial statements, such as sales per employee, further made the model appealing to include in this thesis due to its differentiating factor.

Table 1

For explanation of variables, see table 7 in appendix.

Authors	Industry	Number of Variables Analysed	Variables Analysed	Accuracy in Original Study
Sharma & Mahajan (1980)	Retail	2	CA_CL, EBIT_TA	Years prior: 1 2 3 4 5 92% 78% 74% 73% 77%
McGurr & DeVaney (1998)	Retail	7	EBIT_TA, SALES_EMP, EBIT_SALES, CHLTD, CA_CL, LTL_TA, CHWC	70%-75% one year ahead of failure in numerous validation samples
Ohlson (1980)	Industrial	9	SIZE, TL_TA, WC_TA, CL_CA, OENEG, NI_TA, FFO_TL, NITWO, CHIN	Years prior: 1 2 1 or 2 96% 96% 93%
Altman (2000)	General adjusted for private firms and nonmanufacturers	5 and 4* in separate models *When also adjusting for nonmanufacturers	WC_TA, RE_TA, EBIT_TA, BV_TL, SALES_TA* *Sales/TA only included in the model adjusted for solely private firms	91% one year ahead of failure

Method

Data source

The data on private companies was derived from the Serrano Database. The Serrano Database is an extensive database on Swedish private (nonlisted) firms, containing financial statement information and information on bankruptcy filings. The financial statement data is based on financial statement data from the Swedish Companies Registration Office (Bolagsverket) and includes a firm's income statement and balance sheet items. In addition, there is history with general company data from Statistics Sweden (SCB), bankruptcy information from the Swedish Companies Registration Office, and group data from Bisnode's group register. The Serrano Database covers most legal forms in the Swedish business community. As of April 31st 2021, there are more than 1,200,000 firms in the total data set, with data spanning back to 1932 and up until 2018.

There are several advantages to the database, with the main one being the high level of cleaning and organization performed. There is one data entry per calendar year for the respective field of each company, allowing for relatively easy analysis of business trends and the application of statistical methods. The database also has an established framework for how to register and treat the gathered data to ensure comparability between companies and calendar years. The database is thus adjusted for phenomena such as broken accounting periods and omissions and gaps in submitted financial statements, which could otherwise pose challenges when analysing private firms.

Research design

The study looked at one economic event as the indicator of financial distress: bankruptcy filings. A bankruptcy filing in Sweden means management and shareholders immediately lose control over the firm, which is either liquidated or merged into a bankruptcy auction buyer's receiving company, effectively ending its operations (Thorburn, 2000). This is the event bankruptcy prediction models such as Altman's or Ohlson's attempt to predict, using financial ratios of varying nature and amount. A predictive model needs to have the information used to predict an outcome well before said outcome occurs, to ensure the model does not capture current events and thus becomes an explanatory model. As such, the research design was constructed to allow the models to analyse one-two years' worth of financial data, before making a prediction of financial distress two years into the future. The reason for two years and not the following year is because when a firm is that close to bankruptcy, the financial data is likely already affected, and for private firms, an impending bankruptcy can be so disruptive that a firm simply ignores to report financial statements, resulting in unreliable data one year before bankruptcy. Being a study focusing on adding to the current field of bankruptcy prediction, it was only natural to use a database with the most up-to date financial figures. The Serrano database, as of April 31st 2021, contains financial data up to 2018. This implies that the models will be able to look at financial performance up to 2016 before making their predictions. To ensure this study does not capture outdated business trends and dynamics, the tail of the sample was limited to ten years, meaning the first financial statements are from 2006 and the first predictions of financial distress from 2008.

In terms of cleaning the dataset, firstly all nonretailers were excluded in the sample, using the denoted SNI-codes as a guide. Firms with SNI-codes starting with "47", the code for general retailing in Sweden excluding motorized vehicles, were included. Unfortunately, this means excluding retailers who have failed to or reported an incorrect code, but given the sheer size of the sample regardless, it would not have been worth the effort to adopt a more manual approach. Subsidiaries were also excluded, as their bankruptcies often can be the result of a strategic decision rather than poor financial performance and would thus be misleading. Finally, firms with missing data that caused any incomplete model to generate were dropped, to ensure the models all accessed the same sample. After cleaning, our sample contained 14,550 retailers, of which 511 (3.51%) had filed for bankruptcy in the period 2008-2018, and 14,039 had no such filings. After testing for outliers, it was determined a handful of firm entries significantly deviated from the mean, possibly skewing the data. To reduce their impact, data below the 1st percentile and above the 99th underwent transformation through winsorization, which means to transform the outliers above and below the interval to values just inside it. The advantage of winsorization over trimming (excluding outliers) is that the outliers can be kept in the sample, without having them skew the data.

Methodology and variable definitions

We first estimated five logistic distress prediction models from our data. Since our sample consisted of private firms, we followed Altman's recommendations and used the private-firm version of Altman's Z-model, where the book value of equity is used instead of the market value. (Altman, 2000). Therefore, we estimated the following logistic regressions:

Altman's (five variable)

$$logit\,(P(X_{it+2} = 1|\,X_{it})) = \alpha_0 + \beta_1 WC\_TA_{it} + \beta_2 RE\_TA_{it} + \beta_3 EBIT\_TA_{it} + \beta_4 BV\_TL_{it} + \beta_5 SALES\_TA_{it}$$

Altman's (four variable)

$$logit\,(P(X_{it+2} = 1|\,X_{it})) = \alpha_0 + \beta_1 WC\_TA_{it} + \beta_2 RE\_TA_{it} + \beta_3 EBIT\_TA_{it} + \beta_4 BV\_TL_{it}$$

Ohlson's

$$logit\,(P(X_{it+2} = 1|\,X_{it})) = \alpha_0 + \beta_1 SIZE_{it} + \beta_2 TL\_TA_{it} + \beta_3 WC\_TA_{it} + \beta_4 CL\_CA_{it} + \beta_5 OENEG_{it} + \beta_6 NI\_TA_{it} + \beta_7 FFO\_TL_{it} + \beta_8 NITWO_{it} + \beta_9 CHIN_{it}$$

Sharma & Mahajan's

$$logit\,(P(X_{it+2} = 1|\,X_{it})) = \alpha_0 + \beta_1 EBIT\_TA_{it} + \beta_2 CA\_CL_{it}$$

McGurr & DeVaney's

$$logit\,(P(X_{it+2} = 1|\,X_{it})) = \alpha_0 + \beta_1 NI\_TA_{it} + \beta_2 CHWC_{it} + \beta_3 SALES\_EMP_{it} + \beta_4 EBIT\_SALES_{it} + \beta_5 CHLTD_{it} + \beta_6 CA\_CL_{it} + \beta_7 LTL\_TA_{it}$$

where:

- X_{it+2} is $BANKRUPTCY_{it+2}$ (and $BANKRUPTCY_{it+2}$ is an indicator variable equal to one if firm *i* filed for bankruptcy in year *t + 2* and otherwise equal to zero);
- WC_TA_{it} is working capital divided by total assets for firm *i* in year *t*;
- RE_TA_{it} is retained earnings divided by total assets for firm *i* in year *t*;
- $EBIT\_TA_{it}$ is earnings before interest and taxes divided by total assets for firm *i* in year *t*;
- BV_TL_{it} is book value of equity divided by total liabilities for firm *i* in year *t*;
- $SALES\_TA_{it}$ is sales divided by total assets,
- $SIZE_{it}$ is the ln(Total Assets/GDP price level index) for firm *i* in year *t*;
- TL_TA_{it} is total liabilities divided by total assets for firm *i* in year *t*;
- CL_CA_{it} is current liabilities divided by current assets for firm *i* in year *t*;
- NI_TA_{it} is net income divided by total assets for firm *i* in year *t*;
- FFO_TL_{it} is earnings before interest and taxes divided by total liabilities for firm *i* in year *t*;

- $NITWO_{it}$ is an indicator variable equal to one if firm *i* reported a negative net income for the last two years, zero otherwise;
- $OENEG_{it}$ is an indicator variable equal to one if firm *i*'s total liabilities exceed total assets in year *t* and otherwise equal to zero;
- $CHIN_{it}$ is a scaled change in net income [(NIt - NIt-1)/(|NIt| + |NIt-1|)] for firm *i* in year *t,*
- CA_CL_{it} is the reverse of CL_CA;
- $CHWC_{it}$ is a change in working capital, equal to one if working capital increased from the previous year and zero otherwise;
- $SALES\_EMP_{it}$ is sales divided by number of employees;
- $EBIT\_SALES_{it}$ is earnings before interest and taxes divided by sales,
- $CHLTD_{it}$ is the scaled change in long-term debt [(LTDt - LTDt-1)/LTDt-1] for firm *i* in year *t*;
- LTL_TA_{it} is long-term liabilities divided by total assets.

A dummy variable tracking firm size was also created, to track how the models performed over different segments. The three sizes were micro firms, SMEs, and large firms. A firm was denoted a micro firm if it, during the measured period, never exceeded 3MSEK in annual turnover or 1.5MSEK in total assets.[1] A firm was denoted large if it was among the top 3% in annual turnover or total assets. Finally, firms not fitting either of the previous conditions were denoted as SMEs.

The models predicted bankruptcy with logistic regressions using training and testing samples. Specifically, the models were first trained (i.e., its parameters estimated) using the in-sample (training sample). The in-sample was a subset containing half of randomly selected bankrupt and nonbankrupt firms from the original sample. Then, the estimated parameters were used to predict bankruptcy on the out-of-sample (testing sample) firms, calculating the probability of bankruptcy within two years for each model. The results were then compared between the models by analysing the classification of actual bankruptcy cases, both through standard confusion matrices using a specific cut-off point and through ROC curves, showing performance on all cut-off points.

[1] Under Swedish regulation, a firm is classified as a micro firm if it meets two out of the following three characteristics: total assets lower than 1.5 million SEK, total revenues lower than 3 million SEK, and average number of employees lower than 3.

Preliminary data analysis

Table 2 Characteristics of bankrupt and nonbankrupt firms in their last reported financial statements

	Variable	Bankrupt firms Mean	Nonbankrupt firms Mean
	$BANKRUPTCY_{it+2}$	1	0
Debt	BV_TL_{it}	0.574	1.716
	TL_TA_{it}	0.938	0.661
	CL_CA_{it}	0.964	0.667
	$OENEG_{it}$	0.319	0.171
	CA_CL_{it}	2.181	3.993
	$CHLTD_{it}$	-0.094	-0.073
	LTL_TA_{it}	0.257	0.156
Profitability	RE_TA_{it}	-0.178	0.055
	$EBIT\_TA_{it}$	-0.096	-0.011
	NI_TA_{it}	0.278	0.095
	$NITWO_{it}$	-0.076	0.001
	$CHIN_{it}$	-55318.97	91421.54
	$EBIT\_SALES_{it}$	-0.068	-0.031
Turnover	WC_TA_{it}	0.171	0.359
	$SALES\_TA_{it}$	2.857	2.401
	$CHWC_{it}$	0.415	0.524
	$SALES\_EMP_{it}$	1157301	1053683
Other	$SIZE_{it}$	6.933	6.881
	FFO_TL_{it}	-0.131	-0.029
	N=	511	14039

The difference in means were significant beyond the 0.05 level for all variables tested.

A simple mean analysis allows for an insight into the potential difference in financials between a solvent and a bankrupt firm. Looking at Table 2, it is clear that the solvent firms have an overall healthier balance sheet with higher solvency, a better current ratio, and positive trends. Perhaps the most apparent difference is the change in net income (CHIN) which on average sharply decreases for the bankrupt firm while it increases for the solvent firm. All mean differences between the bankrupt and nonbankrupt group were significant beyond the 0.05 level.

Empirical Results

Logistic regression results of estimating the distress prediction models

The first step was to train the models on the in-sample data. The results indicated that the financial ratios used in the models were of varying statistical significance. For Altman's five-variable model, variables WC_TA, RE_TA, and BV_TL were significant, while EBIT_TA and SALES_TA were not. These results were replicated in the four-variable version which, with variable SALES_TA excluded, showed significance for all variables except for variable EBIT_TA. These variables, showing turnover / total assets & EBIT / total assets respectively, capture a snapshot in time of a firm's profitability. The problem when looking at private firms and especially as the book value of equity becomes a factor, is that while a poor performing firm likely sees a decrease in turnover and EBIT, the denominator in total assets might very well also see a decrease. Therefore, as both sides of these ratios move for a firm approaching bankruptcy, they offer less predictive power than for listed firms, where the emphasis is put on market value.

The Ohlson model yielded significant results for five of the nine variables used. Of the significant variables, all showed a positive coefficient except for the variable CHIN: the scaled change in net income YoY. The variables not showing significance were variables WC_TA, CL_CA, OENEG & FFO_TL. Of these, variables WC_TA & OENEG denoted a negative coefficient, while variables CL_CA & FFO_TL denoted a positive coefficient. The Sharma model showed negative significance for both its measures. Finally, the McGurr model yielded significance for four of its seven variables. The significant variables were CA_CL, CHWC, SALES_EMP, and LTL_TA, where only SALES_EMP showed a positive coefficient. But it should be denoted that one of McGurr's measures has been recoded to better fit the sample, possibly taking away from its original performance. EBIT_SALES was originally gross profit divided by total sales in McGurr's study but had to be transformed because the Serrano database had very few reported values on the gross profit level.

While some models indicated a better fit because all their variables had significant explanatory power, this does not necessarily mean the model was better trained. Looking at the log-likelihood, which is a better measure of goodness of fit as it represents the best combination of variables, the Ohlson model had the best fit at -1011, followed by McGurr at -1020, the Altman models at -1043, and lastly Sharma at -1061. As a check, tests were made where variables with low significance were dropped, but this only resulted in the overall performance worsening.

Table 3 Results of in-sample estimation of the logistic regression using the five bankruptcy prediction models

	Altman 5V (Z-stat)	Altman 4V (Z-stat)	Ohlson (Z-stat)	Sharma (Z-stat)	McGurr (Z-stat)
Intercept	-3.15	-3.07	-5.29	-2.94	-4.27
	-(26.79)*	-(37.16)*	-(9.65)*	-(32.21)*	-(16.96)*
WC_TA	-0.46	-0.48	-0.08		
	-(2.39)*	-(2.52)*	-(0.21)		
RE_TA	-0.51	-0.51			
	-(3.93)*	-(3.93)*			
EBIT_TA	-0.19	-0.17		-0.56	
	-(0.95)	-(0.87)		-(2.94)*	
BV_TL	-0.17	-0.18			
	-(3.26)*	-(3.39)*			
SALES_TA	0.03				
	(1.00)				
SIZE			0.12		
			-(2.33)*		
TL_TA			1.45		
			-(4.73)*		
CL_CA			0.01		
			-(0.05)		
OENEG			-0.06		
			-(0.23)		
NI_TA			0.81		0.13
			-(2.23)*		-0.45
FFO_TL			0.08		
			-(0.37)		
NITWO			0.39		
			-(2.55)*		
CHIN			0.000001		
			-(5.67)*		
CA_CL				-0.15	-0.05
				-(5.09)*	-(1.89)*
CHWC					-0.33
					-(2.21)*
SALES_EMP					0.00000007
					-(1.38)*
EBIT_SALES					-0.05
					-(0.18)
CHLTD					-0.002
					-(0.01)
LTL_TA					-1.6
					-(6.64)*
N	7275	7275	7275	7275	7275
Log likelihood	-1043	-1043	-1011	-1061	-1020

* significant beyond the 0.05 level

Distress-scores and out-of-sample predictions

Having trained the models on the in-sample data, the next step was to use the estimated parameters to predict bankruptcies on the out-of-sample data. Specifically, the sample was randomly split in half, making for a training set and test set of 50/50 proportions. Each model predicted the probability of bankruptcy using the logarithmic function [e^{score} / (1 + e^{score})]. Then, a decision rule was introduced, where firms with a predicted probability in the top 5% of the distribution were classified as likely going bankrupt, while the other 95% were classified as unlikely to go bankrupt. The purpose of this decision rule is to better capture the proportion of bankruptcies present in a real business environment, where there typically only is a limited number of firms each year filing for bankruptcy. Moreover, the reality is that many firms headed for bankruptcy seal their fates in different ways, either by acquisition or debt restructuring. Thorburn (2000), reported a survival rate of approximately 75% for firms subjected to a bankruptcy auction. Hence, many firms that are rightfully predicted to go bankrupt based on their financial performance escape it, and this decision rule attempts to account for that. The decision rule is then as follows, $Prob(B) = \begin{cases} 1, \text{if Score } \leq 0.95 \\ 0, \text{if Score } > 0.95 \end{cases}$

Running a *t*-test comparing the mean classification score (0-1) between bankrupt and nonbankrupt firms (0-1), the null hypothesis was rejected at the 0.001 significance level, indicating that all five models are in fact able to predict bankruptcies of nonlisted, Swedish retail firms.

Table 4 Descriptive statistics for bankruptcy prediction measures

Model	Status	Mean	SE	SD	[95% conf. int.]	t-value	p-value
Altman 5V	Nonbankrupt	0.047	0.003	0.211	0.042	-6.972	<0.001
	Bankrupt	0.142	0.022	0.350	0.100		
Altman 4V	Nonbankrupt	0.047	0.003	0.211	0.042	-6.388	<0.001
	Bankrupt	0.135	0.021	0.342	0.093		
Ohlson	Nonbankrupt	0.045	0.002	0.208	0.041	-9.321	<0.001
	Bankrupt	0.173	0.024	0.379	0.127		
S & M	Nonbankrupt	0.048	0.003	0.213	0.043	-4.930	<0.001
	Bankrupt	0.115	0.020	0.320	0.076		
M & D	Nonbankrupt	0.045	0.002	0.208	0.040	-10.183	<0.001
	Bankrupt	0.185	0.024	0.389	0.137		
N=	Nonbankrupt	7015					
	Bankrupt	260					

Table 4 reports summary statistics on the bankruptcy probabilities predicted by each model separately for bankrupt and nonbankrupt firms. We use the t-test to test whether the mean values of predicted bankruptcy probabilities are different between bankrupt and nonbankrupt firms.

Having deemed all the models to be able to predict financial distress, the next step was comparing their performance by evaluating their classification accuracy. While the log-likelihood gave an indication of the model's fit in-sample, this does not immediately translate into the accuracy of prediction, for instance, because of the risk of overfitting. Since bankruptcy is a dichotomous event (a firm is not more or less bankrupt, it either is or is not) the accuracy can be visualised using a confusion matrix, showing accurate predictions (positive and negative) as well as type I and type II errors. A potential flaw of the classical confusion matrix is that it oversimplifies the nuances in the models' bankruptcy prediction. Around the cut-off point, there will likely be healthy firms just over it, erroneously classified as bankrupt and struggling firms just below, erroneously classified as healthy. Introducing a grey zone, like Altman (1968), or a multi-category confusion matrix-like Kallunki & Pykkö (2013) would mitigate the problem. But, for the sake of transparency and the conscious decision to be more consistent with reporting than studies that report accuracy but include a grey zone, these features were excluded. Instead, nuances in classification were captured using ROC curves.

Table 5 Classification results

Model	N	True Positive	True Negative	Type I Error	Type II Error	Accuracy	Sensitivity	Specificity
Altman's 4-variable model	7275	37	6688	327	223	92.44%	14.23%	95.34%
Altman's 5-variable model	7275	35	6686	329	225	92.38%	13.46%	95.31%
Ohlson's model	7275	45	6696	319	215	92.66%	17.31%	95.45%
Sharma & Mahajan's model	7275	30	6681	334	230	92.25%	11.54%	95.24%
McGurr & DeVaney's model	7275	47	6697	318	213	92.70%	18.08%	95.47%

Table 5 reports classification results using the decision rule where firms with a score above the cut-off of 95% are predicted as likely going bankrupt. **"True Positive"** denotes accurately predicted bankruptcies, **"True Negative"** denotes accurately predicted nonbankruptcies, **"Type I Error"** denotes nonbankrupt firms falsely predicted as bankrupt, **"Type II Error"** denotes bankrupt firms falsely predicted as nonbankrupt, **"Accuracy"** is the share of accurate predictions out of all observations, **"Sensitivity"** is the share of predicted bankruptcies out of all bankruptcies and **"Specificity"** is the share of predicted nonbankruptcies out of all nonbankruptcies.

As the model parameters were reestimated using a training sample, the model will inherently be fitted to the observations in the test set. As such, the reestimated models will be sensitive to skewness in the data and, if firms of a particular size are overrepresented, the models might be overfitted to those observations. More specifically, if our sample contains a large share of micro firms, which might be subject to more volatility and inconsistency than larger firms, the models might capture this volatility and essentially capture noise. To account for this, classification performance was compared on the three different firm size variables previously constructed, to investigate any signs of overfitting.

When isolating micro firms, all models displayed an increase in sensitivity except Ohlson's model. The increased sensitivity did result in a corresponding loss of specificity as more firms were falsely predicted as positive (i.e., as going bankrupt), leading to a decrease in overall accuracy. Ohlson's model instead decreased in sensitivity for an increase in specificity and an overall increase in accuracy. This effect was reversed for both the SME and large firm subsets. Here, the Ohlson model increased in sensitivity for both but at the cost of a relatively lower specificity and accuracy, while the other four models decreased in sensitivity and increased in specificity and accuracy. A general observable trend is that as firm size decreases, so does the overall predictive accuracy of the models. However, our results did not indicate any strong signs of better performance on one segment as a result of overfitting.

Table 6 Classification results between different firm sizes

Model	N	True Positive	True Negative	Type I Error	Type II Error	Accuracy	Sensitivity	Specificity
Micro firms								
Altman's 4-variable model	3824	29	3401	285	109	89.70%	21.01%	92.27%
Altman's 5-variable model	3824	30	3394	292	108	89.54%	21.74%	92.08%
Ohlson's model	3824	21	3537	149	117	93.04%	15.22%	95.96%
Sharma & Mahajan's model	3824	21	3412	274	117	89.78%	15.22%	92.57%
McGurr & DeVaney's model	3824	27	3486	200	111	91.87%	19.57%	94.57%
SMEs								
Altman's 4-variable model	3347	8	3184	41	114	95.37%	6.56%	98.73%
Altman's 5-variable model	3347	5	3189	36	117	95.43%	4.10%	98.88%
Ohlson's model	3347	24	3065	160	98	92.29%	19.67%	95.04%
Sharma & Mahajan's model	3347	9	3166	59	113	94.86%	7.38%	98.17%
McGurr & DeVaney's model	3347	8	3184	41	114	95.37%	6.56%	98.73%
Large firms								
Altman's 4-variable model	104	0	103	1	0	99.04%	n.m.	99.04%
Altman's 5-variable model	104	0	103	1	0	99.04%	n.m.	99.04%
Ohlson's model	104	0	94	10	0	90.38%	n.m.	90.38%
Sharma & Mahajan's model	104	0	103	1	0	99.04%	n.m.	99.04%
McGurr & DeVaney's model	104	0	103	1	0	99.04%	n.m.	99.04%

Table 6 reports classification results using the same decision rule as in **Table 5** but groups classification results based on firm sizes. A firm was denoted a "M**icro firm"** if it, during the measured period, never exceeded 3MSEK in annual turnover or 1.5MSEK in total assets. (Swedish GAAP). A firm was denoted a **"Large firm"** if it was among the top 3% in annual turnover or total assets. Finally, firms not fitting either of the previous conditions was denoted an **"SME".**

The performance analysis was completed by computing ROC (receiver operating characteristic) curves for the different models. The usefulness of an ROC curve is in its ability to display model performance on all possible cut-off values of bankruptcy probability, making it an excellent complement to the standard 2 x 2 confusion matrix (in which we only displayed classification results for the cut-off value of .95). The ROC displays the model's performance by plotting a two-dimensional relationship between the true positive rate (recall/sensitivity) and false positive rate (1-specificity). A perfect model would predict no false negatives and no false positives, which would be plotted in the upper, left corner, while a random classifier follows a diagonal line from the left bottom to the top right. Hence, the more a model's ROC curve tends towards the upper left corner, the better.

Figure 1 ROC Curve for Altman's five variable model

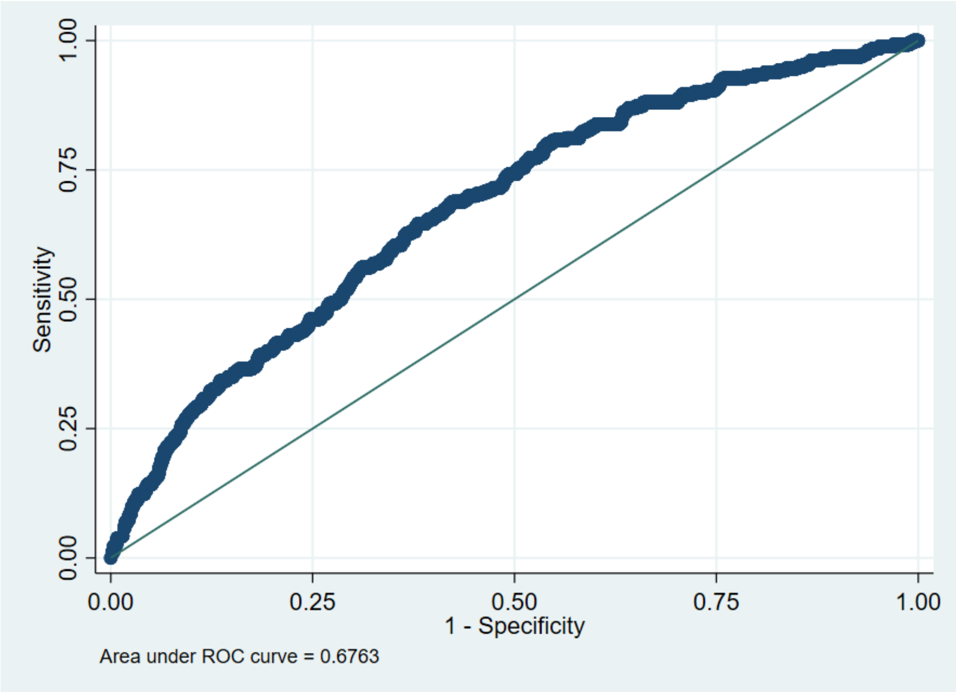


Figure 2 ROC Curve for Altman's four variable model

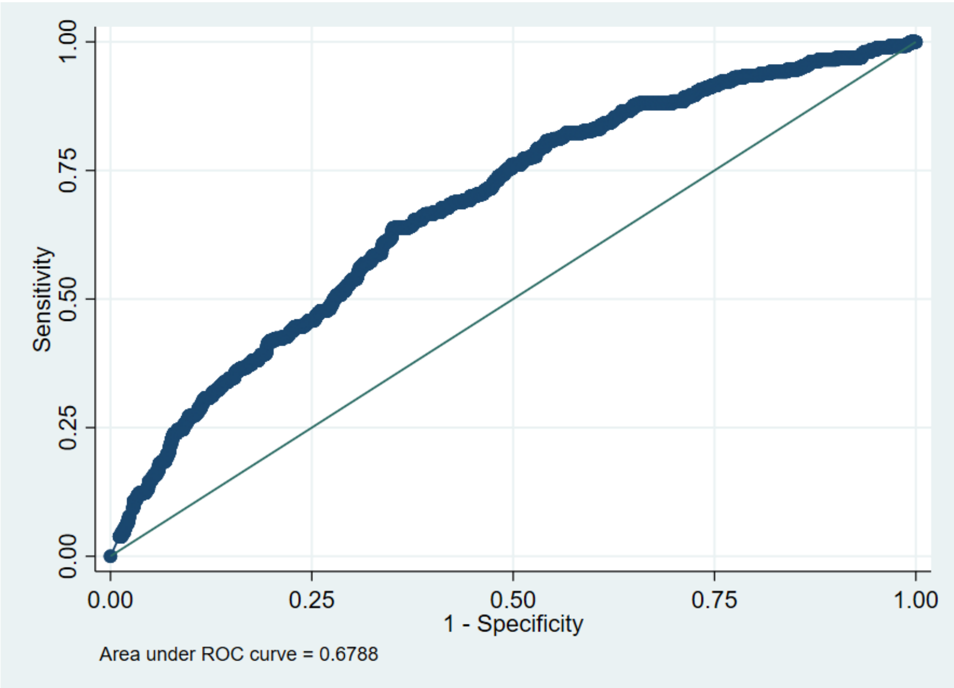


Figure 3 ROC Curve for Ohlson's model

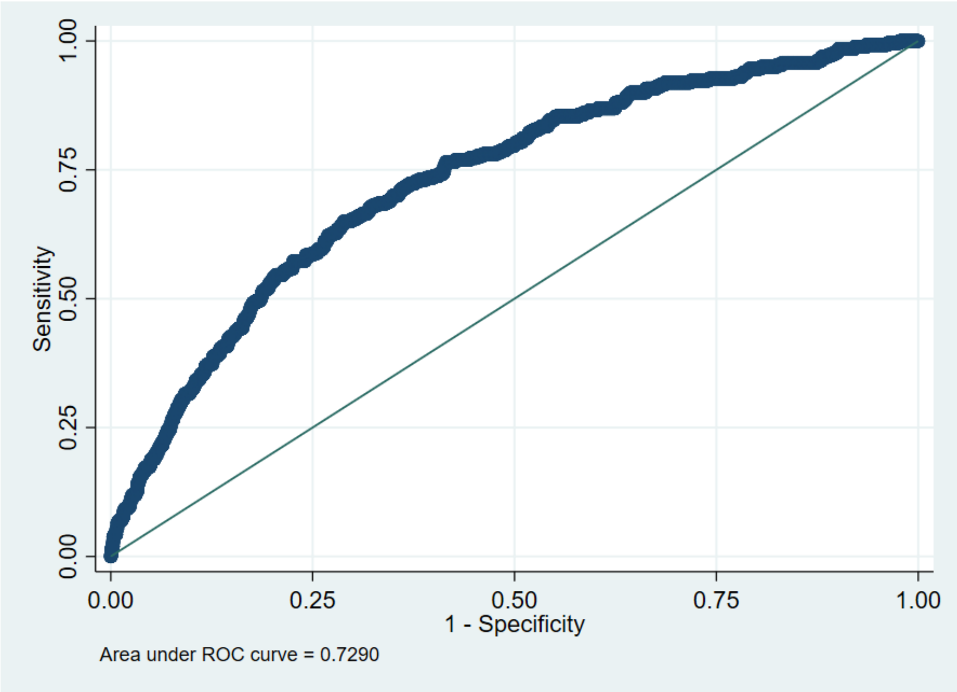


Figure 4 ROC Curve for Sharma & Mahajan's model

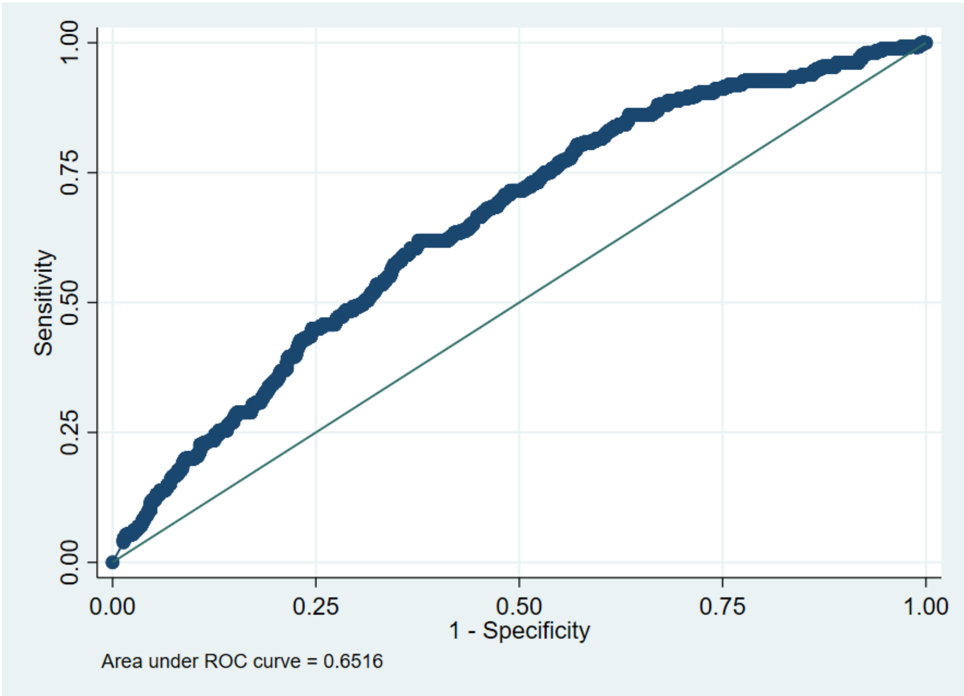
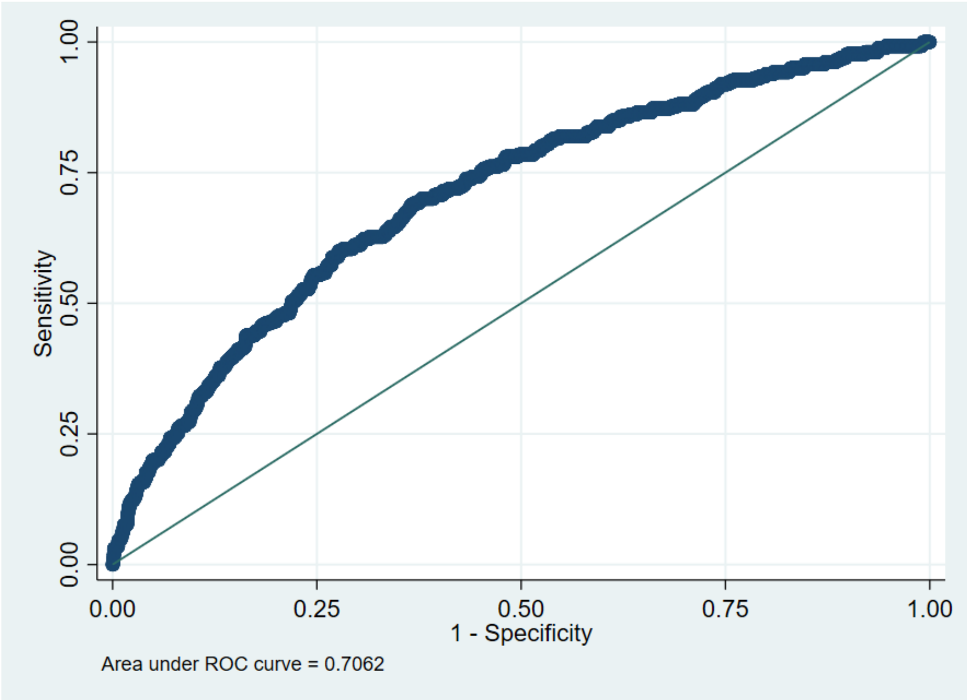


Figure 5 ROC Curve for McGurr & DeVaney's model



Results from ROC curves (Figs. 1-5) show that Ohlson's model (Fig. 3) performed the best overall with an AUC-score of 72.9%, followed by McGurr (Fig. 5) at 70.6%, Altman's 4-variable model (Fig. 2) at 67.9%, then his 5-variable model (Fig. 1) at 67.6% and lastly Sharma's (Fig. 4) at 65.2%. The ROC curves illustrate the possibility to tailor the models' predictions to a given purpose. An analyst looking to identify the most likely bankruptcies in a given sample could match the cut-off point of .95 chosen for the confusion matrix, as it reflects the normal occurrence of bankruptcy in a normal population. A manager, looking to assess the health of their firm, could increase the sensitivity, to reduce the risk of a false negative and, if given a positive prediction, could investigate what metric might have caused the classification and whether it is cause for concern or not. To summarize, the models show a consistent predictive ability for most levels of sensitivity, meaning they are viable regardless of if the aim is to minimize type I or type II errors.

Discussion & Conclusion

Bankruptcy prediction on nonlisted retailers is arguably a difficult feat. Not the least due to the structural issue of inconsistent reporting, but one would also expect the noncapitalization of leases (since our studied sample does not need to adhere to IFRS16) to pose a challenge for the prediction models. Not only would the noncapitalization of asset risk inflating the value of assets for retailers owning their facilities rather than leasing them, but it could also induce a false sense of security, as the inflated balance sheet gives the impression of tools available to sell, which more often than not may be tied up in unsellable assets. Also, looking beyond capitalized leases, inventory often constitutes a major share of retailers' balance sheets and is unlikely to reflect insolvency as a retailer often has as much, if not more, inventory due to low inventory turnover when approaching bankruptcy. Inventory is therefore also likely to play a role in complicating bankruptcy prediction for the retail sector. Beyond intrinsic accounting difficulties, the difference in incentives between a Swedish bankruptcy filing and the U.S. Chapter 11 makes bankruptcy all the more important but also equally difficult. If managers are incentivized to withhold any signs that the firm is headed to bankruptcy, the task falls to the models to accurately predict bankruptcy well before it occurs.

Despite the obstacles, our study shows that the models perform surprisingly well at predicting bankruptcy for nonlisted Swedish retailers—casting doubt on the predictive hinder posed by an inflated and unyielding balance sheet. This study shows that through the mere refinement of reestimating all the model parameters using a training set, our tested models show applicability on the previously untested segment. In our study, we find that Ohlson's and McGurr's models perform best overall, followed by the two Altman models and lastly Sharma's. With some differences found while examining the predictive ability between different firm sizes, the argument can be made that some variables are of higher importance than others for certain firm sizes. More specifically, the Ohlson model, appeared more sensitive to SMEs and large firms while the other four models indicated a higher sensitivity towards the micro firm subset. But as this study focused on comparing model performance on a previously untested sample, little effort was made to compare the individual usefulness of factors on a univariate basis. We leave it for future research to investigate further.

What our study did find, however, was a common feature and likely explanation as to why multivariate prediction models tend to outperform univariate models. Two factors were carried by all working bankruptcy prediction models; one independent variable to assess the amount of debt for the firm and one to assess its profitability. In its simplest form, we find models such as that developed by Sharma & Mahajan, which carry one independent variable to assess a firm's debt capacity (Current ratio) and one to assess its profitability (Return on assets). From this, we then see models aim to carry a higher predictive power by including more independent variables, such as the model developed by McGurr & DeVaney, which carries both Current ratio and Return on assets to assess the aforementioned risk-indicators, but also more intricate variables such as sales per employee and percent change in long-term liabilities. We believe that assessing both a firm's ability to generate income and its capital structure is a contributing factor to multivariate models being superior to their univariate counterparts. This, as otherwise successful businesses may thrive despite a low e.g., EBIT to revenue or high debt to equity—possibly making a model capturing only one of these aspects misleading.

Furthermore, in the study, we found that models carrying more intricate measures of assessing profitability and bankruptcy (I.e., Ohlson's- and McGurr & DeVaney's model) are more likely to accurately predict bankruptcy. Circling back to performance, a bankruptcy prediction model which never misses any bankrupt firms is more useful than a model which misses bankrupt companies but never misclassifies a nonbankrupt firm. Because of this, any bankruptcy prediction model should seek to minimize their type II errors, i.e., when a model predicts solvency, but the firm goes bankrupt. Furthermore, this also means that models can quite often afford to have a high number of type I errors. i.e., when the model predicts bankruptcy, but the company remains solvent. Our chosen cut-off point of .95 for the confusion matrices unfortunately does not seem to be sensitive enough. The chosen point is logically sound, as it reflects the average distribution of bankruptcies in a population, but the choice to exclude a grey zone like Altman proposes seems to result in too many misclassifications. Instead, to capture more of the actual bankruptcies one should increase the model sensitivity. Looking at the ROC curves, most models were able to attain a 75% sensitivity while still retaining a specificity of 50%, which might be a more favourable cut-off point given the purpose of a bankruptcy prediction model.

By aiming to use the models as signals of financial distress rather than binary bankruptcy predictors, we see a great value in a sensitive model that sacrifices specificity for sensitivity. While a 50% specificity hardly characterizes an accurate model—it is likely a model which is good at signalling financial distress as all the firms deemed bankrupt had financial profiles below the standards of their peers. What this implies is that if the models display a high risk of bankruptcy for a given firm, the risk of insolvency remains low, but it is still a cause for concern. However, if the models signal solvency, there is a very low risk of the firm actually becoming insolvent. Here the discussion around type I & type II errors resurfaces, where the models undoubtedly carry a large number of type I errors, but it is not necessarily something negative as it signals financial distress within the company. What would have been extremely harmful is if the percentage of firms deemed solvent which actually went bankrupt were to be higher—as the models would have signalled that there is no financial distress while there in reality was cause for concern.

Furthermore, while a theoretically perfect model would not have misclassified the bankrupt firms as solvent, trying to reach 100% accuracy out of sample when using a large sample size can almost be deemed impossible as some companies will display a misleadingly healthy financial profile even in an impending bankruptcy. This point was also explained in Ohlson's study wherein he stresses the fact that none of the bankrupt firms in his study which were falsely deemed solvent could have reasonably been predicted as bankrupt because their financial profile appeared so healthy. Here, an improvement to quantitative models might be to include a qualitative evaluation of the firm, to enhance the predictive accuracy—or at least make the identification of bankrupt firms easier. However, this point is also largely disproven by Ohlson as none of the audit reports of the bankrupt firms he deemed solvent showed any signs of financial distress, and some firms even paid dividends a year prior to their bankruptcy—casting doubt on the predictive ability of qualitative assessments.

While Ohlson's qualitative inclusions might prove futile in trying to identify and correct false negatives, we believe a qualitative feature can be very useful in false positive predictions. As mentioned, when a sensitive model signals for bankruptcy, its assessment has a large chance of being incorrect. But it can be very valuable for managers to understand why their firm was classified as likely to go bankrupt. Since all those firms were classified as likely to go bankrupt, their financials must have resembled those of bankrupt firms more than a solvent. As such it would be important to understand what caused the resemblance; if it is because the firm operates with an unusual business model, have an aggressive expansion plan, or if there actually is cause for concern. Our point is that a positive prediction will not be either true or false, but always something to look further into, especially for a retail manager looking to avoid falling victim to the retail apocalypse.

In our study, we help bridge the gap between the current state of bankruptcy prediction research and the private retail sector—thereby adding to the body of literature dedicated to foreseeing financial distress. Despite the many theoretical challenges with predicting bankruptcies for private firms and retail firms, we shed light on a previously unexplored frontier and determine that, in fact, accurate prediction is possible, both using industry agnostic- and sector-specific models. For future research, we find that a key element in revising and appropriating models to untested samples is reestimating the model parameters to the samples, using a training and test set. For practical users of the bankruptcy prediction models, we find that the models evaluated should contain intricate financial data to provide high accuracy and we advise practitioners to make qualitative assessments to investigate all positive predictions before taking a conclusive stance on bankruptcy.

Limitations

The study research design comes with a couple of inherent limitations. The model comparison results in the inclusion of a large number of variables. This, combined with the choice to look at private firms, who tend to report with inconsistent quality, meant that a large share of observations was lost due to missing values. One could theorize that insufficient reporting could be a sign of a poorly managed firm, thereby having a correlation to bankruptcies. But, as the intention was to consistently compare all the five models, this relationship could not be investigated and remains a subject for further research.

Another limitation was the noncapitalization of leases on our sample. While it can be viewed as an achievement that the models were able to accurately predict bankruptcy despite the challenges posed by the noncapitalization, as previously discussed, the fact is that most literature points to the fact that capitalizing leases increases accuracy for retail bankruptcy prediction. Our study focused on private firms not having to adhere to IFRS16, but future research could make interesting findings by either separating the sample into an asset-owning group and an asset-leasing group or manually capitalize leases for a smaller sample. Such research could tell us whether there is even more room for improvement on the private retail bankruptcy prediction frontier.

A limitation also arises in our choice of using financial data dating back 12-24 months prior to the bankruptcies predicted. While this allows us to have a higher degree of accurately reported firm data and predict rather than explain bankruptcies, we also limit ourselves in not looking at the most recent data prior to firm failures. Other studies have found value in using data reported 4-12 months prior to the bankruptcy. However, as these studies often are conducted on public firms, the risk of data loss or inaccurate reporting is minimized due to mandatory auditing.

Looking at the models evaluated in the study, another limitation arises in that we only consider logit and MDA-based bankruptcy prediction models. Thereby, we overlook possibly superior computer-based models. However, we find great value in establishing the functionality of MDA- and Logit-based models as these are advantageous for managers and analysts in determining underlying causes for bankruptcy. Tying this to our goal of helping managers foresee financial distress, the choice of models finds further support as all the evaluated models can give clear indications of what drives their bankruptcy risk.

Finally, because of our choice to introduce a decision rule for classifying a firm as bankrupt, any comparisons to other studies should be made with caution. The decision rule serves multiple purposes, it both reflects the normal occurrence of bankruptcies in a sample and provides a standardized method for comparing the models, where the original methodologies differed greatly. But, as a result, directly comparing our accuracies to other studies can be misleading and is therefore cautioned against.

Appendix

Table 7 Variables used in the analysis and their descriptions

Dependent variable	**Description**
$BANKRUPTCY_{it+2}$	An indicator variable equal to one if firm *i* filed for bankruptcy in year *t + 2* and otherwise equal to zero
Independent variables	
WC_TA_{it}	Working capital divided by total assets for firm *i* in year *t*
RE_TA_{it}	Retained earnings divided by total assets for firm *i* in year *t*
$EBIT\_TA_{it}$	Earnings before interest and taxes divided by total assets for firm *i* in year *t*
BV_TL_{it}	Book value of equity divided by total liabilities for firm *i* in year *t*
$SALES\_TA_{it}$	Sales divided by total assets
$SIZE_{it}$	The ln(Total Assets/GDP price level index) for firm *i* in year *t*
TL_TA_{it}	Total liabilities divided by total assets for firm *i* in year *t*
CL_CA_{it}	Current liabilities divided by current assets for firm *i* in year *t*
NI_TA_{it}	Net income divided by total assets for firm *i* in year *t*
FFO_TL_{it}	Earnings before interest and taxes divided by total liabilities for firm *i* in year *t*
$NITWO_{it}$	An indicator variable equal to one if firm *i* reported a negative net income for the last two years, zero otherwise
$OENEG_{it}$	An indicator variable equal to one if firm *i*'s total liabilities exceed total assets in year *t* and otherwise equal to zero
$CHIN_{it}$	A scaled change in net income [(NIt - NIt-1)/(|NIt| + |NIt-1|)] for firm *i* in year *t*
CA_CL_{it}	The reverse of CL_CA
$CHWC_{it}$	A change in working capital, equal to one if working capital increased from the previous year and zero otherwise
$SALES\_EMP_{it}$	Sales divided by number of employees
$EBIT\_SALES_{it}$	Earnings before interest and taxes divided by sales
$CHLTD_{it}$	The scaled change in long-term debt [(LTDt - LTDt-1)/LTDt-1] for firm *i* in year *t*
LTL_TA_{it}	Long-term liabilities divided by total assets

References

Aktiespararna. (5 June 2021). Börstabeller. www.aktiespararna.se/analysguiden/borstabeller#

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. The journal of finance, 23(4), 589-609.

Altman, E. I., Haldeman, R. G., & Narayanan, P. (1977). ZETATM analysis A new model to identify bankruptcy risk of corporations. *Journal of banking & finance*, *1*(1), 29-54.

Altman, E. I. and Levallee, M. L. (1981) Business failure classification in Canada. Journal of Business Administration 12(I), 147-164.

Altman, E., 1993. Corporate Financial Distress and Bankruptcy. Wiley, New York.

Altman, E. I., & Saunders, A. (1997). Credit risk measurement: Developments over the last 20 years. Journal of banking & finance, 21(11-12), 1721-1742.

Altman, E. I. (2000). Predicting financial distress of companies: Revisiting the Z-Score and Zeta® models.

Beaver, W. H. (1966). Financial ratios as predictors of failure. Journal of accounting research, 71-111.

Begley, J., Ming, J., & Watts, S. (1996). Bankruptcy classification errors in the 1980s: An empirical analysis of Altman's and Ohlson's models. Review of accounting Studies, 1(4), 267-284.

Bellovary, J. L., Giacomino, D. E., & Akers, M. D. (2007). A review of bankruptcy prediction studies: 1930 to present. Journal of Financial education, 1-42.

Bhargava, M., Dubelaar, C., & Scott, T. (1998). Predicting bankruptcy in the retail sector: an examination of the validity of key measures of performance. Journal of Retailing and Consumer Services, 5(2), 105-117.

Chaganti, R. S., Mahajan, V., & Sharma, S. (1985). Corporate board size, composition and corporate failures in retailing industry [1]. Journal of management studies, 22(4), 400-417.

Clement, C. Machine Learning in Bankruptcy Prediction–A. Editorial Board, 178.

Fejér-Király, G. (2015). Bankruptcy prediction: A survey on evolution, critiques, and solutions. Acta Universitatis Sapientiae, Economics and Business, 3(1), 93-108.

Harrison, N., & Thomas-Dupuis F. (2018). The Future of Retail. Oliver Wyman Insights.

Helm, S., Kim, S. H., & Van Riper, S. (2020). Navigating the 'retail apocalypse': a framework of consumer evaluations of the new retail landscape. Journal of Retailing and Consumer Services, 54.

Horváthová, J., & Mokrišová, M. (2018). Risk of bankruptcy, its determinants and models. *Risks*, *6*(4), 117.

Kallunki, J. P., & Pyykkö, E. (2013). Do defaulting CEOs and directors increase the likelihood of financial distress of the firm?. *Review of Accounting Studies*, *18*(1), 228-260.

Klepáč, V., & Hampel, D. (2016). Prediction of bankruptcy with SVM classifiers among retail business companies in EU. Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis, 64(2), 627-634.

McGurr, P., & DeVaney, S. (1998). Predicting business failure of retail firms: An analysis using mixed industry models. Journal of Business Research, 43(3), 169-176.

McGurr, P. T., & Devaney, S. A. (1998). A retail failure prediction model. The International Review of Retail, Distribution and Consumer Research, 8(3), 259-276.

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. Journal of accounting research, 109-131.

Pang, J., & Kogel, M. (2013). Retail bankruptcy prediction.

Shankar, V., Kalyanam, K., Setia, P., Golmohammadi, A., Tirunillai, S., Douglass, T., ... & Waddoups, R. (2021). How technology is changing retail. Journal of Retailing, 97(1), 13-27.

Sharma, S., & Mahajan, V. (1980). Early warning indicators of business failure. Journal of marketing, 44(4), 80-89.

Thorburn, K. S. (2000). Bankruptcy auctions: costs, debt recovery, and firm survival. *Journal of financial economics*, *58*(3), 337-368.