Magisteruppsats

Master's thesis – one year

Företagsekonomi

Business Administration

Time's up!

The impact of multi-year accounting data on bankruptcy prediction

Dennis Hedback

Mittuniversitetet
MID SWEDEN UNIVERSITY

Campus Härnösand Universitetsbacken 1, SE-871 88. **Campus Sundsvall** Holmgatan 10, SE-851 70 Sundsvall.
Campus Östersund Kunskapens väg 8, SE-831 25 Östersund.
Phone: +46 (0)771 97 50 00, Fax: +46 (0)771 97 50 01.

MID SWEDEN UNIVERSITY
Department of Business, Economics and Law

Examiner: Peter Öhman, peter.ohman@miun.se
Supervisor: Darush Yazdanfar, darush.yazdanfar@miun.se
Author: Dennis Hedback, dehe1500@student.miun.se
Main field of study: Business Administration
Semester, year: Spring, 2019

Abstract

Theoretical bankruptcy processes spanning multiple accounting periods have been proposed. At the time of writing, most existing bankruptcy prediction studies do not take these failure processes into consideration, even though they are proposed to increase prediction performance. This thesis takes a novel approach, taking the years preceding corporate bankruptcy into account into prediction models, by using artificially intelligent random forests to assess the impact of the time dimension of bankruptcy on the prediction performance metrics accuracy (i.e. the probability of correctly classifying any firm), recall (i.e. the probability of correctly classifying a bankrupt firm), and precision (i.e. the probability that a firm classified as bankrupt, is actually bankrupt). The experimental findings indicate that estimating prediction models from multiple years of data at the firm level raises accuracy and recall of prediction models. These findings suggest that accounting for the time dimension of bankruptcy indeed increases bankruptcy prediction accuracy through an increased recall rate. However, this thesis cannot show that precision is affected by using multiple years of data at the firm level. Thus, further research conducted in other socio-economical contexts, as well as using other performance metrics, is called for.

Keywords: bankruptcy prediction, failure process, time dimension

Sammanfattning

I skrivande stund tar de flesta prediktionsmodeller i litteraturen inte hänsyn till teoretiska konkursprocesser över flera redovisningsperioder, trots att de föreslås ge bättre prediktionsprestanda. I denna uppsats används ett anpassat tillvägagångssätt för att, i prediktionsmodeller, ta hänsyn till de år som närmast föregår bolagskonkurser. Genom att använda artificiellt intelligenta random forest-modeller, bedömer uppsatsen hur tidsaspekten hos konkurser påverkar prediktionsprestandamåtten träffsäkerhet (d.v.s. sannolikheten för att göra en korrekt förutsägelse), recall (d.v.s. sannolikheten för att göra en korrekt förutsägelse för ett bolag som faktiskt går i konkurs) och precision (d.v.s. sannolikheten för att ett bolag som förutsägs gå i konkurs, faktiskt går i konkurs). Resultatet från experimentet indikerar att om prediktionsmodeller estimeras från flerårsdata på bolagsnivå, ökar deras prediktionsprestandamåtten träffsäkerhet och recall. Det tyder på att då hänsyn tas till tidsaspekten hos konkurser, görs träffsäkrare förutsägelser, främst genom att modellerna blir mer säkra på vilka bolag som faktiskt kommer att gå i konkurs. Denna uppsats kan dock inte påvisa att prediktionsmodellers precision påverkas av bolagsspecifika flerårsdata. Således krävs ytterligare forskning i andra socioekonomiska sammanhang, men också ytterligare forskning som använder andra prestandamått.

Nyckelord: konkursprognos, konkursprediktion, konkursprocess, tidsaspekt

This page intentionally left blank.

Preface

While I hold no dogmatic position of whether quantitative or qualitative research is the superior method of inquiry in the social sciences, the sometime obvious struggle between these disciplines amuses me. Rather, I believe that interplay between the two is essential to make sense of the world. Perhaps that's why I've taken a pragmatic approach to research, where society is just the thing contradicting your, more or less, naïve theory. This thesis may be a good example of such interplay, where the polar opposites of exploratory quantitative findings and theory induced from various case studies are taken together to deduce hypotheses tested through experiment. The choice of method should, in and of itself, be quite unusual for a master's thesis in business administration.

Wading through piles of articles of operational research has sometimes felt like I was in way over my head for a BA student. While frustrating, the satisfaction of each aha-moment increased with the level of technical obscurity, and also made me think back to all moments of high school math where I thought that it would be of no practical use in the future. It turns out that seconds-guessing your future self is a waste of time. As Simone Gertz put it: "Pursuing stupid things is just a humble acknowledgement that you don't know what the best answer is". These words have pretty much been the soundtrack of the making of this thesis.

I would like to thank my children, Alde-Marie and Signe for being relatively understanding that dad needs "to sit and look at a lot of numbers". I know the amount of work I put into this has been hard for you. You are amazing every day. I would also like to thank my supervisor, Darush Yazdanfar, for superb critique and interesting philosophical discussions, as well as my examiner, Peter Öhman, for getting deep into the fine details of this thesis, spotting problems others don't. Other people I would like to thank are Lisa Eklöf for very insightful discussions during the final seminar, and Christer Sandvik for his exquisite pep talking ability and his pragmatic view of research. Last, but not least, I would like to thank Johan Krantz for supplying bankruptcy data, and making dealing with the Swedish Enforcement Authority a really good experience, especially considering the predicament I could be in, if my reasons for contacting them would have been the same as people usually have.

Signed: ______________________________

Dennis Hedback

2019-08-01, Sundsvall, Sweden

Contents

Tables

Figures

1 Introduction

1.1 Background

The assumption that a business will continue its operations for the foreseeable future is a fundamental principle of accounting, and perhaps the most important one for stakeholders. This assumption is crucial in order to prepare financial reports in accordance with both International Financial Reporting Standards and local accounting principles (IASB, 1997). Due to the costs associated with bankruptcy (Branch, 2002), measures are taken by stakeholders to ensure that a business of interest does not fail unexpectedly. For example, shareholders employ auditors to quality-assure accounting information (Öhman et al., 2006), while creditors model the risk of clients defaulting on their loans (Altman & Sabato, 2007). In fact, the task of anticipating bankruptcy is of importance to anyone making decisions based on financial reports.

For these reasons, researchers have been interested in predicting bankruptcy since the 1960s (Altman, 1968), with contributions from numerous fields of research (Cleary & Hebb, 2016; du Jardin & Séverin, 2011; Lennox, 1999). This combined body of research has mainly rendered accounting ratios as predictors of bankruptcy. A sizeable portion of these bankruptcy prediction models are highly accurate[1] (Abellán & Mantas, 2014; Jabeur & Fahmi, 2018; Khademolqorani et al., 2015; Zięba et al., 2016). In the face of highly accurate models, the task of predicting bankruptcy then seems clear cut.

1.2 Problem discussion

However, the costs associated with bankruptcy are still substantial (Quintiliani, 2017), indicating that bankruptcy prediction is not a solved problem. As suggested by both Balcaen and Ooghe (2006) and Appiah et al. (2015), this may, in part, be due to theoretical neglect of the time dimension of business failure, meaning bankruptcy prediction studies do not consider any earlier accounting information, at the firm level, than the latest annual accounts. Indeed, firm failure processes spanning multiple accounting periods have been proposed (Amankwah-Amoah, 2016; Crutzen & van Callie, 2008; Ooghe & de Prijcker, 2008). In theory, firms fail at different rates, often exhibiting different accounting characteristics depending on which type of trajectory towards bankruptcy they follow.

While such failure processes have begun to be empirically observed (Lukason & Laitinen, 2016, 2019; Nummela et al., 2016), it is not obvious that these would have an impact on bankruptcy prediction performance. On the contrary, the last two decades has seen an unprecedented international move towards regulation stipulating more time-relevant accounting information, with the rationale that it leads to better decisions (De George et al., 2016). This notion has been supported by some studies (e.g. Houqe et al., 2016; Yeh et al., 2019), but rejected by others (e.g. Braga, 2017; Gordon et al., 2017).

Nevertheless, historical accounting information seems to have some bearing on bankruptcy prediction performance. Few such studies have been done, and further empirical results are called for (Appiah et al., 2015; Balcaen & Ooghe, 2006). French studies indeed indicate that considering the time dimension of bankruptcy might lead to more accurate prediction models when the predictive horizon exceeds one year (du Jardin, 2015; Mselmi et al., 2017). In theory however, accounting for the time

[1] In excess of 90 per cent accuracy.

dimension of bankruptcy should also improve prediction performance in the short term (Balcaen & Ooghe, 2006), which is important for stakeholders who make their decisions with shorter predictive horizons.

Sweden differs from other countries in terms of taxation level, legislation, and the financial environment of firms, while being described as a small economy based on exports (Öhman & Yazdanfar, 2017). Specifically, Swedish bankruptcy law, said to be more in line with bankruptcy laws of the UK and USA, differs from that of France in that it is more creditor than debtor oriented (Kammel, 2008). Firm failure processes may also differ between countries, while predictors of bankruptcy may vary between exporting and non-exporting firms (Laitinen & Lukason, 2014; Lukason & Laitinen, 2018). Through statutory annual publication of final accounts, the Swedish business context provides public accounting data. The Swedish context then brings an opportunity to assess the general usefulness of the time dimension of bankruptcy on prediction models, and to extend the bankruptcy prediction literature by fitting prediction models to an additional financial environment.

1.3 Purpose statement

Taking the discussion under the previous subsection into account, the purpose of this thesis is to assess the impact of time dimensionality on bankruptcy prediction performance.

2 Theoretical framework and previous findings

As noted by Crutzen and van Callie (2008), the bankruptcy literature can be broadly split into two perspectives; the predictive and the preventive. This thesis holds the predictive perspective, while also drawing from the time dimension of bankruptcy from the preventive literature in general, and the firm failure process literature in particular. As the name suggests, the predictive literature concerns itself with the development of bankruptcy prediction models. The comparatively smaller preventive literature on the other hand, primarily explores while firms go bankrupt.

The remainder of this section is structured as follows; in subsection 2.1, the bankruptcy concept is established, and, for comparison's sake, the reader is given a primer on the Swedish legal context. Then, subsection 2.2 gives a non-technical overview of previous bankruptcy prediction model development in the predictive literature. Lastly, subsection 2.3 discusses firm failure processes from the preventive literature, as well as relevant predictive studies, in order to establish hypotheses.

2.1 The bankruptcy concept and the Swedish context

Previous literature uses a plethora of concepts when describing a failing firm. These include, among others, corporate failure, business failure, firm failure, financial failure, bankruptcy, and financial distress (Amendola et al., 2017; Doumpos et al., 2017; du Jardin, 2016; du Jardin & Séverin, 2012; Erkens et al., 2012; Geng et al., 2015; Gepp & Kumar, 2015; Jones et al., 2017; Kosmidis & Stavropoulos, 2014; Lakshan & Wijekoon, 2012; Mselmi et al., 2017; Revilla et al., 2016; Scherger et al., 2017; Serrano-Cinca et al., 2018; Van Peursem & Chan, 2014). Depending on the researcher, concepts can be distinct from, overlap with, or be interchangeable with concepts used by other researchers, that is, the definitions of these concepts are often arbitrary.

Appiah et al. (2015) support the notion that these concepts are used as synonyms in the literature, but go on to argue that financial distress is the odd one out, and should be viewed as truly distinct from the others. *A dictionary of finance and* *banking* (Law, 2018) states that financial distress is "a situation in which the activity of a firm is influenced by the possibility of impending insolvency". Thus, financial distress is the state which precedes legal bankruptcy. Viewing bankruptcy as a synonym of firm failure, this thesis will make a point in distinguishing between bankruptcy and financial distress.

What then operationally constitutes a bankrupt firm in the data varies. Previous research use definitions ranging from stock exchange delisting, through resignation of director, to profits being lower than forecasted (Appiah et al., 2015). For example, Chinese studies sometimes define a bankrupt firm as one which has received a "special treatment" designation from financial inspectors (Kim et al., 2016). A bankrupt firm can also be defined as one which is marked as inactive in the researcher's database of choice (Le & Viviani, 2018). However, most of the previous research define a bankrupt firm as one in the state of legal bankruptcy (Pompe & Bilderbeek, 2005), something Appiah et al. (2015) argue allows for objectively determining the date of failure, and hence straightforwardly classifying firms as either failed or non-failed.

Legal bankruptcy is the state of having one's property submitted to external management under applicable bankruptcy law. Another option to an insolvent corporation is that of debt restructuring. Compared to, for example, the USA, the Swedish context differs somewhat in that debt restructuring is not a form of bankruptcy, but rather a period of protection from bankruptcy. Consequently, and for comparison's sake, this thesis only concerns itself with bankruptcies comparable to chapter 7 bankruptcies in American federal law, but not chapter 11 bankruptcies. Defining bankruptcy as such is reasonable considering the implied ability of a restructuring firm to improve financially. Since failed restructurings lead to bankruptcy, a firm in the process of restructuring should consequently be viewed as distressed, but not bankrupt.

In Sweden, bankruptcy law states that an insolvent debtor shall be filed for bankruptcy by itself or a creditor, to a district court [tingsrätt]. If the applicant is the debtor, insolvency is presumed by the court, declaring the debtor to be bankrupt. If the applicant is a creditor, the applicant must prove the insolvency of the debtor. These two types of applications correspond well to international practice, where bankruptcy can either be filed for by the debtor or petitioned for by creditors. Concerning corporate bankruptcy specifically, filing as a debtor is done by the board of directors. Internationally, failure to do so in time may lead to litigation against board members. In Sweden, the potential personal liabilities of board members are conditioned on whether there are reasons to believe that half of the shareholders' permanent equity has been spent. If so, the board must compile a verification balance sheet for liquidation purposes, submit this balance sheet to auditing, and announce a special meeting of the shareholders. The shareholders may then decide to either liquidate the firm or continue its operations. In the latter case, the firm has eight months to restore the shareholders' permanent equity by either earnings or equity issuance, else it must be liquidated. Failure from the board to perform any of these steps results in board members becoming personally liable for the debts of the corporation. (Swedish National Tax Board, 2019)

This brings into question if the concept of bankruptcy should be extended to liquidation. Appiah et al. (2015) point out that, in some studies, the bankruptcy definition is indeed widened to one of "economic bankruptcy", thus including more firms than those just legally bankrupt. Generally, the presumable danger in this should be the possibility of using a definition that is more easily predicted, rather than one which is comparatively more useful to predict. Specifically, Altman et al.

(2017) point out that liquidated firms could as well have ceased their business activities for other reasons, for example the merger of two firms, or the discontinuance of a daughter firm. Hence, this thesis defines a failed firm as one that is legally bankrupt.

2.2 A brief review of bankruptcy prediction models

2.2.1 An overview of model development

The predictive literature begins with the work of Altman (1968), whose multi-discriminant analysis (MDA) of accounting ratios is regarded to be the seminal multivariate model in the field. As such, it has been employed in numerous academic and applied settings over the decades. The other traditional statistical method is that of logistic regression (LR), used in the literature since the 1970s (Martin, 1977). MDA and LR models often show comparable accuracy, as shown in Table 1. The main advantage of LR models is that they are more robust in estimating the individual importance of predictor variables, thus indicating something about each predictor variable's contribution to bankruptcy (Press & Wilson, 1978). Both MDA and LR models have become the de facto measuring sticks in the predictive literature, with modern model development still being compared to these traditional methods (Iturriaga & Sanz, 2015; Zhou et al., 2014).

With the advent of increased computing power, such modern development is often done using artificial intelligence (AI) models, making their entrance into the predictive literature with Odom & Sharda (1990). Since 2010, most bankruptcy prediction models have concerned themselves with AI models. The merits of these models are that they do away with a lot of assumptions implied on the data, compared to MDA and LR. For example, AI models such as neural networks and support vector machines are not sensitive to data skewness, outliers or multicollinearity (Alaka et al., 2017). At the same time, Table 1 shows AI models to have quite consistently better prediction accuracy than the traditional statistical models in the 31 recent bankruptcy prediction models reviewed for this thesis, a trait also asserted by Barboza et al. (2017).

While AI models offer promising opportunities to limit the problem of bankruptcy prediction, they are no be all and end all solution. Like other research fields, the predictive bankruptcy literature must be viewed holistically in order to assess the state of the art. Some researchers contribute to the field by constructing models tailored to a specific economic environment (Alaminos et al., 2016; Altman et al., 2017; Cultrera & Brédart, 2016; Gavurova et al., 2017; Huang et al., 2017b; Mselmi et al., 2017) or a specific industry (Cleary & Hebb, 2016; Heo & Yang, 2014; Iturriaga & Sanz, 2015; Tserng et al., 2014). By cross-referencing these citations with Table 1, such studies seems more likely to exclusively use traditional statistical methods such as MDA or LR.

Other researchers contribute to model development by proposing novel technical approaches in order to increase classification accuracy, often on the bleeding edge of AI technology (du Jardin, 2016; Wang & Wu, 2017; Yu et al., 2014; Zięba et al., 2016). Others aim to better their models one step at a time, doing experimental studies in order to assess some model trait's impact on prediction performance (Abellán & Mantas, 2014; du Jardin, 2015; Liang et al., 2016). For example, researchers contribute to model development by comparing different model types and model ensembles, that is, different types of models working in conjunction (Chou et

Table 1: Characteristics of previous studies

		Highest obtained accuracy			
Author (year)	Sample	AI	MDA	LR	Time dim.
Gavurova et al. (2017)	690	.94			0
Gordini (2014)	3,100	.72		.67	0
Heo and Yang (2014)	2,762	.77	.51		0
Tsai (2014)	690	.92		.87	0
Tsai et al. (2014)	690	.87			0
Tserng et al. (2014)	87			.79	0
Virág and Nyitrai (2014)	156	.89			0
Wang et al. (2014)	132	.80			0
Yeh et al. (2014)	220	.97			0
Yu et al. (2014)	500	.93	.87		0
Zhou et al. (2014)	2,010	.76	.72	.74	0
du Jardin (2015)	16,880	.81	.80	.81	1
Iturriaga and Sanz (2015)	772	.94	.78	.82	0
Khademolqorani et al. (2015)	180	.94	.77	.80	0
Liang et al. (2016)	688	.93			0
Alaminos et al. (2016)	440			.85	0
Cleary and Hebb (2016)	264		.90		0
Cultrera and Brédart (2016)	7,152			.79	0
du Jardin (2016)	17,540	.84		.84	0
Liang et al. (2016)	478	.81			0
Zięba et al. (2016)	10,503	.96			0
Altman et al. (2017)	5,750,642		.74	.77	0
Barboza et al. (2017)	14,331	.87	.52	.76	0
Chou et al. (2017)	600	.95			0
Gavurova et al. (2017)	700		.77	.85	0
Huang et al. (2017b)	312	.74		.74	0
Mselmi et al. (2017)	212	.94		.92	0
Wang and Wu (2017)	260	.96			0
Jabeur and Fahmi (2018)	800	.93		.47	0
Le and Viviani (2018)	3,000	.81		.81	0
Nehrebecka (2018)	14,191	.63		.67	0

Note: Sample sizes and accuracy values of the 2014 and 2015 studies in the table by Alaka et al. (2017). AI=Artificial intelligence, MDA=multi-discriminant analysis, LR=Logistic regression. Time dim.=0: models estimated from one annual account at the firm level. Time dim.=1: models estimated from multiple annual accounts at the firm level.

al., 2017; Gordini, 2014; Heo & Yang, 2014; Khademolqorani et al., 2015; Tsai, 2014; Tsai et al., 2014; Virág & Nyitrai, 2014; Zięba et al., 2016).

On the one hand, this experimental take on model development seems to render accurate models (see Table 1). On the other hand, it is noted that these studies seldom are grounded in economic theory (Appiah et al., 2015). Rather, they are grounded in AI theory from operational research. Thus, by conducting experiments grounded in economic and financial theory, there is an opportunity to further the development of bankruptcy prediction models, while also assessing the impact of such theory on practical applications.

2.2.2 Discriminant variables

Predictive studies then use differing sets of discriminant variables in order to predict bankruptcy. While there is no optimal set of accepted variables for bankruptcy prediction, the use of accounting ratios as indicators of firm performance has been in practice since the turn of the last century, but it was not picked up by academics until the influential papers of Beaver (1966) and Altman (1968). Accounting ratios, in this sense, are quotients between some accounting item scaled by another, in order to enable comparability of financial performance between firms. The seminal discriminants of Altman (1968) include, among others, working capital/assets, retained earnings/assets and equity/liabilities. The author shows that a number of these accounting ratios are predictive of later bankruptcy, and have since been used in the literature, but also expanded upon by the original author (Altman et al., 2017).

At the same time, bankruptcy prediction studies often make a point in using different ratio selection techniques. Table 2 shows accounting ratios considered in two or more of the studies in Table 1. Typically, predictive studies compile such tables of candidate variables from previous studies, which are then reduced to select more representative variables with better predictive power (Liang et al., 2016). In fact, the aim of some experimental studies in the predictive literature is to assess the impact of different variable selection techniques on prediction performance (Liang et al., 2016, 2015; Wang et al., 2014; Zhou et al., 2014).

Methodological issues can be had with this brute force approach (Appiah et al., 2015). At the same time, choosing ratios which fit the data may capture the underlying characteristics of the population of firms under study, in terms of legislation and economic environment. Since variable selection is often inherent to AI modelling and tailored to the population at hand (Chou et al., 2017; Huang et al., 2017b; Liang et al., 2016, 2015; Nehrebecka, 2018; Tsai, 2014; Yeh et al., 2014; Zhou et al., 2014), practical applications should not necessarily use variables selected in another population, but rather mimic the variable selection process used. This insight should have consequences for studies looking to further the development of bankruptcy prediction models with a grounding in economic theory rather than AI theory. On the one hand, the discriminants of Altman (1968) (R2, R3, R4, R19 and R59 in Table 2) are regarded as theoretically motivated (Lukason & Laitinen, 2019). On the other hand, empirically motivated variables are more prevalent in model development (Huang et al., 2017b; Iturriaga & Sanz, 2015; Nehrebecka, 2018; Yeh et al., 2014; Zhou et al., 2014). Looking to further model development using economic theory, it is far from clear which approach is the more valid. However, the prevalence of empirically motivated variables indicates that it would be more useful for practical applications to assess the impact of theoretical factors using empirical selection of the discriminant variables in Table 2.

2.2.3 Performance metrics of models

As shown in Table 1, classification accuracy is the principally reported performance metric for bankruptcy prediction models, defined as:

$$Accuracy = \frac{\sum True\ postivies + \sum True\ negatives}{\sum Total\ predictions}$$

where true positives is the number of bankrupt firms correctly classified and true negatives is the number of non-bankrupt firms correctly classified. In contrast to these, false positives is the number of incorrectly classified non-bankrupt firms. Analogous to false positives, false negatives is the number of incorrectly classified bankrupt firms. For classification problems in general, this distinction of correct and erroneous classifications into different categories is important, seeing as they can have unequal misclassification costs (Calabrese, 2014). For bankruptcy prediction in particular, while argued that false positives can lead to credit crunches (Gordini, 2014), there is relative consensus that false negatives are the most costly misclassifications when the two categories are viewed in juxtaposition (du Jardin, 2015; García et al., 2019; Lai, 2009; Veganzones & Séverin, 2018).

In practice, a higher number of false positives leads to alternative costs from lost business opportunities, while a higher number of false negatives leads to losses stemming from the bankruptcy of suppliers, customers or other debtors. It goes without saying then, that assuming the survival of all firms under study is a costly prospect, but so is assuming the imminent bankruptcy of all firms. In fact, in real world data where bankruptcy is a rare event, the former strategy would render almost perfect classification accuracy while still missing all bankruptcies. Since accuracy only involves true positive and true negative classifications, it follows that it is hard to judge the usefulness of models which only report this metric.

Powers (2011) notes that different research disciplines use different secondary metrics. For example, medical sciences use the area under the receiver operating characteristics (AUROC) curve, while the behavioral sciences often use specificity and sensitivity. Still, the practice of only reporting accuracy numbers seems prevalent in the predictive bankruptcy literature (Altman et al., 2017; Chou et al., 2017; Cleary & Hebb, 2016; Cultrera & Brédart, 2016; du Jardin, 2015; Le & Viviani, 2018; Mselmi et al., 2017; Tsai et al., 2014; Virág & Nyitrai, 2014; Wang & Wu, 2017; Yeh et al., 2014; Yu et al., 2014; Zięba et al., 2016). The problem also gets compounded if the aim of the study is to compare different types of models, since relative differences in accuracy do not hint at any potential differences in misclassification costs between different types of models (Chou et al., 2017; du Jardin, 2015; Gordini, 2014; Jabeur & Fahmi, 2018; Tsai, 2014)[2].

Other bankruptcy researchers do report secondary performance metrics, such as the respective rates of false positives and false negatives (Barboza et al., 2017; Gavurova et al., 2017; Heo & Yang, 2014; Liang et al., 2016; Tsai, 2014; Wang et al., 2014), AUROC analysis, which compares true positives to false positives (Alaminos et al., 2016; Barboza et al., 2017; Huang et al., 2017b; Khademolqorani et al., 2015; Le & Viviani, 2018; Tserng et al., 2014), as well as two measures called recall and precision (Iturriaga & Sanz, 2015; Khademolqorani et al., 2015; Le & Viviani, 2018; Zhou et al., 2014). Going forward, it should be noted that all the mentioned metrics are interconnected by being mathematically grounded in true positives, true negatives, false positives, and false negatives. Differences between them may therefore be subtle, which might be one possible reason for their apparent arbitrary use in the predictive bankruptcy literature. However, assuming that models cannot be properly evaluated while being agnostic to costs, as well as bankrupt firms being the more costly case to misclassify, this thesis uses the probability of detecting a bankrupt firm, true positive rate, sensitivity or recall:

[2] While du Jardin (2015) does consider misclassification costs by other means, the only reported metric is accuracy.

Table 2: Accounting ratios in previous studies

Ratio	Definition	Occurrences
R1	Current assets/current liabilities	16
R2	EBIT/assets	15
R3	Sales/assets	14
R4	Working capital/assets	14
R5	Liabilities/assets	12
R6	EAT/assets	11
R7	Current assets/assets	10
R8	Equity/assets	10
R9	EAT/equity	9
R10	Retained earnings/assets	9
R11	Quick assets/current liabilities	9
R12	Cash flow/liabilities	8
R13	Liabilities/equity	8
R14	Working capital/sales	8
R15	Cash flow/assets	7
R16	Cash/assets	7
R17	EBIT/equity	7
R18	Sales/accounts receivable	7
R19	Cash flow/sales	6
R20	Cash/current liabilities	6
R21	Current assets/sales	6
R22	Current liabilities/assets	6
R23	EBIT/sales	6
R24	Inventory/sales	6
R25	Long-term liabilities/assets	6
R26	Accounts receivable/sales	5
R27	Cash flow/equity	5
R28	Cash/sales	5
R29	Current liabilities/liabilities	5
R30	EAT/sales	5
R31	EBITDA/sales	5
R32	Sales/equity	5
R33	Sales/fixed assets	5
R34	Quick assets/assets	5
R35	Accounts payable/sales	4
R36	EBIT/invested capital	4
R37	EBITDA/assets	4
R38	Long-term liabilities/equity	4
R39	Assets/liabilities	3
R40	Cash/liabilities	3
R41	Current liabilities/equity	3
R42	EBT/equity	3
R43	EBT/invested capital	3
R44	Equity/fixed assets	3

Table 2 (continued): Accounting ratios in previous studies

Ratio	Definition	Occurrences
R45	Fixed assets/assets	3
R46	Sales/current assets	3
R47	Quick assets/sales	3
R48	(Quick assets – accounts receivable)/sales	2
R49	Cash flow/current liabilities	2
R50	Cash/current assets	2
R51	Current assets/liabilities	2
R52	Current liabilities/current assets	2
R53	Current liabilities/sales	2
R54	EAT/liabilities	2
R55	EBIT/interest expenses	2
R56	EBITDA/permanent equity	2
R57	EBT/assets	2
R58	EBT/net sales	2
R59	Equity/liabilities	2
R60	Equity/permanent equity	2
R61	Financial expenses/assets	2
R62	Financial expenses/EAT	2
R63	Financial expenses/EBITDA	2
R64	Financial expenses/sales	2
R65	Sales/cash	2
R66	Sales/invested capital	2
R67	Sales/working capital	2
R68	Working capital/equity	2
R69	Working capital/liabilities	2

Note: Showing all accounting ratios considered for use as determinants of bankruptcy in the previous studies of Table 1. EBIT=earnings before interest and tax, EAT=earnings after tax, EBITDA=earnings before interest, tax, depreciation and amortization, EBT=earnings before tax.

$$Recall = \frac{\sum True\ postivies}{\sum True\ positives + \sum False\ negatives}$$

From a stakeholder standpoint, the recall rate reveals the probability of incurring losses when presented with a business opportunity towards a firm which will subsequently become bankrupt. As previously mentioned however, this is not the only cost of misclassification. Recall is therefore commonly paired with precision (Iturriaga & Sanz, 2015; Le & Viviani, 2018; Zhou et al., 2014):

$$Precision = \frac{\sum True\ postivies}{\sum True\ positives + \sum False\ positives}$$

From a mathematical viewpoint, recall and precision only differ in terms of false negatives and false positives used as the right term in the denominator. In practice however, the precision rate corresponds well to the question "out of all firms predicted as bankrupt, how many are actually bankrupt?" Larger precision then indicates fewer alternative costs incurred from missed business opportunities. These two metrics are not without critique however. On the one hand, Powers (2011) argues that biases are introduced by ignoring the classification performance of surviving firms. On the other hand, misclassification costs are biased toward the bankrupt case (du Jardin, 2015; García et al., 2019; Veganzones & Séverin, 2018), suggesting that recall and precision are indeed practical for model evaluation.

2.3 The time dimension of bankruptcy

2.3.1 Firm failure processes and previous empirical findings

This subsection gives a superficial introduction to the time dimension of bankruptcy in the context of failure processes. For brevity, it should be noted that the time dimension of bankruptcy refers to considering accounting data from multiple annual accounts preceding the date of prediction, at the firm level. That is, the firm-specific process leading to bankruptcy over time. This should be held in contrast to other temporal aspects of bankruptcy in general, and bankruptcy prediction specifically, such as modeling the expected time to bankruptcy using longitudinal data (Gepp & Kumar, 2015; Laitinen, 2005), comparing prediction performance of longitudinal versus cross-sectional datasets (Berg, 2007; Chou et al., 2017), and assessing prediction performance as the predictive horizon increases beyond one year (Alaminos et al., 2016; du Jardin, 2015; Tserng et al., 2014). Also, a predictive horizon of one year is denoted as t in this thesis.

In the predictive bankruptcy literature, the probability of bankruptcy is indeed often modeled from firms' most recent annual accounts (Iturriaga & Sanz, 2015; Nehrebecka, 2018; Wang & Wu, 2017). In Table 1, this is shown as 0 in the time dimensionality column. As such, bankruptcy prediction models can be said to be predominantly static, something also noted in previous literature reviews (Appiah et al., 2015; Balcaen & Ooghe, 2006). The merits of static modeling is the natural fit of data; financial reports are indeed static snapshots and newer information should be more relevant in describing the financial state of a firm. This idea is also seen in the last two decades of accounting regulation and practice, with focus being shifted from reliability to relevance through the adoption of International Financial Reporting Standards (De George et al., 2016). Previous findings also show that bankruptcy prediction model performance deteriorates as the time to bankruptcy increases (Alaminos et al., 2016; Tserng et al., 2014).

The preventive side of the literature, on the other hand, is exploring and describing the underlying reasons why firms fail in the first place (Amankwah-Amoah & Debrah, 2014; Hamilton, 2006). Starting with Crutzen and van Callie (2008) and Ooghe and de Prijcker (2008), theoretical failure processes are induced from literature review and case studies of bankrupt firms, respectively. Due to the relative novelty of the field, researchers have not converged on any widely accepted typology of such processes (Amankwah-Amoah, 2016). For example, one distinct type of failure process described by Ooghe and de Prijcker (2008), is that of the apathetic established firm. These firms go bankrupt because management is apathetic to gradual changes in the competitive environment, thus failing to respond to the loss of a former strategic advantage. Initially, this can be observed financially as declining sales, leading to the following chain of financial distress; the lowered sales lead to lower profits, which eventually lead to liquidity problems, forcing firms to take on additional liabilities. In turn, this raises financial expenses causing further liquidity problems, eventually leading to bankruptcy. In fact, Ooghe and de Prijcker (2008) shows that, while the root cause of financial distress differ, some variation of this circle of death is inherent to all failure processes, be it the aforementioned apathetic established firm or, for example, the unsuccessful startup or an overly ambitious growth firm. Recently, the preventive literature has produced empirical findings from larger samples, supporting the existence of a finite number of failure processes (Lukason & Laitinen, 2016, 2019; Nummela et al., 2016).

Describing all proposed typologies of firm failure processes is out of scope for this thesis. One important point though, is the different financial characteristics exhibited depending on the temporal location of a firm within its failure process. Lukason and Laitinen (2019) argue that firm failure processes can not only be distinguished based on managerial root causes of financial distress, but rather on temporal failure risk. The authors then go on to show that theoretical short, medium, and long failure processes are empirically distinct: Short failure processes are those where firms show no obvious signs of financial distress until shortly before bankruptcy. The only hints seen are lowered profitability at *t-2*, and high losses at *t-1*, leading to critically low liquidity. In medium failure processes, firms show losses in *t-3*, getting more severe at *t-2* and *t-1*. Consequently, liquidity is lowered already at *t-3* and becomes worse as time goes on, eventually leading to bankruptcy. Furthermore, Lukason and Laitinen (2019) show that firms finding themselves in long failure processes, exhibit similar signs of financial distress in terms of profitability, liquidity, and equity as early as *t-5*.

Firms may then be financially distressed for years without going bankrupt. Several researchers suggest that this can lower bankruptcy prediction performance (Appiah et al., 2015; Balcaen & Ooghe, 2006; Lukason & Laitinen, 2019). Therefore, one alternative to static prediction models is to view a firm's trajectory towards bankruptcy as a discrete process, meaning a process with finite number of states, as conceptualized by Lensberg et al. (2006). Viewing the time dimension of bankruptcy as several discrete steps towards bankruptcy preserves the merit of the natural fit of data as subsequent snapshots of firm performance. However, matching empirical failure processes to theoretical ones is evidently a complicated task (Lukason & Laitinen, 2019). In order to be useful, prediction models should arguably be easy to implement and use. This might be a reason why bankruptcy prediction studies often aim to maximize prediction performance by cherry-picking accounting ratios which best discriminates bankrupt firms from non-bankrupt firms in the used dataset (Wang et al., 2014).

There is a point then, in testing if the mere inclusion of multi-year data in models can improve bankruptcy prediction performance, all else being equal. Even if the approach is sufficiently simple and technically feasible for most applications, there is some doubt to whether it should work. In fact, one might even doubt that accounting for theoretical failure processes would improve bankruptcy prediction. For example, while the belief that past financial performance is indicative of future performance might be common among practitioners, research demonstrates the difficulty for investors to predict stock price movements any better than if modeled as a stochastic process (Fama, 1995; Mishra et al., 2015; Moosa & Vaz, 2015). Seeing as future improvement or deterioration of firms' financial conditions might well be random, it might make more sense to model the risk of bankruptcy as a function of the current level of financial distress, as is commonly done with static models (Iturriaga & Sanz, 2015; Nehrebecka, 2018; Wang & Wu, 2017). The firm failure process literature is indeed lacking turnaround analysis (Lukason & Laitinen, 2019). Other preventive literature does explain turnaround of financial distress, but mainly in terms of qualitative factors (Trahms et al., 2013).

Of the 31 recent bankruptcy prediction models reviewed for this thesis, only one considers the time dimension of bankruptcy in accordance with the aforementioned definition, as indicated by a 1 in the time dimensionality column of Table 1: du Jardin (2015) aims to improve bankruptcy prediction accuracy beyond a 1-year predictive horizon in French firms, by designing time dimensional prediction models based on failure processes. This is done by assuming that different industries have inherently different failure processes, consequently fitting one each of a time dimensional and a non-time dimensional model per industry and financial year and comparing the two. du Jardin (2015) finds that considering the time dimension of bankruptcy does not improve classification accuracy for bankruptcies occurring at time *t*, but that it does for those occurring at time *t+1* and *t+2*. The findings are partially supported by another French study by Mselmi et al. (2017). Although not a time dimensional predictive study per se, these authors show that different accounting ratios discriminate bankrupt firms from non-bankrupt firms depending on the chosen predictive horizon (*t* vs *t+1*).

However, matching a sample of bankrupt firms to a control sample of non-bankrupt firms by the basis of industry and year comes with two issues. Firstly, as Appiah et al. (2015) summarize, matching by industry may lead to models being over represented by industries suffering from recession. Secondly, estimating models on cross-sectional data can make models oblivious to underlying macro-economic factors in the environment which are subject to change. For classification problems in general, this change in the target variable over time is called concept drift (Sun et al., 2017). Indeed, longitudinal data is shown to increase prediction accuracy (Berg, 2007; Chou et al., 2017). Adding longitudinal data at the firm level could then achieve the same effect, but also give the false impression of time dimensionality being the cause of accuracy improvements compared to static models, when this effect is in fact due to the increased amount of data.

2.3.2 Hypotheses

Taken together, there is a case for comparing the accuracy of time dimensional prediction models to that of static models in a more general way. Also, the findings of du Jardin (2015) seems inconsistent with the preventive research of Lukason and Laitinen (2019), who are asserting that bankrupt firms predominantly follow short failure processes, especially in developed countries. Thus, accounting for the time dimension of bankruptcy should theoretically also improve prediction accuracy for bankruptcies occurring at time *t*. The following hypothesis is formulated:

Hypothesis 1 (H1). *Time dimensional bankruptcy prediction models have higher accuracy than static prediction models, ceteris paribus.*

While accuracy is the principal performance metrics of bankruptcy prediction models, other metrics might also be impacted by the time dimensionality of models. It seems probable that firms in the end stages of longer failure processes can be more comfortably classified as bankrupt. At that point, firms have drained their capital reserves and lowered their accumulated profitability to a point where they cannot take on additional liabilities to uphold liquidity (Lukason & Laitinen, 2019; Ooghe & de Prijcker, 2008). The hypothesis is as follows:

Hypothesis 2 (H2). *Time dimensional bankruptcy prediction models have higher recall than static prediction models, ceteris paribus.*

In the same vein, firms in temporary financial distress may be incorrectly classified as bankrupt (Balcaen & Ooghe, 2006), that is, lowering the precision of prediction models. Likewise, firms exhibiting signs of financial distress which are in fact in the beginning stages of a longer failure process as described by Lukason and Laitinen (2019), might also be classified as false positives in static models. The hypothesis is as follows:

Hypothesis 3 (H3). *Time dimensional bankruptcy prediction models have higher precision than static prediction models, ceteris paribus.*

3 Method

3.1 Research philosophy and method justification

In this thesis, knowledge is viewed as what statements about reality are reasonably useful to make (Peirce, 1992). This pragmatic position is in accordance with the emphasis within Swedish graduate studies that method choice should be tailored to solve a specific problem. However, the personal experience of this author is that students then feel obligated to take on an epistemological position which suits the chosen method, and implicitly also take on some ontological position about the nature of reality. Specifically, this manifests itself as a choice between a volatile socially constructed reality and an eternally objective reality. It can be argued that choice of research problem and, in turn, method should instead be influenced by the world of the researcher (Arbnor & Bjerke, 2008). At the same time, such an approach risks discrediting research conducted by advocates of the opposite position. This thesis instead subscribes to the position of Corbin and Strauss (2008); that choice of method is a practical consideration and that interplay between quantitative and qualitative inquiry is needed to advance the state of a given research field.

The absolute truth of something then, is what researchers in a field would conclude given an eternity of inquiry. From this statement, it follows that the current truth comes from the argument best supported by the current state of the art. Therefore, this thesis surrenders all notions about the nature of reality by recognizing that it is not something which can be determined by practical inquiry. Moreover, all methodological considerations conducted in making this thesis are concerned more with practical implications of potential findings, rather than theoretical. In summary, this means that if findings do not have practical value, then they have no value. This is not to say that theoretical implications do not have any practical value in the end. Rather, theory cannot have any value until proven useful in a practical setting. (Peirce, 1992)

Within this pragmatic framework, a quantitative method was chosen to fulfill the purpose of the thesis. Considering the quantitative nature of prediction models, the performance of these is arguably most useful if quantified. Specifically, because testing the derived hypotheses involved comparing two groups of models on measures generated by a process which the author was in control of, an experimental method was chosen (Abellán & Mantas, 2014; du Jardin & Séverin, 2012; Heo & Yang, 2014; Kutyłowska, 2015; Liang et al., 2016; Sun et al., 2017; Wang et al., 2014). Consequently, causality could be established, as implied by the wording of the purpose (Cox & Reid, 2000).

3.2 Literature selection process

3.2.1 General inclusion criteria

In this thesis, journals and publishers of scientific level 1 or 2, as defined by the Norwegian Centre for Research Data at the time of writing, were considered for citation. Exceptions to this rule, regarding publications, were only made for "classic" book titles and older articles whose journals are now discontinued. Regardless, only peer reviewed articles have been cited when referring to previous findings. Websites were considered for citation on a case-by-case basis.

3.2.2 Structured review of bankruptcy prediction models

Initially, the predictive bankruptcy literature was surveyed haphazardly, rendering two recent literature reviews of the field; one highlighting methodological issues (Appiah et al., 2015), and one aiming to develop a framework for model selection (Alaka et al., 2017). It was determined that a research problem could be established by systematic review of bankruptcy prediction models compiled by the latter, in terms of the issues highlighted by the former. To increase relevance, the models reviewed by Alaka et al. (2017) were also supplemented by additional predictive studies from 2016 and later, by using the same search string used by the two mentioned literature reviews: ("Forecasting" OR "Prediction" OR "Predicting") AND ("Bankruptcy" OR "Insolvency" OR "Distress" OR "Default" OR "Failure").

This search was done with Google Scholar and the library search tool (Primo) of Mid Sweden University, and filtered to only include results from 2016 and later. The Google Scholar search results were deemed too large to be manageable, hence a cut-off value of 30 articles was chosen for consideration due to time constraints for the completion of the thesis. Primo, on the other hand, rendered largely irrelevant results compared to Google Scholar. Hence, a cut-off value of 20 articles was chosen for the Primo results, totaling 50 articles from both searches.

Figure 1 illustrates this process. The 50 articles from search results were superficially reviewed in terms of topic and content. Any article not producing at least one bankruptcy prediction model was dropped (cf. Alaka et al., 2017). As Figure 1 shows, this procedure excluded 33 articles, for a total of 17 remaining. These were then compiled into a complete selection together with 14 articles which were already reviewed by Alaka et al. (2017). While Alaka et al. (2017) reviewed more articles, the 14 articles were selected due to being the ones published most recently. Here, 2014 was arbitrarily chosen as a cut-off in order to maintain temporal relevance of the reviewed models.

Figure 1 shows that the whole selection process, rendered 31 articles with publication dates between 2014 and 2018. These were then structurally reviewed in terms of methodological issues highlighted by Appiah et al. (2015). While sample sizes, accuracy metrics and model types for the 2014 and 2015 articles in Table 1 had already been compiled by Alaka et al. (2017), the 17 articles originating from search results were surveyed for sample sizes, accuracy metrics and model types. All 31 articles in the selection were then surveyed for considered accounting ratios, used performance metrics, model type, aim of study and time dimensionality. Other characteristics highlighted by Appiah et al. (2015) were also surveyed for in the search of a research problem, although being irrelevant to the thesis in the end, such as geographical origin of sampled data, predictive horizon, sample matching technique, variable selection process, etcetera.

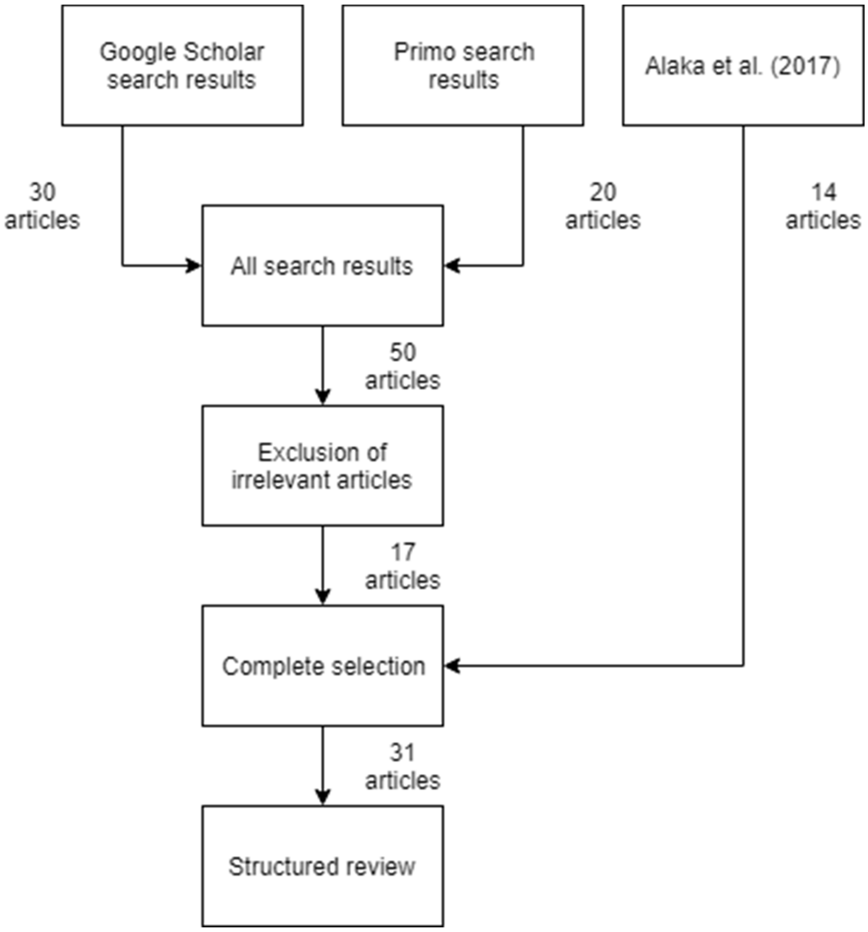

Figure 1: Literature selection process for the structured review.

3.2.3 Unstructured review of the wider literature

By confirming the infrequency of time dimensionality in recent bankruptcy prediction models, the research gap and purpose of the thesis was established through the structured review. Doing so revealed a requirement for appropriate theory and non-predictive empirical findings on the time dimension of bankruptcy. Due to the lack of uniform nomenclature and the relative lack of literature, searching was done more haphazardly by snowball selection of literature either cited in, or citing the only time dimensional predictive study identified in the structured review (du Jardin, 2015). By using such a selection process, one must acknowledge the possibility of selection bias from stepwise citation of sources which are in support of previous findings (Kicinski et al., 2015). However, the relatively few citations in recent articles by researchers familiar with the field (du Jardin, 2015; Lukason & Laitinen, 2019), indicate that most relevant preventive literature has been considered in this thesis.

In total, 101 peer reviewed articles are cited in the thesis, compared to the 31 used in the structured review. While 11 of these are attributable to the selection process of the preventive literature described in the previous paragraph, most are not. Some are statistical pieces used to justify or explain methodological choices (Breiman, 2001), some are predictive or preventive bankruptcy studies used to illustrate some specific point (Quintiliani, 2017; Sun et al., 2017), and some are drawn from the accounting and auditing literature for the same reason (Tagesson & Öhman, 2015). These articles were identified either through prior knowledge of their existence, customized search for a specific topic or snowball selection.

3.2.4 Criticism

The bankruptcy research field is not without criticism. The predictive literature offers many potential combinations of variable selection processes, model types, model ensemble combinations, and model parameters, possibly inviting researchers to try different combinations haphazardly. Consequently, the high accuracy metrics of the previous studies presented in Table 1 could conceivably be a sign of publication bias, wherein researchers and journals choose to publish interesting or impressive results rather than any results. Researchers also seem to misunderstand each other. For example, some predictive studies are cited by others as being time dimensional in the sense used in this thesis, but were revealed to be longitudinal upon review (Berg, 2007). In fact, when finding predictive studies ambiguous on whether they are time dimensional or not (Chou et al., 2017), they were treated as longitudinal.

3.3 Data

3.3.1 Financial data

The impact of time dimensionality on bankruptcy prediction performance was assessed using financial data from the annual accounts of Swedish small and medium sized enterprises (SMEs). Specifically, SMEs where chosen due to being drivers of job creation and economic growth, while reportedly facing a higher probability of failure than larger firms (Yazdanfar & Öhman, 2018). Alternatively, micro sized firms could have been included, but seeing as most predictive and preventive bankruptcy literature focus on either large firms or SMEs (Chou et al., 2017; Cultrera & Brédart, 2016; Khademolqorani et al., 2015; Lukason & Laitinen, 2019; Mselmi et al., 2017), applying theory and previous findings on even smaller firms is not necessarily valid (Appiah et al., 2015).

In this thesis, SMEs are defined are defined as corporations with staff headcounts between 10 and 249, and either balance sheets totaling less than 43 million euros or turnover less than 50 million euros (European Commission, 2017). At the time of writing, the population size can be estimated to 44,920 Swedish SMEs, if using only staff headcount as criteria (Statistics Sweden, 2019). Financial data were collected from the Retriever Business database. The sample size, as well as the range of financial years under study, were dependent on availability of data for firms matching the SME criteria, using the following search parameters:

1. Number of employees: 10 to 249, inclusive.

2. Total assets: Less than SEK 450,000,000[3], or turnover: Less than SEK 523,000,000[3].

3. Form of enterprise [bolagsform]: Limited liability company [aktiebolag].

Since some staff headcount data were missing, it was decided that firms participating in the sample had to have staff headcounts between 10 and 249 during the whole

[3] EUR 1 = SEK 10.463 as of March 19th, 2019, rounded to the nearest million SEK.

period. 209 firms with more than 249 employees and 19,386 firms with less than 10 employees had slipped through the search parameters and were removed. Several studies show that the variables discriminating bankrupt firms from non-bankrupt firms differ between financial firms, such as banks, and other industries (Betz et al., 2014; Climent et al., 2018; Iturriaga & Sanz, 2015; Le & Viviani, 2018). Hence, all firms with industry codes [SNI-koder] 64, 65, 66, and 68 were dropped from the data. After removing empty and duplicate records, these corrections resulted in a main sample of 19,627 firms and 158,654 financial year observations between 2009 and 2017.

3.3.2 Bankruptcy data

As previously stated, *t* is defined as a period of one year following the date of prediction. However, time usually passes between the account closure and the publication of financial reports. In order to be useful, a bankruptcy prediction model should arguably be estimated from data known at the time of prediction, meaning *t* should start at the date of financial report publication. To estimate this, prediction periods were offset by half a year from the date of account closure (Barboza et al., 2017).

Accordingly, bankruptcy data for the period 2010-06-01 to 2019-04-24 were collected from the Swedish Enforcement Authority [Kronofogden] and matched to financial data by comparing corporate identity numbers [organisationsnummer] and date of bankruptcy. It should be noted that one month of data is missing (2019-04-25 through 2019-05-31), due to time constraints in the completion of this thesis. In Sweden, firms can also choose to have their financial year offset from calendar year, although choosing calendar year as financial year is most common. Lacking this information, it was assumed that all firms in the data were using calendar year as financial year. Methodological implications of these two deviations are discussed in subsection 3.7. Using the matching criteria, 724 bankruptcies could be matched to the 19,627 firms in the main sample, indicating that bankruptcy is a rare event in Swedish SMEs.

3.4 Random forests

In order to be useful, firm failure processes should preferably be automatically identified. In doing so, parametric statistics offers a disadvantage compared to AI models. Merely dumping multi-year variables into a logistic regression or multi-discriminant analysis would probably lead to a poor fit due to multicollinearity (Asar, 2017).

One model type fitting these criteria is the decision tree, visualized in Figure 2. Besides being intuitive and comparatively easy to understand, a decision tree is an artificial representation of reasoning such as “if last year’s profitability was below X and this year’s profitability is lower than Y, then…”, analogous to the theoretical characteristics of a failure process. The demerit of the decision though, is its sensitivity to noise and relatively bad performance when classifying observations it was not fitted from. (Safavian & Landgrebe, 1991)

A more robust way of using decision trees for classification is that of the random forest, proposed by Breiman (2001). Random forests are ensembles of decision trees such that each tree in the forest is estimated from a bootstrapped sample, meaning each observation in the data used for fitting the forest can be seen more than once in each tree. The forest then makes a prediction probability of bankruptcy using the average of all predictions in the forest. This approach is insensitive to noise and more robust than individual decision trees, while still preserving the aforementioned traits of decision tree models.

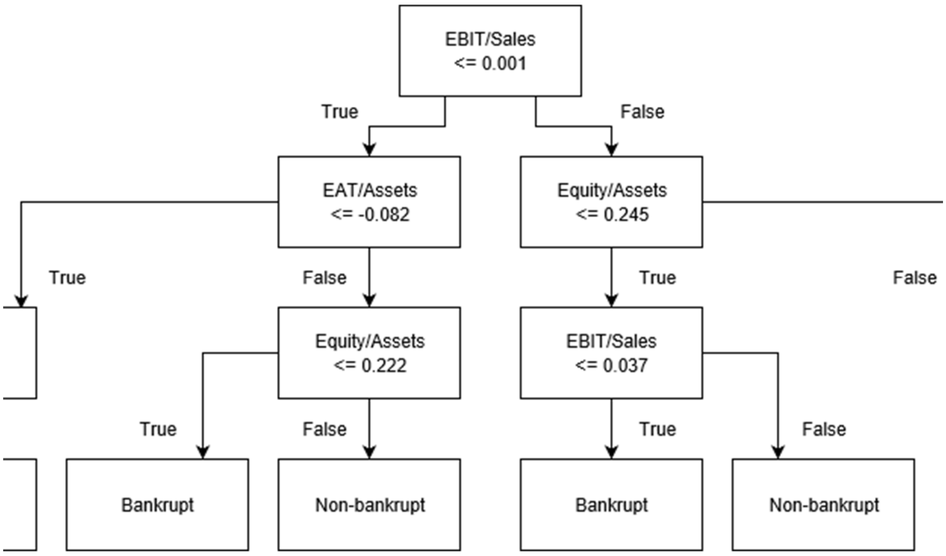

Figure 2: Anatomy of a decision tree.

Previous bankruptcy prediction studies have also used random forests in order to model complex non-linear relationships between predictor variables and the possible bankruptcies of firms (Jabeur & Fahmi, 2018; Yeh et al., 2014; Zięba et al., 2016). In order to fulfil the purpose of the thesis, the impact of time dimensionality on bankruptcy prediction should reasonably be assessed on model types used in the literature. In doing so, it can be argued that other models than random forests, such as support vector machines and neural networks, also have the proper characteristics for this purpose, or that several model types should be assessed simultaneously. However, due to time constraints, it was determined that only one type of model could be used in the experiment. With the added benefit of being sufficiently simple to explain in the context of a business administration thesis, random forests were chosen for this purpose.

3.5 Accounting ratio selection

With the same rationale as with model choice, it was determined that the impact of time dimensionality on prediction performance should be assessed using exploratory variable selection, as is common in the literature. Accounting ratios used in the studies of Table 1 were gathered from structured literature review, as described in section 3.2. This resulted in an initial set of 998 accounting ratios presented in tables in said literature. The remainder of this section describes the method of selecting which of those ratios to use in this thesis.

3.5.1 Ratio definition normalization

Due to differing nomenclature in the gathered ratios (e.g. Chou et al., 2017 vs. Tserng et al., 2014), they were linguistically normalized in order to avoid duplicates and to be able to count the number of occurrences. The following words were replaced with their corresponding arithmetic sign: "over", "to", "divided by", "per", "plus", "minus", "times", "multiplied by". Natural language processing was employed to normalize the inflection of nouns (e.g. "asset" vs. "assets"), where instances of the less common form were replaced with the more common form. Spelling errors and some linguistic synonyms were corrected and normalized, respectively, by splitting each accounting item into n-grams and manually inspecting pairs of n-grams whose edit distances were below three characters (e.g. "long term" vs. "long-term", "operations income" vs. "operating income" and "pretax" vs. "pre-tax").

Ratios defined by name rather than mathematical expression were normalized manually. Ratios defined as the turnover or the margin of some accounting item were taken to be sales divided by the item, and the item divided by sales, respectively. Commonly used ratios defined by name, such as ROI (return on investment) and "current ratio" were redefined as shown in Table 3. All other named ratios were omitted due to being ambiguously defined (e.g. "earnings ratio").

Synonymous accounting items were also normalized manually. Examples include "turnover" vs. "sales", "stockholder" vs. "shareholder", "net profit" vs. "EAT" (earnings after tax). Lastly, constant mathematical terms were removed from ratio definitions, and ratios deemed uncalculatable from the collected data were omitted (e.g. "value added/sales"), resulting in 337 unique ratios, of which the 69 occurring in two or more structurally reviewed studies are shown in Table 2.

3.5.2 Exploratory factor analysis

While random forest models are capable of handling a large number of collinear variables, they typically see diminishing performance gains for additional variables when all dimensions of the data has already been accounted for in the model. Maximizing the number of variables can still be a valid approach in maximizing model performance. However, since the purpose of this thesis is not to create a maximally accurate model, too many accounting ratios would make for an unnecessarily complicated analysis. Still, there is a practical perspective of usefulness in assessing the impact of time dimensionality of empirically selected variables. Therefore, the dimensions of the data were reduced through exploratory factor analysis (cf. Shie & Chen, 2012).

First, ratios R1-R69 were calculated from the data (see Table 2). In using 69 accounting ratios, it is reasonably expected that some are linearly interdependent, meaning some ratios can be calculated from others. This leads to a non-positive definite correlation matrix, in turn disabling robustness tests of the factor solution (Huang et al., 2017a). To circumvent this problem, the factor analysis was done iteratively by extracting factors of ratios occurring in 10, 9, 8 (and so on) studies until the correlation matrix was non-positive definite. At this point, saturation of ratios was deemed to be achieved, the offending ratios were omitted from the analysis, and an oblique factor solution was extracted.

The commonly cited explained variance (eigenvalue) threshold of 1.0 for determining how many of these factor to keep was considered (Chen et al., 2006), but defeated the aim of the factor analysis by keeping too many factors. Instead the scree test criterion was used, as suggested by Hair et al. (2014). While arguably being subjective and arbitrary, the scree test has the merit of identifying factors with a higher proportion of common variance than unique variance, done by plotting eigenvalues against the number of factors and choosing a cut-off point where the slope of the curve levels off. The eigenvalue for each extracted for each extracted factor is plotted against the number of extracted factors in Figure 3. The total number of factors, as indicated by the x-axis, is 23, meaning the iterative extraction reached ratios occurring in six or more previous studies before saturation (see Table 2). Note that current liabilities/assets and long-term liabilities/assets are linearly dependent on some other variable in the correlation matrix. Thus, they are omitted from the scree plot, diagnostic tests and factor solution. The slope of the curve in Figure 3 levels off at

Table 3: Named ratios and their corresponding definitions

Ratio	Common name	Replacement
R1	Current ratio	Current assets/current liabilities
R2	ROA	EBIT/assets
R11	Quick ratio	Quick assets/current liabilities
R17	ROE	EBIT/Equity
R36	ROI	EBIT/Invested capital
R36	ROIC	EBIT/Invested capital

Note: ROA=return on assets, EBIT=earnings after before interest and tax, ROE=return on equity, ROI=return on investment, ROIC=return on invested capital.

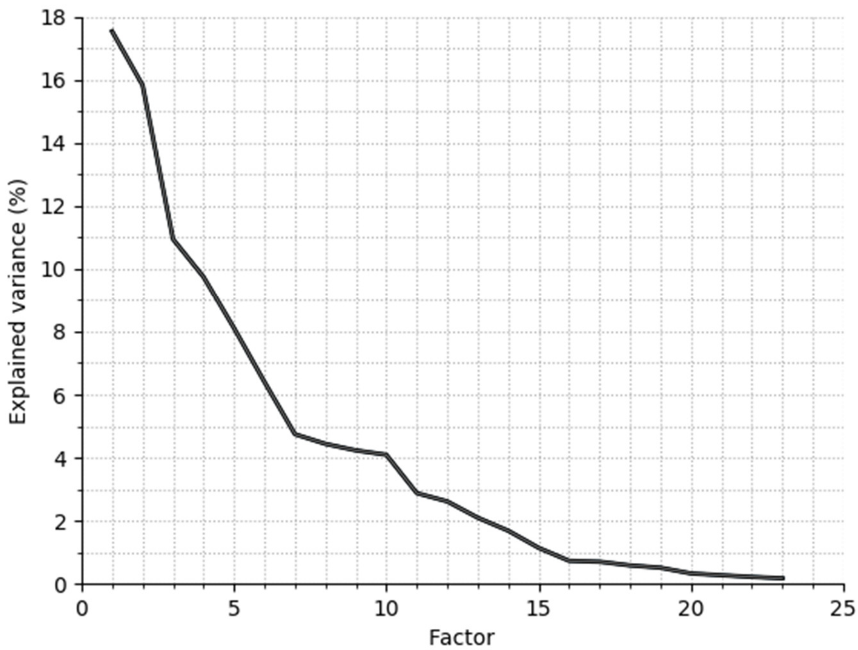

Figure 3: Explained variance per extracted factor.

seven and sixteen factors, indicating that the proportion of unique to common variance of factors changes at these cut-off points. Since sixteen factors was deemed impractical for analysis, the former cut-off point of seven was used in the following experiment. Adding all eigenvalues together in Figure 3 up until and including the seventh factor indicates that these seven factors represent 73.3 per cent of the variance of ratios R1-R25.

Before proceeding with the factor extraction, the Keiser-Meyer-Olkin (KMO) test for sampling adequacy and Bartlett's test of sphericity were performed to determine the factorability of a seven-factor solution of ratios R1-R25. The results of these tests are presented in Table 4. The KMO test indicates adequate sampling for factor analysis, with the test statistic being larger than the recommended value of 0.6 (Hair et al., 2014). Likewise, Bartlett's test is significant at $p<0.001$, meaning the included

Table 4: KMO and Bartlett's test

Test	Statistic	Value
Kaiser-Meyer-Olkin test for sampling adequacy	KMO-statistic	0.652
Bartlett's test of sphericity	Approx. chi-square	2,424,821.108
	df	253
	Significance	0.000

Note: KMO=Kaiser-Meyer-Olkin.

ratios are probably related and in turn factorable (Hair et al., 2014). Given these test statistics, the oblique seven-factor solution of 23 ratios is shown in Table 5.

For each factor kept, the ratio with the largest absolute value of its factor loading was selected as the variable representing that factor and the factors were labeled as theoretical concepts with respect to the highest loading ratio. The highest loading ratio in factor 1 is eat/assets, designated as profitability. Except for working capital/assets, factor 2 contains leverage ratios. Working capital/assets is more commonly seen as a liquidity ratio, but possibly loading well in this factor anyway, due to the common denominator. As shown in Table 5, the ratio representing this factor is equity/assets. Factor 3 is designated as profit margin, simply due to its representing ratio EBIT/sales, and the interpretation that other ratios in the factor probably covary because of the common denominator. Factor 4 can be interpreted as cash flow normalized by some firm size measure, with cash flow/assets loading as the relatively best representation of the factor. Factor 5 contains liquidity ratios and the representative ratio quick assets/current liabilities is designated as such. Factor 6 is represented by current assets normalized by assets. The co-loading ratios of sales/assets and cash/assets is probably due to the common denominator and the factor is designated current assets based on the highest loading ratio. Factor 7 is designated inventory margin.

For each financial year observation in the data, the selected ratios were also calculated for the two previous financial years, except for cash flow ratios which, requiring two years of data to be calculated, were only calculated for the previous year. After this, any values of approaching infinity was changed to zero. Additionally, firm size was controlled for using the natural logarithm of total assets, in order to ensure that models were discriminating between bankrupt and non-bankrupt firms, instead of smaller and larger firms.

3.6 Experiment design

3.6.1 Resampling

After calculating time dimensional values for all observations in the main sample, records with empty values were dropped, resulting in a sample of 517 and 118,875 financial year observations with and without subsequent bankruptcy, respectively. As mentioned in subsection 2.2.3, dealing with statistically rare events may bias models towards the negative case. To avoid this, bankruptcy prediction models are instead estimated from matched samples (Alaminos et al., 2016; Mselmi et al., 2017; Tsai, 2014). Matched samples are commonplace in fields such as medicine where

Table 5: Designations and oblique factor solution of ratios R1-R25

		Factor loadings						
Ratio	Definition	1	2	3	4	5	6	7
Factor 1: Profitability								
R6	EAT/assets	.946	.025	.001	.003	.032	.027	-.024
R9	EAT/equity	.933	.061	.000	-.008	-.032	-.077	.011
R2	EBIT/assets	.921	.019	.004	.009	.024	.072	-.037
R17	EBIT/equity	.834	.054	.001	-.007	-.036	-.081	.002
R19	Retained earnings/assets	-.680	.604	.004	-.007	-.039	-.058	-.041
Factor 2: Leverage								
R8	Equity/assets	.001	.959	.001	-.016	.006	-.111	-.077
R5	Liabilities/assets	-.023	-.953	-.004	.014	-.048	.162	.101
R4	Working capital/assets	.020	.728	.008	.002	.204	.324	.149
R13	Liabilities/equity	-.028	-.124	-.001	.002	.021	-.087	-.054
Factor 3: Profit margin								
R23	EBIT/sales	.017	.006	.973	-.010	.002	-.006	-.005
R21	Current assets/sales	.005	-.004	-.896	.015	.026	-.009	.101
R14	Working capital/sales	-.007	-.003	.868	.027	.031	-.003	.092
Factor 4: Cash flow								
R15	Cash flow/assets	.004	-.015	-.024	.957	.015	.003	.016
R12	Cash flow/liabilities	.001	-.001	-.027	.828	-.051	-.002	.045
R19	Cash flow/sales	-.009	-.031	.051	.791	-.010	-.033	-.016
Factor 5: Liquidity								
R11	Quick assets/current liabilities	-.005	-.055	.006	-.037	.969	-.060	.083
R1	Current assets/current liabilities	-.008	-.019	.009	-.037	.952	-.073	.185
R20	Cash/current liabilities	.012	.076	-.010	.045	.725	.136	-.306
Factor 6: Current assets								
R7	Current assets/assets	-.024	.112	.011	.007	.054	.837	.230
R3	Sales/assets	-.017	-.381	.012	-.067	-.137	.653	-.076
R16	Cash/assets	.032	.158	-.017	.109	.247	.550	-.444
Factor 7: Inventory margin								
R24	Inventory/sales	-.006	-.005	-.026	-.003	.068	-.036	.786
R18	Sales/accounts receivable	.004	-.003	-.004	-.014	-.008	-.032	-.208
Omitted ratios								
R22	Current liabilities/assets	-	-	-	-	-	-	-
R25	Long-term liabilities/assets	-	-	-	-	-	-	-

Note: EAT=earnings after tax, EBIT=earnings before interest and tax.

control samples are used to determine the effectiveness of drugs when controlling for fixed factors such as gender and age in experiments. In bankruptcy prediction, bankrupt and non-bankrupt firms are likewise matched by similarity in variables such as year, industry and firm size. Appiah et al. (2015) criticizes such matching techniques as being arbitrary and less representative of the main sample.

Therefore, and in order to fulfill the ceteris paribus part of hypotheses, a randomly matched sample technique was used for this experiment when controlling for the fixed effect of time dimensionality. While the most common matching ratio used in the literature is 1:1 between bankrupt and non-bankrupt firms (e.g. du Jardin, 2015), there is also a case to be made for uneven matching ratios, which better represents real world data. Thus, 50 matched samples were created by holding the bankrupt sample of 517 observations constant, matching it 1:3 by randomly drawing 1,551 observations from the non-bankrupt sample for each of the 50 matched samples (cf. Chou et al., 2017). Naturally, being able to also have the bankrupt component of each sample randomly drawn from a larger sample would be preferable. However, the presented approach is analogous to most bankruptcy prediction studies, where the bankrupt sample is often the limiting factor in terms of sampling (see Table 1).

3.6.2 Cross validation

The 50 samples were randomly split into three folds for cross validation as illustrated in Figure 4 (cf. Chuang, 2013; du Jardin, 2010). For each split, two random forest models, one time dimensional and one non-time dimensional, were fitted from two of the folds and performance metrics were calculated by validating the models on the third fold. As an illustrative example, during Split 1 in Figure 4, all observations in Fold 2 and Fold 2 are used to fit one model using three years of accounting data at the firm level. The same observations are then used to fit a second model, but now with only one year of firm level accounting data. Both models are then validated on the unseen observations in Fold 1 in terms of accuracy, recall and precision., and the calculated performance metrics are stored. This process is then repeated for Split 2 and Split 3.

While cross validation can be performed with an arbitrary number of folds, three was chosen to keep the fold sizes comparatively large. The advantage of this approach is that all available bankruptcy data is used in model fitting while also confirming models' predictive performance on new samples. At the same time, two identical sets of bankrupt firms used to estimate models are not likely to occur twice. When accounting for the matched sample of surviving firms, the diversity increases even more, making samples more independent of each other. With 50 samples and 3 folds each, the experiment yielded 150 measurements each of accuracy, recall and precision for each of the two model types, totaling 300 measurements.

3.6.3 Test of between-models effects

Hypotheses were tested through multivariate analysis of variance (MANOVA), with performance metrics as dependent variables and model group as a fixed factor. The mean of each performance metric was compared between the time dimensional random forests, and the non-time dimensional random forests. In order to confirm hypotheses *H1-H3*, it was required that the mean of the respective metric was larger in the time dimensional group compared to the non-time dimensional group, and that the mean difference was statistically significant at $p<0.05$.

3.6 Validity and reliability

Data were collected and matched automatically, which lowers the risk of random errors due to the human factor, but in turn raises the risk of systematic errors due to potentially flawed logic when implementing automatic routines. That is, increasing reliability but potentially lowering validity. To mitigate this, prepared data was manually confirmed to be coherent with each respective data source.

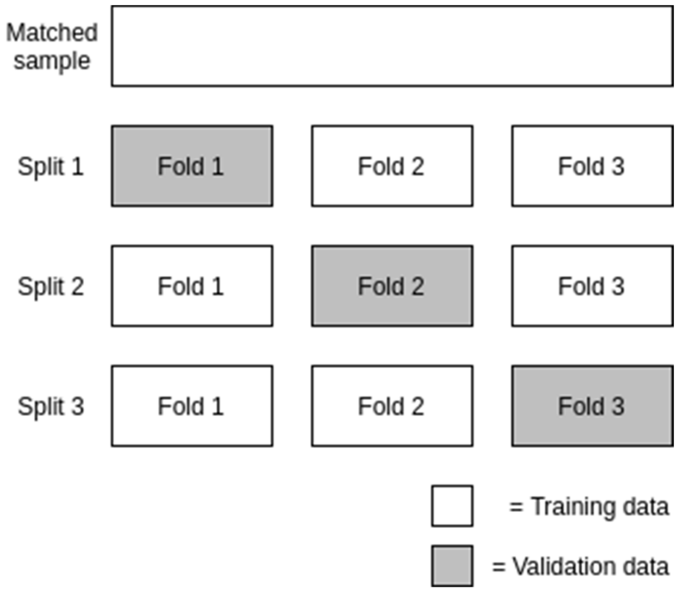

Figure 4: The concept of 3-fold cross validation.

Random forests are, as the name implies, fitted randomly, suggesting that exact reproduction of results is impossible, thereby lowering reliability. However, computers are inherently deterministic and unable of generating true randomness, instead generating only apparently random distributions from a given seed value. Knowing this value makes results reproducible. Hence, seed values were chosen haphazardly and noted.

Another potential reliability issue is that of accounting periods of firms under study not matching calendar years, as mentioned in subsection 3.3.2. However, in assuming that firms most likely do have calendar years as accounting periods, effects stemming from firms closing their accounts prior to the new year should be cancelled by firms closing their accounts after the new year. Related to this, another potential data related problem is that of one month of missing bankruptcy data, lowering the validity of the performed experiment. Missing only approximately 1/12 of the bankruptcies of one year, however, should arguably not lead to significant effects on findings.

Due to ratio selection being empirical, coherency between theoretical definition (extracted factor designation) and operational indicator (chosen ratio) may be lacking. This is part of a wider methodological issue in the bankruptcy prediction literature (Appiah et al., 2015; Balcaen & Ooghe, 2006; Lensberg et al., 2006). At the moment though, empirical evidence from operational research suggest that exploratory variable selection is the most useful method of well-fitting models (Chou et al., 2017; Wang & Wu, 2017; Yeh et al., 2014). Nevertheless, steps were taken to minimize the potential impact on validity by drawing accounting ratios from the predictive bankruptcy literature. In addition, KMO and Bartlett's tests were performed as factor analysis diagnostics, to ensure the sanity of extracted factors (Hair et al., 2014).

Diagnostics were also performed on the estimated MANOVA model using Levene's test for each of the dependent performance metrics (Hair et al., 2014). An assumption implied on the data when performing a valid MANOVA, is equality of error variances. Testing for this assumption then becomes critical in evaluating the robustness of hypothesis tests. For this task, Box's M test was also considered, but

Levene's chosen as the more robust alternative, seeing as Box's test could be too sensitive to assumptions imposed on the data generated by the experiment (Manley, 2004).

3.8 Ethical concerns

In making this thesis, actions were taken to ensure compliance with the ethical guidelines of the Swedish Research Council (2017). The thesis only concerns itself with publicly available data. Thus, the collected data needed not be treated with secrecy. Nevertheless, bankruptcy has the potential of being a negatively life altering event for people involved. Due to this, only data from limited liability companies, as opposed to enterprise forms legally tied to physical people, were considered for analysis. Additionally, no firms nor people (such as CEOs, board members and auditors) are mentioned by name or other identifying information.

4 Findings

4.1 Descriptive statistics and univariate analysis

Relevant descriptive statistics of the extracted factors, as well as univariate analysis of variance between bankrupt and non-bankrupt observations at one, two and three years prior to bankruptcy are shown in Table 6. The control variable firm size, here represented by ln(assets), is also included. According to Table 6, firm size does indeed discriminate bankrupt firms from non-bankrupt firms in univariate analysis, supporting previous assertions of bankruptcy being a more common event in smaller firms. Likewise, the mean values of profitability, cash flow, leverage, and liquidity significantly differ between bankrupt and non-bankrupt firms at *t-1*. This is consistent with the considered accounting ratios in previous studies (Chou et al., 2017; du Jardin, 2015; Liang et al., 2016; Mselmi et al., 2017; Tserng et al., 2014; Wang & Wu, 2017; Yu et al., 2014; Zhou et al., 2014). Profit margin and inventory margin, however, do not discriminate the groups of firms at *t-1*, inconsistent with previous studies (du Jardin, 2015, 2016; Jabeur & Fahmi, 2018; Yu et al., 2014; Zięba et al., 2016).

As also shown in Table 6, the significance value of ratios calculated for *t{-2,-3}* indicate mean differences between bankrupt and non-bankrupt firms for the same ratios as for *t-1*, with the exception of profit margin, which becomes significant at *t-2* and then again insignificant at *t-3*. This is probably due to data snooping; many different comparisons increase the probability of at least one being falsely significant. Thus, it cannot be comfortably stated that the effect exists, especially considering the large difference in p-values in relation to the other two measurements of the same ratio. For the other variables however, the p-values are largely consistent across measurements. Variables showing significant mean differences between firm groups also exhibit smaller mean differences as time to bankruptcy increases, consistent with previous observations of diminishing predictive power of discriminant variables, the earlier predictions are made (Alaminos et al., 2016; du Jardin, 2015; Iturriaga & Sanz, 2015; Tserng et al., 2014).

Table 6: Descriptive statistics and 1-way ANOVA

	Bankrupt		Non-bankrupt			
Factor	Mean	St. dev.	Mean	St. dev.	F	Sig.
One annual account prior to forecasting period (t-1)						
Firm size	9.339	1.267	10.046	1.245	165.867	.000
Profitability	-1.106	21.892	0.064	0.291	327.121	.000
Leverage	-1.058	25.111	0.293	0.381	327.588	.000
Profit margin	-0.505	9.215	-0.015	8.097	1.882	.170
Cash flow	-0.205	4.066	0.000	0.274	148.829	.000
Liquidity	0.887	0.627	1.491	2.226	37.944	.000
Current assets	0.752	0.247	0.731	0.250	3.705	.054
Inventory margin	0.086	0.152	0.068	0.407	1.020	.312
Two annual accounts prior to forecasting period (t-2)						
Profitability	-0.045	0.214	0.065	0.169	219.098	.000
Leverage	0.124	0.365	0.292	0.319	142.429	.000
Profit margin	-0.291	4.755	0.016	3.121	4.952	.026
Cash flow	-0.016	0.129	0.004	0.185	5.858	.016
Liquidity	1.015	1.007	1.465	1.345	57.659	.000
Current assets	0.753	0.237	0.728	0.250	5.069	.024
Inventory margin	0.079	0.123	0.068	0.406	0.389	.533
Three annual accounts prior to forecasting period (t-3)						
Profitability	-0.013	0.160	0.063	0.165	108.718	.000
Leverage	0.155	0.248	0.288	0.259	134.306	.000
Profit margin	-0.071	0.917	-0.016	7.728	0.026	.872
Liquidity	1.017	0.892	1.450	1.328	55.069	.000
Current assets	0.750	0.238	0.724	0.251	5.481	.019
Inventory margin	0.076	0.115	0.077	2.533	0.000	.995

Note: N(Bankrupt)=517, N(Non-bankrupt)=118,875.

4.2 Experiment results and hypothesis tests

Table 7 shows mean performance metrics across 50 iterations of 3-fold cross validation for time dimensional and non-time dimensional random forests, respectively, when fitted from the extracted factors in the previous subsection. Indeed, there are mean differences in all of accuracy, recall, and precision between the two model groups. According to Table 7, time dimensional models perform comparatively better than non-time dimensional models, all else equal. Note that weighted means are not included in Table 7 as is common with MANOVA analysis, since sample sizes are equal, meaning weighted means are equal to true means.

Before testing if the mean differences between model groups are significant, Table 8 shows that the assumption of equality of variances between the groups is fulfilled. Levene statistics are insignificant at $p>0.05$ for all metrics, indicating that the variances are indeed suited for MANOVA. Table 9 shows the estimated 1-way MANOVA test of between-models effects. As can be seen, the intercepts have large F-statistics, suggesting performance metrics of random forest bankruptcy prediction models are better explained by other factors than time dimensionality alone. This is expected given small mean differences shown in Table 7. Coming as no surprise, the small mean difference of between-model total accuracy confirms that time-relevant

Table 7: Mean performance metrics

Metric	Time dimensionality	Mean	St. dev.	N
Accuracy	0	.826	.019	150
	1	.832	.013	150
	Total	.829	.013	300
Recall	0	.520	.035	150
	1	.553	.036	150
	Total	.537	.039	300
Precision	0	.706	.040	150
	1	.712	.038	150
	Total	.709	.039	300

Note: Time dimensionality=0: models estimated from one annual account at the firm level. Time dimensionality=1: models estimated from three annual accounts at the firm level.

Table 8: Levene's test of equality of error variances based on mean

Metric	Levene statistic	df1	df2	Sig.
Accuracy	.006	1	298	.939
Recall	.000	1	298	.994
Precision	.359	1	298	.550

Table 9: 1-way MANOVA test of between-models effects

Source	Dep.	Sum of squares	df	Mean sq.	F	Sig.	Partial eta sq.
Intercept	Accuracy	206.090	1	206.090	1,216,411.691	.000	1.000
	Recall	86.352	1	86.352	69,436.924	.000	.996
	Precision	150.948	1	150.948	97,655.840	.000	.997
Time dimensionality	Accuracy	.003	1	.003	19.085	.000	.060
	Recall	.087	1	.087	69.620	.000	.189
	Precision	.003	1	.003	1.638	.202	.005

accounting information is indeed more useful than historical information for decision making. As suggested by Lukason and Laitinen (2019), this may be due to SMEs in developed countries having short failure processes, wherein few signs of financial distress emerge until just prior to bankruptcy. However, the fixed factors of time dimensionality do have some bearing on model performance. As shown in Table 9, the MANOVA test reveals significant mean differences at $p<0.001$ for the performance metrics accuracy and recall, but not for precision. Accordingly, hypotheses *H1* and *H2* are supported, while *H3* is not supported.

Consistent with the rejection of *H3*, precision is the performance metric least affected by model time dimensionality in the chosen sample, in terms of effect size. As shown in the partial eta squared column of Table 9, time dimensionality only accounts for 0.5 per cent of the variance of precision mean difference. The rejection of *H3* is inconsistent with theory in that identifying firms in temporary financial distress should lead to fewer false positives (Balcaen & Ooghe, 2006). Likewise, it is hypothesized that identifying firms in the early stages of longer failure processes would lead to better precision. One explanation is that a random forest fitted from time dimensional data is not up to these tasks, despite its artificial reasoning; if SME failure processes are predominantly short (Lukason & Laitinen, 2019), it might be more accurate in the end, to classify a temporarily distressed firm as bankrupt.

Another explanation is that of the Swedish legal context. Firms in the end stages of longer failure processes might opt for liquidation instead of bankruptcy. At this point, years of negative profitability have eaten into firms' equity to a point at which they are legally obliged to petition for liquidation, making models misclassify these firms as bankrupt. The rejection of *H3* as opposed to the support of *H1* and *H2* might also be due to methodological issues. Particularly, the chosen method of variable selection might render variables which are inadequate of distinguishing surviving firms. Considering the critique by Powers (2011), the relative increase in true positives might not be enough to tip the scales when using the precision metric, since it does not take the absolute number of true negatives into account.

Accuracy and recall metrics have comparatively larger partial eta squared values at 6.0 and 18.9 per cent, respectively. These statistics suggest that taking into account firms' financial history when predicting bankruptcies may improve classification accuracy mainly because of improved recall, or put in other words; a reduced number of false negatives. The impact of a better recall rate is visualized in Table 10, showing the aggregate classification of all 3-fold verifications performed for each model group over 50 iterations, wherein time dimensional random forests have a relatively larger proportion of correct classifications of bankrupt firms.

By contrast, the proportion of false positives is a bit larger in time dimensional models, perhaps for reasons discussed in the previous paragraphs. The support of *H1* and *H2* is consistent with theory (Balcaen & Ooghe, 2006; Lukason & Laitinen, 2019; Ooghe & de Prijcker, 2008), but inconsistent with previous empirical results; du Jardin (2015) can not show any impact of time dimensionality on classification accuracy at time t. This discrepancy might be due to differences in the economic environment of Swedish and French firms, as well as different legal traditions in bankruptcy law (Kammel, 2008). It might also be due to methodological differences in model type used, or how the prediction period, t, is defined.

Table 11 shows the relative importance of all factors in non-time dimensional and time dimensional random forests, sorted by relative factor importance. The relative importance between factors seems preserved between model groups. Discrepancies between the importance of factors in Table 11 and the univariate analysis in Table 6 can be noted; profit margin is the second most important discriminant in the former, while insignificant in the latter. However, comparison between these two analyses is poor at best. A random forest is non-linear, and the importance of its factors is calculated by the number of observations reaching a given node. By contrast, ANOVA is based on significantly differing mean values between groups. Thus, the only meaningful comparison is that between factors in the random forest models.

Table 10: Aggregate classification table of 3-fold test data

		Bankruptcy predicted	
Time dimensionality	Bankruptcy observed	0	1
0	0	71,946	5,604
	1	12,434	13,416
1	0	71,749	5,801
	1	11,558	14,292

Note: Time dimensionality=0: models estimated from one annual account at the firm level. Time dimensionality=1: models estimated from three annual accounts at the firm level. Bankruptcy observed=0: firm not bankrupt at time t. Bankruptcy observed=1: Firm bankrupt at time t. Bankruptcy predicted=0: Firm predicted as non-bankrupt at time t. Bankruptcy predicted=1: firm predicted as bankrupt at time t.

Consistent with failure process analysis by Lukason and Laitinen (2019), the most important discriminant in Table 11 is profitability. As expected, profitability, profit margin, leverage, and liquidity become less important as the time to bankruptcy increases. At the same time, current assets, cash flow, and inventory margin does not exhibit the same diminishing importance. Curiously, the importance of firm size decreases as multi-year data is included in the models, suggesting that time dimensionality might enable models to partially compensate for the inherent larger risk of bankruptcy in smaller firms.

5 Concluding remarks

5.1 Main conclusions and contributions

The purpose of this thesis is to assess the impact of time dimensionality on bankruptcy prediction performance. The performed experiment with random forests on Swedish SME data suggests that accounting for the time dimension of bankruptcy significantly increases model performance in terms of increased recall, in turn increasing prediction accuracy. At the same time, accounting for the time dimension of bankruptcy does not seem to impact prediction precision significantly, suggesting that it is relatively easier to identify a subsequently bankrupt firm than a subsequently surviving firm using multi-period accounts. However, it cannot be ruled out that this is due to methodological issues in variable selection or inherent flaws in the precision metric itself (Powers, 2011).

This thesis contributes to the emerging firm failure process literature by establishing that accounting for the time dimension of bankruptcy probably raises prediction accuracy, as suggested (Appiah et al., 2015; Balcaen & Ooghe, 2006; Lukason

Table 11: Aggregate relative factor importance in random forests

Time dimensionality	Factor	Relative importance	Normalized rel. importance
0	Profitability $_{t-1}$	30.32	.202
	Profit margin $_{t-1}$	28.49	.190
	Leverage $_{t-1}$	22.52	.150
	Firm size $_{t-1}$	17.29	.115
	Liquidity $_{t-1}$	16.38	.109
	Current assets $_{t-1}$	12.71	.085
	Cash flow $_{t-1}$	12.59	.084
	Inventory margin $_{t-1}$	9.71	.065
1	Profitability $_{t-1}$	18.07	.120
	Profitability $_{t-2}$	10.40	.069
	Profitability $_{t-3}$	6.19	.041
	Profit margin $_{t-1}$	18.05	.120
	Profit margin $_{t-2}$	10.03	.067
	Profit margin $_{t-3}$	5.76	.038
	Leverage $_{t-1}$	11.29	.075
	Leverage $_{t-2}$	7.79	.052
	Leverage $_{t-3}$	6.53	.044
	Firm size $_{t-1}$	7.99	.053
	Liquidity $_{t-1}$	6.10	.041
	Liquidity $_{t-2}$	4.86	.032
	Liquidity $_{t-3}$	4.66	.031
	Current assets $_{t-1}$	4.62	.031
	Current assets $_{t-2}$	4.20	.028
	Current assets $_{t-3}$	4.33	.029
	Cash flow $_{t-1}$	4.91	.033
	Cash flow $_{t-2}$	4.75	.032
	Inventory margin $_{t-1}$	3.22	.021
	Inventory margin $_{t-2}$	3.06	.020
	Inventory margin $_{t-3}$	3.19	.021

Note: Time dimensionality=0: models estimated from one annual account at the firm level. Time dimensionality=1: models estimated from three annual accounts at the firm level. Relative factor importance=the aggregate percentage of tree nodes reached before classification over 150 estimated models. Normalized rel. importance=average percentage of tree nodes reached before classification.

& Laitinen, 2019). Specifically, this also seems to be the case with short predictive horizons. The unique contribution, however, is the analysis on how the time dimension of bankruptcy impacts the performance metrics of recall and precision. Moreover, the literature is extended by fitting time dimensional models in a new economic environment. As any given field of research cannot be summarized by one study alone, the act of doing so is furthering the process of completing the picture of the impact of time dimensionality on prediction performance. Seeing as this niche is comparatively incomplete, the marginal knowledge gain should be relatively large. Furthermore, a problem inherent to all prediction modeling is that of concept drift, where the characteristics of the target variable change over time (Sun et al., 2017). By adding 2011 through 2017 data, this thesis affirms the relevance of previous findings (du Jardin, 2015).

5.2 Implications

Drawing from these findings, the immediate practical implications are that decision makers looking to decrease the rate of false negative classifications may want to consider both recent and older accounting information when predicting bankruptcy or modeling credit risk. These insights should be valuable to any stakeholder making decisions based on financial reports, including creditors such as collection agencies, banks, and suppliers, as well as other stakeholders, such as insurance companies and owners. The findings presented may be specifically interesting to auditors, seeing as they normally make decisions based on the most recent annual accounts (Tagesson & Öhman, 2015).

In turn, the betterment of stakeholders has societal implications. The level of overall economic risk in society decreases, increasing consumption and investments (Lensberg et al., 2006). Eventually, this leads to increased economic growth, lowering unemployment rates, and lifting people out of poverty. Thus, to the ability to predict bankruptcy is another cog in the wheel of society at large, important in order to stabilize the economy through risk aversion and less bankruptcy costs.

As a side note, since the findings presented in this thesis indicate that previous financial performance is indicative of future performance, they challenge the notion that future financial performance can be best understood as a random walk (Fama, 1995; Mishra et al., 2015; Moosa & Vaz, 2015). Therefore, bankruptcy modeling should perhaps not be seen as the probability of bankruptcy as a function of the severity of current financial distress, but rather as a set course along a trajectory that may be possible to deviate from. Granted, such view is controversial, but may have the merit of more accurate predictions.

5.3 Limitations and suggestions for further research

The thesis is limited in terms of some methodological issues highlighted in the literature. Bankruptcy is a rare event in real world data; much rarer than in the matched samples commonly used to fit bankruptcy prediction models. Also, in order to test the effect of time dimensionality in isolation, the performed experiment was cross validated. In order to be useful, models must reasonably perform well in rare events data ex ante, in which model precision is likely to deteriorate, since it will most likely be subjected to a larger number of false positives.

The thesis is also limited in that the used method cannot comfortably be said to identify failure processes as they are described in theory. Failure processes, as conceptualized, have root causes in corporate governance, not in accounting ratios (Amankwah-Amoah, 2016; Crutzen & van Callie, 2008; Ooghe & de Prijcker, 2008). Instead, in aiming to mimic common processes in model estimation by modeling bankruptcy prediction from financial symptoms, accounting ratios were selected empirically. These variables were modeled using non-linear logic. Due to the opacity of models however, the findings do not explain prediction performance in terms of theoretical time dimensional factors such as firm failure processes. Another suggestion for further research then, although challenging, is to identify theoretical failure processes prior to bankruptcy. Including such factors in transparent models could enable a more thorough explanation of time dimensionality's impact on classification performance.

Lastly, and as mentioned, this thesis cannot show that time dimensional models have better prediction performance in terms of prediction precision. It is suggested that this is due to the Swedish legal context or methodological issues stemming from either variable selection or the use of a flawed performance metric. Also, in less developed countries, SMEs follow different failure processes compared to SMEs in more developed countries (Lukason & Laitinen, 2019). The use of time dimensional models in such economic environments may therefore differ in terms of performance. Consequently, the impact of the time dimension of bankruptcy should be assessed in other socio-economical contexts, as well as using other performance metrics.

References

Abellán, J., & Mantas, C. J. (2014). Improving experimental studies about ensembles of classifiers for bankruptcy prediction and credit scoring. *Expert Systems with Applications*, *41*(8), 3825–3830.

Alaka, H. A., Oyedele, L. O., Owolabi, H. A., Kumar, V., Ajayi, S. O., Akinade, O. O., & Bilal, M. (2017). Systematic review of bankruptcy prediction models: Towards a framework for tool selection. *Expert Systems With Applications*, *94*, 164–184.

Alaminos, D., Del Castillo, A., Fernandez, M. A., & Fernández, M. Á. (2016). A global model for bankruptcy prediction. *PLoS ONE*, *11*(11).

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, *23*(4), 589.

Altman, E. I., Iwanicz-Drozdowska, M., Laitinen, E. K., & Suvas, A. (2017). Financial distress prediction in an international context: A review and empirical analysis of Altman's Z-score model. *Journal of International Financial Management and Accounting*, *28*(2), 131–171.

Altman, E. I., & Sabato, G. (2007). Modelling credit risk for SMEs: Evidence from the U.S. market. *Abacus*, *43*(3), 332–357.

Amankwah-Amoah, J. (2016). An integrative process model of organisational failure. *Journal of Business Research*, *69*(9), 3388–3397.

Amankwah-Amoah, J., & Debrah, Y. A. (2014). Air Afrique: The demise of a continental icon. *Business History*, *56*(4), 517–546.

Amendola, A., Giordano, F., Parrella, M. L., & Restaino, M. (2017). Variable selection in high-dimensional regression: A nonparametric procedure for business failure prediction. *Applied Stochastic Models in Business and Industry*, *33*(4), 355–368.

Appiah, K. O., Chizema, A., & Arthur, J. (2015). Predicting corporate failure: A systematic literature review of methodological issues. *International Journal of Law and Management*, *57*(5), 461–485.

Arbnor, I., & Bjerke, B. (2008). *Methodology for creating business knowledge* (3rd ed.). Thousand Oaks: SAGE.

Asar, Y. (2017). Some new methods to solve multicollinearity in logistic regression. *Communications in Statistics-Simulation and Computation*, *46*(4), 2576–2586.

Balcaen, S., & Ooghe, H. (2006). 35 years of studies on business failure: An overview of the classic statistical methodologies and their related problems. *British Accounting Review*, *38*(1), 63–93.

Barboza, F., Kimura, H., & Altman, E. I. (2017). Machine learning models and bankruptcy prediction. *Expert Systems with Applications*, *83*, 405–417.

Beaver, W. H. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research*, *4*, 71–111.

Berg, D. (2007). Bankruptcy prediction by generalized additive models. *Applied Stochastic Models in Business and Industry*, *23*, 129–143.

Betz, F., Oprică, S., Peltonen, T. A., & Sarlin, P. (2014). Predicting distress in European banks. *Journal of Banking and Finance*, *45*(1), 225–241.

Braga, R. N. (2017). Effects of IFRS adoption on tax avoidance. *Revista Contabilidade & Finanças*, *28*(75), 407–424.

Branch, B. (2002). The costs of bankruptcy: A review. *International Review of Financial Analysis*, *11*(1), 39–57.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32.

Calabrese, R. (2014). Optimal cut-off for rare events and unbalanced misclassification costs. *Journal of Applied Statistics*, *41*(8), 1678–1693.

Chen, J., Marshall, B. R., Zhang, J., & Ganesh, S. (2006). Financial distress prediction in China. *Review of Pacific Basin Financial Markets and Policies*, *09*(02), 317–336.

Chou, C. H., Hsieh, S. C., & Qiu, C. J. (2017). Hybrid genetic algorithm and fuzzy clustering for bankruptcy prediction. *Applied Soft Computing Journal*, *56*, 298–316.

Chuang, C. L. (2013). Application of hybrid case-based reasoning for enhanced performance in bankruptcy prediction. *Information Sciences*, *236*, 174–185.

Cleary, S., & Hebb, G. (2016). An efficient and functional model for predicting bank distress: In and out of sample evidence. *Journal of Banking and Finance*, *64*, 101–111.

Climent, F., Momparler, A., & Carmona, P. (2018). Anticipating bank distress in the Eurozone: An Extreme Gradient Boosting approach. *Journal of Business Research*.

Corbin, J. M., & Strauss, A. L. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (3rd ed.). Thousand Oaks: SAGE.

Cox, D. R., & Reid, N. (2000). *The theory of the design of experiments* (1st ed.). London: Chapman & Hall/CRC.

Crutzen, N., & van Callie, D. (2008). The business failure process. An integrative model of the literature. *Review of Business and Economics*, *53*(3), 287–316.

Cultrera, L., & Brédart, X. (2016). Bankruptcy prediction: The case of Belgian SMEs. *Review of Accounting and Finance*, *15*(1), 101–119.

De George, E. T., Li, X., & Shivakumar, L. (2016). A review of the IFRS adoption literature. *Review of Accounting Studies*, *21*(3), 898–1004.

Doumpos, M., Andriosopoulos, K., Galariotis, E., Makridou, G., & Zopounidis, C. (2017). Corporate failure prediction in the European energy sector: A multicriteria approach and the effect of country characteristics. *European Journal of Operational Research*, *262*(1), 347–360.

du Jardin, P. (2010). Predicting bankruptcy using neural networks and other classification methods: The influence of variable selection techniques on model accuracy. *Neurocomputing*, *73*(10–12), 2047–2060.

du Jardin, P. (2015). Bankruptcy prediction using terminal failure processes. *European Journal of Operational Research*, *242*, 286–303.

du Jardin, P. (2016). A two-stage classification technique for bankruptcy prediction. *European Journal of Operational Research*, *254*, 236–252.

du Jardin, P., & Séverin, E. (2011). Predicting corporate bankruptcy using a self-organizing map: An empirical study to improve the forecasting horizon of a financial failure model. *Decision Support Systems*, *51*(3), 701–711.

du Jardin, P., & Séverin, E. (2012). Forecasting financial failure using a Kohonen map: A comparative study to improve model stability over time. *European Journal of Operational Research*, *221*(2), 378–396.

Erkens, D. H., Hung, M., & Matos, P. (2012). Corporate governance in the 2007-2008 financial crisis: Evidence from financial institutions worldwide. *Journal of Corporate Finance*, *18*(2), 389–411.

European Commission. (2017, August). What is an SME? Retrieved March 19, 2019, from Internal Market, Industry, Entrepreneurship and SMEs - European Commission website: https://ec.europa.eu/growth/smes/business-friendly-environment/sme-definition_en

Fama, E. F. (1995). Random walks in stock market prices. *Financial Analysts Journal*, *51*(1), 75–80.

García, V., Marqués, A. I., & Sánchez, J. S. (2019). Exploring the synergetic effects of sample types on the performance of ensembles for credit risk and corporate bankruptcy prediction. *Information Fusion*, *47*, 88–101.

Gavurova, B., Packova, M., Misankova, M., & Smrcka, L. (2017). Predictive potential and risks of selected bankruptcy prediction models in the Slovak business environment. *Journal of Business Economics and Management*, *18*(6), 1156–1173.

Geng, R., Bose, I., & Chen, X. (2015). Prediction of financial distress: An empirical study of listed Chinese companies using data mining. *European Journal of Operational Research*, *241*(1), 236–247.

Gepp, A., & Kumar, K. (2015). Predicting financial distress: A comparison of survival analysis and decision tree techniques. *Procedia Computer Science*, *54*, 396–404.

Gordini, N. (2014). A genetic algorithm approach for SMEs bankruptcy prediction: Empirical evidence from Italy. *Expert Systems with Applications*, *41*(14), 6433–6445.

Gordon, E. A., Henry, E., Jorgensen, B. N., & Linthicum, C. L. (2017). Flexibility in cash-flow classification under IFRS: Determinants and consequences. *Review of Accounting Studies*, *22*(2), 839–872.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). Exploratory factor analysis. In *Multivariate Data Analysis* (pp. 89–150). Harlow: Pearson.

Hamilton, E. A. (2006). An exploration of the relationship between loss of legitimacy and the sudden death of organizations. *Group and Organization Management*, *31*(3), 327–358.

Heo, J., & Yang, J. Y. (2014). AdaBoost based bankruptcy forecasting of Korean construction companies. *Applied Soft Computing*, *24*, 494–499.

Houqe, M. N., Monem, R. M., Tareq, M., & Van Zijl, T. (2016). Secrecy and the impact of mandatory IFRS adoption on earnings quality in Europe. *Pacific-Basin Finance Journal*, *40*, 476–490.

Huang, C., Farewell, D., & Pan, J. (2017). A calibration method for non-positive definite covariance matrix in multivariate data analysis. *Journal of Multivariate Analysis*, *157*, 45–52.

Huang, J., Wang, H., & Kochenberger, G. (2017). Distressed Chinese firm prediction with discretized data. *Management Decision*, *55*(5), 786–807.

IASB. (1997). *IAS 1 Presentation of financial statements*. International Accounting Standards Board.

Iturriaga, F. J. L., & Sanz, I. P. (2015). Bankruptcy visualization and prediction using neural networks: A study of U.S. commercial banks. *Expert Systems with Applications*, *42*(6), 2857–2869.

Jabeur, S. Ben, & Fahmi, Y. (2018). Forecasting financial distress for French firms: a comparative study. *Empirical Economics*, *54*(3), 1173–1186.

Jones, S., Johnstone, D., & Wilson, R. (2017). Predicting corporate bankruptcy: An evaluation of alternative statistical frameworks. *Journal of Business Finance and Accounting*, *44*(1–2), 3–34.

Kammel, A. J. (2008). The law and economics of corporate insolvency – Some thoughts. In P. J. Omar (Ed.), *International insolvency law: Themes and perspectives* (1st ed., pp. 61–83). Aldershot: Ashgate.

Khademolqorani, S., Zeinal Hamadani, A., & Mokhatab Rafiei, F. (2015). A hybrid analysis approach to improve financial distress forecasting: Empirical evidence from Iran. *Mathematical Problems in Engineering*, *2015*, 1–9.

Kicinski, M., Springate, D. A., & Kontopantelis, E. (2015). Publication bias in meta-analyses from the Cochrane Database of Systematic Review. *Statistics in Medicine*, *34*(20), 2781–2793.

Kim, J., Jung, S., & Ha, M. (n.d.). Consolidated financial reporting and the effect of mandatory IFRS adoption on the information content of earnings announcements: evidence from Korea. *Asia-Pacific Journal of Accounting & Economics*.

Kim, M. H. Y., Ma, S., & Zhou, Y. A. (2016). Survival prediction of distressed firms: evidence from the Chinese special treatment firms. *Journal of the Asia Pacific Economy*, *21*(3), 418–443.

Kosmidis, K., & Stavropoulos, A. (2014). Corporate failure diagnosis in SMEs: A longitudinal analysis based on alternative prediction models. *International Journal of Accounting and Information Management*, *22*(1), 49–67.

Kutyłowska, M. (2015). Neural network approach for failure rate prediction. *Engineering Failure Analysis*, *47*(PA), 41–48.

Lai, K. W. (2009). Audit opinion and disclosure of audit fees. *Journal of Accounting, Auditing and Finance*, *24*(1), 91–114.

Laitinen, E. K. (2005). Survival analysis and financial distress prediction: Finnish evidence. *Review of Accounting and Finance*, *4*(4), 76–90.

Laitinen, E. K., & Lukason, O. (2014). Do firm failure processes differ across countries: Evidence from Finland and Estonia. *Journal of Business Economics and Management*, *15*(5), 810–832.

Lakshan, A. M. I., & Wijekoon, W. M. H. N. (2012). Corporate governance and corporate failure. *Procedia Economics and Finance*, *2*, 191–198.

Law, J. (Ed.). (2018). Financial distress. In *A Dictionary of Finance and Banking* (6th ed.). Oxford: Oxford University Press.

Le, H. H., & Viviani, J. L. (2018). Predicting bank failure: An improvement by implementing a machine-learning approach to classical financial ratios. *Research in International Business and Finance*, *44*, 16–25.

Lennox, C. (1999). Identifying failing companies: a re-evaluation of the logit, probit and DA approaches. *Journal of Economics and Business*, *51*(4), 347–364.

Lensberg, T., Eilifsen, A., & McKee, T. E. (2006). Bankruptcy theory development and classification via genetic programming. *European Journal of Operational Research*, *169*(2), 677–697.

Liang, D., Lu, C.-C. C., Tsai, C.-F. F., & Shih, G.-A. A. (2016). Financial ratios and corporate governance indicators in bankruptcy prediction: A comprehensive study. *European Journal of Operational Research*, *252*(2), 561–572.

Liang, D., Tsai, C. F., & Wu, H. T. (2015). The effect of feature selection on financial distress prediction. *Knowledge-Based Systems*, *73*(1), 289–297.

Lukason, O., & Laitinen, E. K. (2016). Failure processes of old manufacturing firms in different European countries. *Investment Management and Financial Innovations*, *13*(2), 310–321.

Lukason, O., & Laitinen, E. K. (2018). Failure of exporting and non-exporting firms: do the financial predictors vary? *Review of International Business and Strategy*, *28*(3–4), 317–330.

Lukason, O., & Laitinen, E. K. (2019). Firm failure processes and components of failure risk: An analysis of European bankrupt firms. *Journal of Business Research*, *98*, 380–390.

Manley, B. F. J. (2004). Tests of significance with multivariate data. In *Multivariate statistical methods: A primer* (3rd ed., pp. 35–58). London: Chapman & Hall/CRC.

Martin, D. (1977). Early warning of bank failure: A logit regression approach. *Journal of Banking and Finance*, *1*, 249–276.

Mishra, A., Mishra, V., & Smyth, R. (2015). The random-walk hypothesis on the Indian stock market. *Emerging Markets Finance and Trade*, *51*(5), 879–892.

Moosa, I., & Vaz, J. (2015). Why is it so difficult to outperform the random walk? An application of the Meese–Rogoff puzzle to stock prices. *Applied Economics*, *47*(4), 398–407.

Mselmi, N., Lahiani, A., & Hamza, T. (2017). Financial distress prediction: The case of French small and medium-sized firms. *International Review of Financial Analysis*, *50*, 67–80.

Nehrebecka, N. (2018). Predicting the default risk of companies. Comparison of credit scoring models: Logit vs support vector machines. *Econometrics*, *22*(2), 54–73.

Nummela, N., Saarenketo, S., & Loane, S. (2016). The dynamics of failure in international new ventures: A case study of Finnish and Irish software companies. *International Small Business Journal: Researching Entrepreneurship*, *34*(1), 51–69.

Odom, M. D. D., & Sharda, R. (1990). A neural network model for bankruptcy prediction. In *1990 IJCNN International Joint Conference on Neural Networks* (pp. 163–168). New Jersey: IEEE.

Öhman, P., Häckner, E., Jansson, A. M., & Tschudi, F. (2006). Swedish auditors' view of auditing: Doing things right versus doing the right things. *European Accounting Review*, *15*(1), 89–114.

Öhman, P., & Yazdanfar, D. (2017). Short- and long-term debt determinants in Swedish SMEs. *Review of Accounting and Finance*, *16*(1), 106–124.

Ooghe, H., & de Prijcker, S. (2008). Failure processes and causes of company bankruptcy: A typology. *Management Decision*, *46*(2), 223–242.

Peirce, C. S. (1992). *The essential Peirce, volume 1: Selected philosophical writings* (N. Houser & C. J. W. Kloesel, Eds.). Bloomington: Indiana University Press.

Pompe, P. P. M. M., & Bilderbeek, J. (2005). The prediction of bankruptcy of small- and medium-sized industrial firms. *Journal of Business Venturing*, *20*(6), 847–868.

Powers, D. M. W. (2011). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness & correlation. *Journal of Machine Learning Technologies*, *2*(1), 37–63.

Press, S. J., & Wilson, S. (1978). Choosing between logistic regression and discriminant analysis. *Journal of the American Statistical Association*, *73*(364), 699–705.

Quintiliani, A. (2017). Financial distress cost of Italian small and medium enterprises. *The Journal of Risk Finance*, *18*(5), 564–580.

Revilla, A. J., Pérez-Luño, A., & Nieto, M. J. (2016). Does family involvement in management reduce the risk of business failure? The moderating role of entrepreneurial orientation. *Family Business Review*, *29*(4), 365–379.

Safavian, S. R., & Landgrebe, D. (1991). A survey of decision tree classifier methodology. *IEEE Transactions on Systems, Man, and Cybernetics*, *21*(3), 660–674.

Scherger, V., Terceño, A., & Vigier, H. (2017). The OWA distance operator and its application in business failure. *Kybernetes*, *46*(1), 114–130.

Serrano-Cinca, C., Gutiérrez-Nieto, B., & Bernate-Valbuena, M. (2018). The use of accounting anomalies indicators to predict business failure. *European Management Journal*.

Shie, F. S., & Chen, M.-Y. (2012). Prediction of corporate financial distress: An application of the America banking industry. *Neural Computing and Applications*, *21*(7), 1687–1696.

Statistics Sweden. (2019). Index - Beräkning. Retrieved March 24, 2019, from Företagsregistret website: https://www.foretagsregistret.scb.se/Rakna?RakneTyp=RAKNA_FORETAG

Sun, J., Fujita, H., Chen, P., & Li, H. (2017). Dynamic financial distress prediction with concept drift based on time weighting combined with Adaboost support vector machine ensemble. *Knowledge-Based Systems*, *120*, 4–14.

Swedish National Tax Board. (2019). *Aktiebolagslagen | Rättslig rådgivning | Skatteverket*.

Swedish Research Council. (2017). *Good research practice*. Stockholm: Swedish Research Council.

Tagesson, T., & Öhman, P. (2015). To be or not to be - Auditors' ability to signal going concern problems. *Journal of Accounting and Organizational Change*, *11*(2), 175–192.

Trahms, C. A., Ndofor, H. A., & Sirmon, D. G. (2013). Organizational decline and turnaround: A review and agenda for future research. *Journal of Management*, *39*(5), 1277–1307.

Tsai, C. F. (2014). Combining cluster analysis with classifier ensembles to predict financial distress. *Information Fusion*, *16*(1), 46–58.

Tsai, C. F., Hsu, Y. F., & Yen, D. C. (2014). A comparative study of classifier ensembles for bankruptcy prediction. *Applied Soft Computing*, *24*, 977–984.

Tserng, H. P., Chen, P.-C., Huang, W.-H., Lei, M. C., & Tran, Q. H. (2014). Prediction of default probability for construction firms using the logit model. *Journal of Civil Engineering and Management*, *20*(2), 247–255.

Van Peursem, K., & Chan, Y. C. (2014). Forecasting New Zealand corporate failures 2001-10: Opportunity lost? *Australian Accounting Review*, *24*(3), 276–288.

Veganzones, D., & Séverin, E. (2018). An investigation of bankruptcy prediction in imbalanced datasets. *Decision Support Systems*, *112*, 111–124.

Virág, M., & Nyitrai, T. (2014). Is there a trade-off between the predictive power and the interpretability of bankruptcy models? The case of the first Hungarian bankruptcy prediction model. *Acta Oeconomica*, *64*(4), 419–440.

Wang, G., Ma, J., & Yang, S. (2014). An improved boosting based on feature selection for corporate bankruptcy prediction. *Expert Systems with Applications*, *41*, 2353–2361.

Wang, L., & Wu, C. (2017). Business failure prediction based on two-stage selective ensemble with manifold learning algorithm and kernel-based fuzzy self-organizing map. *Knowledge-Based Systems*, *121*, 99–110.

Yazdanfar, D., & Öhman, P. (2018). Growth and job creation at the firm level: Swedish SME data. *Management Research Review*, *41*(3), 345–358.

Yeh, C. C., Chi, D. J., & Lin, Y. R. (2014). Going-concern prediction using hybrid random forests and rough set approach. *Information Sciences*, *254*, 98–110.

Yeh, Y. M. C., Yao, W.-R., & Chin, C. L. (2019). Do local leads deliver contracting benefits? Evidence from emerging market syndicated loans. *Asia-Pacific Journal of Accounting & Economics*, *26*(3), 301–319.

Yu, Q., Miche, Y., Séverin, E., & Lendasse, A. (2014). Bankruptcy prediction using extreme learning machine and financial expertise. *Neurocomputing*, *128*, 296–302.

Zhou, L., Keung, K., & Yen, J. (2014). Bankruptcy prediction using SVM models with a new approach to combine features selection and parameter optimisation. *International Journal of Systems Science*, *45*(3), 241–253.

Zięba, M., Tomczak, S. K., & Tomczak, J. M. (2016). Ensemble boosted trees with synthetic features generation in application to bankruptcy prediction. *Expert Systems with Applications*, *58*, 93–101.