Kristian Karlsen
Nils Mathisen

Capital Structure and Machine Learning Techniques in Scandinavia

Master's thesis in Economics and Business Management
Supervisor: Stein Frydenberg
May 2021

Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Economics and Management
NTNU Business School

NTNU
Norwegian University of Science and Technology

Kristian Karlsen
Nils Mathisen

Capital Structure and Machine Learning Techniques in Scandinavia

Master's thesis in Economics and Business Management
Supervisor: Stein Frydenberg
May 2021

Norwegian University of Science and Technology
Faculty of Economics and Management
NTNU Business School

NTNU
Kunnskap for en bedre verden

Acknowledgements

This thesis marks the end of our master's degree in Economics and Business Administration at the Norwegian University of Science and Technology (NTNU).

We want to thank associate professor Stein Frydenberg at the NTNU Business School for his constructive feedback on this thesis during our final semester.

We take full responsibility for the content of this thesis.

Trondheim, 27th May 2021

Kristian Karlsen

Nils Mathisen

Abstract

The composition of equity and debt has been an extensive debate in the corporate finance literature over the last decades. Although, a reliable estimation method for target leverage and the adjustment speed towards it is not yet agreed upon. In prior research, evidence point towards nonlinear relations between capital structure determinants and leverage. In this master's thesis, we apply machine learning techniques to detect complex patterns in the data and test whether the techniques provide increased predictive performance relative to traditional methods. Further, we explore capital structure determinants in Scandinavia, including Finland, and compare the differences between them.

The data set comprises 1294 Scandinavian and Finnish firms listed on the respective stock exchanges during the time period 1990 to 2019. We employ the random forest and the least absolute shrinkage and selection operator model (LASSO) as machine learning techniques to compare against the traditional ordinary least squares model (OLS). Machine learning properties are proven to be particularly helpful in previous literature improving predictive performance as complex patterns in training samples are revealed. The best performing model, random forest, provides more accurate predictions measured by a 5% to 49% reduction in the RMSE compared to the OLS model. The improved predictions increase the speed firms adjust towards the target leverage by two and three times compared to the LASSO and OLS models. Finally, we study differences between cross-country determinants and found two main differences. The high number of fishing and shipping firms in Norway ensures a skewed distribution of the industries, leading to a higher importance of tangibility in Norway compared to the rest of Scandinavia. The strong involvement of the financial institution in the Finnish economy causes a higher sensitivity for macroeconomic fluctuations. Therefore, increasing the importance of the term spread and tax rate determinants in Finland.

In this thesis, we find evidence on the importance of z-score, market-to-book, tangibility, and cash, however, importance is also given to additional variables. Our results suggest that no simple model can be applied to predict capital structure, rather a composition of multiple variables is considered important in predicting leverage. Overall, we find that machine learning techniques can improve predictive performance on capital structure and that cross-country determinants differ between Scandinavian countries.

Sammendrag

Sammensetningen av egenkapital og gjeld har vært en omdiskutert debatt i bedriftsfinans litteraturen de siste tiårene. Til tross for stor diskusjon, er det ikke oppnådd enighet om en pålitelig modell for å beregne gjeldsmålet eller justeringhastigheten mot det. Tidligere forskning finner bevis for ikke-linære forhold mellom determinantene i kapitalstruktur og gjeldsgraden. Vi bruker derfor maskinlæringsteknikker for å oppdage komplekse mønstre i dataene og tester om maskinlæring predikerer kapitalstruktur med høyere nøyaktighet, sammenlignet med tradisjonelle metoder. Videre undersøkes determinanter i Skandinavia, inkludert Finland, og forskjeller mellom disse.

Datasettet består av 1294 skandinaviske og finske børsnoterte selskaper i perioden 1990 til 2019. Vi bruker maskinlæringsteknikkene ≪random forest≫ og LASSO til å sammenligne med OLS-modellen. I tidligere litteratur har maskinlæringsegenskaper vist seg å være spesielt nyttig til å forbedre ≪out-of-sample≫ prediksjon i datasett med komplekse mønstre. Den beste modellen, ≪random forest≫, predikerer med høyere nøyaktighet sammenlignet med OLS-modellen. Dette måles med en reduksjon i RMSE på henholdvis 5% til 49%. Mer nøyaktige ≪out-of-sample≫ prediksjoner i ≪random forest≫ modellen fører til en justeringshastighet som er to og tre ganger høyere sammenlignet med LASSO og OLS-modellene. Videre sammenligner vi forskjellene mellom determinanter på tvers av land i Scandinavia og kommer frem til to hovedfunn. Det høye antallet rederier og fiskebedrifter i Norge gir en skjev fordeling av industriene. Dette øker viktigheten av ≪tangibility≫ i Norge. I Finland har finansinstitusjonene mindre begrensninger for å inntre på eiersiden av selskaper. Sterkere posisjon blant finske banker fører til økt følsomhet for makroøkonomiske svinger, og medfører økt viktighet av determinantene ≪term spread≫ og ≪tax rate≫ i Finland.

I denne masteroppgaven finner vi fremtredende bevis på viktigheten av ≪z-score≫, ≪market-to-book≫, ≪tangibility≫ og ≪cash≫, men legger også vekt på viktigheten av flere determinanter. Resultatene antyder at ingen enkel modell er tilstrekkelig til å forutsi kapitalstruktur, derimot anses en sammensetning av flere variabler som viktige for å predikere gjeldsgraden. Samlet sett finner vi at maskinlæringsteknikker kan forbedre ≪out-of-sample≫ prediksjon på kapitalstruktur, og at viktigheten av ulike determinanter på tvers av de skandinaviske landene er forskjellig.

Contents

1 Introduction **1**

1.1 Research questions and main contributions 2

2 Theory and literature **4**

2.1 Theory 4

2.1.1 Miller and Modigliani 4

2.1.2 Pecking order 4

2.1.3 Trade-off theory 5

2.1.4 Market timing 5

2.2 Literature 6

2.2.1 Capital structure 6

2.2.2 Machine learning 7

2.2.3 Speed of adjustment 8

3 Methodology **10**

3.1 Ordinary least squares 10

3.2 Least absolute shrinkage and selection operator 11

3.3 Random forest 12

3.4 Predictive performance 14

3.5 Speed of adjustment 15

4 Data **16**

4.1 Data selection 16

4.2 Data processing 16

4.3 Dependent variables 17

4.4 Independent variables 17

4.5 Descriptive statistics 19

4.6 Regression model 21

5 Empirical findings **23**

5.1 Predictive performance 23
5.2 Variable presentation 24
5.3 Speed of adjustment 28
5.4 Cross-country determinants and speed of adjustment in Scandinavia 30
5.5 Robustness tests 36

6 Conclusion **38**
6.1 Limitations and further research 39

References **40**

Appendix **43**
A Variable definition and sources 43
B Correlation matrix 46
C OLS assumptions 47
D OLS regression 48
E Testing OLS assumptions in Scandinavia 49
F Variable importance plot industries 51
G Variable importance plot limited 54

List of Figures

1 Importance plot Scandinavia, determinants predicting market and book leverage . . . 25

2 Importance plot Scandinavian countries, determinants predicting market leverage . . . 31

3 Importance plot Scandinavian countries, determinants predicting book leverage . . . 32

4 Importance plot industries, determinants predicting market leverage in Scandinavia using the random forest model. . . . 51

5 Importance plot industries, determinants predicting book leverage in Scandinavia using the random forest model. . . . 52

6 Importance plot excluding the z-score, determinants predicting market and book leverage in Scandinavia using the random forest model. . . . 54

List of Tables

1 Descriptive statistics . . . 19

2 Out-of-sample predictions . . . 23

3 LASSO coefficients Scandinavia . . . 26

4 Speed of adjustment Scandinavia . . . 29

5 LASSO coefficients in Scandinavian countries . . . 33

6 Speed of adjustment in Scandinavian countries . . . 35

7 Industry distribution in Scandinavian countries . . . 36

8 Variable specification, abbreviation, definition, and sources . . . 43

9 Cross-correlation matrix . . . 46

10 Core factor regression . . . 48

11 Hausman test . . . 49

12 Test for heteroskedasticity . . . 49

13 Test for autocorrelation . . . 49

14 Test for multicollinearity . . . 50

1 Introduction

Capital structure has been an active research field for financial economists over the last decades. The debate centres around how firms estimate leverage targets and whether they adjust towards the desired leverage[1]. Capital structure is an important research field because the composition of debt and equity contribute to maximizing a firm's financial value (Binsbergen et al., 2010). The composition can be altered to finance operations internally or externally issuing equity or debt. Although many studies have been performed in the corporate finance literature, a reliable estimation method for target leverage and the speed at which they adjust to target is not yet agreed upon.

Modern research on capital structure began when Modigliani and Miller (1958) introduced the irrelevance proposition, describing how a firm's value will not be affected by capital structure in a perfect capital market. The theory was a simplified version of reality and has been criticized by other theories aiming to disprove the irrelevance theorem. The pecking order theory developed by Myers and Majluf (1984) argues how asymmetric information between investors and firms affects capital structure decision making. Furthermore, the market timing theory states that firms choose between debt and equity based on market situations (Baker & Wurgler, 2002). Finally, the trade-off theory claim that a firm's target leverage is determined by balancing the costs and benefits of debt (Kraus & Litzenberger, 1973). Fischer et al. (1989) introduced the dynamic framework claiming that moving towards target leverage depends to a large extent on the adjustment cost, especially at a sub-optimal debt level. We find that firms move towards target leverage when changing their capital structure. Hence, we find support for the dynamic trade-off framework in Scandinavia.

Prior research finds nonlinear relations between common variables used to measure capital structure dynamics and the dependent variable (Graham & Leary, 2011). We employ the machine learning models random forest and least absolute shrinkage and selection operator (LASSO) to detect these functional forms. Machine learning properties are particularly helpful in improving predictive performance as complex patterns in training samples are revealed. Additionally, an improved prediction of target

[1] Leverage refers to a firms ratio of debt to total capital.

leverage allows more precise estimates on the speed of adjustment. The machine learning methods are specialized in prediction tasks and, therefore, well suited for capturing capital structure dynamics.

1.1 Research questions and main contributions

Inspired by Amini et al. (2021) and their related empirical study, we shed light on how machine learning models can improve accuracy in predicting leverage. To examine capital structure in Scandinavia[2], we use a data set of 1294 firms listed on the Oslo Stock Exchange, Nasdaq Copenhagen, Nasdaq Stockholm, and Nasdaq Helsinki during the time period 1990 to 2019. As a measure of leverage, debt is scaled on the market and book value of assets. Market leverage is useful because of its future-oriented properties while book leverage is a common and frequently applied measure of target leverage (Barclay et al., 2006). To benefit from market and book leverage properties, we apply both measures to our study. The results regarding the most important determinants are similar for both market and book leverage across countries. When national differences are compared, independent of leverage measures, differences become more comprehensive. Institutional factors tend to vary between countries, which is why we compare the determinants between them (Bancel & Mittoo, 2004).

Our thesis contributes to the ongoing research on capital structure and furthers the use of machine learning in this research field. Thus, the research question for this thesis being:

“Can machine learning predict capital structure with higher performance than traditional linear methods in Scandinavia? Are there differences in capital structure determinants across Scandinavian countries?“

Since machine learning techniques are better than linear methods at fitting data to complex forms, we can include a higher number of determinants than allowed by the ordinary least squares (OLS) without running into problems of overfitting (Amini et al., 2021). By estimating capital structure with a larger set of variables, the machine learning techniques perform more accurate predictions than the OLS. Looking at RMSE for book leverage, the random forest estimates an 18% reduction while LASSO estimates a

[2]We choose the wider definition of Scandinavia including Finland, as well as Norway, Sweden, and Denmark.

7% reduction compared to OLS. Random forest outperforms the linear methods when predicting market leverage with 38% lower RMSE than OLS, indicating more nonlinear interactions for market leverage than book leverage. The increase in prediction accuracy led to a higher speed of adjustment towards a leverage target. The random forest estimates increase the speed firms adjust towards target leverage by two and three times, compared to the LASSO and OLS estimates, respectively. The increase in adjustment speed supports Amini et al. (2021) and their findings claiming that linear estimates have lower adjustment speed than nonlinear methods.

The data set comprises firm-, industry-, and macro-level determinants used to predict leverage. The best performing model is random forest estimating z-score, market-to-book, and tangibility as the most important determinants for market leverage. For book leverage: z-score, market-to-book, tangibility, cash, industry leverage, and research and development are the most important determinants. The random forest model does not perform variable selection, thus, the remaining determinants are included in the prediction showing less than 20% importance compared to the most prominent determinant. LASSO selects tangibility, cash, z-score, and dividend as determinants for market leverage prediction. For book leverage: tangibility, cash, z-score, dividend, market-to-book, and industry leverage are selected, discarding all other determinants in the regression. Our data analysis suggests that multiple variables drives and determines leverage. Overall, we conclude that no simple model can be applied to predict capital structure, rather the composition of multiple variables is considered important to predict both market and book leverage.

Capital structure differences between countries in Scandinavia are smaller than in comparison to other non-Scandinavian nations (Bancel & Mittoo, 2004). However, some differences exist and the main findings from this thesis are in Norway and Finland. The tangibility determinant is estimated as most important in Norway due to heavily represented industries, where tangibility is central, affecting the representative sample of listed firms. In Finland, term spread and tax rate are more important compared to their Scandinavian neighbours. This institutional difference is a result of a financial system in Finland where financial institutions stand stronger to interact on the ownership side in firms.

2 Theory and literature

In this section, we present the modern capital structure theories trade-off, pecking order, and market timing. We also provide a brief overview of relevant finance literature reviewing cross-country determinants, speed in which firms adjust towards target leverage, and related empirical work on the research field using machine learning techniques.

2.1 Theory

2.1.1 Miller and Modigliani

Modigliani and Miller (1958) irrelevance proposition argues that leverage does not affect a firm's market value. The theorem is built upon strict assumptions and criticized in a multitude of papers (Frank & Goyal, 2008)(Baker & Wurgler, 2002)(Myers, 1984). The theory is considered the foundation of modern capital structure and the first generally accepted. It greatly influenced the development of new theories such as the trade-off- and pecking order theory (Frank & Goyal, 2008).

2.1.2 Pecking order

Myers and Majluf (1984) investigate how asymmetric information between investors and firms affects capital structure decisions. They developed the pecking order theory stating that firms prefer internal financing over external, and when external financing is needed, they prefer issuing debt over issuing new equity (Myers, 1984).

Firms prefer internal financing because fewer costs are involved, and no information is communicated to the market. If there are no asymmetric information Myers (1984) argues that debt and equity cost are about the same. However, with asymmetric information, the cost of issuing equity increase. Assuming a firm is maximizing the current stockholders' profit, they will not issue stocks when undervalued. Following this logic, issuing stocks signals to the market that the stock is overpriced, and stock prices tend to fall (Myers & Majluf, 1984). As a result of asymmetric information, issuing new debt is preferred over issuing new equity.

2.1.3 Trade-off theory

According to the trade-off theory, a firm's target leverage is determined by balancing the costs and benefits of debt (Myers, 1984). The benefit of debt comes from tax deduction from interest paid, which protects future earnings from tax, while the cost of debt comes from bankruptcy risk and associated financial distress. Agency costs occur when decisions made by managers conflict with the interests of the shareholders. These costs can be reduced by restricting capital, issuing debt mitigate overinvestments due to free cash flow (Jensen, 1986).

If there were no costs associated with moving towards target leverage, the static trade-off framework assumes that each firm's observed target leverage should be its optimal leverage. No such assumption is realistic as costs adjusting to the optimum is anticipated. High adjustment cost could explain variation in leverage, but nothing in the static framework suggests adjustment costs are a first-order concern (Myers, 1984).

Myers (1984) claimed that variation in actual leverage could be explained by high adjustment cost since firms are forced further away from their initial debt ratio. This dynamic trade-off framework describes capital structure over multiple periods, where the benefits of adjusting exceed the downsides. Fischer et al. (1989) present a dynamic framework where minor recapitalization in small and riskier firms, as well as lower-tax and lower-bankruptcy costs, more easily lead to swings in the debt ratio.

2.1.4 Market timing

The market timing theory of capital structure states that firms choose between debt and equity based on the stock prices, interest rates, and other market situations. When a firm needs capital, it should choose whatever is the most beneficial between debt and equity at the given time. Among other things, the firm should issue stocks when the stock price is high. According to market timing theory, ”capital structure evolves as the cumulative outcome of past attempts to time the equity market” (Baker & Wurgler, 2002, p.27). Both Baker and Wurgler (2002) and Welch (2004) find that fluctuations in stock prices and stock returns have an impact on capital structure. Furthermore, Baker and Wurgler (2002) find that firms raising funds when their valuation was high generally has lower leverage, while the opposite is true for firms raising funds when their valuation was low.

2.2 Literature

2.2.1 Capital structure

Frank and Goyal (2009) studied which factors explain capital structure for US-listed firms from 1950 to 2003. They identified six core factors using a linear model selection approach with Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). The determinants with a positive effect on market leverage are median industry leverage, tangibility, log of assets, and expected inflation, while market-to-book, assets ratio, and profit have a negative effect. Regarding book leverage, the results are somewhat similar. However, the factors log of assets, market-to-book, and the effect of inflation are not reliable.

Harris and Raviv (1991) present previous capital structure literature and relate them to empirical evidence based on agency costs, asymmetric information, market interactions, and corporate control. The theories they present are mostly complementary, and most of the potential factors emerging are small and ”general principles”, leaving the context in which the factors are important an unanswered empirical question. Their conclusion was that size, assets, growth opportunities, and non-debt related tax shield increases with leverage, while profitability, research and development, uniqueness, bankruptcy probability, advertising expenditures, and volatility decreases with leverage.

Fama and French (2002) test predictions in the trade-off- and pecking order theory regarding dividends and debt. They find that more profitable firms are less leveraged, consistent with the pecking order theory and contradicting the static trade-off framework. They also find that short-term fluctuations in earnings and investments are generally absorbed in debt, again consistent with the pecking order theory. The trade-off theory assumption from the static framework claiming that firms issue new debt to benefit from the tax shield is rejected. On the other hand, they find that small high growth firms prefer to issue new equity over debt, supporting the market timing theory and going against the pecking order. Frank and Goyal (2009) also find that firms paying dividends has a lower leverage. A paper from Cotei and Farhat (2009) finds evidence for both the pecking order and trade-off theory and imply that the theories are not mutually exclusive.

The capital structure theories differ in their views on target leverage. The pecking order and market timing theories argue that there is no target leverage, while it is an essential part of the trade-off theory. In a survey study, Graham and Harvey (2001) find that 81% of firms have some sort of target leverage and that large firms tend to have a more strict leverage target. Flannery and Rangan (2006) explain that leverage also varies between industries because firms use an industry median debt ratio as a benchmark on how equity and debt should be distributed.

Bancel and Mittoo (2004) conducted a cross-country analysis to find determinants of capital structure in European firms. They found differences across countries, especially between Scandinavian and non-Scandinavian countries. Scandinavian managers tend to have a significantly different view on capital structure, such as equity, convertible debt, and raising foreign capital, compared to their fellow peers in other countries. They speculate that differences come from random chance, the population of firms, or other institutional effects such as moral or ethical norms. Further, Öztekin and Flannery (2012) confirm their hypothesis that institutional features such as transactions costs are associated with a firm's adjustment speed. Their results are consistent with the dynamic trade-off framework.

2.2.2 Machine learning

Amini et al. (2021) use several machine learning techniques to find the determinants of capital structure for US listed firms and compare their findings to an OLS model using the core factors identified by Frank and Goyal (2009). They found that the random forest (RF) machine learning model performed best predictions with an R^2_{os} of 56% compared to the OLS prediction of 39% R^2_{os}. Graham and Leary (2011) found a nonlinear relationship between determinants and the dependent variables in capital structure. Amini et al. (2021) argue that one of the reasons RF predicts more accurately is because it includes nonlinear relationships. According to Frank and Shen (2019), identifying an accurate leverage target has a significant effect on the speed of adjustment. They find that more accurate leverage targets increase the speed of adjustment and argue that the speed of adjustment, on average, is faster than reported in the capital structure literature. Amini et al. (2021) compare the speed of adjustment when target leverage is estimated with different methods and finds that the more accurate predictions made by random forest lead to a higher speed of adjustment.

Using a random forest model, Amini et al. (2021) find that market-to-book, industry leverage, cash, z-score, profitability, stock returns, and firm size are the most important determinants for market leverage. The same determinants are most important for book leverage; however, industry leverage and cash are more important than market-to-book. Further, their linear machine learning model least absolute shrinkage and selection operator (LASSO) also pick cash and industry leverage as the most important factors. Sohrabi and Movaghari (2020) used LASSO to find determinants of capital structure in Iran. Their model produced slightly better estimates in- and out-of-sample compared to the core factor model by Frank and Goyal (2009), estimating tangibility, industry leverage, and profitability as three stable determinants.

2.2.3 Speed of adjustment

Studying the adjustment speed towards a leverage target goes beyond the discussion if a target exists and highlights the importance of reaching the leverage targets. Hovakimian et al. (2001) claim that speed of adjustment changes across firms and time, depending on the transaction costs relative to the changes in leverage. According to Leary and Roberts (2005), firms adjust in clusters, or more precisely once a year on average. Flannery and Rangan (2006) claim that share price fluctuations impact the market debt ratio for a short time period and that the effort to reach target leverage offset these within few years. Further, they found that firms respond to equity issuances and equity price shocks by rebalancing their leverage over the next one to four years.

Fama and French (2002) note that a firm's debt ratios adjust slowly towards target leverage. However, there are conflicting assessments. The literature presents a high variation on the adjustment speed estimates within one time period, varying from 7% to

35% between observed leverage and target leverage. The lowest adjustment speed by Fama and French (2002) is only 7% pr year. Flannery and Rangan (2006) claim that firms close their leverage gap at the rate of more than 30% per year, which is considered a high adjustment speed.

Öztekin and Flannery (2012) found that institutional features such as low transaction costs influence the speed at which the firm adjust towards desired leverage, consistent with the dynamic trade-off theory. More narrowly, they found that the legal origin of Scandinavia, in general, has higher adjustment benefits and lower adjustment cost as a result of the impact from legal and financial traditions compared to other European countries.

3 Methodology

In this section, we present ordinary least squares (OLS), least absolute shrinkage and selection operator (LASSO), and random forest (RF) with associated tuning parameters and specifications. The section also contains a methodological review of predictive performance and the speed of adjustment.

3.1 Ordinary least squares

The ordinary least squares (OLS) technique is used to estimate linear regressions by explaining observed data points by minimizing the sum of squared residuals. Regression analyses use a theoretical model and a set of data to estimate coefficients from a sample of the population. A coefficient indicates the change in the dependent variable with a one-unit increase in the independent variable, while other independent variables remain constant. To estimate valid coefficients, OLS needs to fulfil the seven classical assumptions listed in appendix C. T-values are used for each estimated coefficient in the equation to test the hypotheses on individual regression slope coefficients (Studenmund, 2016). R^2 is used to evaluate the overall fit for each estimated model, representing the variance in the dependent variable explained by the model. The higher R^2, the closer the fit between the estimated regression equation and the sample data.

When combining time-series and cross-sectional data, we get panel data structure. In panel data, observations on the same variable from the same cross-sectional sample are found in more than one time period (Studenmund, 2016). Panel data can be applied with three estimation techniques, pooled OLS, fixed effects, and random effects models. Pooled regression ignores the panel data structure assuming independent observations. This assumption is not likely realistic due to heterogeneity within the panels of our data. In a random effect model the error term does not correlate with the dependent variable. Whereas in the fixed effects model, panel data is used to estimate changes within each unit. Here no correlation between the independent variable and the error term exists, consisting of time-specific and unit-specific effects (Studenmund, 2016). The formula used to calculate fixed effects regression is:

$$Y_{it} = \beta_0 + \beta_1 X_{1i} + \cdot \cdot \cdot + \beta_N X_{Ni} + V_{it} \tag{1}$$

3.2 Least absolute shrinkage and selection operator

Least absolute shrinkage and selection operator (LASSO) is a machine learning regression method performing both shrinkage and variable selection (Tibshirani, 1996). LASSO minimizes residual sum of squares based on a penalty term L_1. The tuning parameter lambda (λ) is used to control the amount of shrinkage. A higher lambda pushes the coefficients towards zero. When a coefficient hits zero it is emitted from the model, this is the model selection part of the LASSO. The LASSO estimate is defined as:

$$\hat{\beta}(lasso) = \arg_{\beta} \min \left\| y - \sum_{j=i}^{p} x_j \beta_j \right\|^2 + \lambda \sum_{j=1}^{p} |\beta_j| \tag{2}$$

Using LASSO instead of OLS you trade lower variance for higher bias, which can lead to better prediction accuracy (Hastie et al., 2015). Reducing the number of variables can establish a model that is easier to interpret. LASSO regressions tend to include one variable from a set of highly correlated variables and set the others to zero leading to higher bias but avoiding multicollinearity.

To reduce bias in the predictions and increase stability of the selection process, Zou (2006) proposed the adaptive LASSO. The adaptive LASSO runs multiple LASSO with the same lambda. The coefficients that are set to zero in the first run are removed before the next LASSO. The remaining variables have penalty weights applied to help drive small coefficients to zero. Because the adaptive LASSO produces more parsimonious and stable models, it is applied to our thesis. The adaptive LASSO estimate is defined as:

$$\hat{\beta}^{*(n)} = \arg_{\beta} \min \left\| y - \sum_{j=i}^{p} x_j \beta_j \right\|^2 + \lambda_n \sum_{j=1}^{p} |\beta_j| \tag{3}$$

We use cross validation to find the tuning parameter creating the most efficient model. Cross validation splits the data into K number of random samples. One of the samples is designated as the testing sample, while the rest are designated as training samples. The LASSO is applied multiple times on this sample set for different values of λ, and the MSE is recorded. This process is then repeated K times so that each split of K is used as test sample. Finally, the λ that results in the lowest MSE on average is chosen as the tuning parameter. To estimate a sparser model the ”one-standard-error” rule is applied (Hastie et al., 2015). Following this rule, the highest λ within one CV standard error of the minimum value of MSE is selected.

3.3 Random forest

Random forest is a machine learning technique that creates multiple regression trees on bootstrapping samples and use the mean results of all the trees. A simple regression tree splits the data into smaller groups, based on an if-then statement, with more homogeneous responses. Within each group, a model is used to predict the outcome (Kuhn & Johnson, 2013). The process of splitting and deciding the depth of the tree is commonly referred to as ”growing” the tree. One of the most common methods of ”growing” the tree is the CART (Breiman et al., 1984). When ”growing” a regression tree, the model begins searching the entire data set S for the predictor[3] and split value that partition the data set into two samples that minimises the overall sum of squared errors (Kuhn & Johnson, 2013).

$$SSE = \sum_{i \in S_1} (y_i - \bar{y}_1) + \sum_{i \in S_2} (y_i - \bar{y}_2)^2 \tag{4}$$

Where y_1 and y_2 are the averages of the training set outcome of the split samples S_1 and S_2. This process is then repeated on S_1 and S_2 finding the predictor and split value within one of the samples that produce the highest reduction in SSE, leading to three unique samples. This process is repeated until the reduction in SSE is marginal, or another specified limit is reached (Kuhn & Johnson, 2013).

Single regression trees are known to lack predictive power and have a higher model variance when compared to other methods (Kuhn & Johnson, 2013). To reduce variance, bootstrap aggregation (bagging) can be used. A bootstrap sample is generated by drawing multiple small random samples from the original data set. When a sample

[3]Predictor is the term for determinant used by Kuhn and Johnson (2013).

is drawn into the bootstrap sample it is not removed from the data set. Therefore, a bootstrap sample can include multiple of the same observation. Bagging is the process of ”growing” a regression tree on each bootstrap sample. The predicted value of the bagged model is the average of all the trees. The process of averaging over multiple trees reduces the variance in the model and makes the predictions more stable (Kuhn & Johnson, 2013).

A potential problem with bagging is highly correlated trees from bootstrap samples generated using the same set of predictors, especially if some predictors dominate the trees (Amini et al., 2021). To reduce correlation among the bootstrapped trees, Breiman (2001) introduced random forest. Random forest takes bootstrap samples and generates a tree for each sample. However, at each split only a random selection of k predictors are considered candidates for splitting. This leads to less correlation among trees and can help highlight predictors that were ascribed to little importance in bagging. The prediction value for random forest is also the average of all trees (Amini et al., 2021).

Machine learning models usually have parameters that cannot be directly estimated in the data (Kuhn & Johnson, 2013). The process of finding and selecting these parameters is called tuning. The tuning parameters for random forest are the number of randomly selected predictors k at each split and the number of trees generated (Kuhn & Johnson, 2013). To tune our model, we ran multiple random forest models on the data set with varying number of trees and predictors selected using different time periods to create training and testing data samples. Kuhn and Johnson (2013) recommend using at least 1000 trees, increasing trees beyond this led to minuscule improvements in our model. For k we found that eight predictors randomly selected at each split on average generated the lowest RMSE. As a result of the tuning process, all random forest models presented in this thesis uses 1000 trees and $k = 8$.

Random forest allows nonlinearity between dependent and independent variables and can highlight hidden interactions better than the linear models such as OLS and LASSO (Amini et al., 2021), leading to better prediction estimates. On the other hand, random forest suffers from a common machine learning issue. As the model becomes more efficient and complex, it gets harder to interpret. The random forest model does not produce coefficients describing the relationships between dependent and independent variables (Kuhn & Johnson, 2013). It does however produce importance factors displaying relative importance for each factor, the factor with the highest importance value is then normalized to 1. The remaining importance factors represent the predictors' importance compared to the most prominent predictor.

3.4 Predictive performance

The data set is divided into separated training and test samples. The training sample is used for building models and estimating parameters, while the testing sample is used to assess predictive performance.

To evaluate and compare the model's predictive performance we use root mean squared error (RMSE) and out-of-sample R-squared(R^2_{os}). MSE is defined as the mean squared difference between the predicted and observed value. RMSE is estimated as the squared root of MSE, penalizing large errors by squaring the errors before they are added. Hence, it is recognized as a good estimation indicator (Studenmund, 2016). R^2_{os} compares the estimated results to a historical average. An R^2_{os} of 0.10 is interpreted as the model predicting 10% better than the historical average. The RMSE and R^2_{os} are defined as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} \left(y_{i,t+1} - \hat{y}_{i,t+1}\right)^2}{N}} \tag{5}$$

$$R^2_{os} = 1 - \frac{\sum_{i=1}^{N} \left(y_{i,t+1} - \hat{y}_{i,t+1}\right)^2}{\sum_{i=1}^{N} \left(y_{i,t+1} - \bar{y}_{i,t+1}\right)^2} \tag{6}$$

Where N is the number of observations in the testing sample, y denotes the observed leverage, $\hat{y}$ the predicted leverage, and $\bar{y}$ the historical average.

3.5 Speed of adjustment

In a market with no friction, a firm would always be at its optimal leverage. However, a realistic market contains adjustment costs which involve that a firm continuously adjusts towards its target leverage. Speed of adjustment is the portion of the gap between target leverage and observed leverage that a firm closes in one year. The gap is defined as:

$$GAP_{i,t} = E\left(y_{i,t+1}\right) - y_{i,t} \tag{7}$$

The speed of adjustment λ is a value between 0 and 1, reflecting the market conditions for adjusting leverage and a firm's desire to adjust towards its target leverage. A firm with small adjustment costs moving towards target leverage have a lower λ than a firm with high adjustment costs.

Following Amini et al. (2021) we use the predicted values as a proxy for target leverage and the partial adjustment framework to estimate the speed of adjustment:

$$\Delta y_{i,t+1} = \lambda GAP_{i,t} + \varepsilon_{i,t+1} \tag{8}$$

Equation (8) is a regression calculated using pooled OLS with bootstrapped standard errors to account for the generated regressor (Pagan, 1984). The regression coefficient for GAP is the estimated speed of adjustment λ. To enhance the interpretation of speed of adjustment we calculate half-life in years as $ln(0.5)/ln(1-\lambda)$. This is the time needed to close half the gap between the observed leverage and the leverage target.

4 Data

In this section, a description of the data set and how we process the data is presented. The dependent and independent variables are reviewed, and the belonging descriptive statistics and correlation matrix are discussed. Lastly, we present the specification test for the linear regression and test the classical OLS assumptions.

4.1 Data selection

The data set comprises listed firms in Scandinavia from 1990 to 2019. We choose the wider definition of Scandinavia including Finland, as well as Norway, Sweden, and Denmark. Our data mainly consist of accounting and balance sheet data from WRDS Compustat Global, stock market conditions from Thomson Reuters Eikon, and the Nordic Statistical database. Market and macroeconomic conditions are extracted from OECD and the Tax Foundation.

4.2 Data processing

Within the 29 years, firms that have been delisted as a result of merges, bankruptcy, or other reasons are included to improve the accuracy of the historical data and avoid survivorship bias. It is important to note that publicly traded companies are not an indication of the average firm. Financial firms with SIC-code inside the range of 6000-6999 such as insurance and banks are excluded, to avoid influencing total leverage composition by including firms with strict capital regulations(Frank & Goyal, 2009). Although firms can change sector and SIC code, this is not taken into account unless it is reflected in the Compustat database. We also follow the example of Frank and Goyal (2009) when dropping missing observations on debt, assets, and market value of equity. After processing the data, we are left with 1 294 firms and 12 910 observations which we consider a representative sample for the population of listed firms in Scandinavia.

To account for outliers, firm-level variables are winsorized at the 1st and 99th percentiles. The method replaces outliers with the observations located on the given upper and lower percentiles. Winsorizing is used extensively in the finance literature. Frank and Goyal (2009) and Öztekin and Flannery (2012) winsorize at the 0.5th and 1st percentiles on both upper and lower distribution, respectively. Finally, all explanatory variables are lagged by one year to get more informative results (Frank & Goyal, 2009).

4.3 Dependent variables

The dependent variable in capital structure models is a measure of leverage. Most commonly a measure of debt ratio is used. In this thesis, total debt to book value of assets (TDA) and total debt to the market value of assets (TDM) are used as dependent variables. The literature is divided regarding which debt ratio measure is better. TDA is a more stable measure since the market value fluctuates more than the book value. Graham and Harvey (2001) find that managers tend not to re-balance their leverage according to market movements because of the transaction costs involved. On the other hand, Welch (2004) argues that book value of assets is primarily a ”plug” number used to balance the assets and equity rather than an accurate target measure of leverage. According to Barclay et al. (2006) TDA is backward-looking because accounting reports what has already taken place, while TDM is forwards looking since the estimation of future growth is included in the market value of assets.

4.4 Independent variables

The 21 independent variables used in this thesis closely follow the variables used in Amini et al. (2021) which is based on Frank and Goyal (2009). The independent variables are categorized into three main categories: firm-, industry-, and macro-level. A detailed description of the variables is available in appendix A.

The firm-level factors are size, profitability, growth, nature of assets, deprecation, and risk. Market-to-book is used as a proxy for growth opportunities (Frank & Goyal, 2009). To measure risk, we include Altman (1968) z-score. The z-score uses five weighted financial ratios as a measure of profitability, leverage, liquidity, solvency, and activity to evaluate the probability of bankruptcy. A score higher than 3 is considered a solid firm and a score less than 1.8 indicates a higher risk of bankruptcy. A dummy variable measures firms that pay dividend at a given time period. For industry-level factors, the median industry leverage and median industry growth are used. Lastly, expected inflation, GDP growth, market returns, top tax rate, and term spread are the macro-level factors.

4.5 Descriptive statistics

	Observ.	Mean	Median	Std.	Min.	Max.
Leverage measures						
TDM	12910	0.275	0.200	0.250	0.000	0.961
TDA	12910	0.248	0.223	0.183	0.001	0.853
Profitability						
Profit	12872	0.044	0.091	0.201	-0.922	0.381
Firm size						
Size	12910	6.591	6.533	2.213	2.003	11.885
Mature	12910	0.949	1	0.220	0	1
Growth						
MTB	12910	1.758	1.0136	2.269	0.243	15.103
ChgAssets	11312	0.050	0.039	0.320	-1.438	1.216
Capex	11755	0.050	0.031	0.059	0	0.329
Nature of Assets						
Tang	12907	0.248	0.182	0.233	0	0.906
RnD	12910	0.075	0	0.377	0	3.293
SGA	12321	0.232	0.162	0.260	0	1.306
Cash	12907	0.137	0.084	0.154	0	0.798
Taxes						
TaxRate	12910	0.259	0.263	0.0367	0.2	0.508
Depr	12770	0.045	0.039	0.034	0	0.203
Risk						
Z-Score	12868	3.719	2.640	5.224	-6.076	35.656
Stock market conditions						
MarketRet	12910	0.092	0.086	0.209	-0.425	0.910
Industry						
IndustryLev	12910	0.230	0.224	0.073	0.086	0.514
IndustryGr	12859	0.044	0.043	0.058	-0.197	0.504
Debt market conditions						
TermSprd	12910	0.008	0.008	0.010	-0.036	0.037
Macroeconomic conditions						
InflationExp	12894	0.016	0.018	0.011	-0.005	0.094
GDPGr	11312	0.039	0.036	0.038	-0.108	0.190
Dividend						
Dividend	12910	0.487	0	0.500	0	1

Table 1: Descriptive statistics
The table describes summary statistics including the number of observations, mean, median, standard deviation, min-, and max values for all variables. Sample period is 1990 to 2019. All variables are winsorized at the 1st and 99th percentiles. For a complementary description of the variables, see appendix A

As table 1 reports, the average firm has a total debt to market ratio of 27.5% and debt to book ratio of 24.8%. Standard deviation informs us that market leverage has a higher variance than book leverage and that there are few firms with higher leverage measures than 50%. The average firm has a profit margin of 4.4%. Furthermore, a higher median indicates that a majority of firms pull the average up while a minority of highly unprofitable firms pull the average further down. Size is the natural logarithm of total assets. The mean and median are relatively even, implying a fairly symmetrical distribution on firm size.

The market-to-book ratio measures the book value of assets in comparison to market value, with an average of 1.76. Assets growth and physical investments both end up with a mean of 5%. However, assets growth has a much higher standard deviation indicating larger differences in growth. Research and development, assets tangibility, and non-production cost have 7.5%, 24.8%, and 23.2% as their mean, respectively. Here research and development have the highest standard deviation, meaning activity in developing new products or services varies between the sample firms. Cash holdings have a mean of 13.7% and the top tax rate average for Scandinavia is 25.9%. Z-score is a measure of bankruptcy probability risk with a 2.6 median, this means that most firms are not considered to have a high risk of bankruptcy. Industry leverage range from 8.6% to 51.4% showing that there is a substantial difference between the industries. The term spread is the difference between the long-term and short-term interest rate, it has a mean of 0.8%.

Correlation matrix

A correlation matrix indicates the correlation between variables in the data sample and is measured by strength and direction of the linear relationship (Studenmund, 2016). The pairwise correlation matrix in appendix B shows correlation and significance level between all response variables and explanatory variables. The table shows that a majority of variables correlating with TDM and TDA are significant at 5%-level. Further multicollinearity can lead to issues in the linear model which we account for further in section 4.6.

The variables most correlated to TDM in appendix B are size, market-to-book, tangibility, non-production cost, cash holdings, tax rate, z-score, and industry leverage. These variables explain most of the variance in a linear relationship. For TDA the most correlated variables are tangibility, cash holdings, z-score, and industry leverage. In more detail, we find that profit is significantly positively correlated to TDM and negatively correlated to TDA. This means that more profitable firms have higher market leverage and lower book leverage. The size and tangibility variables inform us that large firms with a high ratio of intangible assets tend to have high leverage. Market-to-book is negatively correlated with TDM, as we would expect. An increase in the market value of assets leads to a higher market capital and reduces the debt ratio. Further growth factors, non-production costs, and research and development expenditures have a negative effect on both TDM and TDA. Cash has a significant negative effect on TDM and TDA, firms with more cash equivalents have lower leverage. Z-score has a high negative correlation with TDM and TDA, indicating higher leverage for firms with a higher risk of bankruptcy. Industry leverage has a positive correlation with TDM and TDA.

4.6 Regression model

To find an appropriate model for the panel data structure, we perform a Hausman test, testing non-systematic differences in coefficients. Table 11 in appendix E presents the Hausman test rejecting the null hypothesis, indicating that differences in the coefficients are systematic (Studenmund, 2016). Based on the Hausman test results, the fixed effects model is applied to the linear regression model.

To achieve favorable estimates in the OLS regression model, we further test the seven classical assumptions listed in appendix C using the respective tests in appendix E. The modified Wald test in table 12 calculates groupwise heteroskedasticity in the residuals of a fixed effects regression model. This test rejects the null hypotheses, and the errors exhibit groupwise heteroskedasticity. The Wooldridge test in table 13 rejects the null hypothesis of no first-order autocorrelation. The VIF-index in table 14 shows signs of multicollinearity, making it hard to distinguish the effect of one variable from the other and causes the t-values to decrease (Studenmund, 2016). As a rule of thumb, VIF values above five can imply problems regarding multicollinearity. As to normality, the central limit theorem ensure normally distributed coefficients in large samples (Stock & Watson, 2015). To reduce variance in the residuals our data is winsorized at the 1st and 99th percentile in the tails of the distribution. Although some of the OLS assumptions are not fulfilled, we obtain unbiased standard errors by utilizing clustered standard errors in the linear regression model.

5 Empirical findings

In this section, we interpret the results obtained from the data. And prove that machine learning techniques lead to a significant improvement in predictive performance accuracy. We further present the determinants of importance, explaining target leverage in Scandinavia and the speed firms adjust towards the target leverage. Lastly, we study the institutional differences between the Scandinavian countries.

5.1 Predictive performance

The R^2_{os} and RMSE measure and compare the predictive performance of the RF, LASSO, and OLS models. The data set is split into a training period 1990-2015 and a testing period 2016-2019. R^2_{os} compares the predictive results to the historical average, while RMSE measures the difference between the predicted and observed value. The results are presented in table 2.

	Norway		Sweden		Denmark		Finland		Scandinavia	
	R^2_{os}	RMSE	R^2_{os}	RMSE	R^2_{os}	RMSE	R^2_{os}	RMSE	R^2_{os}	RMSE
OLS										
TDM	0.205	0.235	0.010	0.190	-0.837	0.271	-1.552	0.332	-0.408	0.251
TDA	0.369	0.192	-0.019	0.174	-0.156	0.208	0.142	0.158	0.067	0.184
LASSO										
TDM	0.152	0.243	0.153	0.176	0.067	0.193	0.251	0.180	0.060	0.205
TDA	0.423	0.184	0.136	0.160	0.181	0.175	0.368	0.136	0.197	0.171
RF										
TDM	0.484	0.188	0.439	0.139	0.507	0.138	0.419	0.157	0.435	0.156
TDA	0.427	0.182	0.247	0.148	0.411	0.148	0.392	0.132	0.340	0.154

Table 2: Out-of-sample predictions
This table shows R^2_{os} and RMSE from OLS, LASSO, and RF, predicting book leverage (TDM) and market leverage (TDA) as dependent variables, in Norway, Sweden, Denmark, Finland, and Scandinavia as a whole. Explanatory variables in the OLS correspond with the core factors in Frank and Goyal (2009), while the LASSO and the RF model further demonstrate the importance of additional factors as in Amini et al. (2021).

The RF model has the most accurate predictions across countries and leverage measures, with performance significantly increasing when predicting the market leverage. Among the linear methods, the LASSO predictions are more accurate than the OLS except for market leverage in Norway. In Scandinavia, the RF predictions provide a RMSE of 0.156 and 0.154 for TDM and TDA, respectively. LASSO and OLS experience a drop off in performance predicting TDM in Scandinavia with RMSE at 0.205 for LASSO and 0.251 for OLS, compared to TDA where the RMSE is 0.171 for LASSO and 0.184 for OLS. RF outperforms the OLS with a RMSE reduction of 16.30% for TDA and a 38.85% reduction for TDM. Similar results are found on a national level. When estimating TDM as a leverage measure in Scandinavia, the OLS R^2_{os} is -0.408. The negative R^2_{os} in the OLS prediction means that the estimate is 40.8% less accurate than the historical average, compared to RF predicting 43.5% better than the historical average in Scandinavia, respectively.

The improvement in RF prediction accuracy for TDM implies more nonlinear and hidden interactions in market leverage than book leverage. Our results are consistent with Amini et al. (2021), who found a greater improvement in accuracy predicting TDM than TDA with the RF model compared to the linear methods. The predictive performance of the LASSO model compared with the OLS model does not match with the ranking of the respective model's performance in Amini et al. (2021). Our LASSO model performed better than the OLS in Scandinavia, while the LASSO model on US-listed firms in Amini et al. (2021) performed worse than the OLS. The results are expected, as Amini et al. (2021) uses a similar data set as the core factor model in Frank and Goyal (2009).

5.2 Variable presentation

In this subsection, we present the variables' importance from the RF model and the coefficients from the LASS0 and OLS model, and further compare the results with theory and literature from section 2.

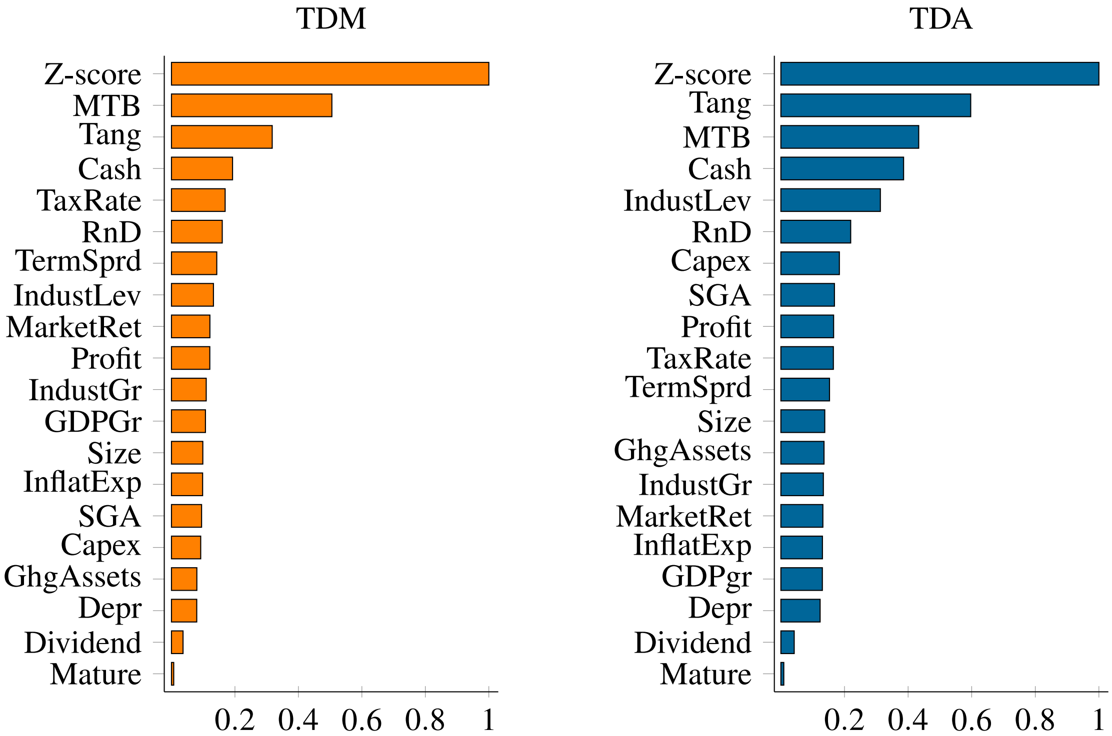


Figure 1: Importance plot Scandinavia, determinants predicting market and book leverage

The figure shows the importance of explanatory variables predicting book leverage (TDM) and market leverage (TDA) in Scandinavia using the random forest model. The variable with the highest importance value is normalized to 1. For a complementary description of variables, see appendix A.

Figure 1 presents variable importance fitted to the training data in Scandinavia. RF chooses z-score as the most prominent determinant for predictive performance. Market-to-book and tangibility are also considered important in predicting TDM, and the determinants have a minimum of 20% weights relative to z-score normalized to one. Tangibility, market-to-book, cash, median industry leverage, and research and development are important determinants for predicting TDA. RF includes 20 variables with individual weights of importance predicting the response variable.

Table 3 presents explanatory variables predicting TDM and TDA using the adaptive LASSO. Tangibility is the only positive coefficient on TDM, while cash, z-score, and dividend are negative. LASSO further predicts positive coefficients on TDA to be tangibility, market-to-book, and industry leverage, while cash, z-score, and dividend are negatively correlated. The goal of the LASSO is to build a prediction model, whereby LASSO does not care about p-values (StataCorp, 2021). LASSO coefficients are penalized for better predictions and should not be interpreted as OLS coefficients. Additional LASSO methodology in section 3.2.

Table 10 in the appendix presents OLS regression results in Scandinavia using a data set combining the training and testing samples. The regression model is based on the core factor model presented by Frank and Goyal (2009). Comparing our model to the core factor model, coefficients show the same direction on TDM for all determinants. The determinants with a positive coefficient on TDM are size, tangibility, industry leverage, and expected inflation, while profitability and market-to-book are negative. All coefficients are significant at a 1%-level. For TDA size, tangibility, industry leverage, and expected inflation are positive and significant at a 1%-level. Profitability is negative and significant at a 1%-level. Lastly, market-to-book is insignificant on TDA as we expected.

	TDM	TDA
Tang	0.334	0.248
Cash	-0.259	-0.218
Z-Score	-0.021	-0.016
Dividend	-0.015	-0.019
MTB	-	0.015
IndLev	-	0.145

Table 3: LASSO coefficients Scandinavia

The table presents determinants selected by the LASSO model to predict market (TDM) and book leverage (TDA) and their associated penalized coefficients. For a complementary description of the variables, see appendix A.

Noticeably LASSO selects few factors, due to variable selection, compared to RF drawing predictive information from a much broader set of covariates (Amini et al., 2021). Tangibility, cash, z-score, market-to-book, and industry leverage are selected in the LASSO model and given high importance in the RF model. However, the importance of the dividend dummy is divided by the LASSO and the RF model. LASSO selects dividend in top six determinants predicting TDM and TDA, while RF estimates that dividend does not exceed 5% importance relative to the most important factor. The low importance estimated for the dividend is consistent with the low importance for dummies in the RF model in Amini et al. (2021).

Z-score is a measure of risk, consisting of multiple factors, and the most important determinant explaining capital structure in our model. The RF model selects z-score as the most important determinant explaining TDM and TDA, supported by the LASSO. The lasso estimate a negative z-score coefficient indicating that financially healthy firms are less leveraged and that firms with high leverage have an increased risk of bankruptcy. Z-score is not included in the OLS regression. According to the trade-off theory, capital structure is a balance between costs and benefits of having debt. From a trade-off theory perspective, firms with high risk and low profit should reduce leverage to gain from lower debt costs and utilize the tax shield more reasonably. The theory supports the importance of the determinant in both machine learning models.

The RF model finds market-to-book as the second and third most important determinant for TDM and TDA. The LASSO model includes market-to-book as a determinant predicting TDA but omits it predicting TDM. Further, market-to-book is negative and significant at a 1%-level predicting TDM in the OLS. The variable is suited looking forward onto market-based leverage. Market-to-book is not significant for TDA in the OLS regression, consistent with Frank and Goyal (2009) explaining how book leverage is backward-looking. However, the coefficient on TDA is positive in the LASSO model and consistent with the trade-off theory claiming how firms take on new debt at a lower cost with a higher market-to-book ratio. From a market timing perspective, firms with a high market-to-book ratio reduce leverage to take advantage of equity mispricing.

According to the RF model, cash is the fourth most important determinant predicting TDM and TDA. The LASSO model also uses cash for prediction, and the coefficient is negative. As cash holdings within firms increase, so do the financial flexibility allowing firms to reduce debt or finance new projects with retained earnings instead of issuing new debt. Consistent with the pecking order theory that firms prefer internal over external financing.

Tangible assets are defined as assets with a real transactional value. All three models estimate tangibility with a high degree of importance predicting TDM and TDA. OLS and LASSO estimate a positive coefficient for the determinant, consistent with Titman and Wessels (1988) claiming that firms with high proportions of tangible assets have a higher target debt ratio and relatively low bankruptcy costs because of the collateral properties of tangible assets. RF chooses tangibility as the second most important determinant for TDA and the third most important for TDM. Because book value excludes growth opportunities and is closely reflected by tangible assets, tangibility is considered more important for predicting TDA. This importance is supported by Amini et al. (2021), claiming that managers are likely to base the debt issuance policy on book leverage.

Median industry leverage proves to be an important variable estimated from the three models. The OLS regression estimates a positive relationship. Industry leverage is more important in predicting TDA than TDM using RF and is only included in the LASSO predicting TDA. These findings are reasonable as the industry leverage is calculated using median book leverage. Median industry leverage does not affect capital structure itself, but firms often use industry leverage as a proxy for target leverage (Hovakimian et al., 2001).

To summarize, figure 1 demonstrates that the importance of determinants varies depending on the economic environment. The top four determinants explaining capital structure are on a firm-level, followed by a large share of macro-level determinants influencing TDM while the fifth most important determinant for TDA is on an industry-level. Overall, results show that many interactions are important in explaining the capital structure, but not all are equally important. The results agree with Amini et al. (2021), demonstrating how their top three most important determinants predicting TDM and TDA stay the top three important after adding the macro-level determinants directly into the model.

5.3 Speed of adjustment

Speed of adjustment refers to how quickly a firm closes the gap between observed leverage and target leverage. Using a partial adjustment framework, the speed of adjustment is a balance between the costs and benefits of reaching the target leverage. The predicted leverage for each model is used as target leverage, resulting in different speed of adjustment for each model. The speed of adjustment in Scandinavia is presented in table 4.

	OLS	LASSO	RF
TDM			
GAP	0.066***	0.099***	0.202***
Observations	2676	2676	2689
Adjusted R^2	0.022	0.031	0.055
Half-life in years	10.225	6.663	3.078
TDA			
GAP	0.156***	0.160***	0.219***
Observations	2676	2676	2689
Adjusted R^2	0.058	0.047	0.066
Half-life in years	4.099	3.974	2.807

Table 4: Speed of adjustment Scandinavia

The table shows the speed of adjustment in Scandinavia, with total debt scaled by the market value of assets (TDM) and the book value of assets (TDA) as dependent variables. The table includes the GAP, the number of observations, adjusted R^2, and half-life in years. Speed of adjustment is the portion of the gap between target leverage and observed leverage that a firm closes in one year, estimated using the following model: $\Delta y_{i,t+1} = \lambda GAP_{i,t} + \varepsilon_{i,t+1}$. The GAP is calculated as $GAP_{i,t} = E\left(y_{i,t+1}\right) - y_{i,t}$ and half-life in years is calculated as $ln(0.5)/ln(1-l)$. The GAP is tagged with *** p<0.01, ** p<0.05, * p<0.1 regarding the significance level.

The speed of adjustment results is closely related to the prediction accuracy of the model. Similarly to the prediction accuracy, differences between the models' speed of adjustment are greater when estimating TDM. Whereby, speed of adjustment estimates for RF is twice as fast compared to the LASSO model and three times faster than the OLS model in table 4. The higher speed of adjustment estimated from a more accurate leverage target is consistent with Amini et al. (2021) and supporting Frank and Shen (2019) theory that errors calculating target leverage in previous literature leads to a slower speed of adjustment. Moreover, when estimating the speed of adjustment for TDA, the difference between the LASSO and the OLS model are small compared to the RF model estimated 33% faster. Lastly, all methods estimate a higher adjustment speed for TDA than TDM. Graham and Harvey (2001) found in their survey that firms tend to use book leverage as target leverage, which aligns with our results that the speed of adjustment is higher for TDA than TDM.

5.4 Cross-country determinants and speed of adjustment in Scandinavia

Figure 2 and 3 show the importance factors from the RF model on TDM and TDA for each Scandinavian country. For TDM the two most important determinants z-score and market-to-book, are consistent across all countries. For TDA, the z-score is the most important determinant in Denmark, Sweden, and Finland. Tangibility is the most important determinant in Norway, followed by the z-score. Further, z-score, tangibility, cash, industry leverage, research and development, and market-to-book are estimated with higher than 20% importance across all Scandinavian countries.

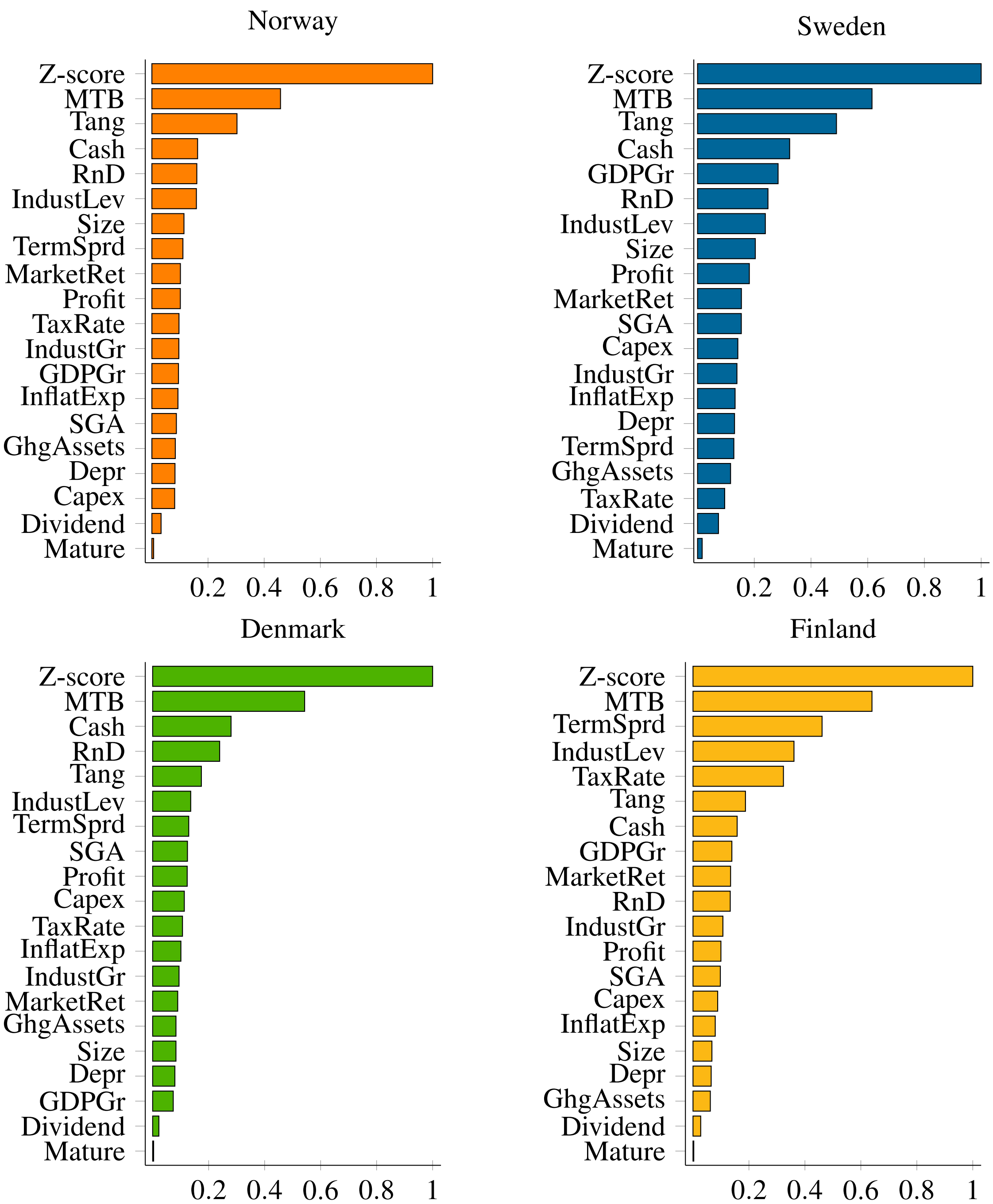


Figure 2: Importance plot Scandinavian countries, determinants predicting market leverage

The figure shows the importance of explanatory variables predicting market leverage (TDM) in Norway, Sweden, Denmark, and Finland using the random forest model. The variable with the highest importance value is normalized to 1. For a complementary description of the variables, see appendix A.

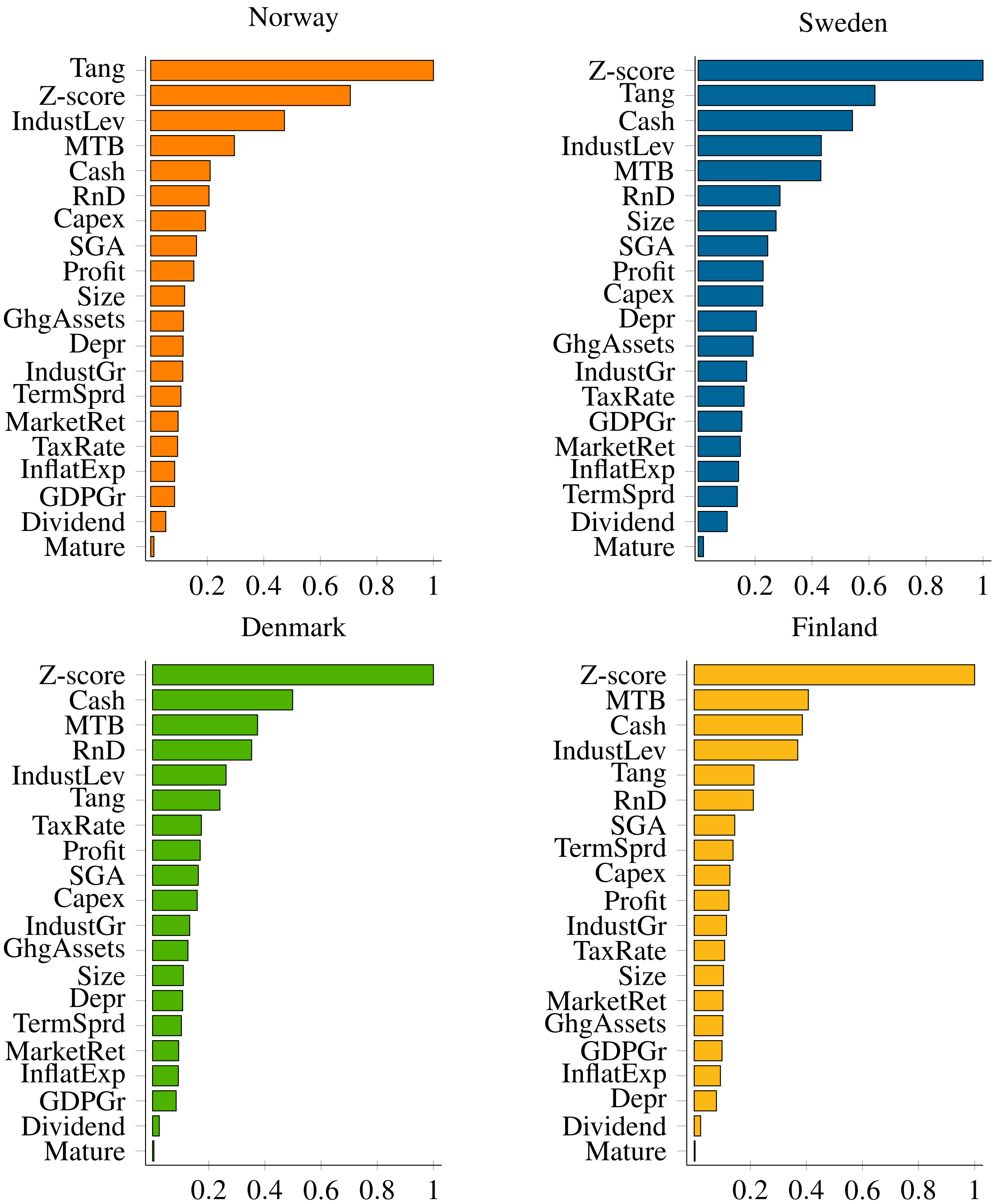


Figure 3: Importance plot Scandinavian countries, determinants predicting book leverage

The figure shows the importance of explanatory variables predicting book leverage (TDA) in Norway, Sweden, Denmark, and Finland using the random forest model. The variable with the highest importance value is normalized to 1. For a complementary description of the variables, see appendix A.

Table 5 presents the LASSO coefficients for TDM and TDA in the Scandinavian countries. Tangibility is selected by the LASSO model predicting TDM and TDA in all countries. Further, the z-score is included in all the LASSO models except TDA in Denmark. Tangibility, industry leverage, and market-to-book have positive coefficients, while z-score, cash, and dividend are negative.

	Norway		Sweden		Denmark		Finland	
	TDM	TDA	TDM	TDA	TDM	TDA	TDM	TDA
Tang	0.200	0.421	0.316	0.151	0.065	0.040	0.259	0.023
Z-Score	-0.011	-0.001	-0.013	-0.005	-0.015	-	-0.058	-0.037
Cash	-	-	-0.203	-0.147	-0.450	-0.399	-	-0.299
IndLev	-	0.142	-	-	-	0.168	0.142	0.350
Dividend	-	-0.018	-	-	-0.036	-	-	-
MTB	-	-	-	-	-	-	-	0.031

Table 5: LASSO coefficients in Scandinavian countries
The table presents the determinants selected by the LASSO model to predict market (TDM) and book leverage (TDA) and their associated penalized coefficients in Norway, Sweden, Denmark, and Finland. For a complementary description of the variables, see appendix A.

The results from the OLS regression are presented in table 10 in the appendix. Market-to-book is significant for TDM on a 5%-level in Finland, Norway, and Sweden. Profitability is insignificant in Norway on TDA but is positive and significant in the other countries. Tangibility is positive and significant on a 1%-level for all countries except Finland. Further, size is positive and significant on a 5%-level for all countries. Industry leverage is positive and significant in all countries on TDM on a 10%-level and in Finland and Denmark for TDA. Lastly, expected inflation over the coming years is positive and significant in Sweden and Denmark, but not significant in Norway and Finland.

In Norway, the RF model chooses tangibility as the most important variable on TDA, while the LASSO and the OLS models estimate positive and significant coefficients. The determinant indicates that having security on debt is more important in Norway than in the rest of Scandinavia, which is driven by a larger share of firms in industries where tangibility is central. We will further address results regarding industry differences in section 5.5. In addition, cash is less important for book leverage in Norway, indicating that liquidity is more important in the rest of Scandinavia.

In Sweden, industry leverage is not included in the LASSO model and only significant at a 10%-level for TDM in the OLS. In the RF model, industry leverage's importance is equivalent to the other countries, indicating a nonlinear relationship between industry leverage and capital structure in Sweden. According to the RF importance plot, cash is an important determinant in Denmark and Sweden. The cash coefficients in LASSO estimate higher values for Denmark, indicating liquidity as an important factor for leverage. Further, research and development have a higher importance in Denmark and Sweden than in Norway and Finland. Denmark and Sweden also have the lowest debt ratios in Scandinavia. This is consistent with Mac an Bhaird and Lucey (2010) findings that because of the intangible nature of research and development, it is not suitable as debt collateral and usually financed by equity.

All three estimation methods OLS, LASSO, and RF, selects industry leverage as more important in Finland than in the other Scandinavian countries. Indicating that firms in Finland focus more on adjusting towards the industry median leverage. Market-to-book is included only for book leverage in the LASSO model in Finland and is the second most important factor in the RF importance plot. The importance of the determinant signals that Finnish firms with a high market valuation tend to have more leverage than the rest of Scandinavia. The RF model, using TDM in Finland, predicts higher importance for term spread and tax rate than in the rest of Scandinavia. La Porta et al. (1999) discussed different financial systems, where Finland has a ”strong bank” system compared to its Scandinavian neighbours' ”weak bank” system. A ”strong bank” system is recognized by financial institutions standing stronger to interact on the ownership side in firms, also affecting a banks' power through lending (La Porta et al., 1999). When expected profitability for banks is low, term spread is expected to be high (Aksoy & Basso, 2014). As the profits decrease, financial institutions issue less debt. Whereby Finland is more sensitive to these macro conditions as financial institutions have fewer restrictions to participate on the ownership side of firms.

Speed of adjustment

Table 6 presents the speed of adjustment in Scandinavian countries using machine learning techniques. The highest speed of adjustment is estimated by RF[4]on TDM in

Finland with a half-life of 2.119 years. Following after is Sweden with a half-life of 2.792 years, Norway with 3.117 years, and Denmark with a half-life of 3.352 years. Finnish firms adjust towards target leverage 50% faster than Danish firms on TDM. For TDA the speed of adjustment is highest in Sweden with a half-life of 2.073 years, followed by Finland with 3.326 years, Norway with 3.901 years, and Denmark with 6.728 years. Swedish firms adjust towards the target leverage more than three times faster than Danish firms, 50% faster than Finnish firms, and nearly twice as fast as Norwegian firms. The overall speed in which firms adjust towards target leverage in Scandinavia, estimated with the machine learning techniques, is relatively high compared to the average of extreme points in the classical literature by Fama and French (2002) and Flannery and Rangan (2006). However, the results are in line with Amini et al. (2021) estimating with the same machine learning techniques.

	Norway		Sweden		Denmark		Finland	
	LASSO	RF	LASSO	RF	LASSO	RF	LASSO	RF
TDM								
GAP	0.125***	0.199***	0.131***	0.220***	0.103***	0.187***	0.159***	0.279***
Observations	415	422	1525	1530	286	286	451	451
Adjusted R^2	0.059	0.066	0.041	0.061	0.052	0.046	0.046	0.092
Half-Life in Years	5.199	3.117	4.943	2.792	6.388	3.352	4.013	2.119
TDA								
GAP	0.137***	0.163***	0.248***	0.284***	0.080***	0.098*	0.144***	0.191***
Observations	415	422	1525	1530	285	286	451	451
Adjusted R^2	0.033	0.035	0.098	0.102	0.019	0.016	0.029	0.028
Half-Life in Years	4.692	3.901	2.427	2.073	8.360	6.728	4.467	3.266

Table 6: Speed of adjustment in Scandinavian countries
The table shows the speed of adjustment in Norway, Sweden, Denmark, Finland, and Scandinavia, with total debt scaled by the market value of assets (TDM) and the book value of assets (TDA) as dependent variables. The table includes the GAP, the number of observations, adjusted R^2, and half-life in years. Speed of adjustment is the portion of the gap between target leverage and observed leverage that a firm closes in one year, estimated using the following model: $\Delta y_{i,t+1} = \lambda GAP_{i,t} + \varepsilon_{i,t+1}$. The GAP is calculated as $GAP_{i,t} = E\left(y_{i,t+1}\right) - y_{i,t}$ and half-life in years is calculated as $ln(0.5)/ln(1-l)$. The GAP is tagged with *** p<0.01, ** p<0.05, * p<0.1 regarding the significance level.

Our results provide evidence that the importance of the industry leverage in figure 2 and figure 3 is related to the speed of adjustment. The RF model estimates industry leverage as a less important determinant in Denmark, and the speed of adjustment in Denmark as

[4] The increased speed of adjustment in the RF model is related to its prediction accuracy. A further description can be found in subsection 5.3.

lower, compared to the rest of Scandinavia. Finland has the highest adjustment speed on TDM, and industry leverage is more important than in the rest of Scandinavia. In Sweden, the speed of adjustment is higher for TDA than TDM, and the importance of industry leverage is also higher for TDA than TDM. These results indicate that industry leverage is a proxy for target leverage. Therefore, it can be a firm's desire to distinguish itself as low financial risk relative to industry competitors (Fitzgerald & Ryan, 2019). Öztekin and Flannery (2012) find institutional differences between Scandinavian countries to be minor, compared to the differences between Scandinavia and other regions of the world. Öztekin and Flannery (2012) further explain that Denmark and Finland have lower bankruptcy costs than Norway, leading to lower adjustment speed. The low bankruptcy cost is consistent with the low speed of adjustment estimated for Denmark but not in Finland. Lastly, we can not miss that Denmark has the lowest speed of adjustment and the lowest number of observations[5].

5.5 Robustness tests

A robustness test is executed to test whether differences in determinants affecting the capital structure are influenced by a skewed distribution of industries in the Scandinavian countries. Table 7 presents the data set split into six industries, where it becomes clear that the relative industry distribution varies in each country.

	Norway	Sweden	Denmark	Finland	Scandinavia
Fishing, mining, agriculture	17.21%	3.88%	0.77%	0.21%	5.14%
Construction	5.18%	2.74%	4.55%	2.83%	3.52%
Manufacturing	37.48%	54.53%	61.20%	59.18%	53.37%
Transport, electronics, comms.	23.59%	6.54%	13.25%	8.71%	11.29%
Wholesome- and retail trade	3.36%	9.31%	6.94%	8.25%	7.59%
Services, public administration	13.18%	23.00%	13.29%	20.82%	19.09%

Table 7: Industry distribution in Scandinavian countries
The table presents industries split into six groups based on their respective SIC codes for each Scandinavian country and Scandinavia as a whole.

Appendix F presents the importance plot from the RF model estimated on each industry. For TDA, tangibility is the most and second most important factor in the transport, communications, and electronics industries and fishing, mining, and agriculture

[5] The very nature of the data set for each country will influence the results.

industries. These industries are heavily represented on the Oslo Stock Exchange compared to the other Scandinavian exchanges. Moreover, tangibility is less important for the manufacturing and trade industries which are less represented on the Oslo Stock Exchange. Noticeably, tangibility is the most important factor in Norway for TDA as a result of the industry distribution. The robustness test on the empirical findings in Finland shows no evidence that the importance of the macro-level determinants is due to the industry distribution. Neither, for higher importance of research and development in Denmark. Lastly, the median industry leverage offers no comparison value for the RF model estimated on industry samples. The reason is that the data set is split into industry samples, and the median industry is homogenized in each sample and given less importance.

The z-score is a combination of factors, as described in appendix A, and the most important determinant predicting TDM and TDA in the RF model in figure 1. We test whether z-score affects the relative importance of the other factors in the full RF model by estimating a new RF model in appendix G, excluding the z-score. The second most important determinant following the z-score in the full RF model predicting TDM is market-to-book. In the new RF model, market-to-book becomes the most important determinant followed by the same three determinants as in the full RF model. For TDA, tangibility is the second most important determinant in the full RF model and becomes the most important determinant in the new model. Cash was the fourth determinant in the full RF model and is second in the new RF model. Market-to-book was second in the full RF model, ending at number four in the new RF model. Liquidity has properties that can be found in the z-score and can explain why cash is given less importance in the full RF model. Furthermore, the rest of the determinants has minor deviations in the new model compared to the full RF model. Thus, this robustness test affirms that the inclusion of the z-score does not majorly affect the order of the other determinants following it and can therefore be included in the full RF model.

6 Conclusion

The objective of this thesis is to use machine learning to predict capital structure in Scandinavia and compare it against the linear method ordinary least squares (OLS). And highlight the differences in capital structure determinants across the Scandinavian countries. Machine learning models such as random forest and least absolute shrinkage and selection operator (LASSO) are applied to take advantage of their attributes suited for capturing capital structure dynamics. The results show that the machine learning models perform more accurate predictions than OLS. Random forest is the best performing model with consistent results throughout the Scandinavian countries, allowing nonlinear and complex interactions, outperforming the OLS model with a 5% to 49% lower RMSE. The increase in prediction accuracy by the random forest model also leads to a higher speed in which firms adjust towards target leverage by two and three times the speed estimated by LASSO and OLS, respectively. Overall, we find evidence that machine learning improves predictions on capital structure in Scandinavia compared to the traditional methods.

Predicting capital structure in Scandinavia, the random forest model estimates z-score, market-to-book, tangibility, and cash as the most important determinants. The LASSO model selects tangibility, cash, z-score, and dividend to predict market and book leverage. The LASSO model performs variable selection, discarding the determinants not selected, while random forest does not discard the low importance determinants. In the LASSO model, the z-score has a negative coefficient predicting leverage, consistent with the trade-off theory perspective claiming firms with high risk and low profit should reduce leverage to utilize the tax shield in a better way. Cash is also consistent with the dynamic trade-off theory as an increase in leverage means utilizing the tax shield more reasonably, reducing agency costs due to restricting capital. Market-to-book is consistent with a market timing perspective, as firms with a high book ratio reduce leverage to take advantage of equity mispricing. Tangible assets can be used as debt collateral, reducing the cost of debt. Following the trade-off theory, optimal leverage increases when the cost of debt decreases. The positive coefficients estimated by the LASSO and OLS models are consistent with the trade-off theory. Overall, no simple model or theory is sufficient to predict the capital structure and its belonging dynamics. Rather a composition of multiple variables is considered important to predict capital structure in Scandinavia.

From the implemented machine learning techniques, we find two main differences in cross-country determinants comparing capital structure in Scandinavia. The first finding is in Norway, where tangibility stands out as most important due to heavily represented industries where tangibility is central. The second finding comprises the Finnish financial system. In Finland, the financial institutions stand stronger to interact on the ownership side of firms. The involvement causes the institutions to be more sensitive to fluctuations from macroeconomic conditions, leading to an increase in importance of term spread and tax rate in our random forest model.

6.1 Limitations and further research

In this thesis, we found that machine learning techniques RF and LASSO can estimate accurate predictions on capital structure. We limited the thesis by using only two machine learning methods. Including other prominent methods such as neural networks or gradient boosting could bring an interesting perspective. Since machine learning techniques avoid problems with overfitting, expanding the number of variables included would be an interesting approach. In that way, we can see if there are any effects our model does not include. Only listed firms are included in this thesis, which would make a similar study with non-listed firms informative. A comparison with such a study would bring a wider perspective and investigate how representative listed firms are for the economy as a whole.

Random forest suffers from the black box problem, whereas the more complex interactions it includes, increasing performance but lowers interpretation. To counteract this problem and increase interpretability a SHAP or LIME model would be exciting to implement on this or a similar study. Another thing to note regarding the RF model is how low importance the dummy variable is assigned.

References

Aksoy, Y. & Basso, H. S. (2014). Liquidity, Term Spreads and Monetary Policy. *The Economic Journal*, *124(581)*, 1234–1278. https://doi.org/10.1111/ecoj.12087.

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance*, *23(4)*, 589–609. https://doi.org/10.2307/2978933.

Amini, S., Elmore., R., Öztekin, Ö. & Strauss, J. (2021). Can machines learn capital structure?, 1–68. http://dx.doi.org/10.2139/ssrn.3473322.

Baker, M. & Wurgler, J. (2002). Market timing and capital structure. *The Journal of Finance*, *57(1)*, 1–32. https://doi.org/10.1111/1540-6261.00414

Bancel, F. & Mittoo, U. R. (2004). Cross-country determinants of capital structure choice: A survey of european firms. *Financial Management*, *33(4)*, 103–132. https://www.jstor.org/stable/3666330.

Barclay, M. J., Smith, J. C. & Morellec, E. (2006). On the debt capacity of growth options. *The Journal of Business*, *79(1)*, 37–60. https://doi.org/10.1086/497404.

Binsbergen, J. H. V., Graham, J. R. & Yang, J. (2010). The cost of debt. *The Journal of Finance*, *65(6)*, 2089–2136. https://doi.org/10.1111/j.1540-6261.2010.01611.x

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32. https://doi.org/10.1023/A:1010933404324.

Breiman, L., Friedman, J., Olshen, R. A. & Stone, C. J. (1984). *Classification and regression trees*. Chapman; Hall.

Cotei, C. & Farhat, J. B. (2009). The trade-off theory and the pecking order theory: Are they mutually exclusive? *North American Journal of Finance and Banking Research*, *3(3)*, 1–16. https://ssrn.com/abstract=1536714.

Fama, E. F. & French, K. R. (2002). Testing trade-off and pecking order predictions about dividends and debt. *The Review of Financial Studies*, *15(1)*, 1–33. http://www.jstor.org/stable/2696797.

Fischer, E. 0., Heinkel, R. & Zechner, J. (1989). Dynamic capital structure choice: Theory and tests. *The Journal of Finance*, *44(1)*, 19–40. https://doi.org/10.2307/2328273.

Fitzgerald, J. & Ryan, J. (2019). The impact of firm characteristics on speed of adjustment to target leverage: A uk study. *Applied Economics*, *51(3)*, 315–327. https://doi.org/10.1080/00036846.2018.1495822

Flannery, M. J. & Rangan, K. P. (2006). Partial adjustment toward target capital structures. *Journal of Financial Economics*, *79(3)*, 469–506. https://doi.org/10.1016/j.jfineco.2005.03.004.

Frank, M. Z. & Goyal, V. K. (2009). Capital structure decisions: Which factors are reliably important? *Financial Management*, *38(1)*, 1–37. https://doi.org/10.1111/j.1755-053X.2009.01026.x

Frank, M. Z. & Shen, T. (2019). Corporate capital structure actions. *Journal of Banking & Finance*, *106*, 384–402. https://doi.org/10.1016/j.jbankfin.2019.07.014.

Frank, M. Z. & Goyal, V. K. (2008). Chapter 12 - trade-off and pecking order theories of debt. In B. E. Eckbo (Ed.), *Handbook of empirical corporate finance* (pp. 135–202). Elsevier. https://doi.org/10.1016/B978-0-444-53265-7.50004-4.

Graham, J. R. & Harvey, C. R. (2001). The theory and practice of corporate finance: Evidence from the field. *Journal of Financial Economics*, *60(2-3)*, 187–243. https://doi.org/10.1016/S0304-405X(01)00044-7

Graham, J. R. & Leary, M. T. (2011). A review of empirical capital structure research and directions for the future. *Annual Review of Financial Economics*, *3*, 309–345. https://doi.org/10.1146/annurev-financial-102710-144821

Harris, M. & Raviv, A. (1991). The theory of capital structure. *The Journal of Finance*, *46(1)*, 297–355. https://doi.org/10.1111/j.1540-6261.1991.tb03753.x

Hastie, T., Tibshirani, R. & M.Wainwright. (2015). *Statistical learning with sparsity: The lasso and generalizations*. Chapman; Hall.

Hovakimian, A., Opler, T. & Titman, S. (2001). The debt-equity choice. *The Journal of Financial and Quantitative Analysis*, *36(1)*, 1–24. https://doi.org/10.2307/2676195.

Jensen, M. C. (1986). Agency costs of free cash flow, corporate finance, and takeovers. *American Economic Review*, *76(2)*, 323–329. http://dx.doi.org/10.2139/ssrn.99580.

Kraus, A. & Litzenberger, R. H. (1973). A state-preference model of optimal financial leverage. *The Journal of Finance*, *28(4)*, 911–922. https://doi.org/10.2307/2978343.

Kuhn, M. & Johnson, K. (2013). *Applied predictive modeling*. Springer.

La Porta, R., Silanes, F. L.-D. & Shleifer, A. (1999). Corporate ownership around the world. *The Journal of Finance*, *54(2)*, 471–517. https://doi.org/10.1111/0022-1082.00115

Leary, M. T. & Roberts, M. R. (2005). Do firms rebalance their capital structures? *The Journal of Finance*, *60(6)*, 2575–2619. https://doi.org/10.1111/j.1540-6261.2005.00811.x

Mac an Bhaird, C. & Lucey, B. (2010). Determinants of capital structure in irish smes. *Small Bus Econ*, *35*, 357–375. https://doi.org/10.1007/s11187-008-9162-6

Modigliani, F. & Miller, M. H. (1958). The cost of capital, corporation finance and the theory of investment. *The American Economic Review*, *48(3)*, 261–297. http://www.jstor.org/stable/1809766.

Myers, S. C. (1984). The capital structure puzzle. *The Journal of Finance*, *39(3)*, 575–592. https://doi.org/10.2307/2327916.

Myers, S. C. & Majluf, N. S. (1984). Corporate financing and investment decisions when firms have information that investors do not have. *Journal of Financial Economics*, *13(2)*, 187–221. https://doi.org/10.1016/0304-405X(84)90023-0

Öztekin, Ö. & Flannery, M. J. (2012). Institutional determinants of capital structure adjustment speeds. *Journal of Financial Economics*, *103(1)*, 88–112. https://doi.org/10.1016/j.jfineco.2011.08.014.

Pagan, A. (1984). Econometric issues in the analysis of regressions with generated regressors. *International Economic Review*, *25(1)*, 221–247. https://doi.org/10.2307/2648877.

Sohrabi, N. & Movaghari, H. (2020). Reliable factors of capital structure: Stability selection approach. *The Quarterly Review of Economics and Finance*, *77*, 296–310. https://doi.org/10.1016/j.qref.2019.11.001.

StataCorp. (2021). *Stata lasso reference manual: Release 17*. Statistical Software. College Station, TX: StataCorp LLC. https://www.stata.com/bookstore/lasso-reference-manual/.

Stock, J. H. & Watson, M. W. (2015). *Introduction to econometrics* (3rd ed.). Pearson.

Studenmund, A. H. (2016). *Using econometrics: A practical guide* (7th ed.). Pearson.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, *58(1)*, 267–288. https://www.jstor.org/stable/2346178.

Titman, S. & Wessels, R. (1988). The determinants of capital structure choice. *The Journal of Finance*, *43(1)*, 1–19. https://doi.org/10.1111/j.1540-6261.1988.tb02585.x

Welch, I. (2004). Capital structure and stock returns. *Journal of Political Economy*, *112(1)*, 106–132. http://dx.doi.org/10.1086/379933.

Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American Statistical Association*, *101 (476)*, 1418–1429. https://doi.org/10.1198/016214506000000735.

Appendix

A Variable definition and sources

Table 8: Variable specification, abbreviation, definition, and sources

Variable	Abbreviation	Definition	Source
Market value of equity	MVE	The stock's fiscal year close price times common shares outstanding. Data source: Eikon	(Amini et al., 2021)
Market value of assets	MVA	Debt in current liabilities (DLC) plus long-term debt (DLTT) minus deffered taxes and investment tax credit (TXDITC) plus the market value of equity (MVE). Data source: Compustat	(Amini et al., 2021)
Leverage measures			
Market leverage	TDM	Debt in current liabilities (DLC) plus long-term debt (DLTT) divided by the market value of assets (MVA). Data source: Compustat.	(Amini et al., 2021)
Book leverage	TDA	Debt in current liabilities (DLC) plus long-term debt (DLTT) divided by total assets (AT). Data source: Compustat.	(Amini et al., 2021)
Profitability			
Profitability	Profit	Operating income before depreciation (OIBDP) divided by total assets (AT). Data source: Compustat.	(Amini et al., 2021)
Firm size			
Total assets	Size	The logarithm of total assets (AT). Data source: Compustat.	(Amini et al., 2021)
Mature firm	Mature	A dummy variable which equals one if the firm has been listed and on the Compustat database for more than three years, zero otherwise.	(Amini et al., 2021)
Growth			
Market-to-book	MTB	Market value of assets (MVA) divided by total assets (AT). Data source: Compustat.	(Amini et al., 2021)
Assets growth	ChgAssets	Change in the logarithm of total assets (AT).	(Amini et al., 2021)

Continued on next page

Table 8: Variable specification, abbreviation, definition, and sources – continued from previous page

Variable	Abbreviation	Definition	Source
Physical investment	Capex	Capital expenditures (CAPX) divided by total assets (AT). Data source: Compustat.	(Amini et al., 2021)
Nature of assets			
Assets tangibility	Tang	Net property, plant and equipment (PPENT) divided by total assets (AT). Data source: Compustat.	(Amini et al., 2021)
Innovation investment	RnD	Research and development expenses (XRD) divided by total revenue (REVT). By following the standard practise in the literature, we set the RnD expenses to zero whenever it is missing in the Compustat database. Data source: Compustat.	(Amini et al., 2021)
Non-production cost	SGA	Selling, general and administrative expenses (XSGA) divided by total revenue (REVT). Data source: Compustat.	(Amini et al., 2021)
Cash holdings	Cash	Cash and short-term investments (CHE) divided by total assets (AT). Data source: Compustat.	(Amini et al., 2021)
Taxes			
Top tax rate	TaxRate	Top statutory tax rate for each country from 1990 to 2019. Data Source: TaxFoundation	(Amini et al., 2021)
Depreciation	Depr	Depreciation and amortization (DP) divided by total assets (AT). Data source: Compustat.	(Amini et al., 2021)
Risk			
Bankruptcy probability	Z-score	Altman (1968) Z-Score is 3.3 times earnings before interest and taxes (EBIT) plus 1.4 times retained earnings (RE) plus 1.2 times the difference in total current assets (ACT) and total current liabilities (LCT) plus revenue total (REVT) divided by total assets (AT). Plus 0.6 times market value of equity divided by total liabilities (LT). Data source: Compustat and Eikon.	(Altman, 1968)
Stock Market Conditions			

Continued on next page

Table 8: Variable specification, abbreviation, definition, and sources – continued from previous page

Variable	Abbreviation	Definition	Source
Market returns	MarketRet	Percent change in average annual overall share price indices from previous year. Data source: Nordic Statistics database	
Industry			
Industry leverage	InduLev	The median of corporate leverage (TDM) by 4-digit SIC code and by year. Data source: Compustat	(Amini et al., 2021)
Industry growth	InduGr	The median of assets growth (ChgAsset) by 4-digit SIC code and by year. Data source: Compustat	(Amini et al., 2021)
Debt market conditions			
Term spread	TermSprd	The difference between 10-year government bond and three-month treasury bills. Data Source: OECD	(Amini et al., 2021)
Macroeconomic conditions			
Expected inflation	InflatExp	The expected change in consumer price index over the coming year. Data Source: OECD	(Amini et al., 2021)
Growth in GDP	GDPGr	Change in logarithm of gross domestic product. Data Source: OECD	(Amini et al., 2021)
Dividend			
Dividend	Dividend	A dummy which equals one if the firm has paid dividends (DVT), and zero otherwise. Data source: Compustat	

B Correlation matrix

Table 9: Cross-correlation matrix

This table presents pairwise correlation coefficients between company specific, industry specific, market conditions and macroeconomic variables with belonging t-statistics for the Scandinavian countries. The variables are defined in appendix A. * equals p<0.05.

Variables	TDM	TDA	Profit	Size	Mature	MTB	ChgAssets	Capex
TDM	1.000							
TDA	0.707*	1.000						
Profit	0.072*	-0.025*	1.000					
Size	0.205*	0.123*	0.396*	1.000				
Mature	0.111*	0.044*	0.225*	0.163*	1.000			
MTB	-0.402*	-0.090*	-0.193*	-0.200*	-0.146*	1.000		
ChgAssets	-0.067*	-0.050*	0.190*	0.099*	-0.011	0.015	1.000	
Capex	0.122*	0.180*	0.093*	0.071*	0.019*	0.039*	0.075*	1.000
Tang	0.423*	0.454*	0.212*	0.259*	0.104*	-0.124*	-0.054*	0.507*
RnD	-0.124*	-0.063*	-0.319*	-0.114*	-0.057*	0.156*	-0.034*	-0.074*
SGA	-0.200*	-0.126*	-0.339*	-0.329*	-0.096*	0.141*	-0.108*	-0.156*
Cash	-0.335*	-0.318*	-0.271*	-0.232*	-0.156*	0.284*	0.054*	-0.093*
TaxRate	0.218*	0.076*	0.153*	0.168*	0.189*	-0.162*	-0.036*	0.206*
Depr	0.086*	0.139*	-0.004	-0.090*	0.012	0.011	-0.209*	0.208*
Z-score	-0.411*	-0.372*	0.127*	-0.088*	-0.096*	0.700*	0.130*	-0.015
MarketRet	-0.054*	-0.031*	0.063*	0.039*	0.019	0.063*	0.053*	0.068*
InduLev	0.274*	0.310*	0.087*	0.200*	0.047*	-0.062*	-0.034*	0.267*
InduGr	-0.077*	-0.028*	0.029*	0.024*	-0.092*	0.105*	0.211*	0.107*
TermSprd	0.008	-0.057*	0.095*	0.025*	0.148*	-0.086*	-0.052*	0.015
InflatExp	0.159*	0.089*	0.027*	0.065*	-0.022*	-0.077*	-0.026*	0.092*
GDPGr	-0.027*	-0.020*	0.048*	0.005	-0.001	0.026*	0.093*	0.080*
Dividend	-0.040*	-0.106*	0.383*	0.410*	0.149*	-0.119*	0.047*	0.053*

Variables	Tang	RnD	SGA	Cash	TAX	Depr	Z-score	MarketRet
Tang	1.000							
RnD	-0.113*	1.000						
SGA	-0.290*	0.018	1.000					
Cash	-0.263*	0.351*	0.042*	1.000				
TaxRate	0.231*	-0.059*	-0.255*	-0.028*	1.000			
Depr	0.272*	-0.027*	0.017	-0.128*	0.108*	1.000		
Z-score	-0.163*	0.081*	-0.004	0.347*	-0.090*	-0.131*	1.000	
MarketRet	0.032*	-0.012	-0.044*	0.021*	0.004	-0.027*	0.064*	1.000
InduLev	0.429*	-0.020*	-0.179*	-0.071*	0.228*	0.097*	-0.093*	-0.053*
InduGr	-0.013	0.003	-0.008	0.054*	-0.040*	-0.060*	0.094*	0.423*
TermSprd	0.047*	-0.026*	-0.072*	-0.051*	0.215*	0.036*	-0.037*	0.314*
InflatExp	0.106*	-0.013	-0.116*	0.002	0.336*	0.055*	-0.070*	-0.273*
GDPGr	0.019*	-0.011	-0.049*	0.019*	0.069*	-0.005	0.037*	0.448*
Dividend	0.106*	-0.151*	-0.150*	-0.136*	0.162*	-0.064*	0.028*	0.064*

Continued on next page

Table 9: Cross-correlation table – continued from previous page

Variables	InduLev	InduGr	TermSpr	InflationExp	GDPGr	Dividend
InduLev	1.000					
InduGr	-0.045*	1.000				
TermSprd	-0.078*	-0.205*	1.000			
InflatExp	0.159*	0.054*	-0.290*	1.000		
GDPGr	-0.028*	0.444*	-0.096*	0.103*	1.000	
Dividend	0.028*	-0.010	0.138*	-0.002	0.064*	1.000

C OLS assumptions

The classical OLS assumptions (Studenmund, 2016):

1. The regression model is linear in the coefficients and the error term
2. The error term has a population mean of zero
3. The independent variables are uncorrelated witht the error term
4. The error term is uncorrelated with each other (no autocorrelation)
5. The error term has constant variance (no heteroskedasticity)
6. The independent variables is not a perfect linear function of any other explanatory variable (no multicollinearity).
7. The error term is normally distributed

D OLS regression

Table 10: Core factor regression

This table show result from the linear regression based on the core factor model by Frank and Goyal (2009) for all the Scandinavian countries. All the control variables are lagged by one year, the variables are defined in appendix A. The table consists of regression coefficients tagged with robust t-statistics and belonging significance level *** p<0.01, ** p<0.05, * p<0.1, and clustered standard deviation in parentheses. At the bottom of the table number of observations and R^2 are displayed.

	Norway		Sweden		Denmark		Finland		Scandinavia	
Variables	TDM	TDA	TDM	TDA	TDM	TDA	TDM	TDA	TDM	TDA
Profit	-0.197***	-0.0926	-0.129***	-0.102***	-0.327***	-0.224***	-0.439***	-0.342***	-0.227***	-0.146***
	(0.073)	(0.064)	(0.025)	(0.024)	(0.070)	(0.065)	(0.084)	(0.068)	(0.024)	(0.022)
Size	0.0522***	0.0187**	0.0388***	0.0318***	0.0699***	0.0343***	0.136***	0.0177**	0.0672***	0.0270***
	(0.011)	(0.007)	(0.007)	(0.005)	(0.014)	(0.012)	(0.015)	(0.009)	(0.006)	(0.004)
MTB	-0.0198***	-0.00175	-0.00701***	0.00347	-0.00515	0.00614	-0.0171**	0.000718	-0.0111***	0.00251
	(0.004)	(0.002)	(0.002)	(0.002)	(0.003)	(0.004)	(0.007)	(0.004)	(0.002)	(0.002)
Tang	0.283***	0.304***	0.276***	0.211***	0.349***	0.256***	0.198*	-0.0678	0.299***	0.203***
	(0.049)	(0.036)	(0.050)	(0.041)	(0.071)	(0.069)	(0.105)	(0.071)	(0.032)	(0.027)
InduLev	0.606***	0.0362	0.252*	0.112	0.681***	0.264*	1.202***	0.836***	0.661***	0.224***
	(0.155)	(0.089)	(0.140)	(0.096)	(0.167)	(0.152)	(0.229)	(0.155)	(0.090)	(0.060)
InflatExp	0.0678	0.390	1.805***	0.404**	3.801***	1.021*	0.728	-0.0567	1.625***	0.496***
	(0.404)	(0.265)	(0.204)	(0.158)	(0.636)	(0.543)	(0.466)	(0.305)	(0.169)	(0.125)
Observ.	2.102	2.102	4.907	4.907	2.025	2.025	2.231	2.231	11.265	11.265
R^2	0.160	0.108	0.095	0.051	0.196	0.082	0.293	0.115	0.157	0.061

E Testing OLS assumptions in Scandinavia

Table 11: Hausman test

	(b) Fixed	(B) Random	(b-B) Difference	sqr(diag(V_b-V_B)) Standard error
Profit	-0.146	-0.144	-0.003	0.005
Size	0.027	0.020	0.007	0.001
MTB	0.003	0.001	0.001	.000
Tang	0.203	0.230	-0.027	0.007
InduLev	0.224	0.221	0.003	0.017
InflatExp	0.496	0.498	-0.003	0.017

b = consistent under the null hypotheses (H_0) and alternative hypotheses (H_A)
B = inconsistent under (H_A), efficient under (H_0)
H_0: difference in coefficients not systematic

chi(6) = 80.15
Prob >chi2 = 0.000

Table 12: Test for heteroskedasticity

Modified Wald test for groupwise heteroskedasticity in fixed effect regression model
H_0: No first-order autocorrelation

chi2(1151) = 2.0e+31
Prob >chi2 = 0.000

Table 13: Test for autocorrelation

Wooldridge test for autocorrelation in panel data
H_0: No first-order autocorrelation

F(1,919) = 33.934
Prob >F = 0.000

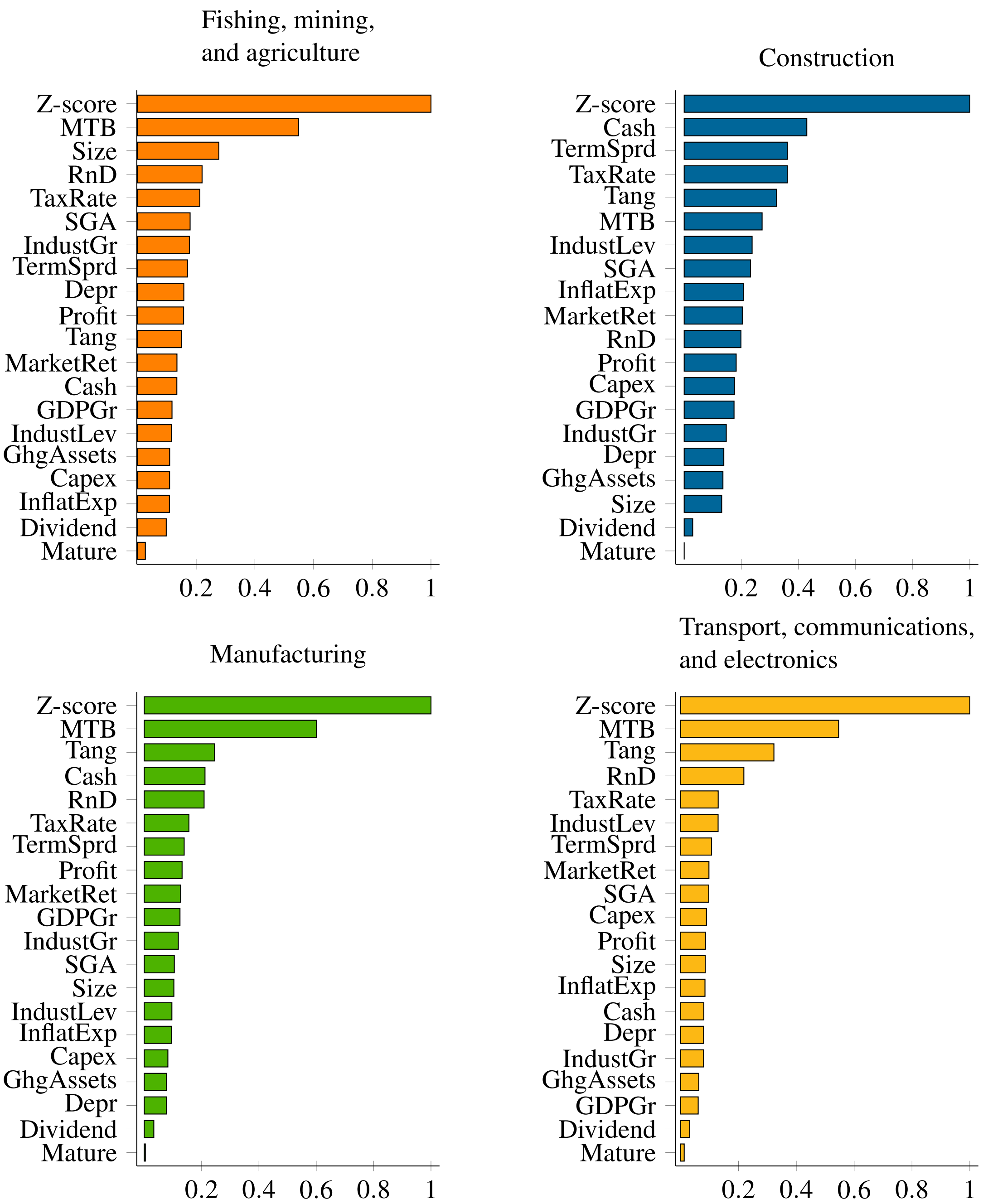
Table 14: Test for multicollinearity

As a rule of thumb, VIF-value over five indicate possible issues with multicollinearity.

	VIF	1/VIF
InduLev	9.98	0.100
Size	8.06	0.124
InflatExp	3.11	0.322
Tang	2.87	0.349
MTB	1.55	0.647
Profit	1.30	0.771
Mean	4.48	

F Variable importance plot industries

Figure 4: Importance plot industries, determinants predicting market leverage in Scandinavia using the random forest model.



Continued on next page

Figure 4: Importance plot industries, determinants predicting market leverage in Scandinavia using the random forest model - continued from previous page.

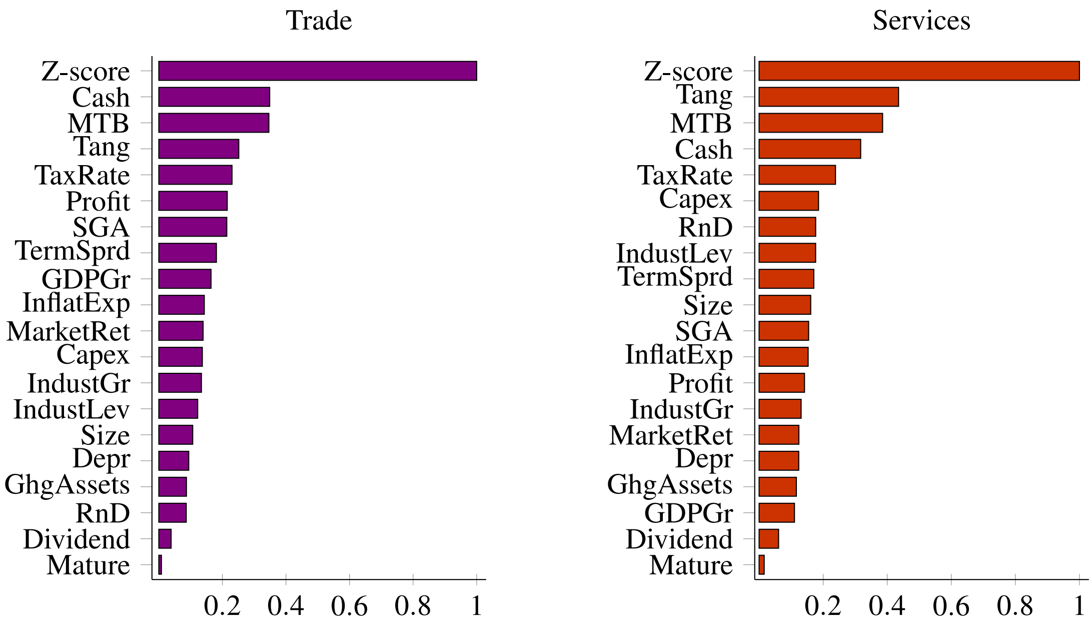
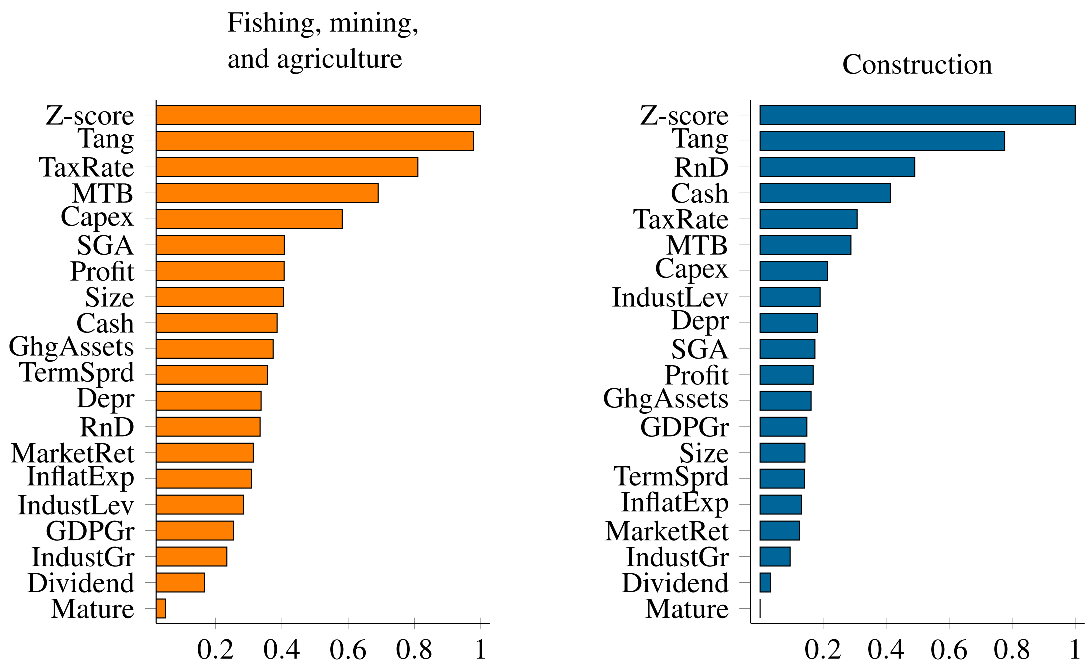
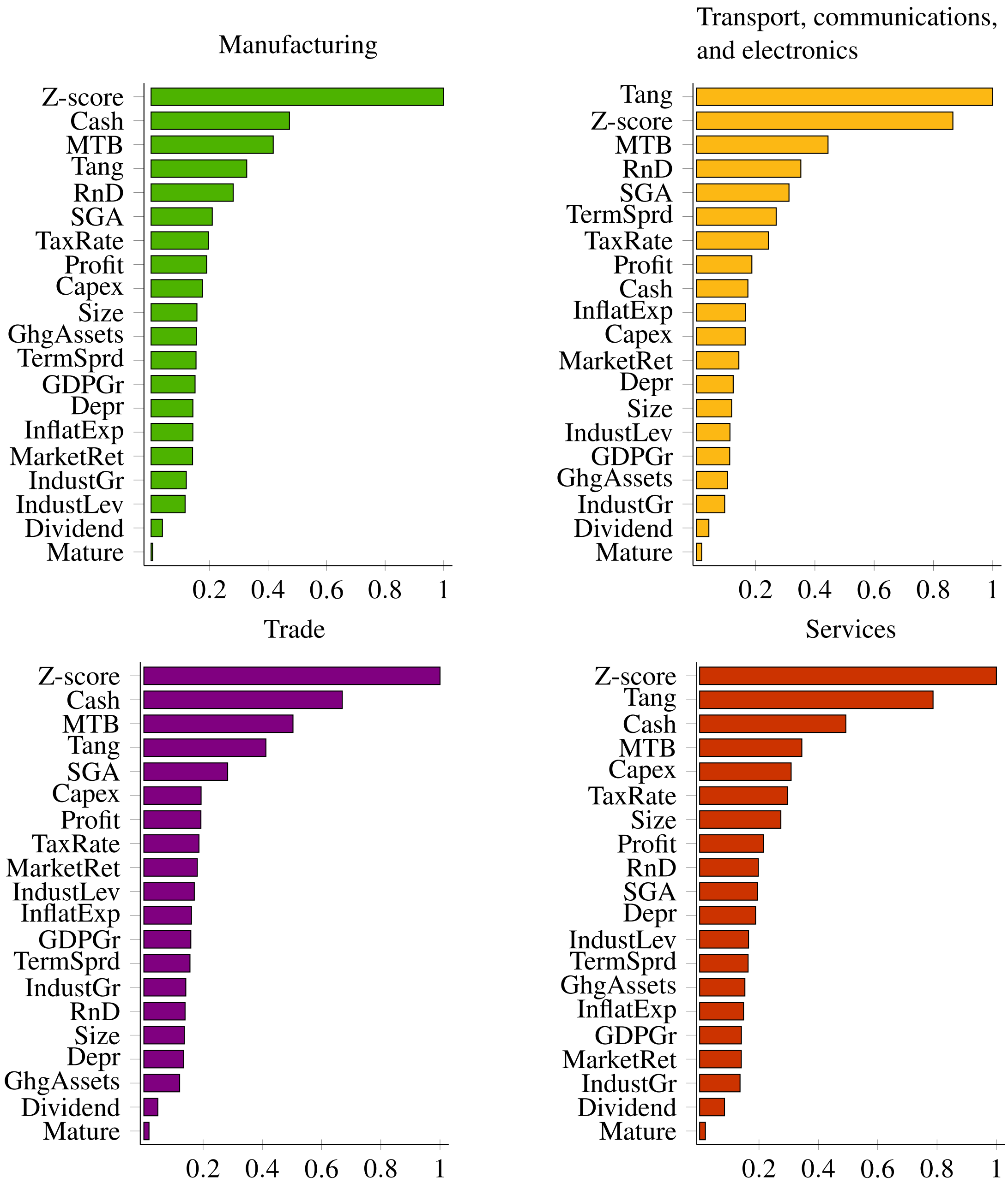


Figure 5: Importance plot industries, determinants predicting book leverage in Scandinavia using the random forest model.



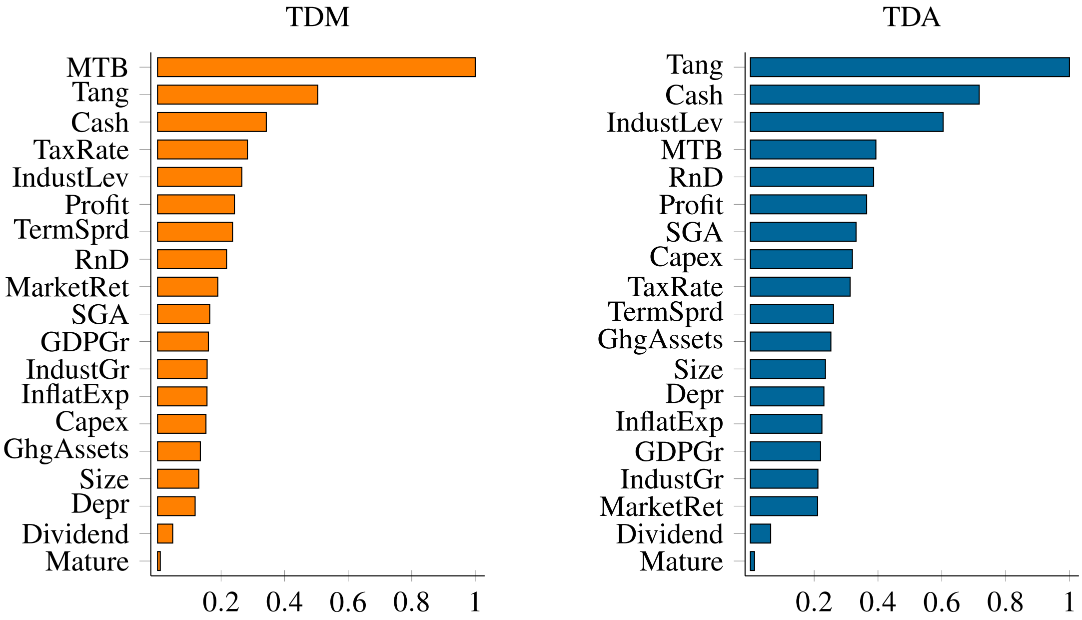
Continued on next page

Figure 5: Importance plot industries, determinants predicting book leverage in Scandinavia using the random forest model - continued from previous page.



G Variable importance plot limited

Figure 6: Importance plot excluding the z-score, determinants predicting market and book leverage in Scandinavia using the random forest model.



NTNU

Norwegian University of
Science and Technology