Norwegian School of Economics
Bergen, Spring 2022

The Norwegian Covid-19 Compensation Scheme

To what extent was compensation distributed in line with the objective of firm viability?

Eivind Emil Andersen and Håvard Legernæs

Supervisors: Sissel Jensen and Hans Jarle Kind

Master of Science in Economics and Business Administration,
Financial Economics & Strategy and Management

NORWEGIAN SCHOOL OF ECONOMICS

Executive Summary

This thesis aims to examine the relationship between bankruptcy risk and received compensation. The Erna Solberg Government enacted the Norwegian compensation scheme to save otherwise viable firms during the Covid-19 pandemic. The compensation scheme was of great political importance and acquired significant public attention through the continuous spotlight of the media. While most focused on changes in bankruptcy frequency and size of compensation, we argue that the distribution within the scheme warrants further attention. In line with literature on resource allocation and forbearance lending, we identify that granting compensation to firms at risk of bankruptcy can distort market mechanisms long-term. Presumably for the same reason, a primary target of the compensation scheme was to exclude unviable firms. Despite the possibility for distribution inefficiencies, there has hardly been any effort to assess the scheme. Our response was to evaluate if compensation was granted in line with the objective of firm viability. We apply bankruptcy prediction to measure firm viability and limit the thesis to the Norwegian compensation scheme during its two first iterations in 2020.

We conduct our research using two stages of methodology. Firstly, we use the random forest machine learning algorithm to estimate the likelihood of near-future bankruptcy prior to the outbreak of Covid-19. The predicted risk of bankruptcy is then used to analyze the trend in distribution of size-adjusted compensation. Our scope is limited to the hospitality industry to achieve increased resolution on within-industry differences.

Our analysis reveals that predicted bankrupt firms received an estimated 54.3 million NOK in compensation. We identify a weak but statistically significant positive relationship between size-adjusted compensation and the predicted risk of bankruptcy. Our thesis concludes that compensation at the aggregate level primarily was distributed in line with the compensation scheme objective of viability. However, compensation distribution within the scheme was moderately asymmetric in that the support was proportionally greater for firms with an elevated probability of bankruptcy.

Acknowledgments

This thesis was written as part of our master's degree in Economics and Business Administration in the spring of 2022.

The Covid pandemic has greatly affected our time at NHH the same as everyone else. However, we fully acknowledge that the extraordinary situation required strict lockdown measures. While we understand the justification for lockdown, throughout the pandemic and our master's degree, we have intensely discussed the rationality of the economic pandemic policies. This discussion has been fueled by a fragmented public debate, lacking empirical research, and inadequate attention to the long-term effects of current policy. With our interest ignited, we wanted to seize the opportunity to answer some of the questions raised in the public debate.

We wish to thank the Centre for Applied Research at NHH for providing data, and our supervisors, Sissel Jensen and Hans Jarle Kind, for providing valuable input and advice. Furthermore, we also extend a thanks to our friends and family who have shown enthusiasm for our thesis topic. Moreover, we also wish to thank The Norwegian Tax Administration for the opportunity to present and discuss our thesis.

Bergen, June 2022

Eivind Emil Andersen

Håvard Legernæs

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1 Covid-19 Pandemic

In the first quarter of 2020, the world was broadly impacted by the spread of the Covid-19 virus. The disease was reported as highly contagious and symptomatic, justifying the WHO to declare it a "Public Health Emergency of International Concern" (Regjeringen, 2022). Governments, including the Norwegian, feared the virological consequences and hastily introduced lockdown measures at the time of detection within their borders. In Norway, this happened on March 12th, 2020. The efforts at the time entailed the lockdown of institutions such as schools, kindergartens, and universities, with the overall effect of limiting movement and social contact (Solberg, 2020). The measures intensified during the following days, as travel restrictions and border control began on March 14th and quarantine and isolation regulations on March 15th (Regjeringen, 2022). Within a few days, lockdown measures included restrictions on international travel, domestic travel, private socializing, and dining, reinforcing the voluntary reductions in these activities. The implementation of lockdown measures had sound argumentation but was expected to plunge economic activity and cause a drastic increase in bankruptcies.

For these reasons, governments around Europe enacted several temporary compensating policies to counter the adverse effects. The primary distinctions in the policies lay in how and what they compensated. For instance, some schemes targeted fixed costs or wage expenses, and loans varied in the share guaranteed by the state. Inevitably, funds were distributed to an array of firms with varying productivity, strategies, physical location, and risk profiles (Altomonte, Demertzis, Fontagné, & Müller, 2021). Furthermore, the magnitude of the policies throughout Europe has varied drastically, measured by the expense as a share of GDP. Countries like Italy, Japan, and Germany range between 40%-45%, compared to Nordic countries at around 10% (NOU, 2021:4, p. 23). The Norwegian pandemic policy approach was split into several independent schemes. Liquidity and employee retention were the primary objectives of the policies collectively (Finans Norge, 2020). Firstly, a temporary layoff scheme (permitteringsordningen) was enacted on the 19th of March, compensating employers for temporary breaks in employment, thereby improving their short-term liquidity during shutdowns (Deloitte, 2020). Secondly, firms were granted a postponement of taxes and fee payments by the Norwegian Tax Administration to increase liquidity. In June 2020, a refined solution was enacted, which granted the postponement of most taxes and fees with monthly installments for six months (NOU, 2021:4). As a part of this policy, the Norwegian Tax Administration could not declare bankruptcies if postponement was granted. Accordingly, the tax authority declared 50% fewer bankruptcies in 2020 compared to 2019 (Skatteetaten, 2021). Thirdly, state-guaranteed loans were introduced 27th of March to ease access to capital for Norwegian enterprises (NOU, 2021:4, p. 113). Fourthly, the most prominent policy in the public debate became the compensation scheme known as "Kompensasjonsordningen." It was intended to compensate firms for fixed unavoidable costs to ease the financial burden of temporary shutdowns.[1]

Among all the schemes, the compensation scheme for fixed costs (hereafter referred to as the compensation scheme) is the only scheme representing a cash transfer with no purpose of repayment.[2] After reviewing the public debate, it also appears to be the most criticized. As Grytten explains (Johnsen, 2020), the compensation scheme is "very generous." Instead of promoting production in the economy, he claims that it incentivizes further closure. Credit analyst Per Einar Ruud, argues that the 50% decline in bankruptcies within the hospitality industry (as of May 2021) was likely fueled by the compensation scheme (Dun & Bradstreet, 2021). We argue that the cash transfer nature of the compensation scheme distinguishes it from the other policy measures' long-term consequences. Moreover, it has been a prominent topic of media attention, and data on the compensation scheme is publicly and readily available, in contrast to the other schemes. For these reasons, we find it appropriate to limit the paper to the compensation scheme.

[1] We define shutdowns as the mandatory or voluntary closing of business due to pandemic restrictions on the affected activity or a general lack of demand.

[2] In contrast to the compensation scheme, we classify the layoff scheme as a reduction of labor costs rather than a cash transfer. While it improves short-term liquidity, it is primarily relevant for the first period of shutdown and more limited in duration. Postponements of taxes and fees, and state guaranted loans are both transfers with a purpose of repayment.

1.2 The Compensation Scheme

The Norwegian Tax Administration (2020) states that the scheme's purpose is to prevent otherwise viable firms from going bankrupt because of the outbreak of the covid pandemic. This way, they argue, the risk of mass unemployment is reduced, resignations are avoided, and the recovery of economic activity is quicker when the temporary crisis is over. The distinction of firm viability is central since direct interpretation means unviable firms are undesired recipients of the compensation scheme.[3] Despite this distinction, the tax authority is unclear in the term's meaning. Based on our interpretation of their statements, firm viability refers to a situation in which a business is surviving. Necessarily, unviable firms should therefore be close to the state of bankruptcy. Therefore, the natural interpretation is that compensation to firms with a high risk of bankruptcy is undesirable.

According to Sticos' advisor Kværnmo (2020), the process of granting compensation was partially automated, as only large deviations between registries mandated manual inspection. The compensation formula is the product of the relative loss in revenue, the sum of fixed costs, and an adjustment factor for a given application period.[4] Variable costs follow activity, meaning that compensating fixed costs would allow more firms to survive without income. In simpler terms, a cash transfer increases firms' liquidity, reducing the likelihood of insolvency and bankruptcy. The comparison period for calculating the loss in revenue is the same month(s) in 2019. If the firm did not exist one year prior, they could use the revenue for January and February 2020 instead. The general formula for granted compensation is summarized below.

$$Granted\ compensation = (Loss\ of\ income\ in\ \%)\ x\ (Fixed\ Costs)\ x\ (adjustment\ factor) \qquad \textit{Eq. 1}$$

As criteria for compensation in the first iteration, being active from March to August 2020, firms were required to document a loss in income equivalent to 30% (20% for March 2020)

[3] We refer to viability both as a scale from low to high, as well as the binary state of being viable depending on the context.

[4] Cost items in the income statement corresponding to the class of fixed costs were provided to assist in defining the group. The general rule was that the cost must be unavoidable (Kvernmo, 2020)

of the same month in 2019, adjusted for growth between 2019 and 2020.[5] A distinction was made between mandatory and voluntary shutdowns, and the former qualified for a more lenient adjustment factor. The deductible was set to ten thousand NOK for March, five thousand for April, and then completely removed. The adjustment factor for the first three months was 0.9 and 0.8, which went to 0.7 for June and July and further reduced to 0.5 for August. In the case of past firm deficits, the compensation would be reduced so that the average monthly deficit was subtracted from the sum of fixed costs in the compensation formulas. The sum of granted compensation could also not exceed the loss of income (Kvernmo, 2020). In the second iteration from September 2020, the criteria changed to make no distinction between mandatory and voluntary shutdowns.[6] Correcting for growth between 2019 and 2020 was also completely removed. Additionally, the adjustment factor was changed to 0.7 for September and October and 0.85 for November to December in 2020. Deficits were still punished the same way as in the first iteration (Hamnes, 2021). While the iterations of the compensation scheme differ in some ways, their primary criteria and structure remain the same.

The official purpose of the compensation scheme indicates that firms' viability would be considered (Skatteetaten, 2020). However, there is little evidence of such a distinction in the criteria for compensation, apart from adjusting for deficits. This is discussed in the Norwegian Government report on the post-Covid economy, which claims the compensation scheme lacked criteria to distinguish between viable and non-viable firms (NOU, 2021:4, p. 48). Subsequently, compensation has likely been delegated to and contributed to the survival of otherwise unviable firms (NOU, 2021:4, p. 8). The same effects are already established for similar schemes in Europe. For instance, Altomonte et al. (2021) find that policies disproportionately benefitted unproductive firms in Italy and Germany. Moreover, they conclude that the design of the schemes matters in avoiding this outcome, which serves as a

[5] **Formula for granted compensation in 1st Iteration** (Kvernmo, 2020)**:** Mandatory shutdown: (Loss of income in %) x (Fixed Costs) x 0,9 & Voluntary shutdown: (Loss of income in %) x (Fixed Costs-deductible) x 0,8

[6] **Formula for granted compensation in 2nd Iteration** (Hamnes, 2021)**:** (Loss of income in %) x (Fixed Costs) x (adjustment factor)

motivation in developing effective schemes. We believe the consequences of misallocation of compensation, if present, could have the primary short-term effect of either wasting public finances or facing the opportunity cost of granting compensation to a more suitable firm. In the long term, we hypothesize that this misallocation can skew the distribution of capital and employees towards less viable firms, limiting economic efficiency. For these reasons, we foresee the scheme's efficiency would benefit greatly from accurately distinguishing viable and unviable firms before granting compensation.

As stated previously, we interpret viability as survivability. It is possible to define the term by measures of productivity. However, bankruptcy has advantages as the viability indicator for its data availability, relatively greater objectivity, and binary nature. This leads us to the field of bankruptcy prediction and methods of classification.[7] The field is usually focused on financial data and ratios. The usefulness of financial data comes from the comprehensive list of variables, the availability of the data, and the signaling of future performance based on past performance. Within this field, high-performing models are necessary since correct classification is difficult to achieve. A wrongful assessment can lead to inaccurate results and misleading conclusions. For this reason, the exploration of multiple models and machine learning can be highly beneficial. We therefore determine that bankruptcy prediction will provide the most accurate quantifiable measure for firm viability, is flexible in results, and allows for discussing the costs of misclassification. As a final point, we argue that a comprehensive methodology will strengthen the robustness of our findings.

1.3 Research Aim

Several compensation policies were implemented to counter the adverse economic effects of lockdown throughout Europe. Among the Norwegian policy response, we assess that the fixed-cost compensation scheme represents the most relevant and feasible motive for our thesis. However, despite its importance and public attention, we have been unsuccessful in identifying research addressing the compensation scheme's distribution. Accordingly, we aim

[7] Section 3.4 describes classification models further. In short, a classification model predicts target variables where the outcome is categorical (e.g bankrupt and non-bankrupt), as opposed to continuous data.

to address this gap and examine the distribution of the compensation scheme with respect to the viability criteria. To enable this evaluation, we apply the field of bankruptcy prediction to quantify firm viability.[8] Furthermore, we limit the compensation scheme to the first two iterations of 2020. For the reasons stated, we narrow our research aim to the following research question:

How can bankruptcy prediction be used to evaluate the Norwegian compensation scheme, and to what degree was compensation distributed in line with the viability objective?

Firstly, we develop a model to predict the likelihood of bankruptcy. The predicted probability among the firms that received compensation is then used to estimate the direct misallocation of funds, found as the sum of compensation granted to predicted bankrupt firms. The purpose is to quantify the magnitude of misallocation.

Secondly, we examine the distribution of compensation for all firms within the scheme to determine if compensation unequally benefited firms with high or low risk of bankruptcy. Therefore, we analyze the relationship between bankruptcy risk and the compensation to revenue ratio, which we refer to as one measure of compensation intensity.[9] We expect a negative relationship between bankruptcy risk and compensation intensity because of our interpretation of the scheme´s objective to exclude unviable firms and the deficit criteria mentioned in section 1.2.

Thirdly, to complement the analysis of compensation to revenue, we perform the same analysis with compensation to employees and labor costs, the other measure of compensation intensity. The dual inspection is performed since the outcome of one relationship can be rationalized through the other. For example, if a high-risk firm has a large workforce, distributing sufficient compensation to save the firm could justify a high compensation intensity. Note that our evaluation compares outcomes against objectives, making the employment issue relevant even though it was not included as a compensation criterion as explained in section 1.2.

[8] We assume and argue that estimated bankruptcy risk pre-pandemic is a suitable indicator before and after the pandemic. While the pandemic changes theoutlook for firms, the bankruptcy risk ex-ante is still the best prior estimate available.

[9] Ratio of compensation over either revenue or labor costs. Inspired by “loan intensity” from Altomonte et al. (2021, p. 7)

1.4 Choice of Industry

Because we aim to analyze compensation intensity across firms, we prefer these firms to be reasonably similar. If they differ considerably, the analysis becomes blurred because there are large variations in bankruptcy frequency, cost structures, profitability, and pandemic impact between industries. For instance, to inspect differences in the compensation to revenue ratio, we require a certain level of homogeneity in the share of fixed costs, as that directly affects the sum granted.

Among others, the hospitality industry saw a sharp decline in sales and economic activity due to pandemic risk and the measures enacted by the government.[10] According to Altomonte et al. (2021), it was the hardest impacted industry in Europe. Therefore, one would expect that the hospitality industry in Norway, consisting of sectors 55 and 56, is among the largest compensation recipients. The distribution for different sectors is shown in Figure 1 and Table 1 (Brønnøysundregistrene, 2022).

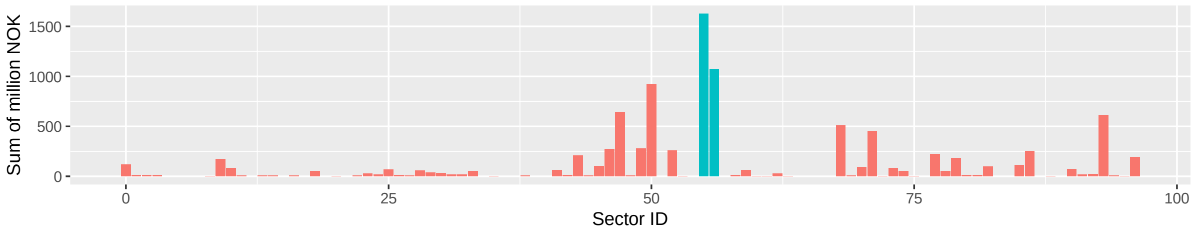


***Figure* 1**: *Descriptive statistics of the compensation scheme. The sum of granted compensation is shown across sector identifiers, limited to 2020 from March to December. The hospitality industry (sectors 55 and 56) are marked in blue. The figure demonstrates that the largest share of compensation was granted to these. The data is gathered from the Brønnøysund Register Center.*

[10] Industry I of NACE 1, corresponding to sector 55 and 56 in NACE 2.

ID	Description	Share of Total Compensation	Weighted Operating Margin (2010-2018)*	Share of bankruptcies 2010-2017
55	Accommodation	17.1%	-0.52%	1.32%
56	Food and Beverage Services	11.2%	2.12%	5.17%
50	Water Transport	9.7%	-2.92%	0.90%
47	Retail Trade exc. Vehicles	6.7%	1.88%	3.23%
93	Sports activities and recreation	6.4%	2.50%	1.15%
68	Real Estate	5.4%	37.6%	0.53%
71	Architectural and Engineering	4.7%	5.63%	0.85%
49	Land Transport and Storage	2.9%	3.73%	2.29%
46	Wholesale exc. Vehicles	2.9%	3.18%	1.60%
52	Warehousing	2.7%	4.80%	0.75%

***Table 1**: Summarizes the granted compensation and average operating margin in the sectors that were granted the most compensation. The sectors are ranked from top to bottom in order of total compensation but are limited to the ten largest recipient sectors. All compensation data is limited to 2020, and the average operating margin is collected from corresponding 2010-2018 data. The operating margin is weighted by revenue. Share of bankruptcies is drawn from the same dataset and the years 2010-2017. This figure uses the definition of bankruptcy from section 3.2.3. No filters were applied.*

The hospitality industry sectors received 2.6 billion NOK in compensation in 2020 alone, constituting 28% of all compensation granted. Sectors 50 (water transport) and 47 (retail) made up approximately 15% of the granted compensation as the 3rd and 4th group of recipients in the share of total compensation. We believe that focusing on heavily impacted industries improves the clarity of results. Furthermore, assessing the sectors with the largest share of compensation allows us to maximize our coverage while retaining specificity. Further interest in hospitality stems from the claims of a high zombie rate and low profitability by Altomonte et al. (2021).[11] Consequently, we have chosen to limit the scope of our thesis to this industry, corresponding to sector codes 55 and 56.

[11] "Zombie firms" are mature firms that have persistent problems meeting their interest payments (Matre & Solli, 2019) and have been found to cause undesirable effects on the economy. In our analysis, we follow McGowan, Andrews, & Millot (2017), such that firms aged 10 years or above with an interest coverage ratio below 1 over a minimum of three consecutive years are defined as zombie firms. Note that young firms therefore are excluded using this definition.

1.5 Outline of Thesis

To examine the distribution of compensation against firm viability, we take a two-step approach. In the first step, we develop a robust bankruptcy prediction model. In the second step, we use that model to predict the likelihood of bankruptcies among the firms that received compensation.

The primary methodological tool in the first step is bankruptcy prediction. Since we equate viability with survival, bankruptcy prediction is the natural quantification of viability. Because the criteria of compensation concern “otherwise viable” firms, we assume that the likelihood of bankruptcy before the pandemic reflects firms´ post-pandemic outlooks. Inspired by existing literature on bankruptcy prediction, we examine multiple algorithms and variables to determine the model with the most predictive power. We review parametric and non-parametric classifiers to ensure robustness and assess their out-of-sample performance using k-fold validation and relevant evaluation measures. Based on the examination of models, we proceed with the one yielding superior performance.

For the second analysis stage, we merge our dataset of 2018 financial statements (SNF) with data on compensation recipients. We use the probability output of the chosen prediction model and analyze the compensation scheme in line with the research question in section 1.3. Firstly, we calculate the sum of compensation granted to predicted bankrupt firms. Secondly, we analyze the compensation adjusted for revenue and employees with respect to the risk of bankruptcy. To infer results, we perform a visual inspection, use ordinary least squares regression (OLS), and perform testing using non-parametric techniques in case of non-linear relationships.

Throughout our thesis, we show how bankruptcy prediction using machine learning can be applied to evaluate the Norwegian compensation scheme. We find that while compensation was delegated primarily in line with the viability criteria at the aggregate level, a higher intensity of compensation was measured among more unviable firms.

2. The Role of Bankruptcy in a Market Economy

2.1 Crisis Policy

Time-limited policies with the purpose of reacting to a crisis are often called reactive policies. While both policy and crisis may vary, the ability to specifically target sudden impacts is a crucial tool in the arsenal of governments. The belief in natural recovery from financial crises was mostly discarded almost a century ago, necessitating some governmental reaction (IMF, 2014). During economic shocks, one of the primary concerns of governments has typically been to retain employment through firm survival.

In previous crises', governmental institutions' interventions have limited bankruptcies through targeted measures such as the provision of solvency support or the acquisition of firms. This was, for instance, the case during the bank crisis of 1987-1992 in Norway (Norges Bank, n.d.). According to Chung and Thewissen (2011), a broad range of doctrines for reactive policies were at play in Sweden, Germany, and the UK during the financial crisis of 2008. According to the authors, each country followed a specific strategy or doctrine determining the emphasis of measures. They describe the German strategy as conservative and note that they were mainly concerned with retaining firm-specific knowledge. In the UK, more emphasis was placed on the market's recovery ability, but interventions such as hiring subsidies were employed. For Sweden, the focus was placed on universal and generous social benefits. However, common for all were government-provisioned credit supplies, guarantees for financial institutions, and nationalizing distressed banks. Still, the main takeaway is the focus on employment during the crisis since mass unemployment appears to be the primary concern with mass bankruptcies. Unlike the crises seen in the 1990s (bank crisis) and 2000s (financial crisis), the role and responsibility of the government during the Covid-19 pandemic has shifted drastically, according to NOU (2021:4). The report argues that while financial institutions carried a more significant share of blame and costs in earlier crises, the government-imposed lockdown has naturally shifted the responsibility towards the government. In response, a natural outcome of this responsibility across Europe is the various policies to ensure economic activity and avoid mass unemployment.

2.2 The Cost of Bankruptcy

The fear of mass bankruptcies and subsequent mass unemployment is understandable since it entails financial, macro-economic, and mental costs. Unemployment raises governmental expenses related to unemployment benefits, central to maintaining a generous welfare system. According to Rotar and Krsnik (2020) countries with lenient unemployment benefits experience a slower return to employment because of increased reservation wages and a decrease in experienced urgency. Subsequently, states with generous welfare, such as Norway, should be acutely aware of the mass unemployment threat. However, regardless of the generousness of welfare, loss of employment causes reduced disposable income, lowering the aggregate demand. In other words, rising unemployment entails lower public income and higher public expenses.

The field of unemployment hysteresis is widely discussed in the literature. In this context, hysteresis refers to the slow recovery of employment after periods of high unemployment, often leading to permanent changes in the natural rate (O'Shaughnessy, 2011). O’Shaughnessy states that the effects were highly present following economic shocks in the 70s and 80s, as spikes in unemployment were followed by a slow recovery. Traits like these indicate hysteresis effects because they keep qualified personnel out of employment. This market failure involves a deadweight loss and keeps economic activity below its optimal point (Røed, 1997). Specifically, Dosi, Pereira, Roventini, and Virgillito (2018) find that several countries struggle with the aftereffects of the 2008 crisis, as employment and GDP growth levels remain below the level of 2008. Furthermore, they claim that wage flexibility strongly affects the presence of hysteresis.

Another aspect of employment is the affiliated mental cost. Employed workers can face severe disutility when losing the foundation for a stable income. This argument aligns with a study by Heggebø and Elstad (2017), investigating the effects of unemployment on self-rated health in 25 European countries with diverging macroeconomic conditions. They conclude that the unemployed have worse health conditions than employed workers throughout Europe, even after adjusting for time-invariant personal characteristics. It was concluded that the unemployed in countries with low unemployment rates faced higher health effects, raising the importance of this aspect in countries like Norway. It could therefore seem plausible that an unemployed person in a society with relatively low unemployment is worse off than those in a society with relatively high unemployment. The study argues that the effect likely stems from an aversion to inequality. The same topic is widely discussed within behavioral economics in much-cited papers such as Engelmann and Strobel (2004). Bankruptcies causing unemployment are therefore likely associated with high mental costs.

Further, bankruptcies entail a total capital loss through the loss of firm-specific accumulation of human capital. As an employee acquires firm-specific knowledge, the person's contribution becomes more valuable to his/her firm than a different one. Therefore, bankruptcy represents a waste in the macro-economic perspective since the firm-specific competence cannot be utilized fully elsewhere. However, it constitutes a more visible loss on an individual basis. Graham, Hyunseob, Si, & Jiaping (2013) found in their study that an employee in a bankrupt firm experiences an average wage loss of 30%. According to the study, that wage loss persisted for at least five years. While the aggregate impact of this effect is difficult to estimate, it provides further reason for the disutility of bankruptcies.

2.3 The Benefits of Bankruptcies

Regardless of the costs associated with bankruptcies, the process plays a significant role in the life cycle of businesses, primarily through the reallocation of resources and role in creative destruction. One example demonstrating the importance of bankruptcies is South Korea throughout the 1990s. According to Hahn and Lim (2004), firms in the South Korean economy during that decade experienced severe exit barriers. The barriers were caused by the standards for bankruptcy proceedings, which allowed for retaining firms that the Court deemed to be of sufficient social value. They found that a drastic reform of bankruptcy proceedings in 1998 lessened the exit barriers and improved total factor productivity among surviving firms (Hahn & Lim, 2004). This is a prime example of the utility of bankruptcies, as they cause the freeing of resources contained within firms. These resources will usually be ineffectively allocated as profitability measures the market's evaluation of the utilization of inputs. This mechanism is central within the resource-based view, in which resources are linked to competitive advantages and superior profit. The framework is frequently cited, and several studies find some empirical support for the central implication between resource and competitive advantage (Barney & Arikan, 2005). This mechanism follows traditional economic theory on allocating resources.

As explained, not all unproductive firms go bankrupt. This deficiency can be partly explained by factors such as firms within corporate umbrellas using or refraining from allowing bankruptcies for strategic reasons. However, another explanation exists within the concept of forbearance lending. Forbearance lending, or so-called “zombie lending,” occurs when a lender supports an otherwise insolvent firm. The term “zombie lending” was first introduced after the Japanese crisis in the late 1990s when stakeholders recognized that these firms existed. There is no universal definition of a zombie firm. However, most agree that these firms are economically unviable and manage to survive by tapping banks and capital markets (Cros & Epaulard, 2021). Motivated by lacking research on Zombie firms in Norway, a master thesis by Matre and Solli (2019) explored their prevalence in Norway. They found that Zombie firms account for 5% of the firms within the hospitality industry and that the overall prevalence in Norway is comparable to other OECD countries. In line with Hofman & Banerjee (2018), it is also argued that zombie firm prevalence typically increases during economic downturns and does not fully recover in the following period. The relevance of zombie firms, and in extension, “zombie-like” firms, is therefore of interest since governmental policies can act as a channel of zombie-lending.[12]

Tracey (2021) from The Bank of England argues that the forbearance lending observed in Japan in the late 1990s shares many similarities with the euro area following the European Sovereign Debt Crisis and the Covid-19 pandemic. Her argumentation is supported by the European Union’s top antitrust official, who in February 2021 claimed that massive Covid-19 bailouts have probably kept companies artificially alive. Compared to the financial crisis and sovereign debt crises, Schivardi, Sette, & Tabellini (2020) argue that banks are not as likely to participate in zombie lending after learning the lesson from the financial crisis. Nevertheless, they highlight that government subsidies, in the place of bank loans, can act as the channel of zombie lending through guarantees and other forms of compensation. If governments, instead of credit institutions, absorb the risk of lending, banks' incentives are

[12] We refer to “zombie”-like firms as firms sharing many of the same characteristics as zombie firms in terms of productivity, profitability, and effect on the economy.

weakened. Credit might therefore accrue disproportionately to non-viable firms. Chief Economist at Econa, Tore Vamraak (2021), further underlines this point. In contrast to the aforementioned authors, he highlights our geographical area of interest and adresses that the governmental compensation policies likely have the potential to increase the zombie firms' prevalence in the Norwegian economy. The potential for increased prevalence necessitates further knowledge on the effects of forbearance lending and zombie firms on the aggregate economy.

Literature on the 1990s Japanese stagnation from Caballero, Hoshi, and Kashyap (2008) underlines the effects of granting credit to otherwise insolvent borrowers. When the 1997 crisis hit Japan, these businesses were made more noticeable. The authors argue that the presence of “zombie” firms harmed the Japanese financial sector and caused higher capital costs. They found that zombie-dominated industries exhibited reduced job creation, destruction, and lower productivity, which contributed to the weak economic growth observed in Japan during their lost decade. This is in line with McGowan, Andrews, & Millot (2017). In their OECD working paper, they found that a high share of industry capital invested in zombie firms is associated with lower employment growth, lower investment, and less productivity-enhancing capital reallocation. In general, the destruction or change of existing structures and firms is a tool for improving said structures and introducing profitable firms. Specifically, voluntary market exits and bankruptcies free resources and allow surplus market demand, facilitating market entries. Compensation policies, on the other hand, reduce the bankruptcy probability, consequently limiting the renewal rate in the business life cycle (NOU, 2021:4).

A concept linked to resource allocation and forbearance lending is Schumpeterian creative destruction. It concerns the dismantling of long-standing practices to make way for innovation and is often described as a driving force of capitalism to ensure economic growth (Kopp, 2021). Therefore, Schumpeterian creative destruction is a highly relevant field concerning bankruptcy alleviation. In the book *The Power of Creative Destruction*, the authors claim that the effect of creative destruction can serve as a lever of growth post-pandemic (Aghion, Antonin, & Bunel, 2021). They also connect this to productivity and resource allocation since reducing inefficient active firms would result in higher average productivity. However, as mentioned in this section, they also comment on the issues regarding mass bankruptcies, such as unemployment, loss of human capital, and loss of usually productive firms. In this way, they claim that choices of compensation policy must be made to accompany creative destruction.

The issues with Zombie firms and forbearance lending are, as explained, caused by poor productivity or the ability to survive independently and have the documented effect of harming the overall economy. Given that governmental policies can act as potential channels of such lending, we interpret this as a motivation for assessing the compensation scheme with respect to firm viability. Accordingly, we find motivation in using bankruptcy prediction.

3. Bankruptcy Prediction Methodology

The purpose of section 3 is to develop a robust and reliable bankruptcy prediction model to quantify firm viability. We use four sets of independent variables, inspired by Altman (1968) and the SEBRA model of the Norwegian Central Bank (Næss, Wahlstrøm, Helland, & Kjærland, 2017). To understand the nature of classification and prediction of binary variables (the state of bankruptcy), we present and discuss evaluation tools used to assess the performance of the models. The mentioned sets of variables are used on four algorithms to develop 16 bankruptcy prediction models. Algorithms are chosen based on established literature and include both parametric- and non-parametric methods of classification. To ensure robustness, evaluate predictive power, and select a model, we use k-fold validation and review corresponding evaluation metrics.[13] We select the model with the highest precision using AUC and the McNemar's test to determine if the given model is significantly better than alternative models in line with Næss et al. (2017).[14] This procedure ensures robust and reliable estimations in the compensation analysis performed later in section 4. Note that the purpose of this section is not to evaluate or interpret the prediction variables. We are solely interested in developing a high-performing and robust bankruptcy prediction model.

3.1 Common Considerations

Bankruptcy prediction is an old field but has increased in complexity with time. Predicting bankruptcies involves different issues and considerations, which we briefly explain and discuss in this section. Firstly, the definition of bankruptcy is debatable, considering that bankruptcy proceedings have multiple outcomes and varying duration. Although the state of bankruptcy can seem objective and well-defined considering its binary nature, a bankruptcy proceeding can also be subjective, following the logic of Dijck et al. (2020), who discuss

[13] We evaluate the performance of the models at a nominal level, against each other, and compare them to results achieved in established literature

[14] In line with Næss et al (2017), who used McNemars test to evaluate if a classification model is significantly better than another by assessing the datapoints that are correctly predicted by the first model but incorrectly predicted by the other. If the p-value of the test is less than 1%, we conclude that the best performing model is significantly better than the other. We present and discuss the issue of threshold optimization in section 3.5.2.

strategic behavior in the interaction between bankruptcy judges and debtors. Ambiguity in the state of bankruptcy can therefore introduce systematic uncertainty in our target variable.

Secondly, bankruptcy represents an extreme event due to its rare occurrence. In consequence, prediction becomes more difficult than in cases where the classes are comparable in frequency. In the literature, the standard circumvention is to overrepresent the share of bankruptcies in the dataset to achieve a balanced dataset. In this way, bankruptcy models are developed and tested using an artificial share of bankruptcies. However, that solution can impede real-world applicability. To our surprise, we found few studies that perform and discuss the real-world applicability of bankruptcy prediction.

Thirdly, most models output the probability of bankruptcy on a scale from zero to one. Therefore, an issue is determining what value on this scale corresponds to a bankruptcy classification. As the threshold value decreases, more firms will be classified as bankrupt, necessarily implying both correct and incorrect classifications across the axis of thresholds. Naturally, an optimization of thresholds can therefore be performed if the costs of misclassifications are known. The literature approaches the issue of threshold optimization with two primary solutions: pre-set threshold values or maximizing an evaluation metric. The third solution, the implementation of misclassification costs, is rarely discussed in the literature.

Fourthly, since firms vary in financial structure and trend towards more heterogeneity (Kinserdal, Hansen, Pelja, & Stemland, 2019), using companies from several sectors and across time introduces a potential source of error. We have identified much discussion regarding this point, but few studies investigate the difference in predictive power for different bundles of industries among Norwegian firms. Chavam and Jarrow (2004) investigate the effects of industry indicators in bankruptcy prediction and find significant results. However, the study is relatively old and does not represent our sample of Norwegian firms. Nevertheless, with the transition toward more heterogeneous firms, one can also argue that bankruptcy prediction has become more complex through an increased manifold of relationships. In consequence, demand has risen for statistical techniques capturing more complex relationships, including interaction effects.

In summary, the field of bankruptcy prediction is widely explored and includes an array of dilemmas. While many issues are covered and discussed in the literature, we still identify lacking discussion regarding class balance, applicability, and threshold optimization.

We find the class balance issue especially relevant, justifying a more in-depth description. The literature outlines two primary solutions. (1) To increase the relative frequency of bankrupt firms until the balance is achieved. (2) To perform no adjustments and instead vary the threshold for classification. Cai and Singenellore (2012) propose to lower the threshold to 10% for classifying any firm as bankrupt as a rule of thumb. Other metrics are also applicable to optimize this threshold, as discussed in section 3.5.2. Altman (1968), Beaver (1966), and Wahlstrøm and Helland (2017) proceed with the first option of class balance, using between 30%-50% of bankrupt firms in their datasets. Ohlson (1980), on the other hand, did not modify the class balance and used the relative bankruptcy frequency of 4.845%, arguing that data adjustments limit applicability. Other researchers, such as Berg (2007) and Cai and Singenellore (2012), agree that dropping observations limits the applicability. Generally, it is easy to get the impression that the literature is highly theoretical in predicting bankruptcy. Considering that applicability is essential in our study, we decided to use unbalanced data. This affects the choices of data collection.

3.2 Data Collection

3.2.1 Source and Software

The Centre for Applied Research at NHH (SNF) provides the dataset for training and application of bankruptcy prediction. The dataset is based on delivered accounts and firm information from the Brønnøysund Register Centre. It includes 197 variables from the income statement, the balance sheet, and other firm characteristics (Berner, Mjøs, & Olving, 2016). With sector filters applied, our dataset consists of 66 639 observations. Because of the delay in registering bankruptcies, we omit the year 2018 for fitting the prediction models. Instead, we use it to perform the prediction and compensation analysis in section 4. Consequently, we only use the years 2010-2017 for fitting our prediction models. We use the statistical software STATA and R for the necessary data modifications. In addition to the SNF dataset, we retrieve the complete dataset of granted compensation within the two iterations of the compensation scheme for 2020. The datasets merge for section 4.

3.2.2 Filters for Data Quality

We apply filters to ensure sufficient data quality and describe the ones applied to increase replicability and transparency in our method. The choice of firms deviates from the norm in the literature, as we use a sample with a wide range in size with less strict filters. Therefore, our restrictions are less rigorus on variables measuring firm size, which could result in more noise in our observations. Noise could be detrimental to achieving sufficient predictive power and is, therefore, one of our primary concerns. Zmijewski (1984) on the other hand argues that the effect of omitting observations should be examined with respect to outcome. For that reason, we investigated the effect of removing the most important filters on asset size and sales in our prediction model. However, we found that the filters made no qualitative difference to the results.[15]

We remove firms with missing values for total assets, inventory, accounts receivables, and accounts payables. Missing values in these variables indicate poor data quality and can be problematic when generating other key figures.[16] Furthermore, we apply filters requiring firms to have at least 100k NOK in sales income and assets. The sum of assets and sales filters follows the reasoning of Kinserdal et al. (2019) and Næss et al. (2017), as firms with assets and sales below 100k NOK are omitted. We examined several financial statements with operating margins beyond 50% on both ends and found that most of the observations suffered from poor data quality. Provisions or reversals of provisions were the most common factor explaining the extreme ratios lowering quality. Since provisions and reversals of provisions have little to do with the operations of a firm, we omit these observations to preserve the quality and predictive power of the data. Further, labor costs must exceed zero, while invested

[15] We investigated the effect of removing asset size and sales filters on predictive performance, choice of classifier, and analysis results. However, we found no qualitative difference in the ranking of classifiers, the predictive power of the non-parametric classifier, and analysis results in section 4. Yet, the parametric classifiers performed noticeably worse in the absence of filters. See section 3.4 for the description of different classifiers.

[16] We followed the same procedure as when examining the effect of our filtering. The examination revealed no qualitative difference in the predictive power and analysis results.

capital and the book value of equity cannot equal zero. We apply all filters with the intent to remove observations with poor data quality and non-active firms. After applying these filters, the number of observations is 45 526.

Examination of the dataset after filters revealed that the data quality of the firms' sector was poor. Ideally, we want to keep observations where the firm's purpose is operations within the hospitality industry. In several cases, financial holding companies used sector codes 55 and 56 when they held stocks of other operating companies within the industry. Since financial holding companies are outside the scope of this paper, we attempt to exclude these observations. We removed observations where the name of a firm included any of the following words: “holding,” “invest,” “eiendom,” “finans,” and “capital,” and where the ratio of financial assets to sales exceeded 5. We applied a less strict ratio filter of 50 to all remaining observations without these keywords. The filter caused the removal of 117 companies. Our final dataset consisted of 45 409 observations where 1893 are identified as bankrupt, representing a bankruptcy share of 4.17%.

We source compensation data from the official sites of the Brønnøysund Register Centre. Each iteration of the compensation scheme had its own dataset with different variables. After merging, filtering for the year 2020, and collapsing observation per firm, 36 691 firms remained across all sectors. After filtering on sectors 55 and 56 with the requirement that the sum of compensation must exceed zero, 5160 observations remained in the dataset, representing all firms that were granted compensation before potential sources of error. The dataset contains variables such as organization number, granted compensation, sector code, and revenue for January and February for 2019 and 2020.

3.2.3 Target Variable

Observations are defined as bankrupt in the last year of financial statement delivery, given that the bankruptcy proceedings started within three years. Additionally, an observation classifies as bankrupt up to 2 years prior to the year of the bankruptcy filing to extend the time horizon of detection. This procedure is in line with Bernhardsen & Larsen (2001) due to the logic that the event of bankruptcy usually requires a time frame beyond one year. It would therefore not be appropriate to define the year of the proceedings as the year of bankruptcy. The registered year of bankruptcy is the one recorded in the SNF dataset. Unfortunately, the variable for bankruptcy includes both bankruptcy and compulsory dissolution, which are not separable within our dataset. However, it should not critically harm the robustness of the models since compulsory dissolutions have held a relatively constant frequency, also during the pandemic (Brønnøysundregistrene, 2022).

The event corresponding to bankruptcy in the dataset typically follows one of the following four patterns; (1) The court can declare bankruptcy due to neglect, for example, by failing to pay creditor´s debt or pay wages to employees. When the court finds the terms for bankruptcies as fulfilled, proceedings will start, and a trustee will gain control over the assets to pay creditors. (2) A company can declare itself bankrupt and start to disband the firm. Companies are required by law to apply for bankruptcy on their initiative if the business fails to generate profits at the expense of creditors. Failure to comply with this requirement is punishable. (3) Companies can fail to comply with the Limited Liability Companies Act (aksjeloven) and therefore be declared dissolved by the court. For example, firms can fail to deliver financial statements within the deadline or fail to appoint an auditor if required. (4) Inactive companies can be disbanded without any legal repercussions (Konkursrådet, 2012).

3.3 Selection of Predictor Variables

Our study uses four sets of relevant ratios identified in established literature. With this consideration, we argue that a broad assessment of new predictor variables is outside the scope of this thesis. Still, we will discuss the importance of the chosen variables and add select atypical variables in our fourth set. The variable sets are described in order of rising complexity. We summarize the mean in each predictor variable, as well as the standard deviations, across the two classes of observations (bankrupt and non-bankrupt) to allow quick inference of the variables' discriminating ability.

Variable Set 1: Altman Z-Score

Altman (1968) developed the well-known Z-score model using multivariate data analysis. He remains one of the most well-known pioneers in bankruptcy prediction. Altman selected five out of 22 reviewed ratios that he found to be most important to assess bankruptcy risk accurately. The ratios are still common in research and provide a well-known model evaluation baseline. We, therefore, use Altman´s set of variables (VS1), which he named x_1 to x_5, resulting in the following ratios:

Altman´s Variables (VS1)		
Predictor Variables	**Mean non-Bankrupt**	**Mean Bankrupt**
$x_1 = \frac{Working\ capital}{Total\ assets}$	-0.03 (0.31)	-0.28 (0.75)
$x_2 = \frac{Retained\ Earnings}{Total\ assets}$	-0.27 (1.58)	-1.14 (2.14)
$x_3 = \frac{Earnings\ before\ interest}{Total\ assets}$	-0.01 (0.55)	-0.44 (0.99)
$x_4 = \frac{Market\ value\ of\ equity}{Book\ Value\ of\ total\ Debt}$	0.5 (2.72)	-0.1 (2.5)
$x_5 = \frac{Sales}{Total\ assets}$	3.55 (3.03)	5.43 (4.52)

Table 2: *Altman Z Score Variables (VS1). Altman´s initial variables are based on his study from 1968. The variables X_1 to X_5 are intended to measure liquidity, firm performance, and solidity. Mean non-bankrupt and mean bankrupt refers to the average value of each variable in observations of the two states of the target variable bankruptcy. Standard deviations are provided in brackets.*

Altman selected the variables to measure critical components of bankruptcy, and all measure either liquidity, solidity, or firm performance. X_1 is meant to measure liquidity due to the logic that working capital is a liquid class of assets.[17] X_2, X_3 and X_5 measure firm performance as retained earnings reveal accumulated previous results, earnings before interest and sales measure the firm's income, and X_4 measures the solidity of the firm given the book value of equity and debt. Altman acknowledged that the market value of equity and debt should be used if available but that the book value serves as a substitute when it is not, as with most SMEs.[18] The summary statistics suggest that bankrupt firms generally have lower working capital, less retained earnings, lower EBIT and book value of equity, and a higher sales ratio to total assets.

[17] Working capital in our study is calculated as inventory plus account receivables minus accounts payables based on a “rule of thumb” procedure. Ideally, one would manually inspect each firm and determine its real working capital, i.e., the capital locked to the running operations of a firm that it cannot be sold off without affecting the operations (Kinserdal, BUS440B 6b Investeringer i driftsmidler og arbeidskapital, 2021).

[18] Small and medium-sized enterprises (SMEs)

Variable Set 2: Alternative Altman Ratios

For the second set of variables (VS2), we have decided to use alternative ratios based on criticism from Kinserdal, et al. (2019) on traditional ratios. We find it reasonable to apply variables from recent literature on bankruptcy prediction since the article raises harsh criticism on the widespread use of traditional ratios such as Altman's. Because the argumentation is convincing, we chose to apply their suggestions and introduce alternative ratios. We want to examine whether the alternative ratios measuring liquidity, solidity, and firm performance yield superior predictions. The motivation for this is primarily to ensure a robust and accurate model and supplement current literature by contributing to the field.

Previous research with Altman´s variables has yielded varying results based on industry, geography, and time. Kinserdal et al. (2019) point out that factors such as the difference in the combination of industries, intangible assets, and book values compared to marked values are among the drivers of the range in results. It is also argued that changing accounting rules and introducing IFRS are essential in impacting the numbers in financial statements, implying that bankruptcy models must be redesigned over time.

Firstly, they argue that working capital is not preferable when measuring liquidity. A better predictor, he argues, is assets that can be liquidated without affecting the operative business, such as investments in stocks, property, and other financial assets. Consequently, he proposes financial assets over short-term debt as a liquidity ratio. Following the naming convention of Altman, we replace X_1 with $K_1 = \frac{financial\ assets}{short-term\ debt}$.

Secondly, he argues that the book value of equity and debt is problematic. The ideal solution would be to find the fair value of all assets and liabilities, and the equity value would then be the remaining residual. However, this is difficult for unlisted firms, and the authors claim they failed to find a good substitute. Still, they comment that financial assets to total debt provide an intuitively better predictor than equity ratio. Moreover, they found that a dummy for negative equity had strong predictive power. Consequently, we replace X_4 with $K_4 = \frac{financial\ assets}{total\ \ debt}$ and a dummy D_4 which takes on the value of 1 if the book value of equity is less than 0.

Thirdly, the authors argue that the remaining variables measuring firm performance, X_2 and X_3 using EBIT and retained earnings, are problematic. Financial statements noise can arise from sources outside operative functions, such as a scenario where a firm can sell a property and recognize the gains that single year. This point is even more apparent for smaller businesses. It is assumed that smaller and less diversified firms are more sensitive to one-time effects and often tend to have lower-quality financial statements. One way to counter this effect is to exclude tiny firms from the sample, but it is not desirable when researching a broad range of firms. For Norwegian companies using NGAAP standards as most small-sized firms, the income statements can be affected by changes in the market value of assets and liabilities without cash flow effects. For example, suppose an owned property rises 10% in value. That gain can be recognized as income in the financial statement even if it does not reflect the operative business. In summary, he proposes using EBITDA instead of EBIT to omit some noise introduced, including depreciation and amortization. In the actual study, EBITDA minus operational investments and changes in working capital was used to isolate cash flow from operations. However, these are not variables available in accounting statements and are unpractical to use unless spending excessive time examining each firm. Consequently, we substitute EBIT with EBITDA in X_3. Over time, the latter seems to better represent actual cash flow, according to Kinserdal et al. (2019).

Although the authors acknowledge that retained earnings take into account the previous year's earnings, the key figure can be affected by dividend payouts and conversions from debt to equity. As a result, they propose using key figures that utilize earnings from the last 2-3 years instead. We substitute X_2 with $K_2 = \frac{average\ of\ EBITDA\ last\ 3\ years}{total\ assets}$. The variable X_5 is kept as it is, as this variable is not commented on in the study.

At last, Kinserdal et al. (2019) argue that the sum of assets can be an inappropriate variable since it includes financial assets when we only want the assets used within the business. They, therefore, propose using operating or invested capital (sum of assets minus financial assets) instead of the sum of assets. Consequently, we substitute the sum of assets in all X_i with invested capital. The variable set contains the following variables after implementing all suggestions:

Alternative Altman Z (VS2)		
Predictor Variables	**Mean non-Bankrupt**	**Mean Bankrupt**
$K_1 = \frac{Financial\ assets}{Short\ term\ debt}$	0.23 (2.07)	0.07 (0.24)
$K_2 = \frac{Average\ of\ EBITDA\ last\ 3\ years}{Invested\ capital}$	0.06 (1.27)	-0.32 (0.91)
$K_3 = \frac{EBITDA}{Invested\ capital}$	0.08 (3.60)	-0.36 (1.08)
$K_4 = \frac{Financial\ assets}{Total\ \ debt}$	0.19 (4.35)	0.16 (1.20)
$K_5 = \frac{Sales}{Invested\ capital}$	3.91 (7.42)	5.99 (5.36)
$D_{negative\ E} = 1\ if\ E\ is\ < 0$	0.27 (0.45)	0.70 (0.46)

***Table** 3: Alternative Altman Ratios (VS2). Based on the criticism of traditional ratios from Kinserdal et al. (2019). The ratios K_1 to K_5 correspond to but are revised forms of the Altman variables of X_1- X_5. $D_{negativeE}$ is a dummy variable for the positive or negative book value of equity. Mean non-bankrupt and mean bankrupt refers to the average value of each variable in observations of the two states of the target variable bankruptcy. The applicability of alternative Altman ratios in the hospitality industry is debatable. Standard deviations are provided in brackets.*

Variable Set 3: SEBRA

The latest and most relevant contribution to our study on bankruptcy prediction is Næss et al. (2017). They performed bankruptcy prediction using several statistical techniques on a large dataset of Norwegian firms and were inspired by the variables used in the SEBRA model by Norges Bank. Variable set 3 (VS3) is extracted directly from their study and summarized in the table below. They propose using seven variables from the SEBRA model with an additional nine variables measuring characteristics of firm performance, solidity, and auditing remarks. We see the choice of these variables as appropriate due to the study's recency and thoroughness. While the variable set is an expansion of SEBRA, we refer to it solely as SEBRA to distinguish it from the fourth variable set. The variable naming convention follows that of Næss et al. (2017), and the variables are shown below.

SEBRA by Næss et al. (VS3)		
Predictor Variables	**Mean non-Bankrupt**	**Mean Bankrupt**
$N_1 = \frac{EBDA}{Total\ debt}$	0.15 (1.18)	-0.28 (7.01)
$N_2 = \frac{Equity}{Sum\ of\ assets}$	-0.04 (1.18)	-0.83 (1.81)
$N_3 = \frac{Working\ captial}{Sales}$	-0.01 (0.42)	-0.08 (0.30)
$N_4 = \frac{Accounts\ payable}{Sum\ of\ assets}$	0.18 (0.31)	0.48 (0.76)
$N_5 = \frac{Debt\ to\ the\ public}{Sum\ of\ assets}$	0.15 (0.18)	0.30 (0.32)
$N_6 = \frac{Debt\ to\ credit\ institutions}{Total\ debt}$	0.14 (0.29)	0.10 (0.20)
$N_7 = \frac{Sales}{Sum\ of\ assets}$	3.55 (3.03)	5.43 (4.52)
$N_8 = \frac{Earnings}{Book\ value\ of\ equity}$	0.33 (30.99)	0.32 (7.68)
$N_9 = \frac{Sales}{Wage\ costs}$	3.05 (3.31)	2.94 (1.85)
$N_{10} = \frac{Inventory}{Sum\ of\ assets}$	0.08 (0.09)	0.12 (0.12)
$N_{11} = \frac{Sales * 1{,}25}{Accounts\ receivable}$	11245 (52891.11)	4344 (4774.50)
$N_{12} = \frac{Current\ assets}{Sum\ of\ assets}$	0.61 (0.30)	0.60 (0.29)
$D_{negative\ E} = 1\ if\ E\ is\ < 0$	0.27 (0.45)	0.70 (0.46)
$D_{audit} = 1\ if\ audit\ remark$	0.07 (0.26)	0.18 (0.39)
$D_i = \begin{matrix} 1\ if\ age\ is\ above\ i\ years \\ 2\ if\ age\ is\ not\ above\ i\ years \end{matrix}, where\ i = 1, \dots, 8$	9.19* (9.51)	4.39* (6.04)

***Table** 4: SEBRA variables (VS3). Based on a recent study performed by Næss et al. (2017) on a large sample of Norwegian firms. Mean non-bankrupt and mean bankrupt refers to the average value of each variable in observations of the two states of the target variable bankruptcy. The means of the dummy for age show the average age of the classes for readability, not the averages of the eight dummies on age used in the analysis.*

The study by Næss et al. (2017) found that the additional variables beyond those used in the SEBRA model improved results for multiple statistical techniques. The variables in set 3 reveal that bankrupt firms have a significantly higher debt ratio to total assets in N_4 and N_5. The latter is interesting for the Norwegian Tax Administration, considering the aim of viability and postponement of payments for taxes and fees. Furthermore, we observe that bankrupt firms are associated with negative book values of equity, auditing remarks, lower sales, and lower earnings.

Variable Set 4: SEBRA Plus

Variable set 4 (VS4) adds value by including variables measuring local market characteristics and macroeconomic variables. The inclusion is based on studies affirming that macroeconomic trends affect bankruptcy frequency (Jacobsen & Kloster, 2005; Nam, Kim, Park, & Lee, 2008). The additional variables beyond VS3 are summarized in the table below.

SEBRA Plus (VS4)		
Predictor Variables	**Mean Non-Bankrupt**	**Mean Bankrupt**
$D_{management} = 1\ if\ changes\ in\ management$	0.03 (0.18)	0.07 (0.26)
$M_{k,j,l} = HHI\ in\ year\ k\ for\ municipality\ j\ \ and\ sector\ l$	2219 (2639.90)	1686 (2260.78)
$MS_{\,i,k,j,l} = market\ share\ for\ firm\ i\ in\ year\ k, municipality\ j\ and\ sector\ l$	7.86 (16.27)	4.89 (12.64)
$S_{j,k} = centrality\ index\ for\ municipality\ j\ in\ year\ k$	4.35 (2.85)	3.93 (2.57)
$SU_k = share\ of\ unemployment\ in\ year\ k$	3.96 (0.48)	3.97 (0.51)
$BNPC_k = relative\ change\ in\ GDP\ in\ year\ k$	1.99 (0.76)	1.96 (0.79)
$KPI_k = change\ is\ consumer\ price\ index\ in\ year\ k$	0.03 (0.80)	0.07 (0.84)

***Table** 5: Variables of SEBRA Plus (VS4) that are not present in variable set 3. I.e., Sebra Plus consists of the SEBRA (VS3) set and the variables in this table. Standard deviations are provided in brackets. Mean non-bankrupt and mean bankrupt refers to the average value of each variable in observations of the two states of the target variable bankruptcy. Standard deviations are provided in brackets.*

We hypothesize that including variables for market structure and macroeconomic factors provides increased predictive power, as it covers exogenous traits of organizational environments. Moreover, the dummy variable for changes in management allows the capture of non-quantifiable endogenous effects, as it has the potential to signal poor performance.

3.4 Classifiers

This section explains the four classifiers used on the four variable sets. We refer to classifiers as the algorithms or methods used to develop models. In general terms, we seek to find the unknown form of f to describe the relationship between X and Y. The true function f is unknown, such that we establish $y = \hat{f}(x)$ in an attempt to best describe f. In brief, the two categories of statistical modeling that describe these relationships are parametric and non-parametric methods. They are typically grouped in two depending on whether the goal is to estimate a qualitative or quantitative target variable. I.e., when estimating models, we need to consider whether we want to use a parametric or non-parametric classifier and whether we seek to estimate a qualitative or quantitative target variable.

Parametric methods involve a more straightforward approach to estimating the unknown form of f because the problem is narrowed down to assuming a functional form by estimating a set of coefficients. For example, both MDA and GLM are linear in X and predict Y by estimating a set of coefficients. This approach is relatively simple, computationally cheap, and implies underlying assumptions about linearity and distribution of the data. The underlying assumptions impose a disadvantage in parametric methods because they typically will not describe the true form of f if the assumptions are not met.[19]

Non-parametric methods like random forest and GAM on the other hand, differ in how they describe the relationship between X and Y by fitting an entirely arbitrary function that allows for many different possible functional forms for f. Essentially, there are no underlying assumptions about the linearity and distribution of the data, such that more flexible patterns can be detected. This major advantage is also the basis for criticism of non-parametric methods. Since the classifiers are more flexible in the functional form of f, non-parametric approaches are vulnerable to overfitting in that they can detect patterns that do not exist, reducing the out-of-sample predictive power (James, Witten, Hastie, & Tibshirani, 2021). As discussed in section 3.6.4, random forest has built-in features that help to counter this.

[19] We note that recent literature often points out that models can easily handle deviations from assumptions, given that the purpose is prediction and not interpretation of causality. This argumentation supported by Berg (2007) and Barnes (1982).

As explained, there is a division in predicting either quantitative or qualitative target variables. Quantitative target variables represent data that is continuous and numerical. Qualitative target variables, on the other hand, take on the values in one of K different classes. In our case, the event of bankruptcy represents a qualitative target variable because it involves classifying observations into categories, i.e., the state of bankruptcy, such that K=2.

We wish to examine the relationship between the binary target variable Y and independent variables X_i in all of VS1-VS4. Y takes on the value of either 0 or 1 depending on the state of bankruptcy:

$$y_{bankrupt} = \begin{cases} 1 \; if \; declared \; bankrupt \\ 0 \; if \; not \; declared \; bankrupt \end{cases}$$

Although the target variable is binary, our classifiers output probabilities as a foundation for classification instead of directly outputting the classes (James et al., 2021). Therefore, a classification threshold is required to classify a firm as either bankrupt or non-bankrupt.

While several parametric and non-parametric methods exist for predicting binary outcomes, we use MDA, logistic regression, GAM, and random forest (RF). The two first classifiers have been popular for decades in the field of bankruptcy prediction, while GAM and random forest have gained in popularity in later years (Alaka, et al., 2018). MDA and GLM are parametric methods in that estimating f consist of estimating a set of parameters $ß_i$. RF and GAM are a non-parametric methods allowing for a more flexible fit to describe f with few underlying assumptions about the data (James et al., 2021).

The scope of this paper is not to discuss nor derive any of the statistical models in detail, but to briefly describe and compare them. Since the classifiers are used for predictive purposes and applied to empirical data, out-of-sample performance metrics supersede the importance of satisfying assumptions.

3.4.1 Multiple Discriminant Analysis (MDA)

MDA is a classifier to study the differences between groups for multiple variables simultaneously (Klecka, 1980). It was first introduced by Fisher in 1936 (Cohen, Cohen, West, & Aiken, 2002), who proposed a technique for minimizing the differences within the groups while maximizing the differences between groups when estimating coefficients. The method has been widely used in bankruptcy prediction and is the basis for the Altman Z-model.

The applicableness of the model depends on several conditions and assumptions. Firstly, two or more mutually exclusive groups are required such that each observation belongs to either of the groups. Secondly, there must be at least two observations within each group. Thirdly, the number of independent variables cannot exceed n-2 observations. Lastly, it must be possible to distinguish between the groups based on the independent variables, typically ratios or interval data. Our analysis satisfies all these conditions. If the purpose is to study causal effects, more statistical conditions must be satisfied, as Altman (1980) described. Still, MDA analysis is potent even with deviations from statistical assumptions, suggesting that the classifier is more robust than initially believed (Klecka, 1980).

According to Altman (1968), the space dimensionality in MDA is the number of groups minus one. Thus, our study is a one-dimension analysis as our dependent variable can take on the two values of 0 and 1. According to Altman (1968), the space dimensionality in MDA is the number of groups minus one. Thus, our study is a one-dimension analysis as our dependent variable can take on the two values of 0 and 1. MDA in its simplest form is summarized below and shares some similarities with standard regression in the way the model is constructed, although the coefficients are derived differently.

$$Z = V_0 + V_1X_1 + V_2X_2 + V_3X_3 \dots\dots + V_nX_n \qquad \textit{Eq. 2}$$

V_i are the raw coefficients in the model and X_i are the independent variables used to discriminate between the classes bankrupt and non-bankrupt in our case. The coefficients are estimated such that linear combinations of the coefficients maximize the difference between the groups (1968), while minimizing the differences within the groups.

Since MDA is an ordinal ranking (discriminatory device), the coefficients according to Ohlson (1980), yield little intuitive interpretation and the scale is not directly relevant for the applicability of a classification structure. To get around this issue, we calculate the posterior probability that an observation belongs to the predicted class of bankruptcy using the Bayes theorem, as computed in the equation below (James et al., 2021).

$$Pr(Y = k|X = x) = \frac{\pi_k f_k(x)}{\sum_{l=1}^{K} \pi_l f_l(x)} \qquad \text{Eq. 3}$$

Where $Pr(Y = k|X = x)$ denotes the posterior probability that an observation belongs to the k´th class or 0 and 1 in our case, i.e., the probability that an observation is classified as bankrupt given the predictor value of $f_k(x)$, given the initial probability of bankruptcy. It is then possible to apply our classification framework for the purpose of evaluating our models. It is, however, argued by Ohlson (1980) that posterior probabilities are inaccurate when the strict statistical conditions of MDA are not met, such as when we use qualitative independent variables violating the normality condition. These strict conditions of MDA contributed to Ohlson´s motivation to use generalized linear models instead.

3.4.2 Generalized Linear and Additive Models (GLM, GAM)

Generalized linear models are generalizations of standard linear regressions allowing for qualitative target variables. According to James et al. (2021), it is possible to generalize a linear model in several ways. The most used method is logistic regression. For any X_i, we use logistic regression to predict the probability that Y belongs to a particular category, that is $p(X) = Pr(Y = 1|X)$ where Y is an element in {0,1}. The logistic model falls within the category of GLM and is summarized as follows:

$$p(x) = \frac{e^{ß_0+ß_1X_1+....+ß_pX_p}}{1 + e^{ß_0+ß_1X_1+....+ß_pX_p}} \qquad \text{Eq. 4}$$

Where X_i represent our independent variables, e is the natural logarithm and the $ß_p$ are the coefficients we want to estimate from the regression. Due to the nature of the model, it is easy to see that the function $p(x)$ must result in values within the interval [0,1], producing an S-shaped form along the X-axis of independent variables. After manipulating Eq. 4, we get the log odds or logit summarized on the left-hand side in Eq. 5. We see that the function is linear in X in our case as in all components of $ß_pX_p$. In standard regression, a unit increase in X corresponds to a ß increase in Y, whereas in our case, one can observe that a unit increase in

X instead increases the log odds by ß. Consequently, a unit increase in X also impacts Y depending on the current value of X in contrast to the case with standard regression.

$$log\left(\frac{p(x)}{1-p(x)}\right) = ß_0 + ß_1 X_1 + \cdots . + ß_p X_p \quad \textit{Eq. 5}$$

An extension from this model was first proposed by Hastie and Tibshirani (1990) to also allow for non-linear relationships between X and Y while still conserving additivity, resulting in the field of generalized additive models (GAM). The extension of the model is summarized in Eq. 6 below and based on the logit of the initial linear logit model.

$$log\left(\frac{p(x)}{1-p(x)}\right) = ß_0 + ß_1 f_1(X_1) + \cdots . + ß_p f_p(X_o) \quad \textit{Eq. 6}$$

The additivity comes from the non-linear components $ß_p f_p(X_o)$, where the model calculates a separate f_p for each X_p before adding up all contributions, thus resulting in additivity.

The method used to derive the coefficients is based on the maximum likelihood function as opposed to the least-squares approach. This is due to the favorable statistical properties of maximum likelihood when fitting the model for qualitative target variables. Although it is outside the scope of this paper to derive the mathematical formulas, the basic intuition behind maximum likelihood is to seek estimates for $ß_0$ and $ß_p$ such that the predicted probability of $\hat{p}(x_i)$ of default using formula Eq. 7 corresponds as closely as possible to the observed class of default. I.e., we estimate $ß_0$ and $ß_p$ so that the model yield results closer to 1 for observations that defaulted and 0 otherwise. The intuition can be summarized formally and is expressed in Eq. 7 through the likelihood function:

$$\ell(ß_0, ß_1) = \prod_{i:y_i=1} p(x_i) \prod_{i':y_i,t=0} (1 - p(x_{i'})) \quad \textit{Eq. 7}$$

$ß_0$ and $ß_1$ are estimated such that the function is maximized. The major advantage of GLM is its applicability to real-world applications and interpretable coefficients to indicate both the magnitude and direction of a unit increase in X on Y. In GLM, it is easy to interpret the effect of X on Y through the log-odds. Additionally, Ohlson (1980) argue that no assumptions related to prior probabilities of bankruptcy and distribution of independent variables have to be made, promoting the use of GLM in our study. The disadvantage of both GLM on the other hand, is the vulnerability to multicollinearity, where correlated independent variables can result in biased coefficients if we start to infer and interpret the coefficients (Berg, 2007). The major drawback of GLM is still its linearity assumption in X. Although GAM has the major advantage of detecting these non-linear relationships, its disadvantage relates to its weaknesses in detecting interaction effects between the variables (James et al., 2021). The use of decision trees and other machine learning techniques are considered superior in detecting these interaction terms without the need for manually creating and exploring them.

3.4.3 Classification Trees and Random Forest

Decision trees are popular within predictive modeling and are commonly referred to as classification trees when target variables take on a discrete set of values (James et al., 2021). The approach has gained popularity as more computational power has become available. There are several advantages of using decision trees. Firstly, they are simple to visualize and interpret, and the decision-making structure is closer to human processes parametric approaches. Secondly, decision trees can replicate complex patterns, such as interactions, and efficiently handle categorical variables.[20] However, James et al. (2021) argue that decision trees often are empirically inferior to other methods due to overfitting. Moreover, their robustness can be lacking unless methods such as bagging and bootstrapping are used to introduce randomization.[21] By non-robust, they refer to how small changes in the data can drastically change the design of the tree. One approach to ensure robustness is using the random forest algorithm, which uses bagging and bootstrapping.

The random forest algorithm uses decision trees to predict Y. To use a classification tree, the standard procedure is to start at the first predictor (node or branch) and move towards the target variables (leaves), where the leaves or terminal nodes determine a binary score of 0 or 1 (based on the most commonly occurring class of training observations within that terminal node from the training data). An example of classification is attached below, based on our classification problem of predicting bankruptcies using Altman´s initial variables. Note that

[20] Ability to capture complex interactions depends on the number of trees used

[21] Bagging and bootstrapping are explained in section 3.6.4

this classification tree is very simple and is not built upon all principles of the random forest algorithm. The leaves (terminal nodes) indicate the probability of belonging to the given class.

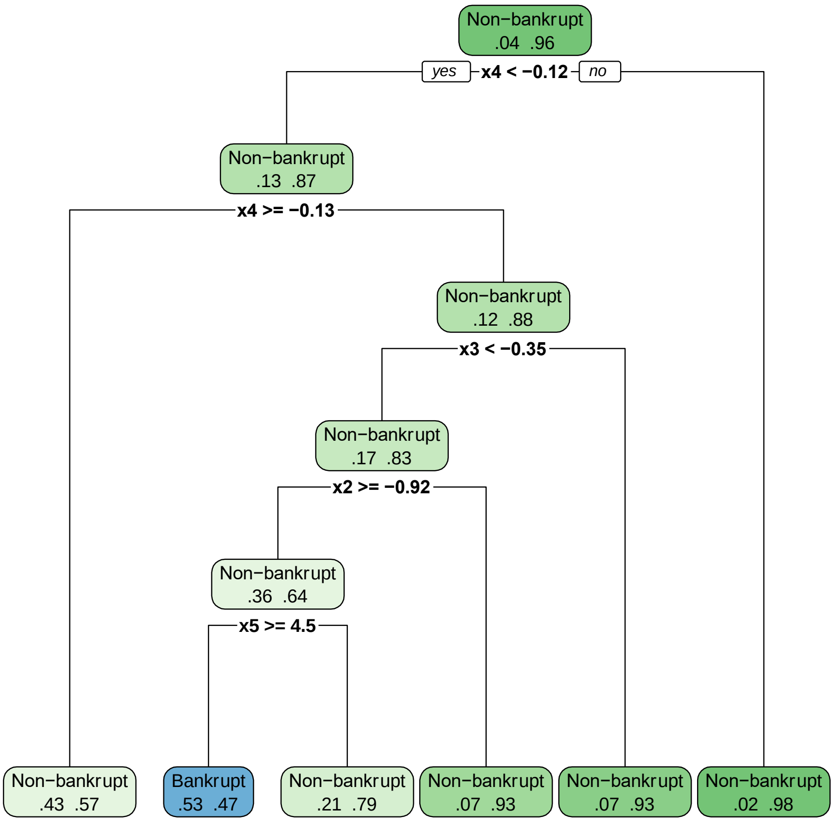


Figure 2: *Simple decision tree example based on our dataset and prediction of bankruptcies using Altman´s ratios (VS1). From the decision tree example, the bottom-most figures display the probability of belonging to the group of bankrupt or non-bankrupt, based on the data used for developing this decision tree.*

Building a classification tree follows the logic of the top-down approach known as recursive binary splitting. The process of recursive binary splitting consists, in each step of the tree building, of making the best splits such that class separation is maximized across the prediction space according to James et al. (2021). This is based on measures of separation ability in the node, commonly referred to as node purity. The purpose is to increase purity, such that as many observations as possible belong to only one class. Each split naturally results in two new nodes further down the tree.

To determine the node purity or separation ability within a node, several measures are suggested in the literature. James et al. (2021) mentions entropy, the classification error rate, and the Gini Index. The latter is suggested in most cases for further procedure. The Gini index is defined as follows:

$$G = \sum_{k=1}^{K} \hat{p}_{mk}(1 - \hat{p}_{mk}) \qquad \textit{Eq. 8}$$

Where $\hat{p}$ represents the proportion of observations in node m from class k. A relatively low value of p is an indication that most observations within that node fall into one node, consequently resulting in what the literature refers to as a highly pure node, which is preferable in decision tree building. Thus, the procedure for decision tree building is to consider the Gini scores for each available variable (including different thresholds when the data is continuous) and select the variable with the lowest Gini score, such that reduction of Gini is associated with higher node purity.

Resulting from the Gini scores, variable importance can be determined for estimating the qualitative target variable using the mean decrease in the Gini index. The index measures a variable's reduction in Gini when making splits averaged over all trees. Since the reduction in Gini is decently low when node purity is high, a lower mean decrease in Gini is associated with lower variable importance. Consequently, variables with high mean reductions in Gini are associated with higher variable importance.

Although the major advantage of using decision trees is its lack of strict assumptions, flexibility, and ability to detect complex patterns (especially if increasing the size of the tree through either depth or available features), concerns arise related to the risk of overfitting and high variance.

To yield predictive power out of sample and mitigate this issue when determining our target variable, one solution is to aggregate results from many trees by using bagging. In bagging, each tree determines a target variable, and the final predicted probability is the average class across all the n-number of trees.

The issue of overfitting is not omitted solely by increasing the number of trees. All trees will be correlated if using the same training data and same set of features to determine y. As such, trees should be de-correlated. Regarding the first aspect, the most common solution is to use bootstrapping techniques where bootstrapped samples are extracted from the training data as a basis for each tree, such that each tree to some extent is de-correlated. To further reduce the risk of correlating the trees, the random forest algorithm can be used. Random forest de-correlates each subtree by imposing restrictions on variable selection in each split. Since the selection of variables is shuffled at random in each iterative split, the technique forces all alternatives to be considered on an equal basis, resulting in a robust ensemble tree where concerns related to correlated trees are mitigated. This variety makes the random forests algorithm more effective than individual decision trees and stands out as an attractive model for our predictions of bankruptcies.

Naturally, since the random forest algorithm uses several trees, the user needs to determine the n number of trees. This decision can be assisted by hyperparameter tuning. Hyperparameter tuning is the practice of optimizing the parameter to control the behavior of the algorithm such that a model reaches a fair balance between bias and variance, as discussed in the previous sub-section. Given a relatively large number of observations, a low value of n can cause the model to fail to predict a significant share of the population so that it is not based on all available information. As the number of p features increases, a higher number of n is typically required to identify complex patterns. We therefore expect that more trees are required to exhaust variance when more information (features) is introduced in the model.

Since the random forest algorithm de-correlates trees such that they are trained independently, overfitting is less of an issue for the purpose of prediction. In fact, the authors of our R package (randomForest) state that random forest cannot overfit data, such that we can run as many trees as we want (Breimann & Cutler, 2022). Still, data tuning is beneficial in determining a satisfactory balance between the computational power required for running the algorithm and predictive power since increasing the number of n causes computational demand to grow exponentially. To determine a proper value of n, we plot the number of trees against error rates, as seen in the figure below for all sets of variables that we use in our study.

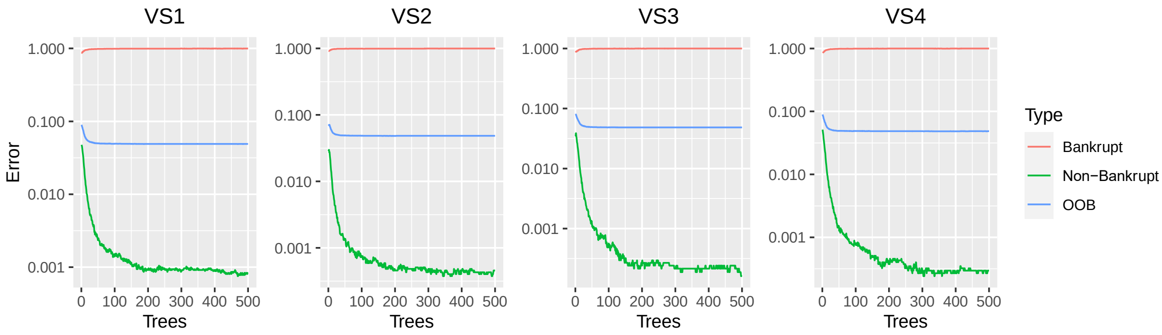


Figure 3: *Data tuning for random forest. We determine a reasonable number of trees by the convergence seen in the graphs. This corresponds to 100 trees for VS1 and VS2 and 200 trees for VS3 and VS4.*

As expected, more complexity in the set of variables requires more trees to exhaust the variance. The Out-Of-Bag error rate (OOB) is a simple method to assess model performance without performing k-fold cross-validation. Recall that bootstrapped samples are used when building the decision tree, such that approximately two-thirds of the dataset is used for building the decision tree, while the remaining third is an Out-Of-Bag sample. As such, we can predict the outcome of the OOB observations using each one of the trees and average the predicted responses, resulting in a classification error for the OOB samples across the n-number of trees. When the predictive power is reduced to the point of no incremental decrease per extra tree, there is little point in increasing the number of trees beyond this point (James, Witten, Hastie, & Tibshirani, 2021). For further procedure in our study, we use $n = 100$ trees for VS1 and VS2, and $n = 200$ trees for VS3 and VS4, as that represents a fair balance between the computational power required and fit of our model considering the variance error rates in the figure.

3.5 Measures of Evaluation and Validation

We apply evaluation tools and metrics to assess the models' predictive ability and ensure validity. Since all the tools used for evaluation are essential for the robustness of our results, we explain the fundamental concepts of evaluation and validation in this section.

3.5.1 Performance Metrics

Previous literature relies heavily on *overall accuracy* to assess classification models in bankruptcy predictions. It is a simple method of measuring the share of correctly predicted observations. The formula for calculating overall accuracy is summarized in the equation below.

$$OA = (TP + TN)/N \quad \textit{Eq. 9}$$

True positives, denoted as TP, is the number of correctly predicted positive states. True negatives, denoted as TN, are the number of correctly predicted negative states. N corresponds to the number of observations. Overall accuracy therefore measures the proportion of correctly predicted classifications. Although this performance metric is simple to understand and interpret, it is easy to manipulate a high overall accuracy when the target variable is unbalanced. Because bankruptcy is a rare event, a model can predict that no firms go bankrupt and achieve an overall accuracy equivalent to the share of non-bankruptcies. A more robust performance metric for evaluating classification models is the confusion matrix. A confusion matrix summarizes the correct and incorrect predictions, as shown in the table below.

Confusion Matrix

Predicted / Actual State	**Bankrupt**	**Non-Bankrupt**
Bankrupt	True positive (TP)	False positive (FP)
Non-Bankrupt	False negative (FN)	True negative (TN)

***Table 6**: Confusion Matrix comparing predictions against the true state of the observations. True positives and true negatives are correct predictions, while false negatives and false positives reflect incorrect predictions.*

As James et al. (2021) discuss, there is a trade-off between predicting true positives and true negatives. A false positive is often called a Type I error, while a false negative is a type II error. In credit analysis, for instance, the cost of making a Type II error can be much greater than the cost of making a Type I error. A Type II error enables the provision of credit to an unviable firm, risking a loss of the entire investment, while the cost of a Type I error is simply the opportunity cost of lost interest income. Although the values in a classification matrix are typically absolute, relative numbers can also be calculated using the matrix summarized in the table below.

Matrix of Error Rates

Predicted / Actual State	Bankrupt	Non-Bankrupt
Bankrupt	True positive rate (TPR)	False positive rate (FPR)
Non-Bankrupt	False negative rate (FNR)	True negative rate (TNR)

Table 7: *Matrix of error rates. Summarizing the share of correct predictions with all observations in that class*

The true positive rate (TPR), also known as sensitivity, is the share of correctly predicted positive states of total positive states. The false positive rate, (FPR) is the share of wrongly predicted positive states of actual negative states.[22] Likewise, the true negative rate (TNR) is the proportion of correctly negative states, while the false positive rate (FPR) measures the remaining misclassifications. For example, the TPR is the share of correctly predicted positive states, as shown below.

$$TPR = \frac{TP}{TP + FN} \qquad \text{Eq. 10}$$

These performance metrics are a crucial element of the assessment of our models. All of our models output probabilities between 0 and 1. MDA uses the predicted posterior probability that an observation belongs to the given class given the initial probability of bankruptcy, GLM and GAM output probabilities that must range between 0 and 1, and random forest (RF) uses bagging to determine the share of responses, essentially resulting in a probability score between 0 and 1. Consequently, if using a threshold rate of 10%, a firm with a predicted probability of default of 7% will not be classified as bankrupt, while a firm with 12% will be predicted as bankrupt. It must be noted, however, that optimizing a proper threshold depends on the cost of misclassification and is not simply to maximize our metrics.

There are several advantages of using continuous probability as opposed to linear scales. Firstly, the probabilities can be interpreted as credit ratings, allowing the generation of

[22] Corresponds to 1-Specificity

categorical variables. Secondly, it allows for the containment of data that provides a further "forwards- perspective". If a firm is classified as a false positive, it will likely have a greatly above-average probability of bankruptcy in the time period following the prediction period. Thirdly, our most insightful assessment tool (AUC scores) can be used since the output is continuous. Fourthly, because there exists a trade-off between the cost of misclassifications of false positives and false negatives, we can measure the total cost of misclassification over the different thresholds, which is advantageous if the cost of misclassification is known.

Several evaluation metrics are applicable to determine the optimal classification threshold value for bankrupt classification. Our study relies on established literature and uses the ROC curve as described in the next section to find an appropriate threshold. The threshold value is optimized by minimizing the Euclidean distance to the perfect point (0,1) on the ROC curve. We also use a second definition where we use a threshold that results in the actual share of bankruptcies reflecting the years 2010-2017.

3.5.2 ROC Evaluation Metric

One of the most common ways to measure the performance of classification models is the Receiver Operating Characteristics curve (ROC-curve) and its corresponding area under the curve (AUC).[23] The ROC curve plots the share of the true positive rate (y-axis) against the share of the false positives (x-axis). When the threshold is reduced, more firms will be classified as bankrupt. However, this increase in predicted bankruptcies must necessary both increase correctly predicted bankruptcies and falsely predicted bankruptcies. Therefore, the ROC curve will capture this payoff, as demonstrated in the figure below.

[23] AUC values in the ranges of [0.5-0.6], [0.6-0.7], [0.7-0.8], [0.8-0.9] and [0.9-1.0] are deemed as "unsatisfactory", "satisfactory", "good", "very good" and "excellent" respectively (James, Witten, Hastie, & Tibshirani, 2021). A value of 0.5 demonstrates no predictive power.

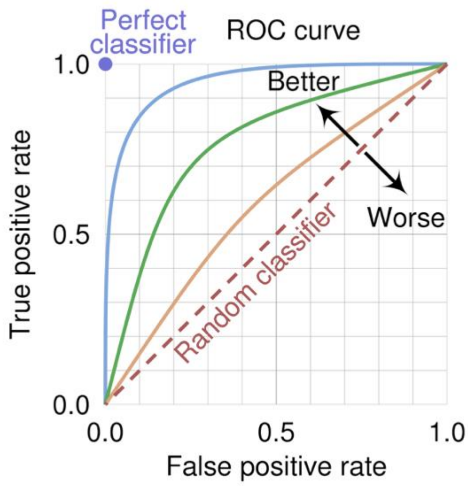


***Figure 4**. ROC curve. It demonstrates the share of true positives against the share of false positives. Source of figure is embedded as hyperlink.*

The ROC curve displays all possible thresholds between 0 and 1 and plots the associated TPR against the FPR. Because the TPR measures the share of correctly predicted positive observations, the measure will score 0% for a threshold of 1 (no positive predictions) and 100% for a threshold of 0 (all observations are predicted positive). Likewise, an FPR of 0% is achieved when no firms are predicted bankrupt, and a rate of 100% is achieved when all predictions are positive. Ideally, the perfect model will be capable of increasing the true positive rate without affecting the false positive rate, thus making no errors in positive predictions. A model therefore has greater predictive/classification ability the closer it is to the upper left corner. Consequently, one would ideally want to maximize the area beneath the curve. A naive approach assuming random guessing will yield a linear curve as displayed with the dotted line, demonstrating no class separation capacity. Models demonstrating AUC scores above 0.5 reflect models with predictive power beyond random guessing, as it surpasses the results of a naive approach. Our aim is therefore to find the model that maximizes AUC scores.

The major advantage of AUC as a performance metric is therefore to assess the models´ ability to measure the trade-off between misclassifications when varying the threshold for classification, which is especially important when the data is unbalanced.

3.5.3 Cross-Validation

Cross-validation is a model validation technique that uses the rotation of data to test models out-of-sample. A proper assessment of statistical models is necessary for robustness. Since our models are used for prediction purposes, as opposed to exploring causal connections, we are interested in how well the model fits previously unseen data. This is related to the variance-bias trade-off explained in section 3.5.4 (James et al., 2021). In short, if statistical models are fitted and tested on the same data, the simple way to ensure good results is by overfitting the model. Although the model would yield good results in-sample, the model would have low predictive power for observations out of sample and tend to identify patterns that do not exist. This is why the rotation of data ensures a more robust evaluation.

To examine how well the model performs on previously unseen data, one can split the full sample of observations into a training and testing set of observations. In the literature this is commonly referred to as a training-test split. The approach is to fit an algorithm such as MDA, GLM or RF based on the training data and apply the model to the remaining observations in the test data to measure how well it predicts a previously unseen sample.

The major advantage of the training-test split is its simplicity and intuitive approach to ensure robustness and proper validation. The disadvantage, however, depends on the methods of performing the initial split. Although the split is typically performed using random sampling, where random observations are assigned to the training and testing samples, there is a probability that either of the samples can be non-representative for the entire population.

A method to counter this disadvantage is k-fold cross-validation (k-fold CV), where the procedure is to vary the data being tested and trained. For any given k, the model will be trained on all k-1 fold(s) and tested on the remaining k^{th} fold. If the selected k is relatively high, the model will be fitted on many combinations of observations, while few model fits are performed if k is relatively low. A demonstrative figure explaining this procedure is attached below.

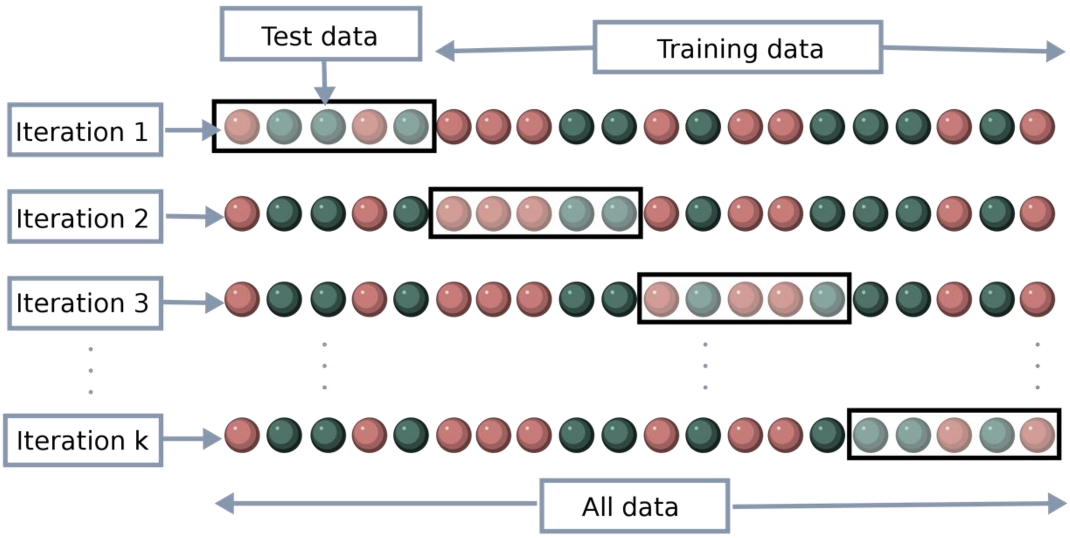


Figure 5. *K-fold validation demonstration. The figure demonstrates that the training- and testing data is varied. In our study, we let 80% of the data predict the remaining 20%. We iterate this process five times (5-fold validation) such that all observations are predicted. Source of figure is embedded as hyperlink.*

To assess the performance and stability of the model, evaluation metrics such as AUC and overall accuracy can be obtained from each of the k-iterations. To obtain a single measure of performance on each model, we calculate the AUC value across all folds simultaneously. We expect this procedure to result in more robust estimates and correct graphs. Using all folds simultaneously provides more information and causes less altering of data for evaluation compared to the average of folds or using one fold.[24] Additionally, this procedure also allows for producing ROC curves across all folds, which entails an evaluation of the whole sample.

[24] We note that we also attempted to use the average across folds but found no qualitative difference

3.5.4 Bias-Variance Tradeoff

The bias-variance is a crucial element in fitting statistical models. Essentially, one can choose the fit of a model to certain data, where it is desirable that it identifies important patterns in the data, while not identifying patterns that do not exist. Thus, too much variance in the output will cause models to identify patterns that do not exist, while too much bias will cause models to tend towards linearity. An optimal bias-variance balance is therefore needed to ensure a proper fit suited for out-of-sample predictions.

With respect to the purpose of validation and k-fold, the bias-variance tradeoff is represented in the choice of k-number of folds. Empirically it has been shown that a choice of k ranging between 5 and 10 has proven to be effective to yield relatively low error rate estimates while not suffering excessively from neither high variance nor bias. Due to the computational power associated with k-fold cross-validation, we utilize the lowest possible value of a reasonable k, namely 5, in accordance with established literature (James et al., 2021).

3.6 Classification Model Performance

In this section, we evaluate the performance of the different models, meaning all combinations of classifiers and variable sets. We structure thissection by the algorithms used: MDA, logistic regression, GAM, and random forest. Within each category of classifiers, performance metrics are reviewed for all the aforementioned variable sets of VS1, VS2, VS3, and VS4 described in section 3.3. We present and discuss every model's performance on all folds simultaneously to avoid excessive tables and figures but summarize the statistics from each fold in Appendix 1-4.[25] In addition to presenting the performance within each fold, we run diagnostics across all folds simultaneously against the true observations. This ensures a robust evaluation of all out-of-sample predictions in our 5-fold validation. We will choose our preferred model based on the highest AUC value from all classifiers and use the McNemar´s test in line with Næss et al. (2017) to examine model superiority.

[25] To improve control of result outputs, we refrained from using packages in R to perform 5-fold validation, producing proprietary functions instead.

3.6.1 Multiple Discriminant Analysis (MDA)

Altman (1968) initially used a linear version of MDA and the variables corresponding to VS1. Motivated by his contribution, we re-estimate the coefficients to examine if the model yields sufficient predictive power. Additionally, we also test the MDA on our remaining sets of variables. We use the package "Mass" in R by Venables and Ripley (2002) to proceed with estimations using the linear approach in line with Altman (1968). Our results demonstrate differences in performance metrics between the choice of variable sets.

MDA Performance Metrics					
Variable Set / Metric	**AUC**	**Accuracy**	**Cutpoint**	**Sensitivity**	**Specificity**
Altman´s Ratios (VS1)	0.762	0.694	0.033	0.721	0.693
Adjusted Altman (VS2)	0.758	0.693	0.062	0.731	0.691
SEBRA Inspired (VS3)	0.815	0.719	0.036	0.77	0.716
SEBRA Plus (VS4)	0.817	0.735	0.038	0.761	0.734

***Table 8**: MDA Performance Metrics. The table displays the performance of MDA on each variable set. AUC refers to the integral of the ROC curve, and accuracy measures the overall share of correct predictions. Cutpoint is the threshold used to classify a bankruptcy based on the Euclidean distance to (0,1) from the ROC-curve.[26] Sensitivity refers to the true positive rate, while specificity is the true negative rate. As observed, VS4 is the highest performing model in terms of AUC and is therefore our preferred model within this category of classifiers.*

The performance metrics from our model demonstrate a considerable difference in performance between variable sets 1 and 2, and variable sets 3 and 4. Altman´s original model yielded a decent performance with an AUC of 0.762 (Altman, 1968). However, we note that the performance metrics are not comparable, considering the difference in underlying data. Moreover, the confusion matrices are summarized visually below and demonstrate similar characteristics across the sets of variables.

[26] Euclidean distance as calculated by $d(x, y) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$

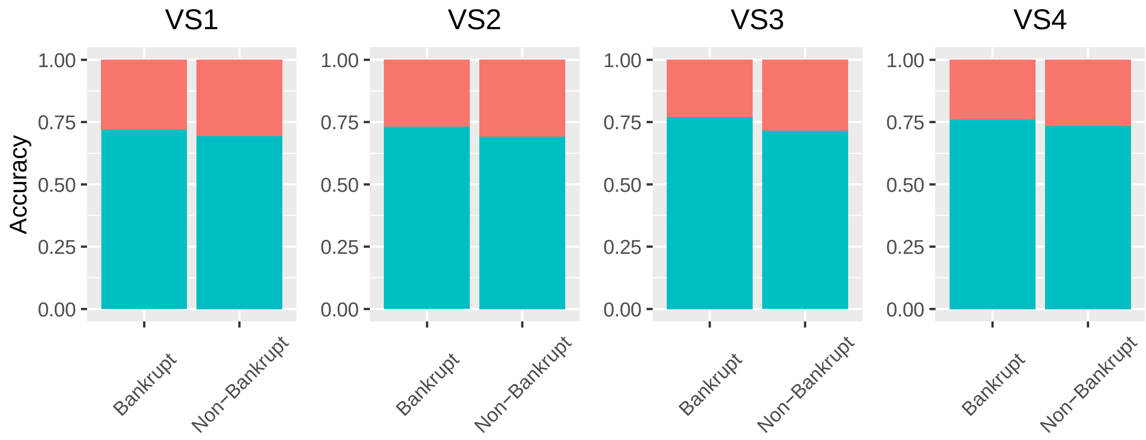


***Figure 6**: MDA Visual Classification Matrices. VS1-VS4 refers to the variable sets described in section 3.3. The columns show the shares of correct classifications in blue and misclassifications in red for actual bankrupt and non-bankrupt observations groups. The AUC-optimizing roc01 function determines the threshold used.*

A comparison between the sets of variables demonstrates that variable set 2 performs worse than Altman´s initial variables based on the AUC value.[27] This can appear surprising given the criticism by Kinserdal et al. (2019). However, it must be noted that we study a specific industry with many small businesses, which contrasts with other contributions to the literature that commonly filter these out. Additionally, much of the criticism addresses traditional liquidity measures, as financial assets are pointed out as a better indicator than working capital. Among the firms in our dataset, holding a large amount of financial assets is far less common because the businesses are negligible in size compared to firms more commonly discussed in the bankruptcy prediction literature. The applicableness of the adjusted Altman´s variables (VS2) variables is therefore debatable for our SMEs. For the interested reader, we also examined the performance across all industries at once.[28] In terms of significance among our MDA models, we determine that the SEBRA Plus model (VS4) outperforms all others in this section based on a p-value of less than 1% when using McNemar´s test. Below, we provide a visual summary of our preferred model.

[27] We note that the difference in performance is not significant. The p-value returned from McNemar´s test is 22.3%.

[28] In Appendix 5, we determine that the alternative ratios outperform Atlman´s initial variables when we review the models across all industries at once. This suggests that the alternative ratios in general seem appropriate, but are not suitable for our industry.

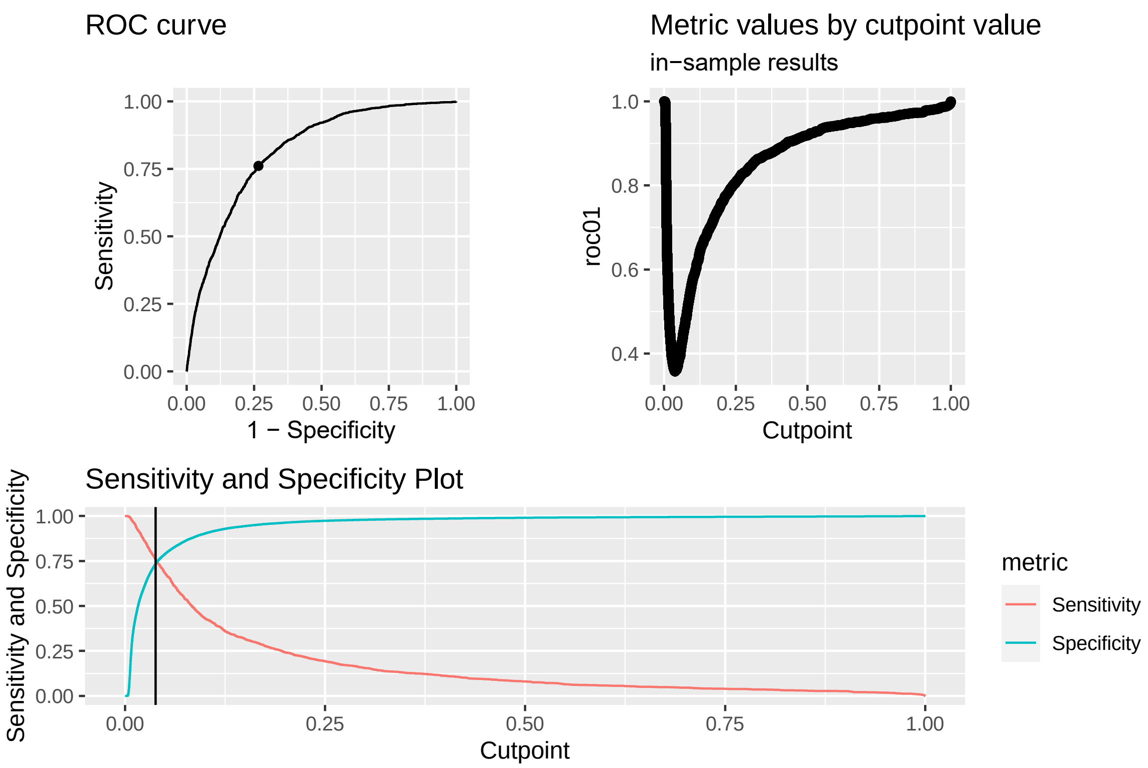


***Figure 7**: MDA Visual Performance Metrics. The ROC curve is the most central performance metric, measuring the share of true positives against the share of false positives. "Metric values by cutpoint value" shows the distance to (0,1) from the ROC curve at each threshold and demonstrates an optimal threshold at 0,038. Cutpoint is the same as the threshold for the classification of bankruptcy. The sensitivity and specificity plot demonstrates the trade-off between true positives and negatives over all possible thresholds. Note that the interaction between the curves in this graph does not necessarily correspond to the optimal threshold based on our metric function to minimize the Euclidean distance. Sensitivity refers to the true positive rate, while specificity is the true negative rate. 1-Specificity is the false positive rate.*

3.6.2 Logistic Regression

Logistic regression to predict bankruptcies gained popularity following Ohlson (1980) and has the advantage of its applicability to real-world applications and interpretable coefficients, although it is vulnerable to multicollinearity in interpreting the coefficients. In line with the logic expressed by Berg (2007), however, this is less of an issue in our study where the purpose is prediction only. Nevertheless, our k-fold validation will reveal any weaknesses in predictive power. The algorithm is gathered from the integrated package "Stats" in R. Our performance metrics are summarized below.

Logistic Regression Performance Metrics					
Variable Set / Metric	**AUC**	**Accuracy**	**Cutpoint**	**Sensitivity**	**Specificity**
Altman´s Ratios (VS1)	0.775	0.726	0.051	0.726	0.726
Adjusted Altman (VS2)	0.755	0.683	0.075	0.741	0.680
SEBRA Inspired (VS3)	0.806	0.697	0.043	0.777	0.693
SEBRA Plus (VS4)	0.808	0.713	0.046	0.763	0.711

Table 9: *Logistic Regression Performance Metrics. The table displays the performance of logistic regression on each variable set. AUC refers to the integral of the ROC curve, and accuracy measures the overall share of correct predictions. Cutpoint is the threshold used to classify a bankruptcy based on the Euclidean distance to (0,1) from the ROC curve. Sensitivity refers to the true positive rate, while specificity is the true negative rate. As observed, VS4 is narrowly the highest performing model in terms of AUC.*

The logistic model demonstrates performance similar to that of MDA. Based on AUC scores, logistic regression is inferior across all sets of variables except for VS1 based on the AUC value. In line with the MDA results, VS4 performs the best while VS2 performs the worst in our 5-fold cross-validation. The optimal threshold is somewhat higher for logistic regression, suggesting that the distribution in bankruptcy probabilities is generally higher. The confusion matrices are summarized visually below.

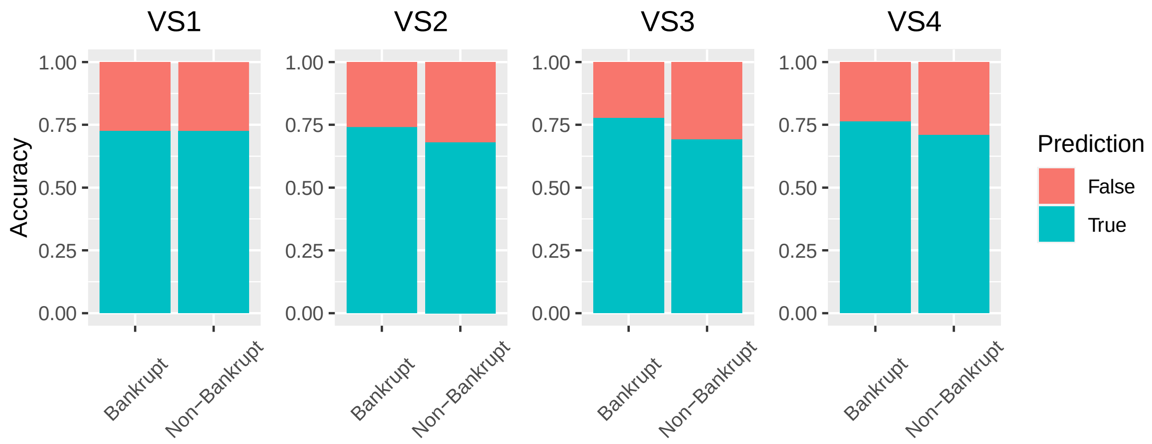


Figure 8*: Logistic Regression Visual Classification Matrices. VS1-VS4 refers to the variable sets described in section 3.3. The columns show the shares of correct classifications in blue and misclassifications in red for actual bankrupt and non-bankrupt observations groups. The AUC-optimizing roc01 function determines the threshold used.*

The results demonstrate very similar characteristics to that of MDA. In terms of significance among our logistic models, we determine that the SEBRA Plus model (VS4) outperforms all others in this section based on a p-value of less than 1% when using McNemar´s test. Below, we provide a visual summary of our preferred model.

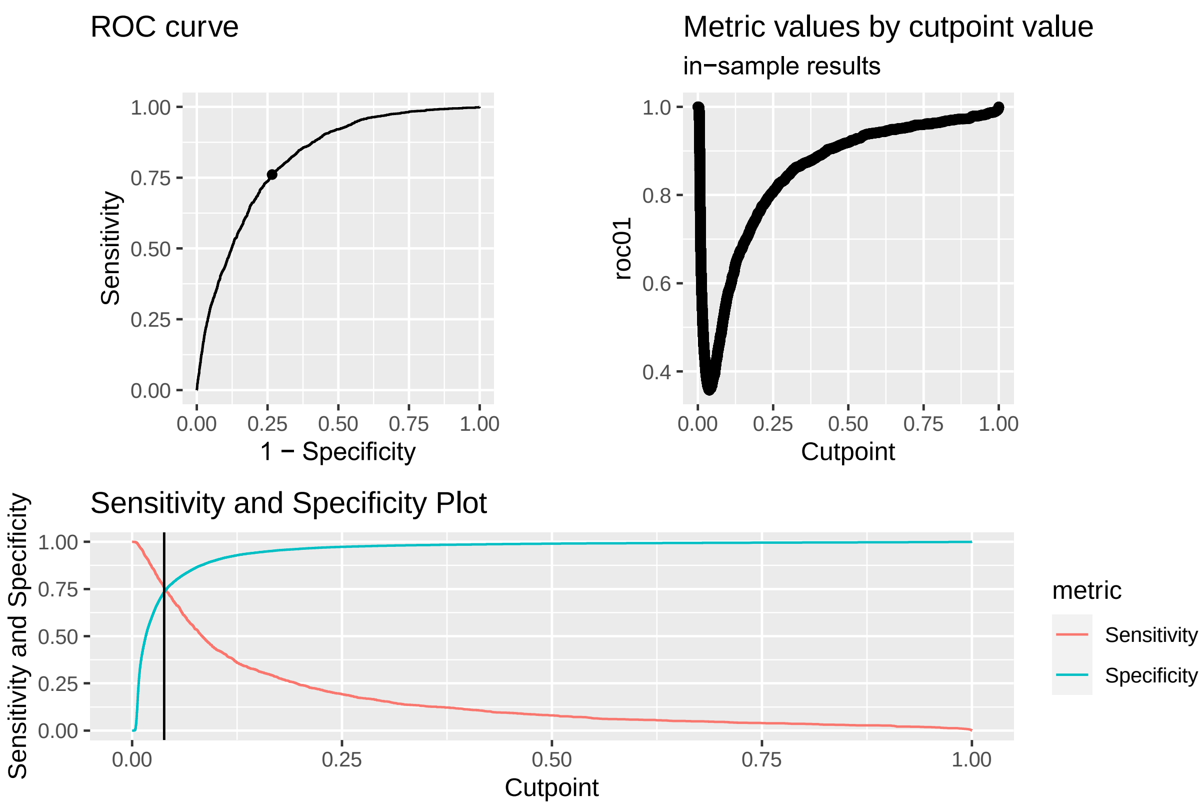


***Figure** 9: Logistic Regression Visual Performance Metrics. The ROC curve is the most central performance metric, measuring the share of true positives against the share of false positives. "Metric values by cutpoint value" shows the distance to (0,1) from the ROC curve at each threshold and demonstrates an optimal threshold at 0,038. Cutpoint is the same as the threshold for the classification of bankruptcy. The sensitivity and specificity plot demonstrates the trade-off between true positives and negatives over all possible thresholds. Note that the interaction between the curves in this graph does not necessarily correspond to the optimal threshold based on our metric function to minimize the Euclidean distance. Sensitivity refers to the true positive rate, while specificity is the true negative rate. 1-Specificity is the false positive rate.*

3.6.3 GAM

GAM has the advantage over logistic regression that it can identify non-linear patterns. Literature within bankruptcy prediction has also demonstrated the superiority of this model (Næss et al., 2017). Continuous variables in our dataset are squared. To treat the data and perform the predictions, we used the package "gam" in R. The results are summarized below.

GAM Performance Metrics					
Variable Set / Metric	**AUC**	**Accuracy**	**Cutpoint**	**Sensitivity**	**Specificity**
Altman´s Ratios (VS1)	0.776	0.74	0.05	0.708	0.742
Adjusted Altman (VS2)	0.769	0.697	0.064	0.744	0.695
SEBRA Inspired (VS3)	0.811	0.749	0.053	0.733	0.75
SEBRA Plus (VS4)	0.815	0.725	0.046	0.776	0.723

Table 10*: GAM Performance Metrics. The table displays the performance of GAM on each variable set. AUC refers to the integral of the ROC curve, and accuracy measures the overall share of correct predictions. Cutpoint is the threshold used to classify a bankruptcy based on the Euclidean distance to (0,1) from the ROC-curve. Sensitivity refers to the true positive rate, while specificity is the true negative rate. As observed, VS4 is the highest performing model in terms of AUC.*

GAM is superior to logistic regression across all thresholds, yielding better AUC scores. The optimal threshold is similar in the distribution, and the ranking between sets of variables is again similar to that of logistic regression and MDA since VS4 performs the best and VS2 performs the worst. The confusion matrices are summarized below.

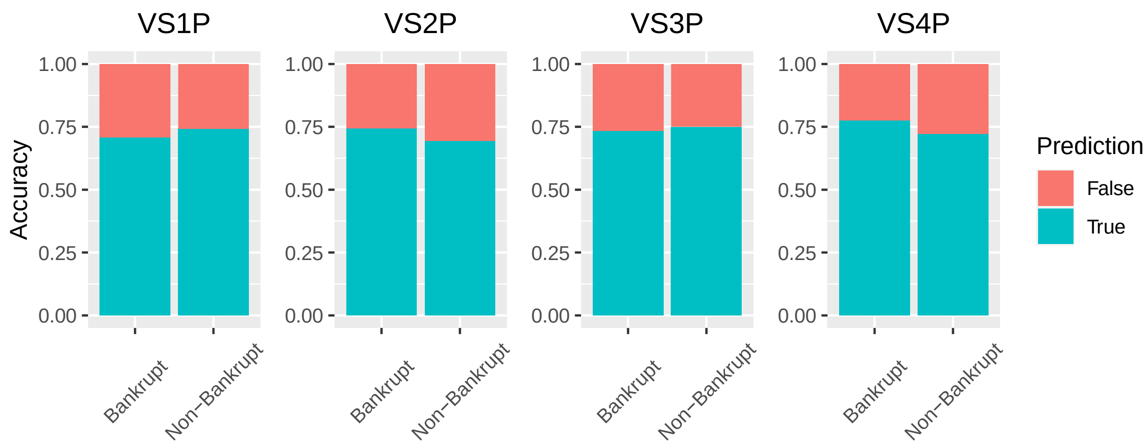


Figure 10: *GAM Visual Classification Matrices. VS1-VS4 refers to the variable sets described in section 3.3. The columns show the shares of correct classifications in blue and misclassifications in red for actual bankrupt and non-bankrupt observations groups. The AUC-optimizing roc01 function determines the threshold used. The variable sets carry the suffix "P" because the variables are squared to allow for non-linear relationships.*

Interestingly, GAM is better in predicting true negatives for Altman´s initial variables, but the other sets of variables share very similar characteristics to that of logistic regression. In terms of significance among our GAM models, the SEBRA Plus model (VS4) outperforms all others based on a p-value of less than 1% when using McNemar´s test. Below, we provide a visual summary of our preferred model.

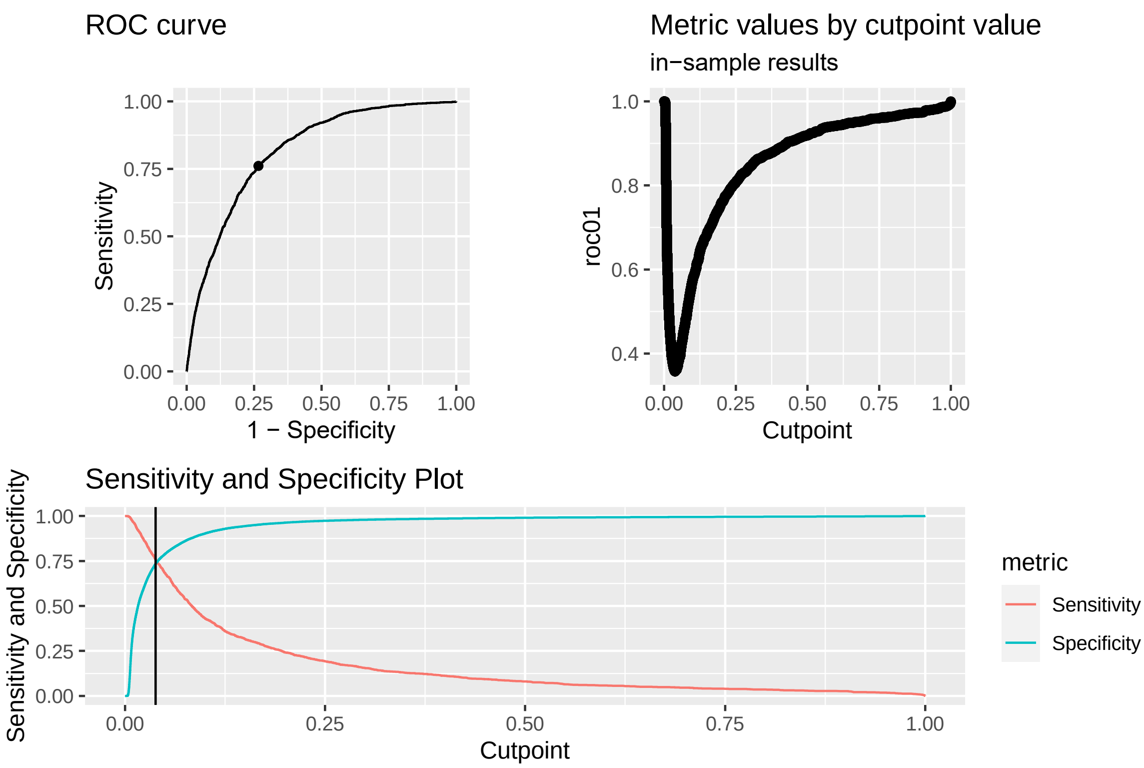


Figure 11: *GAM Visual Performance Metrics. The ROC curve is the most central performance metric, measuring the share of true positives against the share of false positives. "Metric values by cutpoint value" shows the distance to (0,1) from the ROC curve at each threshold and demonstrates an optimal threshold at 0,038. Cutpoint is the same as the threshold for the classification of bankruptcy. The sensitivity and specificity plot demonstrates the trade-off between true positives and negatives over all possible thresholds. Note that the interaction between the curves in this graph does not necessarily correspond to the optimal threshold based on our metric function to minimize the Euclidean distance. Sensitivity refers to the true positive rate, while specificity is the true negative rate. 1-Specificity is the false positive rate.*

Interestingly, the sensitivity of GAM seems somewhat sharper than that of logistic regression, in line with the discussion above. In conclusion, VS4 performs the best among the sets of variables and VS4 and VS3 are superior to VS1 and VS2.

3.6.4 Random Forest

Random Forest serves represents our second non-parametric method. The major advantage of using random forest is its flexibility, ability to detect complex patterns, and treatment of correlation between the variables such that overfitting is less of an issue. In line with the discussion in section 3.4.3, we use n=100 trees for VS1 and VS2, and n=200 trees for VS3 and VS4. To perform the predictions, the R package "randomForest" was used. The performance metrics are summarized below.

Random Forest (RF) Performance Metrics					
Variable Set / Metric	**AUC**	**Accuracy**	**Cutpoint**	**Sensitivity**	**Specificity**
Altman´s Ratios (VS1)	0,769	0,696	0,048	0,713	0,695
Adjusted Altman (VS2)	0,727	0,661	0,03	0,686	0,66
SEBRA Inspired (VS3)	0,824	0,734	0,056	0,768	0,732
SEBRA Plus (VS4)	0,829	0,734	0,062	0,786	0,732

Table 11: *Random Forest Performance Metrics. The table displays the performance of random forest on each variable set. AUC refers to the integral of the ROC curve, and accuracy measures the overall share of correct predictions. Cutpoint is the threshold used to classify a bankruptcy based on the Euclidean distance to (0,1) from the ROC-curve. Sensitivity refers to the true positive rate, while specificity is the true negative rate. As observed, VS4 is the highest performing model in terms of AUC.*

Our RF model is inferior to GAM when using variable sets 1 and 2 but appears superior using variable sets 3 and 4 based on AUC. This aligns with our expectations, as complex patterns and interactions are detected better with more advanced techniques. The threshold is somewhat higher than the other models, suggesting a distribution of bankruptcy predictions different to that of MDA, logistic regression and GAM. The confusion matrices are summarized below.

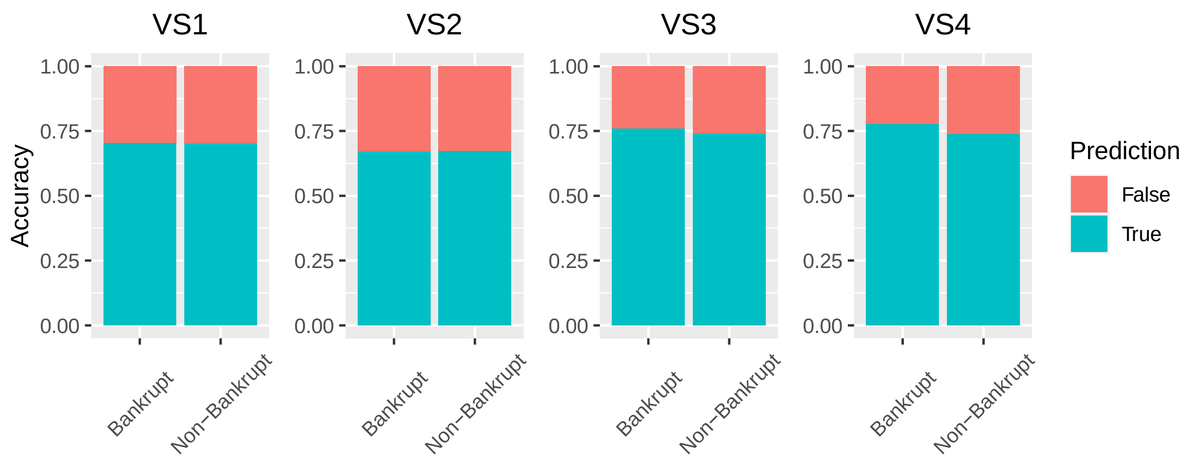


Figure 12*: Random Forest Visual Classification Matrices. VS1-VS4 refers to the variable sets described in section 3.3. The columns show the shares of correct classifications in blue and misclassifications in red for actual bankrupt and non-bankrupt observations groups. The AUC-optimizing roc01 function determines the threshold used.*

Compared to MDA, Logistic Regression and GAM, RF seems to predict a more stable share of true positives and negatives across all sets of variables, based on our threshold determined by minimizing Euclidian distance to (0,1) from the ROC-curve. Regarding significance among our RF models, the SEBRA Plus model (VS4) outperforms all others based on a p-value of less than 1% when using McNemar´s test. Below, we provide a visual summary of our preferred model.

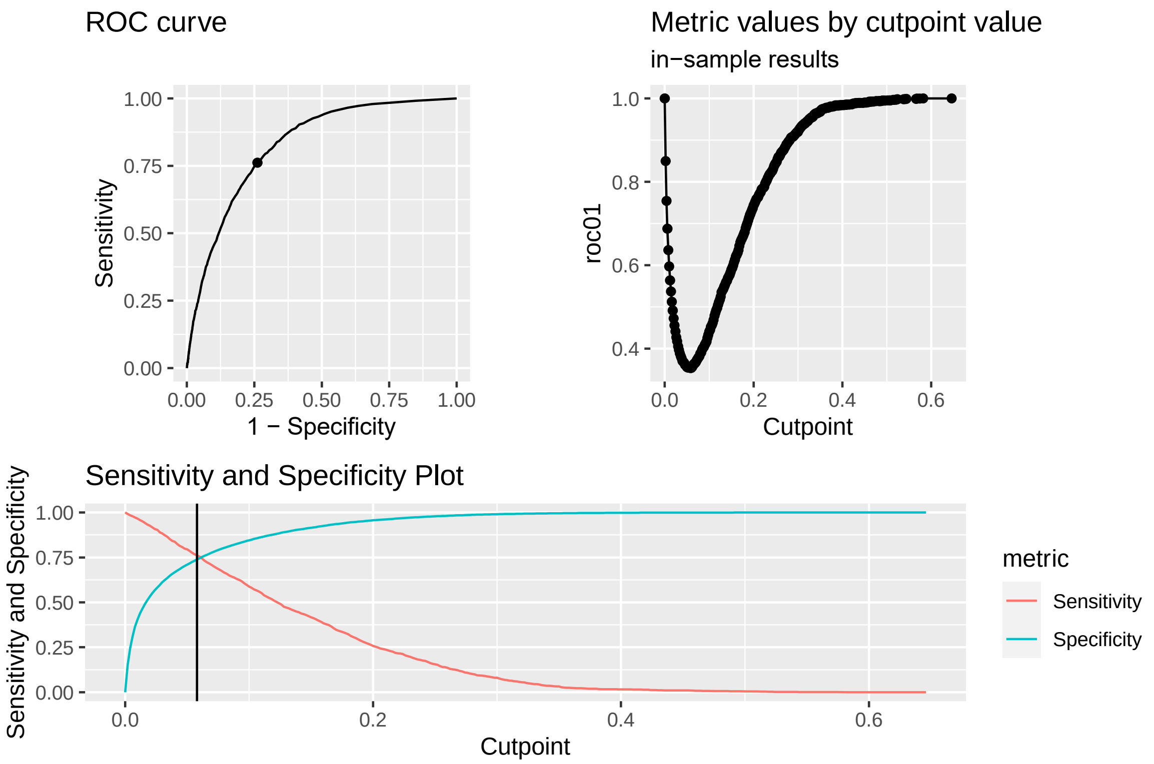


***Figure 13**: Random Forest Visual Performance Metrics. The ROC curve is the most central performance metric, measuring the share of true positives against the share of false positives. "Metric values by cutpoint value" shows the distance to (0,1) from the ROC curve at each threshold and demonstrates an optimal threshold at 0,038. Cutpoint is the same as the threshold for the classification of bankruptcy. The sensitivity and specificity plot demonstrates the trade-off between true positives and negatives over all possible thresholds. Note that the interaction between the curves in this graph does not necessarily correspond to the optimal threshold based on our metric function to minimize the Euclidean distance. Sensitivity refers to the true positive rate, while specificity is the true negative rate. 1-Specificity is the false positive rate.*

3.7 Model Performance Discussion

3.7.1 Performance Comparison to other Studies

In general, our performance metrics demonstrate that increased complexity in variables and classifier improves out-of-sample predictive power. We determine that the non-parametric algorithm random forest outperforms the other classifiers and that the SEBRA Plus variables outperform all other variable sets within each classifier.[29] In line with Næss et al. (2017), we find that the choice of the statistical models matters less than the choice of variables.

Based on nominal values of AUC, we determine that a score of 0.83 falls within the range of “very good,” as explained by James et al. (2021). For this reason, we determine that the predictive ability of our preferred model is satisfactory at a nominal level.

Compared to the literature, Næss et al. (2017) found that GAM yielded the best results with an AUC score of 0.911. They also explain that Bernhardsen and Larsen (2007) from Norges Bank achieved AUC scores of 0.88 and 0.89 using the two respective SEBRA models. Pelja and Stemland’s study (2017) achieved an AUC score of 0.888 using Altman´s variables. Other studies, such as Zhang and Ye (2019), achieved 0.66 for logistic regression and 0.67 for RF, while Messe and Viken (2019) achieved AUC scores between 0.57 and 0.78 with the same classifiers as in this study, but with other variables. In light of these studies, our models and variables indicate a satisfactory overall performance with an AUC of 0.83. Still, it is worth noting that none of the other studies are directly comparable due to the chosen industry, geographical location, and size range among the firms.

Although not examined thoroughly, the differences in data balance might provide an edge to studies using balanced data. Prior studies tend to favor datasets with balanced target variables ranging between a relative share of 30%-50% of bankruptcies in both the training- and testing data. In contrast, our share of bankrupt observations was 4.84%. Another source of disparity is the bankruptcy definition, making it difficult to confirm if studies are directly comparable.

[29] Based on the p-value from the McNemar´s test, we find that the Random Forest model using SEBRA Plus (VS4) variables outperform all other preferred models within each category of classifiers. The p-value was less than 1% in all of these comparisons.

In summary, we conclude that our preferred model performs reasonably well. The strict filters and limitations on size and revenue likely improve performance in some of the established literature. In general, financial statements for small firms are expected to exhibit a large amount of data noise, lowering the predictive power (Kinserdal et al., 2019). Despite the expected noise, we still believe that the models' external validity and utility would decrease by omitting small firms in our dataset. One other concern using random forest was overfitting. Fortunately, our 5-fold cross-validation demonstrates robustness across the folds and sufficient AUC values, and the classifier´s ability to de-correlate the trees seems to counter overfitting.

When comparing our performance metrics to other studies, one significant aspect to consider is our sector limitation in the hospitality industry. On the one hand, one would imagine that firms are more homogeneous within a sector than across sectors. For instance, profitability and working capital vary immensely across industries, so comparing key figures can appear unintuitive. Limiting the scope should therefore improve the accuracy and predictive power of the models in our study. However, the hospitality industry has a more significant share of smaller firms than other industries. Smaller firms are typically vulnerable to more noise in the financial statements, pulling in the direction of lower predictive power. Therefore, the effect of smaller firms pulls in the opposite direction than the industry's homogeneity. The net effect is unclear, and we therefore examined if our predictive models would yield stronger or weaker power within a range of industries.[30] According to these inspections, the hospitality industry is neither the most favorable nor unfavorable for high accuracy in bankruptcy prediction.

3.7.2 Variable Importance

Variable importance is the term for individual variables' contribution in predicting the target variable. As discussed in section 3.4, the random forest algorithm uses decision trees to reduce the impurity of the nodes where all observations belong to a single category. We can generate a measure for variable importance by calculating the mean decrease in Gini for each variable,

[30] For the interested reader, we examined the range of predictive power within different industries using data between 2010 and 2017. Appendix 5 summarizes performance metrics within industries and across all industries simultaneously.

weighted by the number of observations reaching the node (James et al., 2021). Effectively, a practical interpretation is that the mean decrease in Gini measures how important a given variable is to estimating the target variable. A higher mean decrease in Gini is associated with higher importance because it reduces node impurity. A lower mean decrease in Gini is consequently associated with lower importance for estimating our target variable. The variables of the highest importance in the random forest classifier on SEBRA Plus variables (VS4) are provided below.[31]

Variable Importance Random Forest	
Variable and Description	**Mean Decrease in Gini**
$N_5 = \frac{debt\ to\ the\ public}{sum\ of\ assets}$	294.88
$N_2 = \frac{equity}{Sum\ of\ assets}$	289.09
$N_4 = \frac{accounts\ payable}{sum\ of\ assets}$	289.06

Table 12: *Variable Importance, top 3 variables from the SEBRA Plus variables (VS4) using Random Forest. Variable importance is determined through the mean decrease in Gini explained in the current section. A higher mean decrease in Gini is associated with higher importance for predicting the target variable (bankruptcy).*

In line with the previous reasoning in 3.2.3, N_5, $N_2$and N_4 are the most important variables in our random forest model. Interestingly, the most important variable measures the share of public debt to the total assets of a firm. This finding is of value for the Norwegian Tax Administration, as it provides an indication their role as a creditor.[32] Consequently, any lack of initiative from the tax authorities will likely reduce bankruptcies.

3.7.3 Choice of Model

In summary, we determine that the SEBRA Plus variables (VS4) provides the most predictive power based on our performance metrics. This is reflected across all classifiers based on the AUC scores and corresponding p-value from the McNemar´s test, following Næss et al. (2017). Comparing the preferred models from each classifier, we find that the random forest

[31] See Appendix 6 for a complete ranking of variable importance

[32] While the mean decrease in gini implies what variables matter for predicting the target variable of bankruptcy, it is not a causality estimation and must therefore only be interpreted as an indication.

Model using the SEBRA Plus variables outperforms all of MDA, logistic regression and GAM based on AUC-values. This model superiority is supported by calculating the p-value from McNemar´s test, where p-values were less than 1% in all comparisons. In conclusion, we are confident that the random forest classifier using the set of SEBRA Plus variables (VS4) provide a sufficient model for our analysis in section 4.[33]

[33] The Random Forest model used for predicting bankruptcy risk among the compensation recipients is trained on all observations from 2010-2017. Although several models are developed and applied using the k-fold validation, the correct procedure is to use all available data when training the final classifier. This is because k-fold validation is only a technique for validation, not model building (James et al., 2021).

4. Compensation Analysis

As described, our methodology is twofold. From section 3, we have developed a bankruptcy prediction model, and the output of that model is the foundation for the compensation analysis in this section. Specifically, we aim to answer questions from section 1.3 with the quantified viability indicator, to confirm or reject the existence of a relationship to compensation intensity.

4.1 Data Foundation

In line with our research question, we inspect key relationships between compensation, firms, and bankruptcy risk. For this purpose, we use the financial statements from SNF and the dataset of granted compensation from the Brønnøysund Register Centre (2022) to conduct our analysis, as mentioned in section 3.2.1. Ideally, we would be able to assess the bankruptcy risk of all compensation recipients. However, the two datasets are not fully compatible, and we can only assess the bankruptcy risk of firms corresponding to 89% of granted compensation. After merging the data, three main categories of observations remain. (1) Observations solely present in the SNF dataset we call Group 1. This group corresponds to firms we assessed for bankruptcy risk but are not compensation recipients. (2) Observations solely present in the granted compensation dataset are labeled Group 2. These are compensation recipients where we lack financial information from SNF. (3) Observations matched in the two datasets are Group 3. This group is the target of our analysis. However, the existence of the other groups constitutes a limitation for results interpretation.[34] Within the hospitality industry, we identify 5160 firms in the compensation dataset, receiving a total sum of 2.409 billion NOK. From the SNF dataset, we identify 5931 firms. Among these, 3747 received compensation. The intersection of the dataset is illustrated in the table below.

[34] Differences between the groups of observations are included in Appendix 9.

Distribution of Compensation by Dataset			
Dataset	SNF	Intersect	Compensation
N firms	5931	3747	5160
Group ID	Group 1	Group 3	Group 2
Denotation	SNF ∩ Compensation'	SNF ∩ Compensation	SNF' ∩ Compensation
N firms	2184	3747	1413
Compensation in MNOK		2409.7	286.5

Table 13*: Overlap between SNF and Compensation Datasets. It displays the number of firms, corresponding compensation from each dataset, and the intersecting group.*

As stated in section 3.7.3, the chosen model for bankruptcy prediction is random forest using the set of SEBRA Plus variables (VS4). Because of the bankruptcy definition used in model training, we remind that both the binary output and the continuous variable refer to the prediction/likelihood of a firm being declared bankrupt within three years.

We summarize variables in Table 14 grouped by bankruptcy risk in deciles and a binary bankruptcy variable. The binary variable uses a conservative threshold value (bankruptcy definition 1) of 0.169 so that the share of bankruptcies matches the last decade's average. The use of the table is motivated by distinguishing viable and non-viable firms through inspecting the relationship between compensation, bankruptcy risk, and other firm characteristics. We emphasize that the bankruptcy risk score is an approximated bankruptcy probability seen from 2018-12 and that the compensation data is empirical.

Analysis of Compensation grouped by Bankruptcy Risk												
Description / Bankruptcy Group	Deciles of Bankruptcy Risk										Definition 1	
	1	2	3	4	5	6	7	8	9	10	0	1
Environment-specific Information												
Number of Firms	375	375	375	375	375	375	375	374	374	374	3566	181
Sum of Compensation MNOK	280.9	261.8	487.8	546.3	198.3	170.6	138.4	110.1	103.3	112.2	2355.4	54.3
Sum of Sales MNOK	7361.9	5959.3	9955.0	10699.7	5309.1	3867.9	3793.8	2939.9	2405.9	2006.6	53485.2	813.9
Sum of Employees	12167	11006	14730	16644	8890	8976	8602	6769	6443	5793	97421	2599
Averages per Firm												
Average Compensation in TNOK	749.0	698.1	1300.7	1456.8	528.8	455.0	369.1	294.4	276.2	299.9	660.5	299.7
Average Revenue in TNOK	19631.7	15891.4	26546.7	28532.6	14157.5	10314.3	10116.7	7860.6	6433.0	5365.4	14998.7	4496.5
Average Number of Employees	32.4	29.3	39.3	44.4	23.7	23.9	22.9	18.1	17.2	15.5	27.3	14.4
Average Book Value of Equity	4705.2	3393.9	2654.5	4633.9	1301.5	923.0	262.4	-532.6	-564.5	-1158.4	1704.8	-1214.9
Average TNOK in EBITDA (last 3 years)	1731.6	1580.8	2086.5	1892.6	643.6	387.8	444.3	-111.3	-224.9	-379.4	873.0	-514.0
Average Centrality Score	5.0	4.9	4.5	4.1	4.3	4.1	4.1	3.9	3.8	3.7	4.3	3.3
Average Bankruptcy Risk	0.00%	0.14%	0.35%	0.68%	1.20%	2.04%	3.53%	5.77%	9.51%	18.58%	3.21%	23.22%
Average Age	16.63	14.97	13.55	11.30	11.06	8.90	7.56	6.07	5.89	3.94	10.34	3.00
Ratios for Analysis												
Average Compensation Intensity 1	0.037	0.042	0.035	0.041	0.038	0.043	0.041	0.062	0.057	0.117	0.045	0.171
Median Compensation Intensity 1	0.022	0.023	0.020	0.020	0.023	0.023	0.024	0.025	0.024	0.028	0.023	0.030
Average Compensation/employee in TNOK	22.61	23.17	20.34	24.58	21.64	21.50	25.62	18.10	19.47	23.00	21.76	26.91
Average Compensation to Labor Costs	0.09	0.11	0.09	0.09	0.11	0.11	0.15	0.14	0.16	0.25	0.12	0.34
Median Growth Rate*	1.07	1.06	1.07	1.06	1.10	1.07	1.06	1.08	1.09	1.11	1.07	1.11
Pandemic Impact Proxy*	0.18	0.25	0.19	0.19	0.18	0.19	0.17	0.18	0.21	0.21	0.20	0.20
Average Public Debt (balance sheet)	8.67%	9.70%	11.45%	12.48%	12.99%	14.11%	16.97%	17.93%	21.27%	32.05%	14.64%	37.81%
Average Account Payables (balance sheet)	6.59%	7.47%	9.42%	9.93%	11.80%	12.12%	16.18%	18.76%	29.84%	54.52%	15.02%	69.43%

Table 14: *Analysis of Compensation in Deciles. Deciles are ordered from low bankruptcy risk (1) to high (10). Definition 1 refers to the binary classification of bankruptcy using a conservative and historically accurate threshold. The variables denoted with * refer to approximated variables retrieved from the compensation application dataset. These have lower data quality and should be interpreted with caution.*

4.2 Analysis and Results

Compensation Granted

One of the most specific ways of evaluating the distribution of the compensation scheme is by estimating the sum granted to predicted bankrupt firms. This spending contributes to keeping unviable firms alive, harming allocation efficiency, and entailing an opportunity cost of supporting more suitable recipients. In Table 14 we show granted compensation to classified bankrupt firms with the conservative threshold value. 54.3 million NOK of compensation funds were received by classified bankrupt firms based on this estimate. The sum constitutes a 1.58 percent share of total compensation granted among the firms analyzed. Calculating the threshold-independent expected value of compensation to bankrupt firms results in 53.3 million NOK. For these reasons, we find the estimate of 54.3 million NOK convincing.[35]

Since estimates are drawn from specific thresholds, we find it suitable to present the sum of compensation granted to bankrupt firms as a function of the classification threshold. Figure 14 establishes this relationship. The x-axis shows the minimum probability of bankruptcy required by the model to classify it as a binary output bankruptcy. The y-axis summarizes the corresponding sum of compensation to the predicted bankrupt firms at each threshold level.

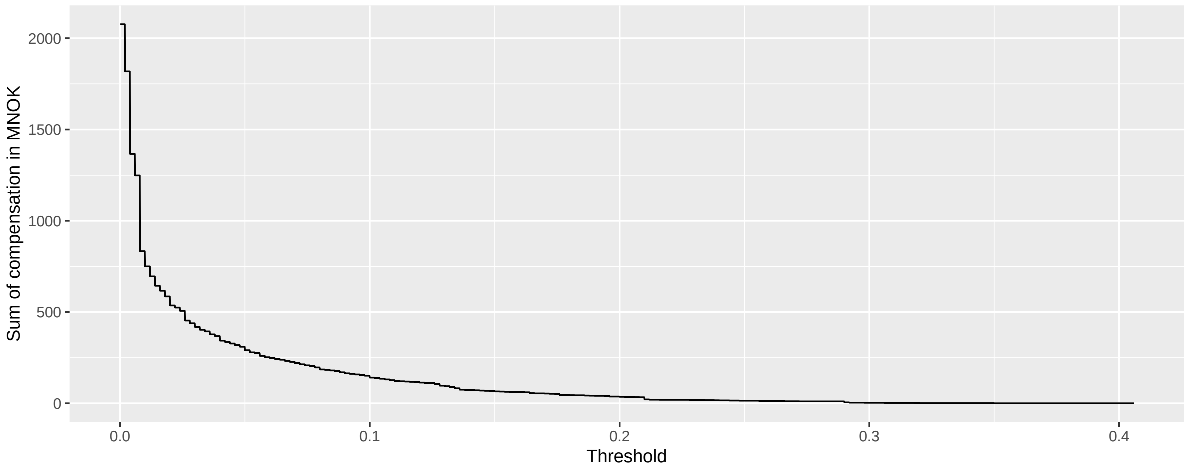


Figure 14: *Compensation over Threshold. The x-axis shows the threshold value necessary for the bankruptcy classification, while the y-axis shows the corresponding total compensation granted to predicted bankrupt firms.*

[35] Sum of individual bankruptcy probability multiplied with compensation granted, so that $E(x) = \sum x_i p(x_i)$

From Figure 14, it is clear that the key figure of granted compensation to predicted bankruptcies is dependent on the threshold used. This dependence introduces the potential for subjectivity in the results because the researcher can choose the threshold. We chose the threshold corresponding to historical shares of bankruptcies to retain a neutral approach.

Compensation Intensity for Revenue

We wish to examine the association between bankruptcy risk and compensation intensity, inspired by the ratio of loan intensity used by Altomonte et al. (2021). In our ratio, revenue refers to sales income in 2018 from the SNF dataset, and high intensity corresponds to more compensation granted per unit of revenue. Patterns in this ratio across deciles of bankruptcy risk will indicate systematic imbalance in compensation. The compensation intensity ratio (CR1) is visualized through average values per decile of bankruptcy risk in Figure 15 below with a 95% confidence interval shown in red.

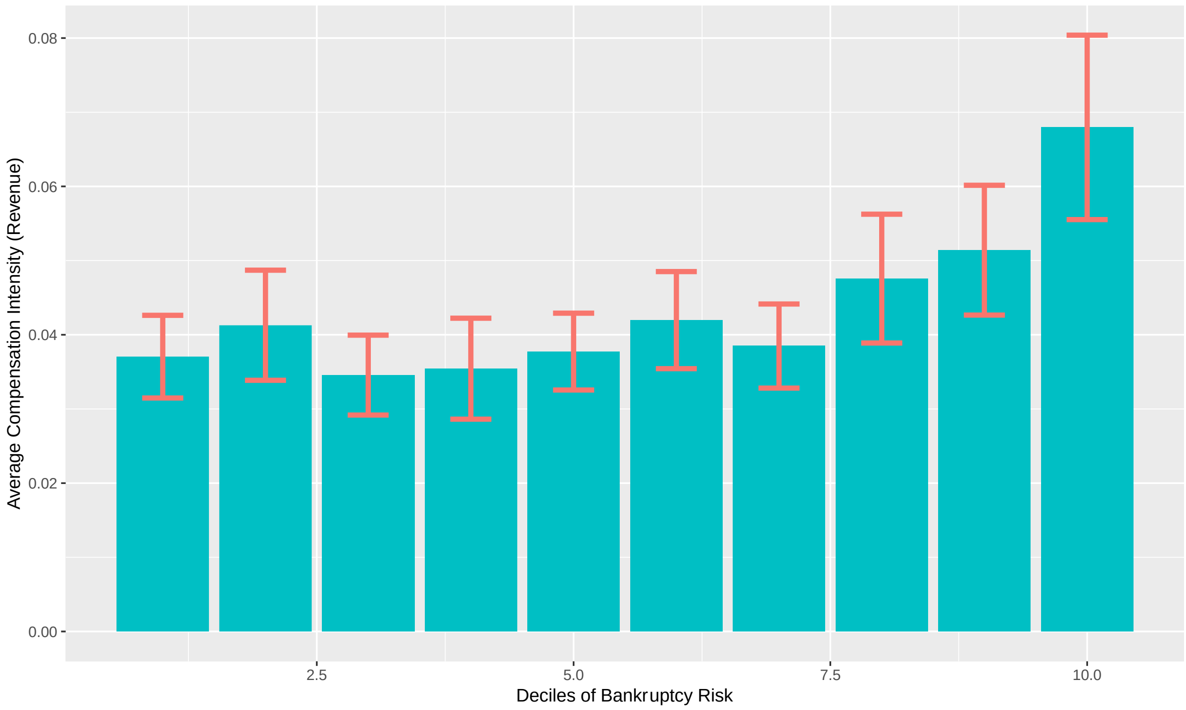


***Figure 15**: Average and Median CR1 (Compensation 2020 to Revenue 2018) ratio over Bankruptcy Risk Decile. The dependent variable was winsorized at the 1% level prior to decile delegation due to extreme values, and is shown per decile, with a 95% confidence interval marked in red. Note that grouped graphs do not necessarily reflect the relationship of individual observations.*

Inspection of the figure reveals a trend between bankruptcy risk and average compensation intensity, where the average compensation intensity is noticeably higher in the 10^{th} decile of bankruptcy risk. In contrast to Altomonte, we perform statistical analysis with our ratios to test the relationship at the individual observation level. Using a ratio introduces issues with interpretations and non-linearity (Barnes, 1982). At the same time, a ratio of compensation over revenue is an easy way to scale all observations for a size. As explained throughout this thesis, the bankruptcy risk is the estimated probability of bankruptcy and serves as our indicator of firm viability. It is a compressed and highly synthetic variable, which also affects interpretability. However, we are not seeking causal inference, as a synthetic random forest estimation cannot cause compensation. Instead, we seek to test association to inspect whether other factors that affect our viability indicator are related to the size-independent magnitude of compensation. We use ordinary least squares (OLS) regression on compensation intensity and bankruptcy risk variables to formalize and test the observed group-level trend at the individual observation level.

	Dependent variable:	
	CR1	log(CR1)
	(1)	(2)
Bankruptcy Risk	0.432^{**}	2.102^{***}
	t = 2.263	t = 5.531
Constant	0.033^{***}	-3.935^{***}
	t = 6.476	t = −162.834
Observations	3,747	3,747
Robust Standard Errors	Yes	Yes

Table 15: *OLS regression with the dependent variable Compensation Intensity for revenue (CR1), referring to the ratio of granted compensation over sales revenue in 2018, is explained by the estimated continuous bankruptcy risk variable. Model 1 is the unmodified dependent variable, while Model 2 uses a log transformation to improve normality and mitigate importance of outliers. Both models use robust standard errors due to the existence of heteroskedasticity. The direction of coefficients and statistical significance indicates a positive association. Reviewing regression (2) graphically in Appendix 7, we see few indications that the coefficients in Model (2) primarily are driven by outliers. Significance codes: 0.001 '***', 0.01 '**', 0.05 '*', 0.1 '.'*

The regression output confirms the visual interpretation, as compensation intensity is positively associated with bankruptcy risk at the 1% level. To determine the OLS model's validity, we inspect the assumptions of linearity, normality, and heteroskedasticity (Wooldridge, 2014).[36] Firstly, we determine that the OLS assumptions are violated in Model

[36] Multicollinearity is not commented on since we use only one explanatory variable.

(1). This is found through the lack of normality in residuals and the presence of homoskedasticity, as seen in Appendix 7. Since we work with cross-section data, heteroskedasticity is a common problem, and the violations are not surprising. As discussed by Barnes (1982), financial ratios rarely follow a normal distribution due to the non-linear nature of ratios. Our model is affected by this problem as the compensation ratio is highly skewed. In response, we apply a log transformation to achieve more normality and linearity in the parameters (Barnes, 1982; Wooldridge, 2014).[37] The transformed dependent variable is used in estimating regression Model (2). Heteroskedasticity can impact the reliability of the coefficients, and we run a White´s test to examine if this is present in (2). White´s test generates a X^2-value of 19.8 with a corresponding p-value of less than 1%. Because of the presence of heteroskedasticity, we use robust standard errors (Wooldridge, 2014). To further investigate the relationship between the variables, we calculate the nonparametric Spearman Rank Correlation Coefficient on the unmodified CR1 from Model (1) and find a significant but weak correlation at the 1% level, supporting the existence of a positive correlation between the ranks of the observations.[38]

This finding contrasts expectations because of the deficit criteria of the compensation scheme. For the objective of compensation, we determine that this positive relationship between bankruptcy risk and compensation intensity indicates that the criteria design was inefficient in distinguishing viable and unviable firms. In a scheme with an effective viability criteria, the direction of the relationship would be the opposite. Since the results were unexpected, we identified and hypothesized three possible explanations for the relation:

(1) We hypothesize that productive and well-run firms have advantages in adapting to shocks. This would explain the relationship in that a pandemic more severely impacts the revenue of less productive firms. For this reason, we generate a proxy variable for the impact of the pandemic. This is calculated by measuring the share of revenue in April 2020 compared to

[37] We note that the normality assumption is vital in parametric hypothesis testing. However, by the central limit theorem, any large sample distribution will be approximated by the normal distribution such that $\bar{Y} \approx N(\mu_y, \frac{\sigma_Y^2}{n})$. Consequently, hypothesis testing can still be used when the data are nonnormally distributed (Wooldridge, 2014). Considering our large sample of firms from the underlying population, we determine that potential nonnormality should not introduce significant bias after logarithmically transforming the dependent variable.

[38] Spearman Rank Correlation Coefficient is the non-parametric alternative to the Pearson coefficient. It detects correlation between the ranks of observations (James et al., 2021).

April 2019. We call this ratio the *impact proxy*, which is presented in Table 15. However, inspecting this ratio reveals no clear trend, indicating that this is not the primary cause.

(2) A second hypothesis is that unproductive firms are more incentivized to exploit the compensation design. For instance, a choice to continue shutdown beyond the most intense periods of lockdown would increase compensation throughout 2020. Since a large share of the firms analyzed have a negative operating margin (40.8% in 2018) , there would be a rationale for continuing the shutdown and receiving compensation for longer. Because we can quantify the number of months between March and August 2020 where an applicant firm has no revenue, we use this as an indicator of shutdown. The only variable we identify having a plausible causal relationship with the impact of lockdown is centrality.[39] The previously mentioned *impact proxy* revealed no clear trend and is therefore not included. For this reason, we review descriptive statistics of firms grouped by months of shutdown.[40]

Months of Shutdown	Average Predicted Risk of Bankruptcy	Centrality Index
0	0.03	4.3
1	0.031	3.7
2	0.036	4.2
3	0.039	4.5
4	0.043	4.4
5	0.048	4.1
6	0.060	4.5

Table 16*: Months of Shutdown. Summarizing descriptive data on the months of shutdown and the average risk of bankruptcy among the firms corresponding to that group. We observe a systematic increase in average risk of bankruptcy when the months of shutdown increase.*

In line with the hypothesis, we find that bankruptcy risk across months of shutdown reveals a clear trend. Simultaneously, the centrality index reveals no clear trend, indicating that virus presence is not the cause. The Spearman Rank Correlation Coefficient between months of shutdown and bankruptcy risk is positive and significant at the 1% level, supporting the existence of a monotonic relationship between the variables. While this is far from conclusive on firms' adapting behavior, it does provide an indicator of the direction of the relationship.

(3) A third explanation is that systematic differences in growth trends between low and high-risk firms cause skewness when using 2018 revenue data. To inspect this hypothesis, we create

[39] Following the common knowledge of greater virus transmission in urban areas.

[40] Centrality Index from high centrality (1) to low (10) per municipality, provided in SNF dataset.

a proxy variable for revenue in 2020 based on reported revenue in the compensation applications dataset for January and February 2020 multiplied by 6. With the estimated revenue, we can generate a second ratio (CR2) of intensity using the correct revenue year. However, one should note that this is vulnerable to seasonal effects that might impact the results. Furthermore, an inspection of the ratios reveals highly noisy data. For this reason, we chose to use the median of decile groups instead of individual variables. In Figure 16 below, the difference between compensation intensity in 2018 (CR1) and 2020 (CR2) is shown.

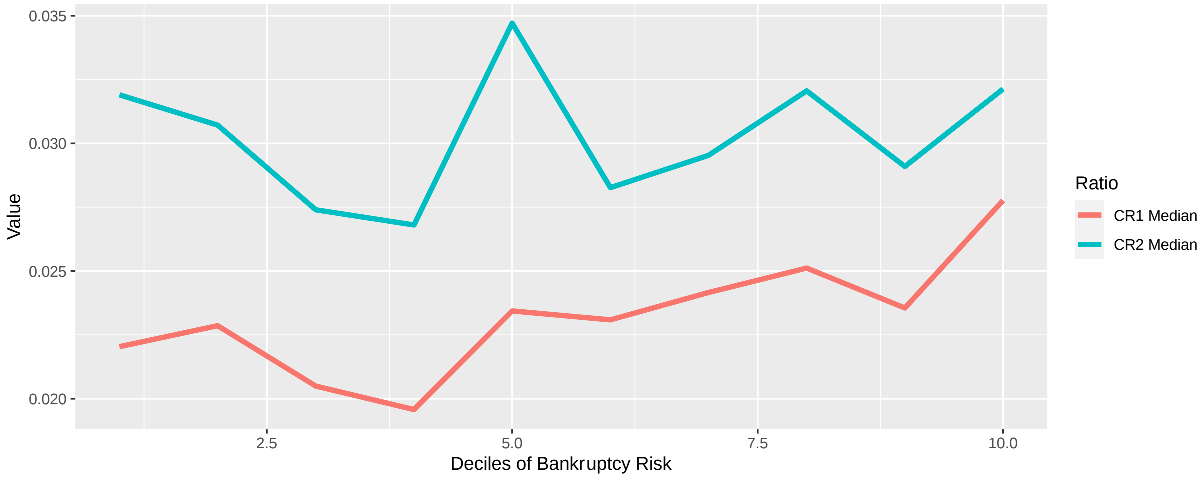


Figure 16: *Diagnostic Graph for the differences between 2018 and 2020 revenue data. The median of compensation intensity (CR1) compared to the median of compensation intensity (CR2). Values are grouped by and plotted against deciles of bankruptcy risk. Note that grouped graphs do not necessarily reflect the relationship of individual observations. The use of median values is done to counteract highly noisy data in CR2 and does not necessarily communicate the trend in individual observations.*

While the ratios mostly follow the same pattern, we observe that the association between compensation intensity and bankruptcy risk appears weaker for the 2020 revenue. While the data robustness of this inspection is weaker than other analyses in this paper, we still find it as an indication that the effect could be weaker than first assumed. Therefore, the identified relationship between compensation intensity and risk of bankruptcy should be interpreted with caution.

Compensation Intensity to Employees and Labor Costs

Following the presented procedure in section 1.3, we analyze the distribution of compensation as a ratio of employees or labor costs. Since we detected skewness in the compensation to revenue ratio, we hypothesize that a ratio of compensation to labor cost moving in the opposite direction can justify an otherwise imbalanced distribution, given the scheme´s aim of avoiding mass unemployment. This would correspond to the example given in section 1.3, where a disproportionally high compensation to unviable firms can be justified in outcome by contributing to the retention of employment.

We use two different ratios to measure the intensity of compensation per employee. Compensation per employee (CE) uses the nominal number in the SNF dataset, while compensation to labor costs (CLC) uses the total wage cost. The inclusion of both is done to correct for variation in the noisy employee variable. The reader should note that the variable of employees refers to the nominal number of employees, not full-time equivalents. Therefore, wage cost can be a better indicator of full-time equivalents. However, we note that individual wages could be higher in low-risk firms.

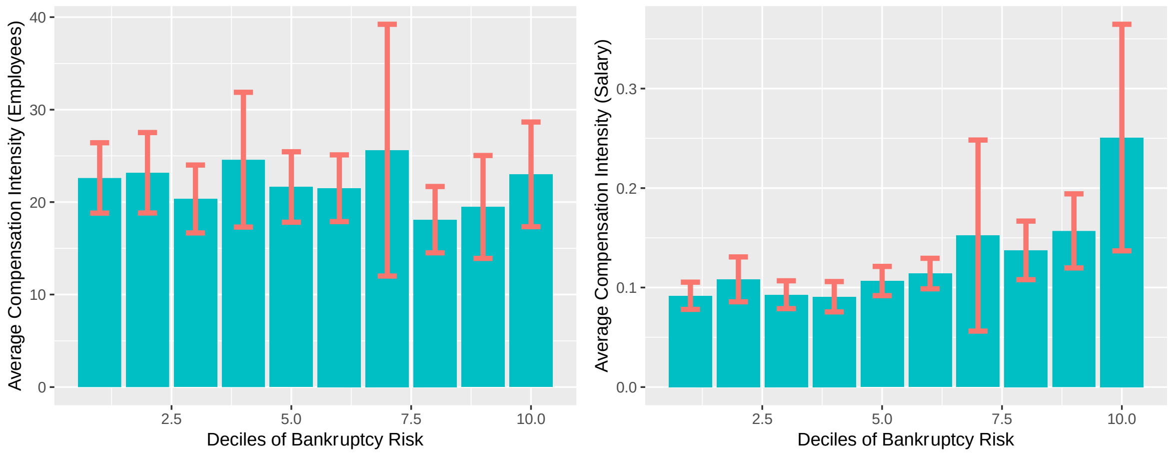


Figure 17*: Two graphs showing compensation to employees (CE) and compensation to labor costs (CLC) over bankruptcy risk in deciles. The dependent variable was winsorized at the 1% level before decile delegation due to extreme values, and is shown per decile, with a 95% confidence interval marked in red. Note that grouped graphs do not necessarily reflect the relationship of individual observations.*

We use two graphs to visualize the relationship between the employment ratio and bankruptcy risk. From observation, there is a higher variance in the CE ratio, which matches our expectations based on the data quality on employee numbers. On the other hand, the CLC ratio shows a more distinct association with bankruptcy risk but in the same direction as the ratio of compensation to revenue. We test the relationships statistically in the table below.

	Dependent variable:		
	CE	CLC	log(CLC)
	(1)	(2)	(3)
Bankruptcy Risk	−0.392	0.780***	1.990***
	t = −0.025	t = 3.075	t = 7.850
Constant	22.022***	0.098***	−2.991***
	t = 19.081	t = 12.616	t = −386.749
Observations	3,747	3,747	3,747
Robust Standard Errors	Yes	Yes	Yes

Table 17: *The table shows OLS regression with the compensation intensity of employees and labor costs explained by the predicted risk of bankruptcy. Model 1 and 2 use an unmodified dependent variable, while Model 3 uses a log transformation to improve normality and mitigate the importance of outliers. All models use robust standard errors due to the existence of heteroskedasticity. The direction of coefficients in model 2 and 3 and statistical significance indicates a positive association. Reviewing regression (3) graphically in Appendix 7, we see few indications that outliers primarily drive the coefficients. Significance codes: 0.001 '***', 0.01 '**', 0.05 '*', 0.1 '.'*

OLS assumptions are reviewed in the same procedure as previously, diagnostic plots can be found in Appendix 7, and the same limitations of ratios and bankruptcy risk persist with respect to hetroskedasticity and normality. The log-transformed dependent variable log(CLC) has a White's test score of 26.9 and a corresponding p-value below 1%. Therefore we use robust standard errors.

In line with the graphs in Figure 17, the CE ratio has no observable relation to bankruptcy risk. However, the CLC ratio is significant, with a coefficient of 0.78. As mentioned, we use log transformation for Model (3) to improve linearity and normality, and this model retains the positive direction with a more significant coefficient. As previously, we use the Spearman Rank Correlation Coefficient on the unmodified CLC in Model (2) to further inspect the existence of a monotonic relationship. We find a weak but significant correlation between the ranks of observations at the 1% level, supporting the existence of a monotonic relationship.

This association is of low value if the average labor cost differences are explained by higher wages and not fewer part-time employees. We still observe an association between bankruptcy risk and compensation intensity to labor costs. Since the two ratios follow the same direction when analyzed over bankruptcy risk, we lack evidence of the hypothesized justification in outcome with respect to employee retention. Moreover, the findings reinforce that high-risk firms have received disproportionally high compensation.

Descriptive Analysis and Summary

While we find the directionality in the analysis robust, we withhold from interpreting coefficients for their economic significance because of the issues regarding the interpretation of ratios and highly synthetic variables. To quantify our descriptive analysis, we compare the bottom and top halves of the dataset, sorted by estimated bankruptcy risk, to capture the economic magnitude of the differences in compensation.

For the figure of granted compensation, we calculate a total compensation of 634,6 million NOK to the top half of bankruptcy risk and 1775,1 million NOK to the bottom half. The interpretation of this division is that the largest share of compensation went to the safer, low-risk firms. However, in contrast to ratios, this figure is highly affected by size. Still, 73,6% of compensation granted to the most viable half does imply a reasonable distribution.

As previously mentioned, we look at compensation intensity to correct for size differences. For the the revenue-based compensation intensity of 2018 (CR1), the top and bottom half of the dataset have average ratios of 0.064 and 0.039, respectively.[41] This corresponds to a 2.5 percentage point difference and means that the weaker half of firms received 66% higher compensation on average, adjusted for its revenue. A similar relation is present in the ratio of compensation to labor costs, as the top half got a ratio (CLC) of 0.162 and the bottom 0.098. This results in a percentage point difference of 6.4 and means that the weaker half received 65% more compensation, adjusted for its revenue. While the differences in compensation between top and bottom are large, we do not see them as critical from a macroeconomic perspective.

In sum, we estimated that at least 54.3 million NOK, constituting 1.58% of the compensation, was granted to predicted bankrupt firms. Our primary finding is the weak but significant association between higher risk of bankruptcy and compensation intensity, as measured by the

[41] Averages of averages are generally bad practice, but the deciles are uniform in size, which makes it equivalent to the deciles used previously.

ratio of compensation to revenue. In addition, we find indications based on Table 16. that riskier firms tended to shut down for longer, either voluntarily or forcibly.

Overall, we observe that the sum of compensation granted to predicted bankrupt firms is relatively low, both nominally and as a percentage of the total compensation. However, the distribution of compensation appears to favor firms with high bankruptcy risk based on the weak but significant relationship in Table 15 and our descriptive analysis in Figure 15. These findings indicate that while the expected direct loss to conservatively estimated bankruptcies is low, compensation was distributed unevenly across the axis of bankruptcy risk.

5. Discussion

5.1 Ambiguous Policy Objective

One of our primary concerns with the compensation scheme is the ambiguity in its criteria. Although the objective stated by the Norwegian Tax Administration at first glance appears lucid (Skatteetaten, 2020), we argue that the issue is more nuanced and requires careful consideration. The official purpose is to save viable firms, limit unemployment rise, and secure rapid post-pandemic recovery in economic activity. One interpretation of the criteria of saving viable firms is that it communicates the aim of not saving unviable ones. We argue that this interpretation is the most natural given the argumentation in official sources as well as the public debate. However, we do acknowledge that there is room for alternative interpretation and that the meaning may have been to discriminate the classes only slightly. While the aim of saving viable firms is logical for ensuring long-term economic efficiency, our data demonstrate that a large share of all employees is employed within firms that can be described as economically unviable. By saving or shielding unviable firms, one would shield the associated employees but hinder the reallocation of the workforce to more productive firms. Essentially, our descriptive statistics in Table 14 shows that there is a conflict between the goals of maintaining employment and simultaneously limiting support to viable firms, since a large proportion of workers are employed in firms with high bankruptcy risk.

One can therefore view this as a trade-off between the sole goals of maintaining overall employment or strictly enforcing the viability criteria. On the one hand, the compensation scheme can be generously designed to save all firms and thereby maintain all employment. On the other hand, one could use strict criteria to minimize compensation granted to unviable firms and thereby cause unemployment among the associated employees. This approach makes more sense considering long-term allocation efficiency but is difficult to implement because of the complex process of determining firm viability with high accuracy. Moreover, this would likely cause more unemployment, resulting in a more severe bankruptcy costs as described in section 2.2. With a foundation in the primary justification for implementation, we must assume that some combinations of employment and productivity-inducement were the goal. The issue with interpreting the objective is that there is no indication of these considerations' weighting.

For this reason, we attempted to infer clarification from the public debate. When NRK confronted former Finance Minister Jan Tore Sanner in May 2020 about compensation granted to unviable firms, he responded that the compensation scheme was an acute measure to help viable firms through the crisis. This was followed by a statement addressing that they, under the scheme's development, prioritized a quick solution rather than perfect accuracy (Kampevoll & Seibt, 2020). To explain why a substantial share of recipient firms had negative results, he argued they were hesitant to use stricter profit criteria since negative results are common among all types of startups. Based on our analysis, we determine this to be a weak explanation since a large share of the firms he described as start-ups constitute a significant proportion of all unviable firms. Innovation is exempt from the objective, and one would therefore expect no difference between innovative startups and other young firms. Essentially, it appears unclear if the conflict between maintaining employment and economic efficiency was explicitly stated as payoffs, if maintaining employment and saving firms were separate foci within the same compensation, or if this conflict was simply not evaluated.

5.2 Outcomes and Explanations

Our analysis demonstrates a weak indication that high-risk firms benefited disproportionally more from the compensation scheme. However, the magnitude of misallocated compensation does not appear to be large enough to cause significant damage. This is because of the low share of 1.58 percent of compensation granted to predicted bankruptcies. On the other hand, we find flaws in the skewed compensation intensity across the range of risk. For instance, the three deciles of highest bankruptcy risk are characterized by negative three-year average EBITDA and book value of equity. We interpret this as a sign of robustness for predicted bankruptcy as the chosen viability indicator. This also underlines the lack of viability in a larger proportion of the firms beyond the conservative threshold of bankruptcy classification, which should alert governmental institutions since this indicates a “greyzone” of unviability among firms. This observation aggravates the compensation estimation for predicted bankrupt firms since a larger share could qualify as unviable with a slightly broader definition. One such definition could be the definition of a zombie firm with a lower age restriction, a reasonable proposal given the high entry and exit rates in the industry. The composition of the industry serves as a warning sign for its general viability. It is a central element for consideration if a similar issue in discrimination of viability emerges.

Compensation intensity holds a statistically significant association with bankruptcy risk in our analysis. Naturally, our results necessitate further discussion on the underlying mechanisms so that we get closer to understanding the relationships between the figures. Association does not necessarily imply faulty distribution of compensation, and we attempt to expand our understanding of the causes. We therefore identify different potential reasons for our results in compensation intensity.

The first mechanism concerns the impact of restrictions and pandemic demand shocks on firms' revenue. If high-risk firms are systematically more affected than firms with a low-risk of bankruptcy, we naturally expect greater compensation intensity because of the compensation formula. This would be a large problem when analyzing several industries at once, and we argue that our industry limitation increases the clarity of the results since we can assume greater homogeneity in pandemic impact. Nevertheless, one can still argue for within-industry effects that explain, or speculate, on how firms are impacted differently.

A second mechanism is the duration of the shutdown if high-risk and low-risk firms adapt differently to the pandemic and the compensation scheme. This is in line with criticism raised in the public debate addressing that weak firms can be more incentiviced for shutdowns since the procpects of compensation from the multitude of policies and shutdown of operations can appear more attractive than running an unproductive firm. Table 16 demonstrates that higher bankruptcy risk is associated with more months of shutdown between March and August 2020. However, it is not possible to determine whether this is caused by firm adaptation or differences in lockdown impact. Even if correcting for levels in pandemic presence, one would still omit demand-side differences based on market positions and firms' different target demographics. For this reason, it is difficult to establish if there are systematic differences in pandemic impact with respect to the loss of revenue.

Systematic differences in firms' cost structures, labor intensity, and productivity across bankruptcy risk could also affect the compensation intensity. If firms with higher bankruptcy risk have more fixed costs, it would explain why compensation to both revenue and labor costs is greater. This is because one of the elements in the compensation formula was fixed costs. Interpretation of results would be more clouded if we included an extensive array of industries, again demonstrating that sector limitation was appropriate. Still, we cannot exclude the possibility that within-industry effects exist if firms with a higher risk of bankruptcy have a greater share of fixed costs. Since fixed costs are not present in the financial statements, we could not examine this further but recognize it as an explanation. All else equal, firms with a higher share of fixed costs are expected to have higher compensation intensity.

While our analysis demonstrates inefficiencies in the scheme, we acknowledge that decision-makers were severely time-constrained. They had to develop a scheme that ensured both ease of implementation, economic efficiency, and remaining within legal constraints. On the other hand, we question the low frequency of reiteration given the magnitude and importance. According to the governmental report (2021:4), the compensation scheme was developed within a period of three weeks. It is argued that the fast implementation was a major success factor since it provided immediate liquidity to firms. Still, we expect that the higher compensation intensity among risky firms entails, at least theoretically, a future efficiency loss, as explained in section 2.3. Naturally, we would then expect a greater effort to improve the accuracy of the scheme in the months and the year following the initial shock.

5.3 Innovation and Adaptation

As the literature indicates, forbearance lending could impede innovation by reducing innovation incentives if survival is secured through other means. As Tracey (2021) argues, we interpret that the effect of forbearance lending on innovation can be applied to both the Norwegian compensation scheme and the other enacted policies explained in section 1.2. Our discussion on the topic is primarily theoretical, considering the unquantifiable nature of innovation. Still, exploring what kind of firms fuel innovation is helpful to understand its relationship to compensation.

Innovation Norway (2021) identified two primary directions in short-term adaptation strategy among businesses during the covid crisis: (1) Proactive strategies characterize responses where companies use crises to develop, often by increasing spending and introducing new products and services. (2) Reactive strategies characterize responses where the focus is cutting costs and limiting operations to maintain solvency.

Reactive strategies in the context of the pandemic would likely correspond to different levels of voluntary shutdown or reduced service. If firms with reactive strategies contributed to innovation, their shutdown would naturally impede such contributions. One could argue that large firms with strong access to capital constitute the firms with the ability to adapt proactive strategies during the pandemic (Sørheim, et al., 2021). Several hotels with better access to capital in the accommodation sector appear to have used the covid pandemic as an opportunity to renovate buildings and train staff. Similarly, restaurants appear to have used the pandemic as an opportunity to introduce more catering and takeaway services while digitizing customer contact. Both examples display elements of proactive adaptation. While strong financials is likely one determinant of distinguishing reactive and proactive firms, there are arguments that the difference is driven by leadership among the smaller firms. One such argument is supported by a qualitative master thesis from NMBU exploring innovation in the hospitality industry (Segerberg, 2018). As to whether innovation occurs in large chains or smaller restaurants, professor Øystein Foros at NHH (Ghaderi, 2019) points out that innovation in larger concerns is more rigid. This rigidity is explained by inflexibility in customers' brand associations. On the other end, it would imply that smaller restaurants can adapt faster to customer preferences. Ultimately, there are theoretical considerations and literature implying that innovation takes place in large firms with strong access to capital and smaller firms with proactive management. While the discussed drivers of innovation are self-explanatory, we argue that the depiction of these connections is relevant because they represent aspects central to decision-makers. For instance, we would expect the designers of the compensation scheme to have an opinion on the impact on innovation and creative destruction through how well it targets different categories of firms.

In light of our results, we highlight that the small and young firms are generally associated with a greater risk of bankruptcy. While Sanner argued in the aforementioned example that start-ups should be sheltered despite poor performance, we believe this is an example of the fallacy that most start-ups become successful companies. Our data and analysis demonstrate that start-ups usually display characteristics of poor viability, often resulting in bankruptcies, and rightfully so.

Ensuring survival for firms regardless of their financial state and future outlook will certainly affect firms' choices. On the one hand, long-lasting compensation policies can reduce incentives to innovate during a crisis, while on the other, firms might raise their risk profile during a regular period. Systematic changes in risk profile is also noted by the government report on the post-pandemic economy (NOU, 2021:4). Increased risk will therefore likely increase the compensation requirements of the government during our next inevitable economic crisis.

5.4 Policy Magnitude

We argue that the distribution and magnitude of the compensation scheme and other policies have impacted the economy and aggregate bankruptcies. While it is unfeasible to quantify the causal effect of the compensation scheme on aggregate bankruptcies, a discussion of the effects provides an understanding of the mechanisms and motivates future research on the topic.[42] As previously discussed, bankruptcies have positive and negative consequences, implying that an optimal level exists. In extension, one interpretation is that policy intervention should aim to approach this optimal level. If that is the historical rate, a deviation will constitute efficiency loss. Stylized, we can illustrate the relationship as follows:

$$Bankruptcy\ Rate\ 2020 = Historical\ Bankruptcy\ Rate + Pandemic\ Shock - Policy\ Effect$$

Optimally, the interventions from enacted policies would therefore equal the effect of the pandemic shock such that the bankruptcy rate remains stable. This would coincidence with the aim of saving "otherwise viable firms." Under the assumption that the historical bankruptcy trend before the outbreak reflects the optimal level, any deviations adjusted for market entries would reveal an inadequate policy adjustment to the pandemic shocks. Data from Brønnøysund Register Centre (2022) limited to our industry reveals that the bankruptcy frequency stayed stable in 2018, 2019, and 2020 with 485, 481, and 479 bankruptcies, but experienced a sharp decline in 2021 at 273 (almost 40% decline). At first glance, it would appear that the enacted policies were largely successful in 2020 but unsuccessful in 2021, considering the vast decline. Adjusted for market entries, however, the picture appears more nuanced. According to data from Brønnøysund Register Centre (2022), market entries were 2485, 2619, 2466, and 2203 in 2018-2021, representing a much higher frequency of entries than exits. As such, one could easily argue that the number of bankruptcies was below an optimal level in line with criticism raised in the public debate. Based on the number of aggregate bankruptcies, we hypothesize that the collective effort of the enacted policies was greater in magnitude than the effect of the pandemic shock.

[42] Ideally, we would want to isolate the effect of the compensation scheme on the aggregate number of bankruptcies. However, with no comparable control group exposed to the same market environment where all else is equal except for the compensation scheme, a differences-in-differences estimator measuring this effect is not possible to estimate with sufficient reliability.

While this indicates redundancy in public spending, one should note that the alternative of subsidizing is not a zero expense. Mass bankruptcies would entail a sudden and potentially unprecedented rise in registered unemployed. The following expense in unemployment subsidies, and its associated hysteresis effect, could quickly surpass the total compensation expense in the absence of other remedying policies. This is clearly an argument for a broad and generous selection of compensation policies.

The efficiency benefit of bankruptcies has been widely discussed and is closely related to resource allocation and forbearance lending. While far from all firms should go bankrupt, the literature discussed in section 2.3 demonstrates that bankruptcies' long-term impacts deserve attention and consideration. In the scope of the compensation scheme, the natural way to ensure sufficient bankruptcies is to ensure bankruptcy of the least viable firms in line with the scheme's objective. To some extent, this is also the outcome of the compensation distribution.

For discussing the compensation scheme's impact on the least viable firms, one can argue that it is limited, given the estimated 54.3 million NOK granted to predicted bankruptcies. As previously stated, several policies were enacted during the pandemic, with different levels of impact on aggregate bankruptcies. However, some argue that the reduction of bankruptcy proceedings by the Norwegian Tax Administration is the most contributing source. This argument harmonizes well with findings on variable importance by our bankruptcy prediction model discussed in 3.7.2, where the variable of public taxes and fees is ranked as the strongest predictor of bankruptcy. In summary, we argue that the magnitude of the compensation scheme is likely limited but that the effect of all policies adds up. For this reason, we argue that the policies should be developed, revised, and evaluated collectively.

6. Limitations

Our methodology is inspired by Altomonte et al. (2021). They performed a similar study across an array of countries in Europe but used labor productivity where we apply bankruptcy risk and kept their analysis purely descriptive. While we find robustness in following the general structure of existing research by the expansion of methodology and complexity, we introduce sources of error. For this reason, we will comment on what we determine to be the greatest limitations to our research.

Firstly, we use 2018 data that does not adjust for revenue growth between 2018 and 2020, which could affect the ratios. While we have attempted a proxy for 2020 revenue through the compensation applications. This variable is noisy and provides limited utility. A review of SNF financial statements and application data reveals noise in the financial statements as well, although to a far lesser extent than the 2020 data. This issue has also been a topic in other thesis and publications using the exact data source. Ideally, we would include all firms belonging to sectors 55 and 56 based on a qualitative evaluation of every firm. However, we expect that firms do not always register the industry identifier correctly, and financial data can be disturbed by non-operational activities and wrongful/manipulative accounting. The extent to which these aspects affect our results is unfortunately unquantifiable. We have not investigated if the removal of the companies not meeting our criteria affects the representativeness of the population of firms because we do not know the actual population.

Concerning the predictive power of our models, we would ideally predict the 2019 ex-ante probability of bankruptcy. More recent data naturally improves the reliability of the results, which is in line with established literature. In contrast to Chava et al. (2004), we use yearly data instead of monthly data. Their study concludes that the use of monthly data considerably increases the predictive power of their models. This is in line with criticism raised by Ohlson (1980), who argues that the time horizon between the time of publishing financial statements and the time of bankruptcy imposes problems in terms of predictive power. Naturally, we cannot use monthly data due to the standard accounting periodization of one year. Nevertheless, using more frequent data would still benefit our study and the precision of our results. It would also aid in the compensation analysis, as monthly data would provide closer compatibility with monthly compensation applications.

In our analysis, we examine the distribution among recipients of the compensation scheme in line with our research question. While our analysis is mostly unaffected by the dataset incompatibility of financial statements and compensation applications, it does constitute a limitation for interpreting the overall effectiveness of the scheme. From Table 13, we see that a large group of firms were non-recipients. Investigating this group, we found that the firms in the group had a higher risk of bankruptcy, as seen in Appendix 9. This characteristic is logical, considering that some of these firms likely went bankrupt before the first quarter of 2020. However, the characteristics associated with these could also result from efficient rejections of applications for unviable firms. At the same time, firms could adapt so well to the pandemic that they do not qualify for compensation. This ambiguity means that while the share of these outcomes affects the scheme's efficiency, it is difficult to identify how. On the other hand, we find that average compensation is low for the observations solely present in the compensation dataset. Furthermore, the natural explanation for lacking in the 2018 financial statements is that they were established after 2018. Given the association between age and bankruptcy risk, we hypothesize that the inclusion of these firms would strengthen the observed association found in our analysis.

We rely on the implicit assumption that firm productivity is the same before and after the pandemic. The covid pandemic represents extraordinary conditions, so the ability to adapt should become crucial in determining “firm viability” during the pandemic. Regardless, we are interested in the post-pandemic future, reducing reliance on the ability to adapt but improving reliance on the untestable assumption that pre-pandemic performance is a reliable indicator of future performance.

The target variable of bankruptcy is, as explained in 3.2.3, vulnerable to noise and subjectivity, both because of bankruptcy procedures and the difference between forced liquidation and declaration of bankruptcy. Owners of firms may also go bankrupt and purchase the bankruptcy estate to restart the same business with fresh financials. The many possibilities introduce noise in the associated financial statements and make the bankruptcy definition vague. It also makes comparisons to other bankruptcy studies less reliable, which remains a source of error in our comparison. Furthermore, our study uses an approximated bankruptcy probability to measure firm viability. That means our findings and conclusions heavily rely on a precise and robust bankruptcy prediction model. While we applied several models and used cross-validation, we still acknowledge that bankruptcy prediction models constitute a limitation in the robustness of the findings in our analysis.

7. Conclusion

There is currently a deficiency of research on the enacted policies resulting from the Covid 19-pandemic in Norway. The policies of the compensation scheme, postponement of taxes, state-guaranteed loans, and the layoff scheme were all intended to prevent mass unemployment and ensure rapid post-pandemic recovery by improving firm liquidity and reducing meaningless bankruptcies (NOU, 2021:4; Skatteetaten, 2020). Since the expected wave of bankruptcies failed to materialize, and the enacted compensation scheme distribution received criticism from multiple fronts, we found it a highly relevant topic for research. In response, we chose to compare the distribution outcome with the purpose behind the scheme's creation, specifically the criteria to exclude unviable firms.

In line with our research question, we have applied bankruptcy prediction using machine learning to quantify firm viability and evaluate the distribution of the Norwegian compensation scheme. For this task, we decided to limit ourselves to the hospitality industry in Norway, motivated by the industry's pandemic impact, the share of received compensation, and the historically high share of bankruptcies. We evaluated sixteen combinations of variable sets and classifiers for bankruptcy prediction, which yielded comparable results. The low variation in performance across models and the inclusion of fivefold cross-validation demonstrate the model's robustness. The performance and robustness make us confident in the model's predictive ability compared to relevant literature.[43] Supporting the model's sufficiency, we find that its use of assessing viability entails a reduction in the cost of error since wrongful predicted bankruptcies still communicate information of low firm viability.[44]

From the objective of distributing compensation to only viable firms, one can claim the government was primarily successful. Our prediction model estimates that only 54.3 million out of 2409.7 million were granted to predicted bankrupt firms. However, we also assessed viability as a continuous variable in addition to the binary bankruptcy criteria. We find that firms with above-average bankruptcy risk received over 60% more compensation, adjusted for revenue, than those with below-average bankruptcy risk. Statistical testing further reveals a

[43] See section 3.3 for explanation of the fourth variable set. The highest performing model is Random Forest on an expanded SEBRA variable set, with an AUC of 0.829. See section 3.5 for explanation of evaluation metrics.

[44] The cost of wrongly classifying an observation. The model outputs probabilities corresponding to the estimated risk of bankruptcy, such that high risk corresponds to financial distress despite a binary classification as "safe".

weak but significant relationship between the estimated risk of bankruptcy and compensation intensity. From the deficit criteria in the scheme, one would expect lower compensation intensity per unit of revenue among high-risk firms, given the average deficits in the deciles of high bankruptcy risk. Therefore, we infer that efforts were insufficient in discriminating for viability. Regarding the scheme's magnitude, we expect other policies to have similar or more significant short-term impacts on bankruptcy frequency. However, concerning the long-lasting altering of resource allocation, we are worried that the identified higher compensation intensity among high-risk firms has the potential to harm economic efficiency.

We have discussed several explanations for the trends and associations identified between the risk of bankruptcy and compensation intensity. These include explanations addressing the exploitation of the scheme, unfortunate incentives provided by the enacted policies, and potential structural differences. Unfortunately, we cannot determine which effects dominate and lead to the observed outcome. While the causal explanation is likely to be a combination of these effects, we have no robust way of isolating the effect of each identified mechanism. For this reason, we have focused on the outcome of the compensation scheme and delegated the task of causal exploration to future academic work.

We identify three primary contributions of this thesis. Firstly, we assess the Norwegian compensation scheme against its objective of only distributing to viable firms. This form of assessment is currently lacking in the literature. We analyzed the relationship between estimated bankruptcy risk and compensation intensity, and identified a direction breaking with the compensation objective. Secondly, this thesis contributes to the literature by demonstrating the application of bankruptcy prediction within a public policy context. The contribution is heightened because of deficiencies in the literature on industry-specific bankruptcy prediction. Thirdly, we believe this thesis answers questions of public interest since the compensation scheme has been widely debated and has become common knowledge across groups in society.

In summary, we have demonstrated how bankruptcy prediction can be used to evaluate the Norwegian compensation scheme in the heavily affected hospitality industry. We found that compensation distribution at the aggregate level was mostly in line with the viability criteria of the compensation scheme objective. At the same time, we determine that compensation was distributed unequally, as firms with higher bankruptcy risk on average received higher intensities of compensation per unit of revenue and labor costs.

8. Research Suggestions

We have identified several possibilities for future research primarily based on the weaknesses of our thesis and further questions of public interest.

In order to correct the discussed limitations, a similar study using 2019 data would provide value through greater accuracy in the analysis. Expanding on our thesis, we recommend researching the collective pandemic policy effect, which allows a more causal evaluation of the policies' impact on bankruptcy frequency. Research in this direction should also aim at better explaining the discussed explanatory mechanisms. Motivated by our descriptive statistics on months of shutdown, we believe qualitative research should be performed to examine the relationship between voluntary shutdown and bankruptcy risk. It is potentially an example of firm adaptation to policy incentives, and exploring the adaptation mechanisms could provide valuable inputs in policy-making. Furthermore, we are interested in our assumption of comparable prodnot founductivity before and after the pandemic. Investigating the future performance of predicted bankrupt firms will clarify this assumption's feasibility. Moreover, we strongly advise a study comparing the speed of compensation policy implementation across countries.

Despite its legal and practical limitations, we are strong proponents of the use of bankruptcy risk or other complex assessments to distinguish recipients of public funding. Bankruptcy prediction considers a range of variables and provides a comparatively objective measurement of firm viability. The difference is considerable compared to proxies such as last year´s earnings, used in the compensation scheme. Misclassification using bankruptcy prediction must necessarily occur in the real world. Knowing the misclassification costs of bankruptcies would allow for an optimal threshold for different practical purposes. This could allow decision-makers to use threshold levels that maximize utility and welfare for the economy.

Regardless, we still believe that other indicators for viability should be assessed. Through the inspection of variable importance, we found that the variable of public debt as a ratio over assets is the most significant contributor to predictive power. Among the three most important variables, none of them directly reflected earnings. For this reason, we suggest that other variables should be researched and evaluated as they might be better one-dimensional criteria than the one used in the compensation scheme. Such research may provide value in case of similar future policies or universal decision-making support where firm viability is relevant.

9. References

Aghion, P., Antonin, C., & Bunel, S. (2021). *The Power of Creative Destruction.* London: The Belknap Press of Harvard University Press.

Alaka, H. A., Oyedele, L. O., Owolabi, H. A., Kumar, V., Ajayi, S., Akinade, O. O., & Bilal, M. (2018). Systematic review of bankruptcy prediction models: Towards a framework for tool selection. *Expert Systems with Applications, 94*, 164-184.

Altman, E. I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance, 23*(4), 589-609.

Altomonte, C., Demertzis, M., Fontagné, L., & Müller, S. (2021, November). COVID-19 financial aid and productivity: has support been well spent? *Bruegel. Policy Contribution 21.*

Barnes, P. (1982). Methodological Implications of Non-Normally Distributed Financial Ratios. *Journal of Business Finance & Accounting, 9*(1), 51-62.

Barney, J. B., & Arikan, A. M. (2005). The Resource-based View: Origins and Implications. In *The Blackwell handbook of strategic management* (pp. 123-182).

Beaver, W. H. (1966). Financial Ratios As Predictors of Failure. *Journal of Accounting Research, 4*, 71-111.

Berg, D. (2007, March/April). Bankruptcy prediction by generalized additive models. *Applied Stochastic Models in Business and Industry, 23*(2), 129-143.

Berner, E., Mjøs, A., & Olving, M. (2016, 12). Dokumentasjon og kvalitetssikring av SNFs og NHHs database med regnskaps- og foretaksinformasjon for norske selskaper. *Regnskapsboka, 10*. Retrieved from NHH.

Bernhardsen, E., & Larsen, K. E. (2001). Modell for analyse av kredittrisiko i foretakssektoren. *Penger og Kreditt, 2*(01), 109-116.

Breimann, L., & Cutler, A. (2022). *Random Forests*. Retrieved from Statistics Berkeley: https://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm#remarks

Brønnøysund. (2022). *Nyregistreringer i Enhetsregisteret*. Retrieved from Brønnøysundregistrene:https://public.tableau.com/app/profile/brreg/viz/NyregistreringeriEnhetsregisteret/NyregistreringeriEnhetsregisteret

Brønnøysundregistrene. (2022). *Innsyn i tildelinger av tilskudd fra kompensasjonsordningen for næringslivet* . Retrieved from Brønnøysundregistrene: https://data.brreg.no/kompensasjonsordning/innsyn/

Brønnøysundregistrene. (2022). *Konkursåpninger og Tvangsavviklinger*. Retrieved from Brønnøysundregistrene:https://public.tableau.com/app/profile/brreg/viz/KonkurspningerogTvangsavviklinger/KonkurspningerogTvangsavviklinger

Caballero, R. J., Hoshi, T., & Kashyap, A. K. (2008, December). Zombie Lending and Depressed Restructuring in Japan. *98*(5), 1943–1977.

Chavam, S., & Jarrow, R. A. (2004). Bankruptcy Prediction with Industry Effects. *Review of Finance, 8*(4), 537–569.

Chung, H., & Thewissen, S. (2011). Falling Back on Old Habits? A Comparison of the Social and Unemployment Crisis Reactive Policy Strategies in Germany, the UK and Sweden. *Social Policy & Administration*, *45*(4), 354–370.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2002). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences.* New York: Routledge.

Cros, M., & Epaulard, A. (2021). Will Schumpeter Catch Covid-19? *CEPR Discussion Papers, 15834*.

Deloitte. (2020). *COVID-19: Temporary layoffs for Norwegian employers*. Retrieved from Deloitte: https://www2.deloitte.com/no/no/pages/legal/articles/Temporary-layoffs-for-Norwegian-employers.html

Dijck, G. v., Hollemans, R., Mas´nicka, M., Frade, C., Benedetti, L., Galanti, L., . . . Pagni, I. (2020). Insolvency judges meet strategic behaviour: A comparative empirical study. *Maastricht Journal of European and Comparative Law*, *27*(2), 158-177.

Dosi, G., Pereira, M. C., Roventini, A., & Virgillito, M. E. (2018). Causes and consequences of hysteresis: aggregate demand, productivity, and. *Industrial and Corporate Change*, *27*(6), 1015–1044.

Dun & Bradstreet. (2021, June 7). *Overraskende få konkurser*. Retrieved from NTB: https://kommunikasjon.ntb.no/pressemelding/overraskende-fa-konkurser?publisherId=10874623&releaseId=17909614

Engelmann, D., & Strobel, M. (2004). Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments. *The American Economic Review*, *94*(4), 857-869.

Finans Norge. (2020, April 20). *Full uavhengighet mellom statens støtteordninger til næringslivet*. Retrieved from Finans Norge: https://www.finansnorge.no/aktuelt/nyheter/2020/korona/full-uavhengighet-mellom-statens-stotteordninger-til-naringslivet/

Ghaderi, H. (2019, October 6). *Restaurantene presses av sterkere konkurranse: – Flere opplever situasjonen som krevende*. Retrieved from E24: https://e24.no/naeringsliv/i/K3kMdy/restaurantene-presses-av-sterkere-konkurranse-flere-opplever-situasjonen-som-krevende

Graham, J. R., Qui, J., Li, S., & Kim, H. (2013, July). Human Capital Loss in Corporate Bankruptcy. *US Census Bureau Center for Economic Studies , 13*(37).

Hahn, C. H., & Lim, Y. (2004, June). Bankruptcy Policy Reform and Total Factor Productivity Dynamics in Korea. *Growth and Productivity in East Asia*, 297-326.

Hamnes, M. (2021, June 09). *Korona: Dette er kompensasjonsordning 2* . Retrieved from Sticos.no: https://www.sticos.no/fagstoff/kategori/regnskap/korona-dette-er-kompensasjonsordning-2

Hastie, T. J., & Tibshirani, R. J. (1990). *Generalized Additive Models.* Chapman and Hall/CRC.

Heggebø, K., & Elstad, J. I. (2017). Is it Easier to Be Unemployed When the Experience Is More Widely Shared? Effects of Unemployment on Self-rated Health in 25 European Countries with Diverging Macroeconomic Conditions. *European Sociological Review, 34*(1), 22-39.

Hofman, B., & Banerjee, R. (2018). *The rise of zombie firms: Causes and consequences.* BIS Quarterly Review.

IMF. (2014, September). *What Is Keynesian Economics?* Retrieved from IMF.org: https://www.imf.org/external/pubs/ft/fandd/2014/09/basics.htm

Innovasjon Norge. (2021). *Innovasjonstalen 2021 Statusrapport: Koronakrisens påvirkning på næringslivet.*

Jacobsen, D. H., & Kloster, T. B. (2005). What influences the number of bankruptcies? *Economic Bulletin*, 191-211.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An Introduction to Statistical Learning.* New York: Springer Science, Business Media.

Johnsen, P. (2020, May 01). *NHH-professor: – Norge står i en tsunami i tre bølger*. Retrieved from Finanswatch.no: https://finanswatch.no/nyheter/bank/article12108098.ece

Kampevoll, F., & Seibt, S. (2020, May 29). *Én av tre bedrifter med koronahjelp var i minus allerede før krisen*. Retrieved from NRK: https://www.nrk.no/norge/en-av-tre-bedrifter-med-koronahjelp-var-i-minus-allerede-for-krisen-1.15015404

Kinserdal, F. (2021). *BUS440B 6b Investeringer i driftsmidler og arbeidskapital.* NHH.

Kinserdal, F., Hansen, A. M., Pelja, I., & Stemland, B. T. (2019). Er Tradisjonelle Regnskapsnøkkeltall Relevante i En «Moderne» IFRS-verden? *Magma*.

Klecka, W. R. (1980). *Discriminant analysis.* Sage Publications.

Konkursrådet. (2012, 01 24). *Innføring i konkurs*. Retrieved from konkursradet.no: https://www.konkursradet.no/innfoering-i-konkurs.305266.no.html

Kopp, C. M. (2021). *Creative Destruction*. Retrieved from Investopedia: https://www.investopedia.com/terms/c/creativedestruction.asp

Kvernmo, B. B. (2020, July 09). *Korona: Dette er kompensasjonsordningen*. Retrieved from Sticos.no: https://www.sticos.no/fagstoff/kategori/regnskap/korona-dette-blir-kompensasjonsordningen

Matre, A., & Solli, A. N. (2019). The Norwegian Living Dead. *Master Thesis, NHH Norwegian School of Economics*.

McGowan, M. A., Andrews, D., & Millot, V. (2017). *The Walking Dead? Zombie Firms and Productivity Performance in OECD Countries.* OECD Economic Department Working Papers(1372).

Messe, E. N., & Viken, T. (2019). Machine Learning in Bankruptcy Prediction. *Master Thesis, NHH Norwegian School of Economics*.

Nam, C. W., Kim, T. S., Park, N. J., & Lee, H. K. (2008). Bankruptcy prediction using a discrete-time duration model incorporating temporal and macroeconomic dependencies. *Journal of Forecasting, , 27*(6), 493-506.

Norges Bank. (n.d.). *Bankkrisen 1988-1993*. Retrieved from Norges Bank: https://www.norges-bank.no/tema/Om-Norges-Bank/historien/Pengepolitikk-finansiell-stabilitet-og-kapitalforvaltning/Finansiell-stabilitet/1990-tallet/

NOU. (2021:4). *Norge mot 2025*. Retrieved from https://www.regjeringen.no/contentassets/612755ca262842329ae0a7968e66351f/no/pdfs/nou202120210004000dddpdfs.pdf

Næss, A. B., Wahlstrøm, R. R., Helland, F. F., & Kjærland, F. (2017). *Konkursprediksjon for norske selskaper – en sammenligning av regresjonsmodeller og maskinlæringsteknikker.* Trondheim: Fagbokforlaget.

Ohlson, J. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research, 18*(1), 109-131.

O'Shaughnessy, T. (2011). Hysteresis in unemployment. *Oxford Review of Economic Policy, 27*(2), 312-337.

Pelja, I., & Stemland, T. (2017). Predicting bankruptcy for Norwegian firms. *Master Thesis, NHH Norwegian School of Economics*.

Regjeringen. (2022). *Tidslinje: myndighetenes håndtering av koronasituasjonen*. Retrieved from regjeringen.no: https://www.regjeringen.no/no/tema/Koronasituasjonen/tidslinje-koronaviruset/id2692402/

Rotar, L. J., & Krsnik, S. (2020). Analysing the relationship between unemployment benefits and unemployment duration. *Society and Economy*, 280–297.

Røed, K. (1997). Den Selvforsterkende Arbeidsledigheten. *SSB: Rapporter 93/19*.

Schivardi, F., Sette, E., & Tabellini, G. (2020, November). Identifying the Real Effects of Zombie Lending. *The Review of Corporate Finance Studies, 9*(3), 569–592.

Segerberg, M. (2018). Innovasjon på restauranter : introduksjon av konseptuell modell og praktiske anbefalinger. *Master Thesis, NMBU*.

Skatteetaten. (2020). *Om ordningen* . Retrieved from Skatteetaten.no: https://www.skatteetaten.no/en/kompensasjonsordning/om-ordningen/?fbclid=IwAR2mI-U_kycb46k19dx62qaPjzp35Sl77Lpg3-GdR4OQiXFBA9XQVDMRS5I

Skatteetaten. (2021, January 26). *Skatteinnkreving under koronakrisen* . Retrieved from Skatteetaten.no:https://www.skatteetaten.no/en/presse/nyhetsrommet/skatteinnkreving-under-koronakrisen/

Solberg, E. (2020, March 12). Statsministerens innledning på pressekonferanse om nye tiltak mot koronasmitte. Retrieved from Regjeringen.no: https://www.regjeringen.no/no/dokumentarkiv/regjeringen-solberg/aktuelt-regjeringen-solberg/smk/taler-og-innlegg/statsministeren/talerogartikler/2020/statsministerens-innledning-pa-pressekonferanse-om-nye-tiltak-mot-koronasmitte/id2693335/

Sørheim, R., Aadland, T., Alsos, G. A., Aspelund, A., Clausen, T., Eide, A. E., . . . Haneberg, D. H. (2021, January). Innovasjon i krisetider II – innovasjon og omstilling under Covid-19. *Kunnskapsoppbygging for Innovasjon Norge*.

Tracey, B. (2021). The real effects of zombie lending in Europe. *Bank of England, Staff Working Paper No.783*.

Vamraak, T. (2021). *Innlegg: Sats på bedrifter med livets rett heller enn på zombie-bedriftene*. Retrieved from Dagens Næringsliv: https://www.dn.no/okonomi/koronakrisen/kompensasjonsordningen/oppstartsbedrifter/innlegg-sats-pa-bedrifter-med-livets-rett-heller-enn-pa-zombie-bedriftene/2-1-976997

Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S-Plus.* Springer.

Wooldridge, M. J. (2014). *Introduction to Econometrics.* Hampshire: Cengage Learning.

Zhang, G. N., & Ye, F. (2019). *Predicting Financial Distress in Norway.* Master Thesis, NHH Norwegian School of Economics.

Zmijewski, M. E. (1984). Methodological Issues Related to the Estimation of Financial Distress Prediction Models. *Journal of Accounting Research*, 59-82.

10. Appendix

Appendix 1: Evaluation of Fold Performance, MDA

Evaluation of Folds: Linear Discriminant Analysis						
	AUC	**Accuracy**	**Cutpoint**	**Roc01**	**Sensitivity**	**Specificity**
			Variable Set 1			
Fold 1	0.762	0.699	0.033	0.408	0.728	0.697
Fold 2	0.732	0.693	0.033	0.445	0.677	0.694
Fold 3	0.774	0.699	0.033	0.402	0.735	0.697
Fold 4	0.771	0.696	0.033	0.408	0.73	0.694
Fold 5	0.779	0.727	0.035	0.396	0.713	0.727
Simultaneously	0.762	0.694	0.033	0.415	0.721	0.693
			Variable Set 2			
Fold 1	0.769	0.671	0.06	0.394	0.792	0.665
Fold 2	0.743	0.689	0.055	0.408	0.738	0.687
Fold 3	0.765	0.697	0.063	0.406	0.732	0.695
Fold 4	0.768	0.698	0.064	0.392	0.754	0.695
Fold 5	0.75	0.671	0.061	0.423	0.738	0.667
Simultaneously	0.758	0.693	0.062	0.409	0.731	0.691
			Variable Set 3			
Fold 1	0.815	0.709	0.034	0.367	0.781	0.705
Fold 2	0.814	0.748	0.041	0.354	0.751	0.748
Fold 3	0.833	0.783	0.048	0.333	0.745	0.785
Fold 4	0.817	0.723	0.036	0.361	0.77	0.721
Fold 5	0.795	0.726	0.038	0.385	0.73	0.726
Simultaneously	0.815	0.719	0.036	0.365	0.77	0.716
			Variable Set 4			
Fold 1	0.818	0.714	0.035	0.362	0.781	0.711
Fold 2	0.817	0.742	0.04	0.355	0.756	0.741
Fold 3	0.834	0.777	0.046	0.324	0.764	0.778
Fold 4	0.818	0.739	0.039	0.346	0.776	0.737
Fold 5	0.8	0.721	0.036	0.375	0.752	0.719
Simultaneously	0.817	0.735	0.038	0.358	0.761	0.734

Appendix 2: Evaluation of Fold Performance. Logistic Regression

Evaluation of Folds: Logistic Regression						
	AUC	**Accuracy**	**Cutpoint**	**Roc01**	**Sensitivity**	**Specificity**
			Variable Set 1			
Fold 1	0.757	0.722	0.050	0.417	0.687	0.724
Fold 2	0.757	0.728	0.050	0.398	0.708	0.729
Fold 3	0.793	0.734	0.051	0.363	0.754	0.733
Fold 4	0.800	0.717	0.049	0.364	0.775	0.714
Fold 5	0.773	0.714	0.051	0.382	0.749	0.712
Simultaneously	0.775	0.726	0.051	0.388	0.726	0.726
			Variable Set 2			
Fold 1	0.759	0.694	0.077	0.394	0.755	0.691
Fold 2	0.732	0.669	0.077	0.444	0.706	0.667
Fold 3	0.774	0.681	0.071	0.391	0.781	0.676
Fold 4	0.768	0.693	0.076	0.389	0.767	0.689
Fold 5	0.756	0.671	0.073	0.405	0.771	0.666
Simultaneously	0.755	0.683	0.075	0.412	0.741	0.680
			Variable Set 3			
Fold 1	0.831	0.708	0.044	0.348	0.819	0.702
Fold 2	0.796	0.704	0.044	0.387	0.753	0.702
Fold 3	0.798	0.689	0.043	0.388	0.772	0.686
Fold 4	0.807	0.734	0.054	0.378	0.731	0.734
Fold 5	0.797	0.710	0.043	0.381	0.754	0.708
Simultaneously	0.806	0.697	0.043	0.380	0.777	0.693
			Variable Set 4			
Fold 1	0.811	0.702	0.044	0.382	0.765	0.699
Fold 2	0.823	0.738	0.049	0.351	0.769	0.736
Fold 3	0.791	0.698	0.044	0.386	0.763	0.695
Fold 4	0.802	0.740	0.051	0.370	0.737	0.740
Fold 5	0.817	0.733	0.052	0.363	0.756	0.731
Simultaneously	0.808	0.713	0.046	0.374	0.763	0.711

Appendix 3: Evaluation of Fold Performance. GAM

Evaluation of Folds: Generalized Additive Model						
	AUC	**Accuracy**	**Cutpoint**	**Roc01**	**Sensitivity**	**Specificity**
			Variable Set 1			
Fold 1	0.774	0.751	0.052	0.387	0.702	0.753
Fold 2	0.775	0.747	0.050	0.392	0.699	0.749
Fold 3	0.780	0.704	0.047	0.386	0.756	0.701
Fold 4	0.776	0.746	0.050	0.375	0.724	0.747
Fold 5	0.774	0.699	0.045	0.396	0.746	0.697
Simultaneously	0.776	0.740	0.050	0.390	0.708	0.742
			Variable Set 2			
Fold 1	0.771	0.711	0.065	0.379	0.757	0.708
Fold 2	0.775	0.697	0.063	0.399	0.744	0.694
Fold 3	0.760	0.679	0.064	0.417	0.736	0.676
Fold 4	0.762	0.684	0.062	0.402	0.756	0.680
Fold 5	0.782	0.686	0.064	0.383	0.787	0.682
Simultaneously	0.769	0.697	0.064	0.399	0.744	0.695
			Variable Set 3			
Fold 1	0.804	0.747	0.053	0.359	0.745	0.747
Fold 2	0.801	0.748	0.052	0.364	0.737	0.748
Fold 3	0.813	0.768	0.059	0.374	0.704	0.772
Fold 4	0.818	0.732	0.047	0.347	0.782	0.730
Fold 5	0.818	0.746	0.053	0.365	0.737	0.747
Simultaneously	0.811	0.749	0.053	0.366	0.733	0.750
			Variable Set 4			
Fold 1	0.822	0.723	0.045	0.343	0.803	0.718
Fold 2	0.819	0.728	0.046	0.344	0.794	0.725
Fold 3	0.816	0.745	0.054	0.348	0.764	0.744
Fold 4	0.815	0.719	0.045	0.349	0.798	0.715
Fold 5	0.808	0.729	0.046	0.376	0.740	0.729
Simultaneously	0.815	0.725	0.046	0.357	0.776	0.723

Appendix 4: Evaluation of Fold Performance. Random Forest

Evaluation of Folds: Random Forest Model						
	AUC	Accuracy	Cutpoint	Roc01	Sensitivity	Specificity
			Variable Set 1			
Fold 1	0.759	0.672	0.044	0.436	0.715	0.669
Fold 2	0.782	0.738	0.062	0.403	0.691	0.740
Fold 3	0.767	0.688	0.046	0.419	0.722	0.686
Fold 4	0.782	0.748	0.062	0.395	0.694	0.750
Fold 5	0.756	0.707	0.048	0.421	0.697	0.708
Simultaneously	0.769	0.696	0.048	0.419	0.713	0.695
			Variable Set 2			
Fold 1	0.721	0.676	0.032	0.456	0.679	0.676
Fold 2	0.743	0.678	0.032	0.441	0.699	0.677
Fold 3	0.728	0.678	0.032	0.462	0.668	0.678
Fold 4	0.723	0.644	0.028	0.469	0.698	0.642
Fold 5	0.720	0.605	0.024	0.480	0.735	0.599
Simultaneously	0.727	0.661	0.030	0.463	0.686	0.660
			Variable Set 3			
Fold 1	0.825	0.736	0.056	0.344	0.781	0.734
Fold 2	0.832	0.731	0.058	0.338	0.799	0.728
Fold 3	0.834	0.742	0.058	0.332	0.794	0.739
Fold 4	0.812	0.746	0.060	0.376	0.722	0.747
Fold 5	0.818	0.749	0.058	0.378	0.716	0.751
Simultaneously	0.824	0.734	0.056	0.355	0.768	0.732
			Variable Set 4			
Fold 1	0.823	0.726	0.062	0.340	0.805	0.722
Fold 2	0.829	0.763	0.072	0.337	0.761	0.763
Fold 3	0.834	0.739	0.060	0.340	0.784	0.737
Fold 4	0.830	0.753	0.068	0.340	0.768	0.752
Fold 5	0.829	0.712	0.056	0.350	0.809	0.707
Simultaneously	0.829	0.734	0.062	0.343	0.786	0.732

Appendix 5: Model Performance Across Sectors

Performance metrics across Industries using VS3 (MDA)

Sector	sector	AUC	n_bankruptcies	n_observations	avg_bankruptcy	acc	optimal_cutpoint	roc01
L - Real estate activities	12	0.856	292	37011	0.008	0.789	0.001	0.291
G - Wholesale and retail trade; repair of motor vehicles and motorcycles	7	0.837	5473	218742	0.025	0.754	0.019	0.327
J - Information and communication	10	0.832	399	42443	0.009	0.757	0.006	0.337
M - Professional, scientific, and technical activities	13	0.827	984	110240	0.009	0.755	0.003	0.344
H - Transportation and storage	8	0.824	721	41501	0.017	0.728	0.008	0.349
C - Manufacturing	3	0.819	1146	65474	0.018	0.749	0.009	0.348
F - Construction	6	0.817	4116	151678	0.027	0.74	0.014	0.346
A - Agriculture, forestry, and fishing	1	0.813	135	17309	0.008	0.763	0.005	0.33
S - Other service activities	19	0.806	336	21605	0.016	0.758	0.006	0.359
N - Administrative and support service activities)	14	0.798	886	41216	0.021	0.73	0.013	0.376
I - Accommodation and food service activities	9	0.795	1918	45476	0.042	0.708	0.037	0.389
R - Arts, entertainment and recreation	18	0.794	189	14652	0.013	0.712	0.006	0.371
K - Financial and insurance activities	11	0.785	38	4081	0.009	0.73	0.002	0.327
Q - Human health and social work activities	17	0.76	158	38254	0.004	0.766	0.001	0.381
P - Education	16	0.749	109	11292	0.01	0.727	0.005	0.421
B - Mining and quarrying	2	0.727	32	4004	0.008	0.652	0	0.448
E - Water supply; sewerage, waste management and remediation activities	5	0.713	42	3398	0.012	0.791	0.006	0.413

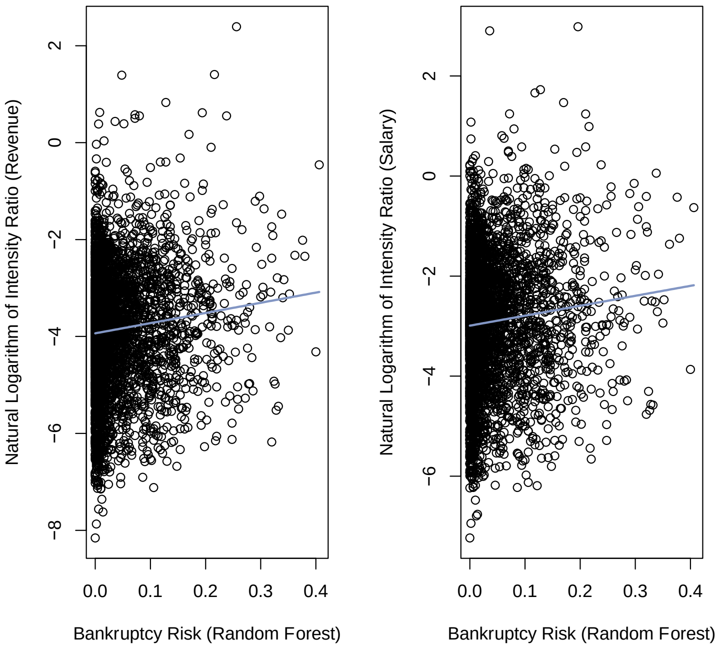
*When trained and tested on all sectors, using 5-fold validation, AUC scores were 0.729, 0.779, and 0.829 using VS1, VS2 and VS3.

Appendix 6: Variable Importance

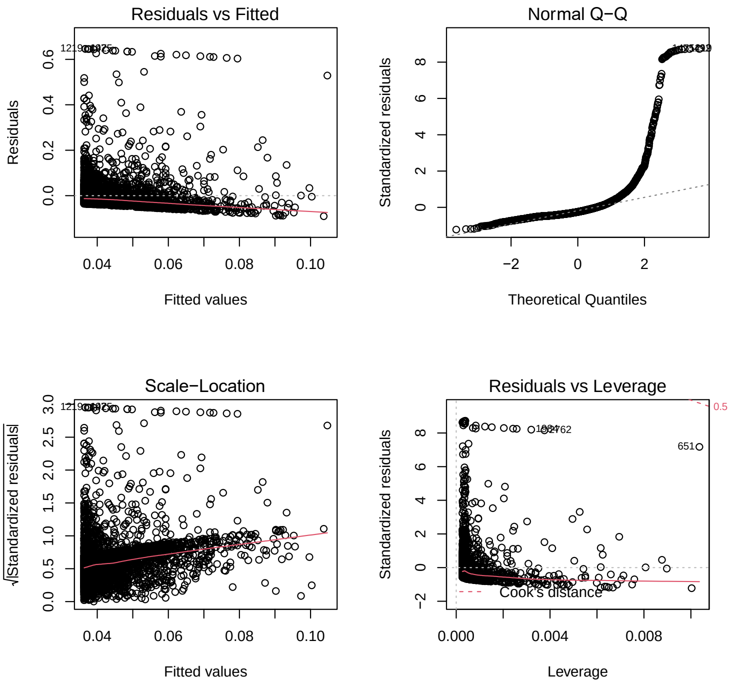
Variable Importance RF	
Variable	**Mean Decrease in Gini**
N_5	294.88
N_2	289.09
N_4	289.06
N_1	279.86
N_3	250.57
N_8	240.93
N_7	233.47
N_9	222.40
$Market\ Share$	221.83
N_{12}	220.00
N_{10}	213.05
N_{11}	210.74
HHI	186.82
N_6	114.64
$Centrality\ Index$	95.95
$Share\ of\ Unemployment$	76.24
$Relative\ Change\ in\ GDP$	75.08
$Changes\ is\ Consumer\ Price\ Index$	72.38
$D_{Negative\ Equity} = 1\ if\ E\ is\ < 0$	27.68
$D_{audit} = 1\ if\ audit\ remark$	26.95
$D_{Age>2\ yrs}$	23.74
$D_{Age>1\ yrs}$	23.40
$D_{management} = 1\ if\ changes\ in\ management$	22.27
$D_{Age>3\ yrs}$	17.52
$D_{Age>4\ yrs}$	17.05
$D_{Age>5\ yrs}$	13.17
$D_{Age>6\ yrs}$	11.45
$D_{Age>7\ yrs}$	11.18
$D_{Age>8\ yrs}$	10.26

Appendix 7: OLS Diagnostic Plots

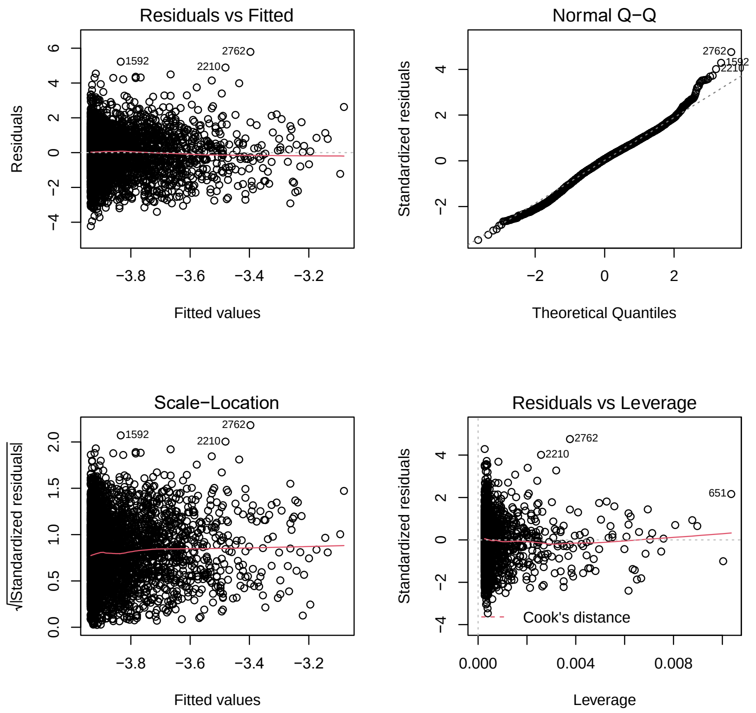
Visual Inspection of Regression Coefficients



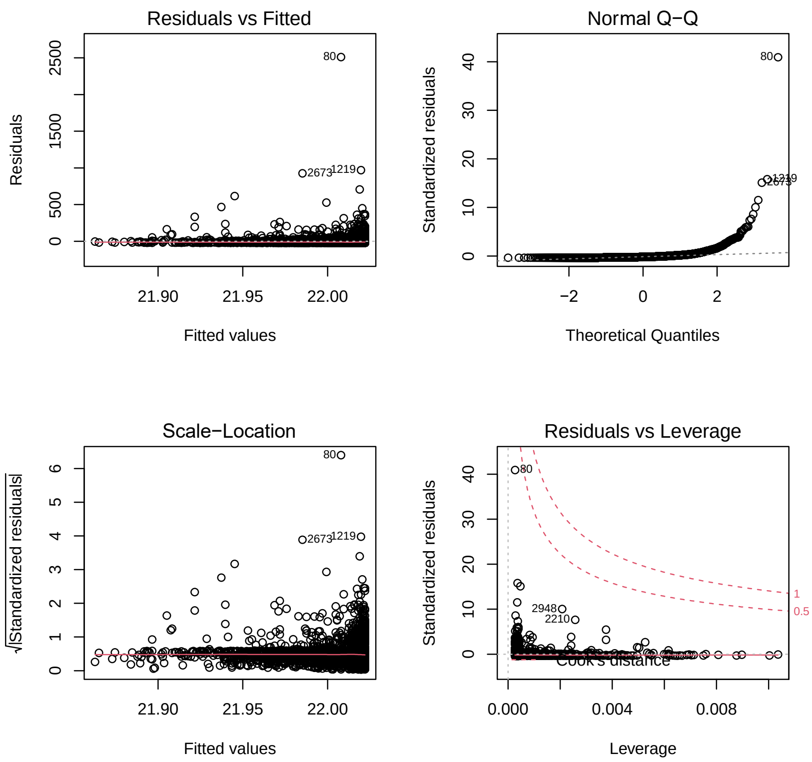
Compensation/Revenue~RF



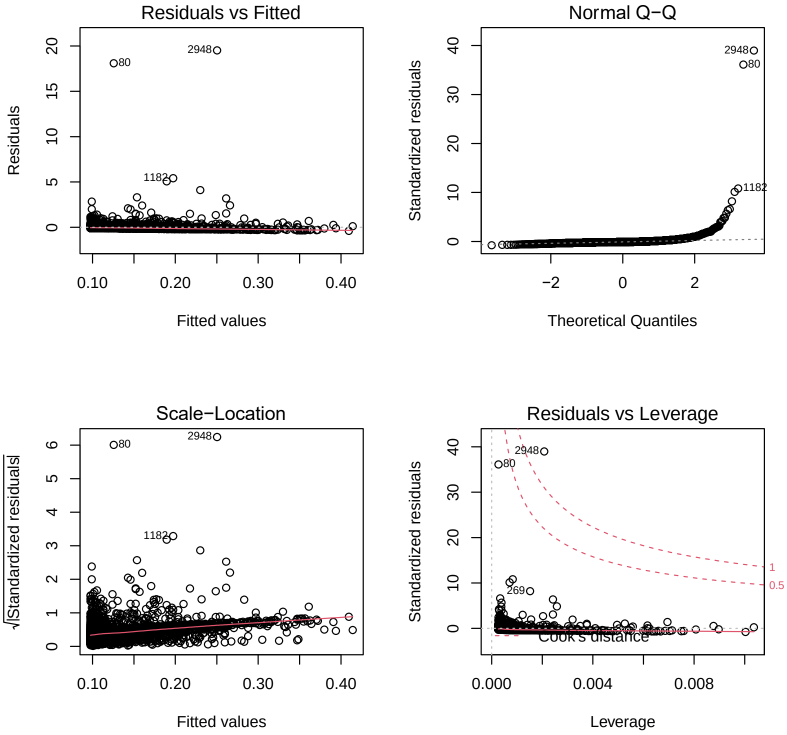
Log transformed Compensation/Revenue (log(CR)) ~RF



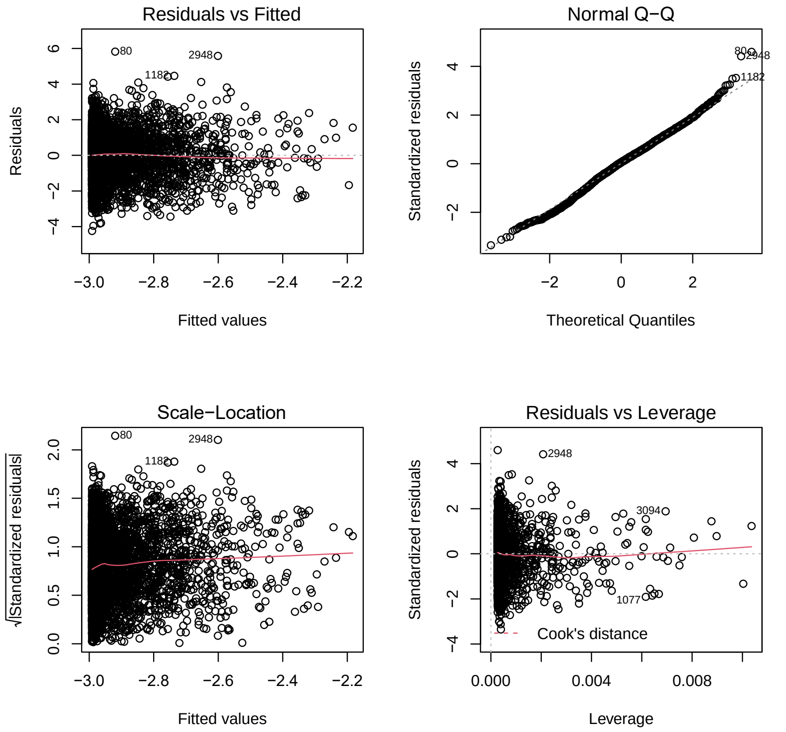
Compensation/Employees (CE) ~RF



Compensation/Labor Costs (CLC) ~RF



Log transformed Compensation/Labor Costs (log(CLC)) ~RF



Appendix 8: Complete Table of Deciles

Analysis of Compensation with Bankruptcy Risk														
	Deciles of Bankruptcy Risk										Definition 1		Definition 2	
Description / Bankruptcy Group	1	2	3	4	5	6	7	8	9	10	0	1	0	1
Environment-specific Information														
Number of Firms	375	375	375	375	375	375	375	374	374	374	3566	181	2933	814
Number of Classified Zombies	5	5	9	8	15	17	12	15	18	9	112	1	84	29
Share of Zombies	1.33%	1.33%	2.40%	2.13%	4.00%	4.53%	3.20%	4.01%	4.81%	2.41%	3.14%	0.55%	2.86%	3.56%
Average Market Share	10.91%	10.70%	8.74%	7.82%	8.21%	11.60%	8.80%	6.98%	6.16%	8.51%	8.95%	6.82%	9.20%	7.57%
Median Market Share	4.98%	3.47%	2.65%	1.59%	1.68%	1.58%	1.12%	1.01%	0.83%	0.61%	1.79%	0.36%	2.11%	0.75%
Average HHI	1441.8	1426.0	1314.9	1290.2	1283.0	1679.5	1454.3	1351.6	1286.2	1472.3	1404.8	1305.7	1397.9	1407.7
Median HHI	1075.2	821.3	777.3	665.0	665.0	723.5	667.7	768.2	665.0	577.8	746.9	448.6	777.3	665.0
Average Centrality Score	5.0	4.9	4.5	4.1	4.3	4.1	4.1	3.9	3.8	3.7	4.3	3.3	4.4	3.8
Aggregated Sums Across Firms														
Sum of Compensation MNOK	280.9	261.8	487.8	546.3	198.3	170.6	138.4	110.1	103.3	112.2	2355.4	54.3	2176.8	232.9
Sum of Sales MNOK	7361.9	5959.3	9955.0	10699.7	5309.1	3867.9	3793.8	2939.9	2405.9	2006.6	53485.2	813.9	49448.7	4850.3
Sum of Employees	12167	11006	14730	16644	8890	8976	8602	6769	6443	5793	97421	2599	86771	13249
Sum of last 3 Years of EBITDA MNOK	649.3	592.8	782.4	709.7	241.4	145.4	166.6	-41.6	-84.1	-141.9	3113.1	-93.0	3249.5	-229.4
Sum of Taxes MNOK	279.1	230.4	361.8	383.7	170.9	162.4	163.8	115.9	106.1	100.5	2031.7	42.9	1851.1	223.5
Averages per Firm														
Average Compensation in TNOK	749.0	698.1	1300.7	1456.8	528.8	455.0	369.1	294.4	276.2	299.9	660.5	299.7	742.2	286.1
Average Revenue in TNOK	19631.7	15891.4	26546.7	28532.6	14157.5	10314.3	10116.7	7860.6	6433.0	5365.4	14998.7	4496.5	16859.4	5958.7
Average Number of Employees	32.4	29.3	39.3	44.4	23.7	23.9	22.9	18.1	17.2	15.5	27.3	14.4	29.6	16.3

Average Labor Cost	262.0	285.1	241.9	242.5	289.7	220.1	206.7	196.3	175.5	163.3	232.7	143.6	245.1	168.0
Average Book Value of Equity	4705.2	3393.9	2654.5	4633.9	1301.5	923.0	262.4	-532.6	-564.5	-1158.4	1704.8	-1214.9	2225.0	-819.1
Average Book Value of Debt	8530.0	6877.2	9017.6	9434.5	5075.8	7375.5	3918.6	4430.2	2812.8	2572.8	6193.6	2325.0	6929.5	2681.5
Average TNOK in EBITDA (last 3 years)	1731.6	1580.8	2086.5	1892.6	643.6	387.8	444.3	-111.3	-224.9	-379.4	873.0	-514.0	1107.9	-281.8
Average TNOK in Outstanding Taxes (balance sheet)	744.3	614.5	964.7	1023.1	455.9	433.0	436.9	309.9	283.7	268.7	569.7	236.9	631.1	274.6
Average Bankruptcy Risk	0.00%	0.14%	0.35%	0.68%	1.20%	2.04%	3.53%	5.77%	9.51%	18.58%	3.21%	23.22%	1.59%	13.47%
Average number of months with shutdown(mar-aug)	3.2	3.3	3.4	3.5	3.5	3.5	3.6	3.7	3.7	3.7	3.5	3.8	3.4	3.7
Average Age	16.63	14.97	13.55	11.30	11.06	8.90	7.56	6.07	5.89	3.94	10.34	3.00	11.40	4.91
Median Age	14	13	11	8	7	5	5	3	3	2	7	2	8	3
Changed Management (%)	0.27%	0.27%	0.53%	1.87%	2.67%	1.87%	4.80%	4.28%	6.42%	5.61%	2.69%	6.08%	2.08%	5.65%
Ratios for Analysis														
Average Compensation Intensity 1	0.037	0.042	0.035	0.041	0.038	0.043	0.041	0.062	0.057	0.117	0.045	0.171	0.042	0.086
Average Compensation Intensity 2	40.214	0.121	607.393	0.114	200.308	0.077	0.167	3.684	0.059	0.056	92.368	0.061	111.616	0.058
Median Compensation Intensity 1	0.022	0.023	0.020	0.020	0.023	0.023	0.024	0.025	0.024	0.028	0.023	0.030	0.022	0.026
Median Compensation Intensity 2	0.032	0.031	0.027	0.027	0.035	0.028	0.030	0.032	0.029	0.032	0.030	0.036	0.030	0.031
Average Compensation per employee I TNOK	22.61	23.17	20.34	24.58	21.64	21.50	25.62	18.10	19.47	23.00	21.76	26.91	22.25	21.12
Average Compensation to Labor Costs	0.09	0.11	0.09	0.09	0.11	0.11	0.15	0.14	0.16	0.25	0.12	0.34	0.11	0.21
Median Growth Rate*	1.07	1.06	1.07	1.06	1.10	1.07	1.06	1.08	1.09	1.11	1.07	1.11	1.07	1.09
Pandemic Impact Proxy*	0.18	0.25	0.19	0.19	0.18	0.19	0.17	0.18	0.21	0.21	0.20	0.20	0.19	0.21
Average Public Debt (balance sheet)	8.67%	9.70%	11.45%	12.48%	12.99%	14.11%	16.97%	17.93%	21.27%	32.05%	14.64%	37.81%	12.94%	25.91%
Average Account Payables (b_sheet)	6.59%	7.47%	9.42%	9.93%	11.80%	12.12%	16.18%	18.76%	29.84%	54.52%	15.02%	69.43%	11.31%	40.49%

Appendix 9: Differences between Observations Groupds

Comparison between the Groups		
Variables	**Received (Group 3)**	**Not Received / Bankrupt (Group 1)**
Environment Specific		
Number of Firms	3747	2184
Average Bankruptcy Risk	4.17%	7.63%
Number of Zombies	113	81
Share of Zombies	3.02%	3.71%
Average Market Share	8.84%	6.22%
Average HHI	1400	1492
Average Centrality Index	4.26	4.24
Aggregated Sums Aross Firms		
Sum of Sales MNOK	54299.07	15750.10
Sum of Employees	100020	33107
Sum of last 3 Years of EBITDA in MNOK	3020.1	404.9
Sum of Taxes MNOK	2074.6	675.2
Average Across Firms		
Average Age	10.0	7.3
Average Revenue in TNOK	14491.3	7211.6
Average sum of Compensation in TNOK	643.1	-
Average Book Value of Equity in TNOK	1563.8	1397.6
Average Book Value of Debt in TNOK	6006.7	4081.2
Average number of Employees	26.7	15.2
Average EBITDA (last three years)	806.0	185.4
Average outstanding taxes in TNOK	553.7	309.2
Average Public Debt (balance sheet)	15.76%	18.53%
Average Account Payables (balance sheet)	17.65%	23.95%
Changed Management (%)	2.86%	4.62%

Firms with no Available Financial Statements, Group 2		
Number of Firms	**Sum Compensation in MNOK**	**Average Compensation in TNOK**
1413	286.87	203.02