LUND UNIVERSITY FINANCE SOCIETY EST 1991

TRADING & QUANTITATIVE RESEARCH REPORT

Systematic Drivers for Spread Difference between Corporate Bonds in SEK vs EUR

In collaboration with:

OQAM

Date: 19.03.2024
Authors: Grigoluns Jaroslavs, Elliot Sjögren, Merriea Mathew
Supervisor: Christian Hjalmarsson

Introduction & Theory

Introduction

In the wake of the Great Recession in 2008, the corporate landscape in the Nordic region underwent a paradigm shift. Faced with uncertainties in traditional banking systems and motivated by opportunities to attract higher returns on their investment, companies began exploring the corporate bond market to diversify their funding sources [1]. The Swedish corporate bond market is considered to be relatively young, with less substantial presence between 2000 and 2008. During this period, the issuance by non-financial companies averaged at $6 billion annually, which doubled to roughly $12 billion in the period between 2009 and 2021. In addition, the share of debt securities, notable corporate bonds, in total debt financing of the companies grew from 9% in 2012 to 16% in 2021.

Until 2010, the bond issuances in Swedish krona were minimal at 4% of the total issuance while euro denominated bonds dominated the market. Between 2000 and 2021, around 57% of the Swedish non-financial bond issues were Euro denominated, 31% were Swedish krona denominated and the rest were denominated in US dollars. As the market grew since 2008, the share of SEK denomination reached 42%, mainly motivated by the growing accessibility of the Swedish bond market for smaller domestic companies. Given the existing Swedish bond market setup and the Eurobond market, there is a need to understand what drives the firms in the Swedish market to issue in SEK and/or foreign currency.

In collaboration with OQAM, based in Malmo, we try to examine the factors that determine the choice of currency in bond issuance. This paper will investigate the systematic drivers influencing the credit spread differences in the secondary market for bonds issued by the same company /issuer in SEK and EUR. The analysis will be conducted using statistical methods such as logistic regression to pick out significant components that drive the spread differences and employ machine learning techniques to predict the scale and direction of the differences.

Background

The growth of the bond market has been continuous since the 2000s and has accelerated since 2008 driven by macroeconomic trends like fall in bank loan supply, loose monetary policy, and increased investor-risk appetite. For the purpose of diversification, many large Nordic companies shifted their source of financing from low-rate government bonds to the corporate bond market, for higher yield on their investments. By the theory of the efficient market, a bond issued in SEK with the same maturity as a bond issued in EUR by the same company should yield the same return. However, in reality, the currency arbitrage phenomenon in the market provides opportunities for investors to capitalize on potentially higher returns when bonds are issued in different currencies.

A firm's choice of currency in bond issuance depends on the macroeconomic environment, the microeconomic features of the firm, and the characteristics of the bonds themselves. Understanding the factors influencing these spread differences is crucial for making informed investment decisions and optimizing credit risk allocation in corporate bond portfolio. This paper uses statistical methods to determine the factors that influence the choice of currency in bond issuance that enables the investor to maximise his returns.

Literature Review

The choice of debt structure, particularly in corporate bonds, has various implications for firms. The decision regarding the choice of currency in bonds can impact the financial health, risk exposure, and strategic positioning of a company. Our paper contributes to the literature by extending the dimensions of the currency choice decision. Earlier research in Maggiori, Neiman and Schreger (2020) underscores a strong home currency bias among global investors to prefer bonds denominated in their own currency. This study reveals that home currency bias does influence portfolio choices at the issuer level despite the assumption that investors could easily hedge currency risk separately. Several researches suggest that there exist a strong pecking order of currencies in bond denomination. Several studies emphasize the significance of interest rate costs and exchange rate dynamics in determining the choice of currency.

Maurizio Michael, Habib Mark Joy (2008) showed that lower interest rates are associated with higher likelihood of issuing debt in the local currency and lower interest rate volatility can increase this probability. Research by Nikolaus Siegfried, Emilia Simeonov, and Cristina Vespro (2007) suggests that countries with longer-government bond curves are more likely to witness the issuance of long-duration bonds. The same paper suggests that firms with higher number of subsidiaries in the global market are more likely to issue bonds in foreign currencies for hedging purposes. Moreover research. suggest that company specific factors like geographic diversification, size of the firm and the proportion of international debt holders influence the likelihood of using foreign currencies.

According to Philippe Bacchetta, Rachel Cordonier, Ouarda Merrouche, (2023) and Nikolaus Siegfried, Emilia Simeonov, Cristina Vespro (2007), macroeconomic factors such as higher capital controls on bond inflow and lower regulatory burden can increase the likelihood of issuing debt in local currency. Maggiori, Neiman, Schreger (2023) paper suggests that the shallowness of the local debt market will increase the likelihood of the firms issuing foreign currency bonds.

2. RESEARCH METHODOLOGY

Overall

The study's empirical analysis focuses on domestic and foreign-currency denominated (Eurobond) issues completed by private and public issuers belonging to different industries, but all domiciled in Sweden. Only currently active and traded bonds were considered throughout the 5-year period from 01.01.2019 to 31.01.2024.

The selection of Eurobonds as the subject of analysis is grounded in two primary considerations:

1.Explicit interest of the supervisory enterprise (OQAM) in the determinants of yield spread between EUR and SEK denominated single-issuer corporate fixed income instruments.
2.Availability of electronically retrievable data on issuer and securities themselves as well as a multitude theory-suggested variables hypothetically having explanatory power in spread direction (sign) and magnitude (absolute value) determination.

Amongst the benefits of quantitative information utilized in the research following aspects are highlighted:

1.Corporate fixed income instruments are issued in substantial volumes (our sample size here) within a highly competitive market, attracting diverse (mostly institutional) investors from various countries. This dynamics fosters market efficiency, liquidity, and reduces the likelihood of price anomalies, thereby facilitating spread comparisons.
2.Eurobond market (in our case only for EUR denominated bonds as SEK is domestic currency) operates with relatively few regulations, eliminating queuing or other procedures possibly having detrimental effect on the quality of data. Listing is only required for a fraction of the issued amount to meet institutional investors' needs, withholding tax does not apply, and bonds are predominantly in bearer form. These factors collectively contribute to a significant improvement in the comparability of bond spreads.

Foundations for correspondence of both currency denominated bonds to be included in the pair are determined by the following factors:

1.Maturity ending period within 1 calendar year from each other.
2.Identical security type (fixed, not subordinated, devoid of option-like features)

Eurobonds' spreads over the corresponding domestic bonds reflect investors' perception of the risk of loss and primary and secondary market efficiency and liquidity conditions. Corresponding to this reasoning, authors' empirical analysis involves regression of the following form:

SPREADi = f(Macro- variables, Micro-influencing variables)

Following this approach, the independent variables used in the empirical analysis can be grouped according to their ability to proxy the above mentioned factors where:

Table 1. List of macro-influencing variables (selected few, full list in appendix).

Explanatory factor	Proxying measure
SEK / EUR sovereign YC steepness	Spread between 10Y - 2Y maturities of Swedish and German Sovereign YC
SEK /EUR YC spreads at different maturities	Spread between 3M, 6M, 1Y, 2Y, 5Y, 10Y, in Swedish and German Sovereign YC maturities
Spot SEK / EUR Exchange Rate	Itself
Multiple currency effective exchange rate	KIX valutakursindex
Level of current inflation in Sweden and Germany	Monthly inflation level at the observation date
Stock market return in SEK /EUR	OMXS 30 Total Weekly Return DAX Total Weekly Return
Stock market volatility in SEK /EUR	OMXS 30 Weekly Volatility DAX Weekly Volatility
Bond market return in SEK / EUR	S&P Sweden IG Corporate Bond Index Weekly Return S&P Eurozone IG Corporate Bond Index

Table 2. List of macro-influencing variables (selected few, full list in appendix).

Explanatory factor	Proxying measure
Issuer Liquidity	Quick ratio, Times Interest earned and other
Issuer Profitability	Gross Margin, EBITDA Margin and other
DuPont/Earning Power	Asset Turnover, Reinvestment rate and other
Issuer Capital Structure	Debt/Equity, (Total Debt - Cash) / EBITDA and other
Operating	A/R Turnover, Avg. A/P days and other

2. RESEARCH METHODOLOGY

Issuer Information	Identifiers
Name CASTELLUM AB	FIGI BBG00Q41F2J4
Industry Other Financial (BCLASS)	ISIN XS2049767598
Security Information	ID Number ZR3638122
Mkt Iss EURO MTN	Bond Ratings
Ctry/Reg SE Currency EUR	Moody's Baa3
Rank Sr Unsecured Series EMTN	
Coupon 0.750000 Type Fixed	
Cpn Freq Annual	
Day Cnt ACT/ACT Iss Price 99.73500	Issuance & Trading
Maturity 09/04/2026 Reoffer 99.735	Amt Issued/Outstanding
MAKE WHOLE @25.000 until 06/04/26/ CALL 06/...	EUR 400,000.00 (M) /
Iss Sprd +125.00bp vs MIDSWAPS	EUR 400,000.00 (M)
Calc Type (1)STREET CONVENTION	Min Piece/Increment
Pricing Date 08/28/2019	100,000.00/ 1,000.00
Interest Accrual Date 09/04/2019	Par Amount 1,000.00
1st Settle Date 09/04/2019	Book Runner JPM,NDASS,SEB
1st Coupon Date 09/04/2020	Exchange EURONEXT-DUBLIN

Issuer Information	Identifiers
Name CASTELLUM AB	FIGI BBG00WMKXHY9
Industry Other Financial (BCLASS)	ISIN SE0013359742
Security Information	ID Number BK9076369
Mkt Iss DOMESTIC	Bond Ratings
Ctry/Reg SE Currency SEK	
Rank Sr Unsecured Series	
Coupon 1.805000 Type Fixed	
Cpn Freq Annual	
Day Cnt ISMA-30/360 Iss Price 100.0000	Issuance & Trading
Maturity 08/19/2025 Reoffer 100	Amt Issued/Outstanding
BULLET	SEK 200,000.00 (M) /
Iss Sprd +168.00bp vs MIDSWAPS	SEK 200,000.00 (M)
Calc Type (364)SWED 4 BUS-DAYS EX	Min Piece/Increment
Pricing Date 08/12/2020	2,000,000/ 2,000,000
Interest Accrual Date 08/19/2020	Par Amount 2,000,000.00
1st Settle Date 08/19/2020	Book Runner DANSKE,NORDEA
1st Coupon Date 08/19/2021	Exchange NOMX STOCKHOLM

Illustration 1. Extract of a bond pair descriptions from Bloomberg Terminal

Z-spread

The possibility to study spread caused as exclusively as possibl only by the currency difference all other principal difference between the bonds in particular pair have been eliminated b applying aforementioned bond selection criteria and utilizatio of a zero-volatility spread (Z-spread) - constant spread tha makes the price of a security equal to the present value of it cash flows when added to the yield at each point on the spc rate Treasury curve where cash flow is received.

$$P = \frac{C_1}{\left(1+\frac{r_1+Z}{2}\right)^{2n}} + \frac{C_2}{\left(1+\frac{r_2+Z}{2}\right)^{2n}} + \frac{C_n}{\left(1+\frac{r_n+Z}{2}\right)^{2n}}$$

- P=Current price of the bond plus accrued interest
- Cx=Bond coupon payment
- rx=Spot rate at each maturity
- Z=Z-spread
- n=Relevant time period

Bloomberg calculated mid Z spreads have been selected at daily frequency for the whole bond coexistence period (trading overlapping) in each pair..

In general, processing algorithm had following order:

1. Filtering out relevant fixed income securities (bonds)
2. Retrieving Z-spreads for EUR and SEK denominated bonds
3. Calculating absolute value of the currency spread
4. Converting spread into a Bernoulli variable based on the rule:

$$\begin{cases} 1, & \Delta \geq |25|bp \\ 0, & \Delta < |25|bp \end{cases}$$

in order to exclude small, shortly-lived "noisy" differences.

5. Applying statistical models on the resulting vector of zeros and ones with two purposes:

- Capturing numerical connection between the factors that might explain the spread (selecting ones that have the most statistical power).

Output - list of the most related variables.

- Constructing the prediction model that would forecast size of the spread based on the input variables.

Output - prediction of the spread for given combination of factors.

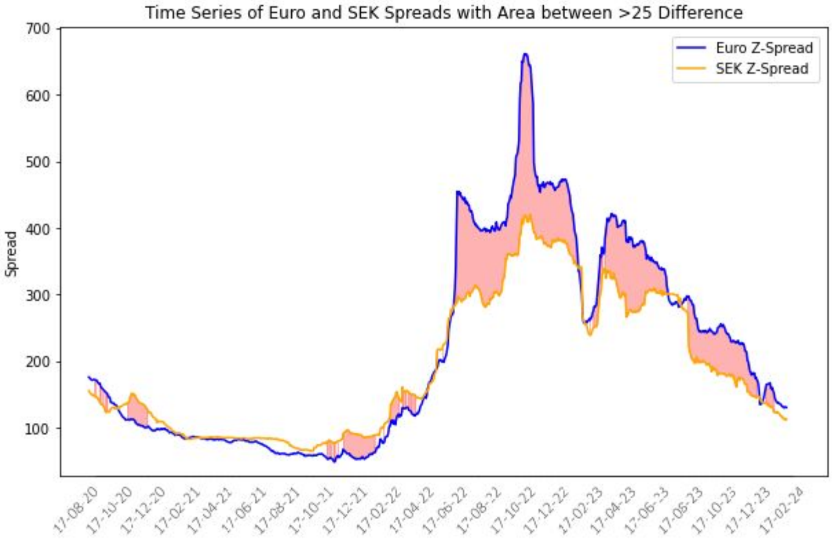

Illustration 2. Visualization of the currency spread for bonds with Bloomberg extracts (see above)

Thus, simply speaking, the objective of the work is characterised in the following activities:

1.Trying to see how combinations of factors explain zeros and ones in the first-target column.
2.How combinations of factors can predict zeros and ones in the first-target column.

Python has been used as the primary data processing and model-building programming language with a minor, mostly connected to data extraction-related procedures undertaken in excel.

2. RESEARCH METHODOLOGY

Outliers

No mechanisms that would anyhow eliminate observations from the data to solve any outlier issues (winsorization) were implemented. Instead, all observations satisfying initially imposed requirements were deemed important in bringing value to the analysis, and denying their place in the dataset was considered as a limiting factor for the level of realism.

Missing Values

All missing values throughout the dataset have been dealt with in the following manner:

1.NA, #VALUE, nan, etc. for dummy-coded variables were substituted with zeros, implying that the analysis criterion for this variable is not satisfied (dummy equal to zero).
2.NA, #VALUE, nan, etc. for other variables were substituted with column-vector averages.

Transformations

A series of transformative changes was done to the variables before the data was “fed” to the model, these transformations were as follows:

1.Dummy-coding of categorical variable (target).
2.Ordering of variables where a logical order with more or less equal increments could be traced.
3.Introduction of interaction terms combining two variables into one.

Date	SEK-EUR	e same p	e same p	R 2020=1	E 2020=	ess SEI
30-01-24	0	-0.42	8.11	117.3	123.29	-1.73
29-01-24	0	-0.42	8.11	117.3	123.29	-1.81
26-01-24	0	-0.42	8.11	117.3	123.29	-1.69
25-01-24	0	-0.42	8.11	117.3	123.29	-1.76
24-01-24	0	-0.42	8.11	117.3	123.29	-1.7
23-01-24	0	-0.42	8.11	117.3	123.29	-1.75

Illustration 3. Extract from the finalized dataframe to be “fed” to the model.

Data merging framework

In correspondence to the previous academic research, for the data aggregation for the consequent analysis data analysis method of pooled OLS regression been chosen as the most appropriate model for the investigation of the relationship between the multitude of predictors and a single predicted variable.

The Pooled OLS model applies the Ordinary Least Squares (OLS) methodology to panel data (which we have) under the assumption of no unobservable entity-specific effects, meaning that all entities in the data set are considered to have the same underlying characteristics.

Table representing the aggregated information on the predicted / dependent variable has the following attributes:

X-axis of time periods over which the unit has been tracked for explanatory factors (time series of Z-spread differences tracked during all available period at the maximum available frequency.

Y-axis - each row represents a unique unit (Z-spread difference for each bond pair)

Z-axis - each column represents the data from the measured variables for that unit (systematic drivers/ explanatory factors)

Company	Bond Pair	Observation date	Spread	Factor 1 value	Factor 2 value	***	Factor N value
Company_1	BP_1.1	OD_1.1.1	1	F1.1	F2.1		FN.1
	***	***	***	F1.2	F2.2		FN.2
Company_1	BP_1.i	OD_1.N.N	0	F1.3	F2.3		FN.3
Company_2	BP_2.1						
***	***						
Company_2	BP_2.j						
***				F1.K	F2.K		FN,K

Table 3. Overall scheme of data aggregation for Pooled OLS

2. RESEARCH METHODOLOGY

Variable selection methods

Previous academic research and industry practitioners' publications resulted in a large list of hypothetical explanatory factors 74 for Real estate and Industrial sectors and 54 for Financial sector. As one of the main research objectives is factor selection, a multitude of advanced methods has been utilised:

1. Pearson correlation with test of significance
2. One-way analysis of variance (ANOVA)
3. Sequation forward variable selection (SFS)
4. Sequential backward variable selection (SBS)
5. Sequential Forward Floating variable Selection (SFFS)
6. Akaike and Bayesian Information criteria (AIC, BIC)

Brief method explanation

Pearson CC - correlation coefficient that measures linear relationship between two sets of data. Has been used iteratively (target with each explanatory factor per time) with every coefficient value tested for statistical significance (C.L. = 0.95)

One-way ANOVA - statistical hypothesis testing framework that assesses whether the means of multiple groups are equal or if there are significant differences among them. Variable selection within this context is driven by the pursuit of identifying the most influential factors that contribute significantly to the observed variance in the dependent variable.

SFS - stepwise algorithm, progressively augmenting the feature subset by evaluating marginal impact of the candidate variable on the chosen fitness metric (R^2). The process commences with an empty feature set and successively incorporates variables until the required number of (15) best is reached.

SBS - identical, but inverse to SFS - the process commences with the complete feature set and successively eliminates variables providing minimal marginal effect until the required number of (15) best is reached.

SFFS - similar to SFS and SBS, it navigates the expansive feature space with a dual perspective – one of forward expansion to consider new variables and another of backward evaluation to reassess the importance of already selected features.

AIC - log-likelihood of the model given the data, adjusted by a penalty term proportional to the number of parameters in the model. A lower AIC value suggests a better trade-off, indicating a model that provides a good fit with a minimal number of parameters:

BIC, -log likelihood imposing a stronger penalty for increasing the number of parameters. BIC incorporates a Bayesian perspective by introducing a logarithmic term that scales with the sample size. BIC generally leads to the selection of simpler models compared to AIC, making it more conservative in terms of model complexity.

Python code wrapped around the pure mathematical calculations was iteratively selecting a subset of best relevant variables from the initially given predictors. This selection process led to enhanced explanatory power, as measured by the adjusted coefficient of determination (R^2).

The methods have been applied to the the total of:

- 6394 observations for Financial Sector
- 7629 observations for Real Estate Sector
- 6304 observations for Industrial Sector

All observations of independent and dependent variables for each sector have been subdivided into 2 cross-validation sections (train and test) in the ratio of 8/10 and 2/10 respectively.

Prediction models

Logistic Regression

Statistical method of logistic regression for estimating the parameters of a logistic model (the coefficients in the linear combination) has been selected as the main model of the research. The logistic regression model is based on the logistic function (sigmoid function) to model the probability of a particular outcome.

Algorithm of model building is as follows:

1. Defining logistic function (Sigmoid):

$$S(z) = \frac{1}{1+e^{-z}}$$

- z - linear combination of predictor variables:

$$z = \beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \ldots + \beta_n \cdot x_n$$

2. Constructing logistic regression model:

$$P(Y = 1) = \frac{1}{1 + e^{-(\beta 0 + \beta 1 X1 + \beta 2 X2 + \cdots + \beta k X k)}}$$

- P(Y=1) - probability of the positive class,
- Y is the binary outcome variable (0 or 1).

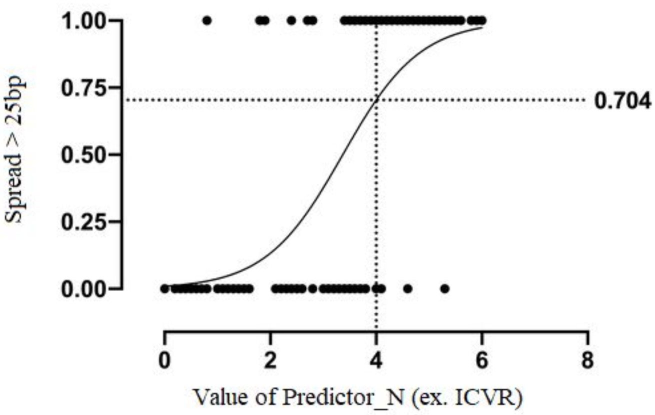

Illustration 4. Principal scheme of Logistic regression

3. Logit Transformation: To linearize the logistic regression model, we apply the logit transformation to the odds ratio:

$$\text{logit}(p) = \ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \ldots + \beta_n \cdot x_n$$

2. RESEARCH METHODOLOGY

4. Instead of maximum likelihood estimation (MLE) cost function and gradient descent have been selected for optimizing logistic regression. The cost function, which is minimized during training, is the negative log-likelihood:

$$J(\beta) = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \cdot \ln(p_i) + (1-y_i)\cdot \ln(1-p_i)\right]$$

The gradient of this cost function with respect to the coefficients β is used to update the coefficients iteratively.

One of the attractive features of logistic regression is possibility to examine the effect of individual variables in isolation, allowing to more thoroughly evaluate particular regression set components, which is of interest for the analyst getting insights from the model.

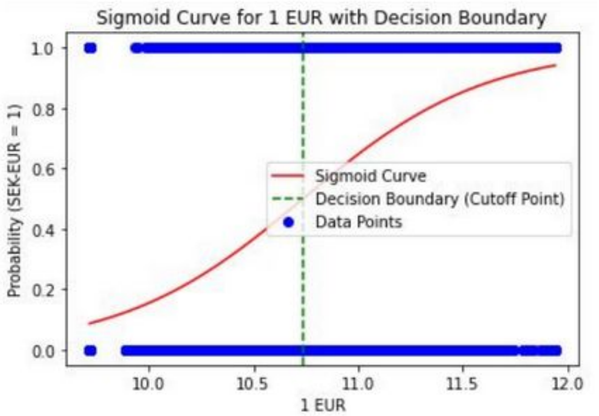

Illustration 5. Fitted sigmoid function for an individual variable (SEK/EUR exchange rate).

Random Forest

Random forest - ensemble learning method based on constructing a multitude of decision trees at training time and on training data has been selected as an addition to the logistic regression in order to obtain an alternative independent

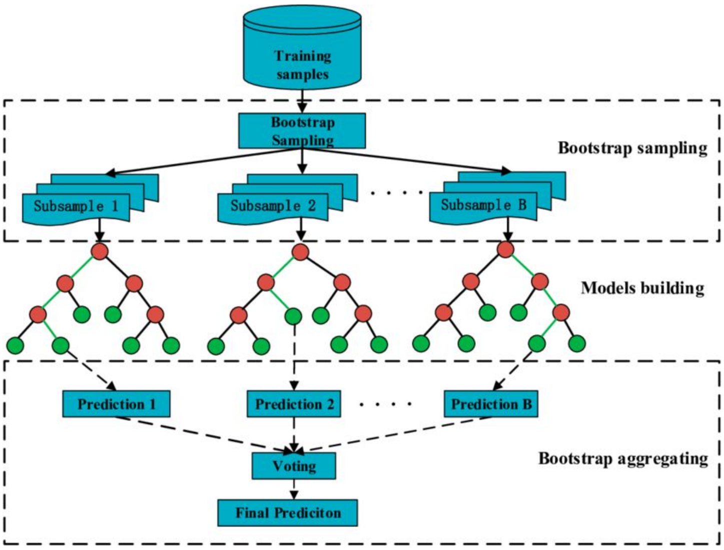

Illustration 6. Principal scheme of Random Forest model

1.Decision Tree Structure - automated decision-making algorithm based on features (factors), with nodes representing decisions and leaves containing final predictions.

2.Recursive Binary Splitting - selecting features and split points to segregate data into subsets until a stopping criterion is met.

3.Splitting Criteria - Gini impurity:

$$Gini(t) = 1 - \sum_{i=1}^{C} p_i^2$$

- t - node
- pi - proportion of samples of class i.

Loss function in the form of Mean Squared Error (MSE) has been used to penalized continuous tree "growth".

$$MSE(t) = \frac{1}{N_t}\sum_{i=1}^{N_t}(y_i - \bar{y}_t)^2$$

Following hyper-parameters to control tree complexity without much adjustment have been used in the model:

Table 4. Main tuning parameters for Random forest model

Feature	Value
Number of trees in the forest	1000
Max. depth of the tree	no limit
Min. number of samples required to split an internal node	5
Min. number of samples required to be at a leaf node	1
Number of rounds to wait for improvement in early stopping	10

The Gaussian Naive Bayes classifier

As a third and final prediction method GNB classifier based on Bayes' theorem has been selected. It makes the naive assumption that features are conditionally independent given the class label.

Here's the fundamental idea and key formulas behind Gaussian Naive Bayes.

1.Bayes' theorem is a fundamental probability formula that relates the conditional and marginal probabilities of random events. For a binary classification task, the formula can be expressed as: $P(y|x) = \frac{P(x|y)\cdot P(y)}{P(x)}$

2. Naive Independence Assumption - "naive" part comes from assuming independence between features given the class label.

$$P(x_1, x_2, \ldots, x_n|y) = P(x_1|y)\cdot P(x_2|y)\cdot \ldots \cdot P(x_n|y)$$

2. RESEARCH METHODOLOGY

3. Gaussian Naive Bayes specifically assumes that the likelihood of each feature given the class label follows a Gaussian (normal) distribution. The probability density function (PDF) of the Gaussian distribution is given by:

$$P(x_i|y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{(x_i-\mu_y)^2}{2\sigma_y^2}\right)$$

- xi - value of feature i.
- μy - mean of feature i for class y.
- σ^2 y - variance of feature i for class y.

4. Decision rule:

During training, the model estimates the parameters for each feature and class based on the training data.

$$\text{Prediction for } x : \hat{y} = \arg\max_y P(x|y) \cdot P(y)$$

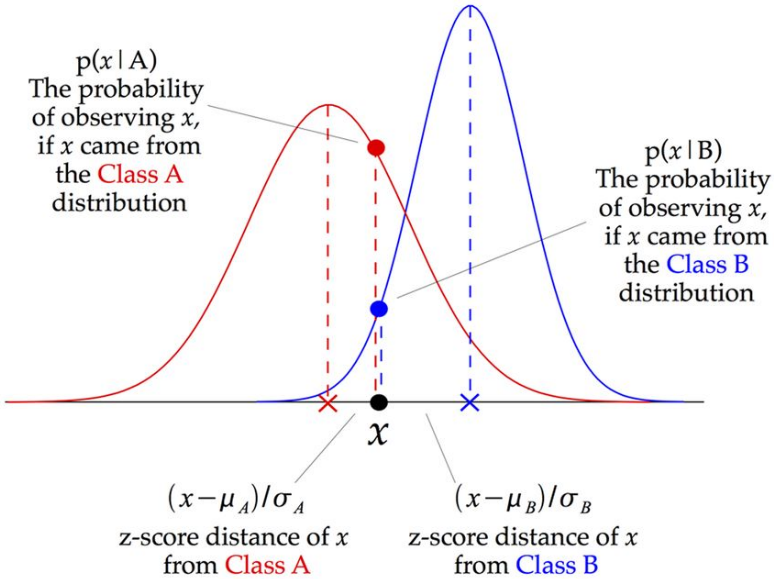

Illustration 7. Principal scheme of GNB classifier, observation X will be assigned to belong to Class A (red distribution)

Data sources and Sample Characteristics

All the data utilized in the empirical part of the research is retrieved from a single source of financial information: Bloomberg Terminal.
All bonds in pairs with special characteristics (e.g. puttable bonds, callable bonds, perpetual bonds, etc.) that would result in them being priced differently have not been included in the sample.

Data availability and Bloomberg reporting methodology resulted in significant limitation which lead to utilization of only currently tradable bonds and their past prices. All the fixed income securities that have already matured were excluded from the sample.

Based on the recommendation and interests of the supervisory firm initial selection has been subdivided into following groups:

1. Nonfinancial public issuers:

Volvo AB, Electrolux AB, Scania AB, Essity AB, Dometic AB, Alfa Laval, AB, Assa Abloy AB, Intrum AB, Securitas AB, Investor AB, Trelleborg AB

2. Financial public issuers:

SEB, Swedbank, Svenska Handelsbanken

3. Real Estate public issuers:

Akelius AB, Castellum AB, Heimstaden AB, Balder AB, Cibus AB

This total sample after filtering out invalid observations has 86 bonds, of which:

- 32 were issued by companies of public Non-Financial sector
- 28 were issued by companies of Financial public sector.
- 26 were issued by companies of Real Estate public sector.

Private FI security issuers have been excluded due to information incompleteness or absence inherent into the nature of accounting and reporting standards of private enterprises.

Governmental/municipal FI security issuers, central banks and supranational institutions have been excluded due to absence of information on financial statement and since they are not exposed to the effect of various economic and financial variables in the same way as privately held or incorporated firms.

4. EMPIRICAL RESULTS - Variable Selection - Pearson Correlation with ToS

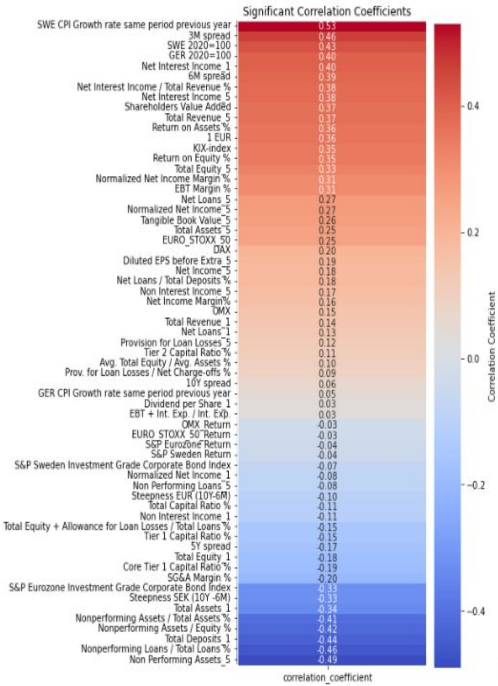

Illustration 8. Correlation for Financial companies

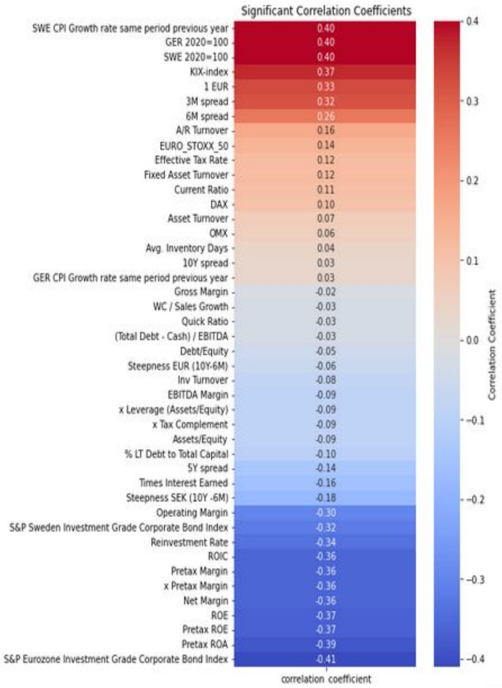

Illustration 9. Correlation for Real Estate companies

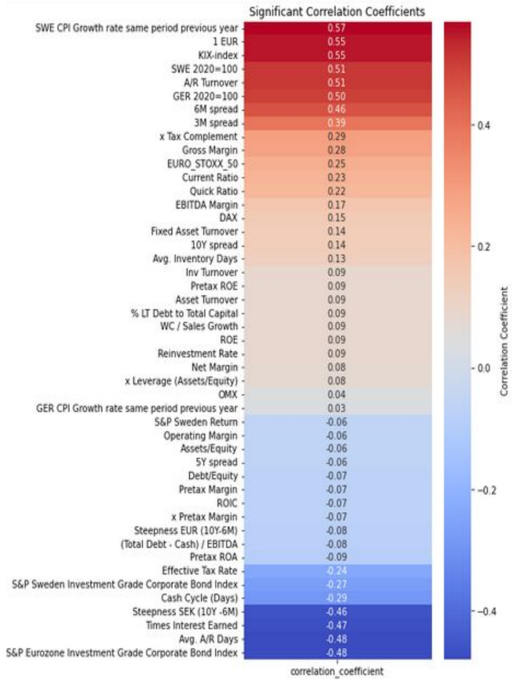

Illustration 10. Correlation for Industrial companies

Correlation measure allows to capture both the magnitude (absolute value) and the direction (sign) of the relationship between explanatory factors and the target variable.

Sign of the correlation coefficient extends the depth of the interpretation and adds dynamism to the results.
If for the analyst the magnitude says “just” about the fact that if the explanatory factor changes - shift in target should be happening as well (accounting for the time leading/lag effects not within the framework of this paper). Then the direction of this relationship helps to see where the target will move.

Taking Swedish CPI change (factor commonly found to be top positively correlated for the spread of all bonds) as an example: observed steady and strong growth (not its level, but its growth - first derivative) of the Swedish inflation (like in the period between Mar 2022 and Jul 2023 when it was above 5% in annual terms) analyst would be able to conclude that the yields on bonds will not only change, they will grow (as they did in this period - see Illustration 2. before). Similar logic applies to the negative correlation.

Different sectors have not equal exposure to explanatory factors, things are made more complicated by the fact, that financial companies are having a different set of key metrics. More specific remarks on every group of companies are as follows:

Positive Correlation

Among macroeconomic factors: measures of inflation (level and growth rate), currency descriptors (spot SEK/EUR and KIX) and interest rate related measures (3 and 6 month spread between sovereign bills).

In regards to the company specific factors for financials the profitability of lending activities (net interest income over 1 and 5 years) as well as Shareholder’s Value Added are the most important ones. For real estate A/R Turnover and Effective tax rates, for industrials A/R Turnover as well, but times Tax complement and Gross Margin levels are pivotal.

Negative Correlation

If positive correlation is dominated by systemically important factors, negative, on the other hand has almost exclusively idiosyncratic measures in the top 10 list.
For financials measures capturing the level of non-performing assets or loans (ratios like non-performing assets/loans over Equity, Total assets or Loans) play more significant role. Real estate companies pay more attention on returns profitability measures (ROA, ROE, ROIC and all types of margins). Industrials are exposed to the mixture of solvency (times interest earned and Net Debt to EBITDA), overall efficiency (cash cycle and Avg A/R days) and some macroeconomic yield curve -related factors (steepness of SEK and EUR sovereign yield curves).

In general, for all studied sectors relatively narrow list of factors stand out as particularly strongly positively correlated. correlated variables for real-estate firms are similar to the non-financial firms. Broad economic factors are more positively correlated to currency spread size, company-specific factors are more negatively correlated.

Full list of variables and more specific information to be found in appendix.

4. EMPIRICAL RESULTS - Variable Selection - 1-Factor ANOVA

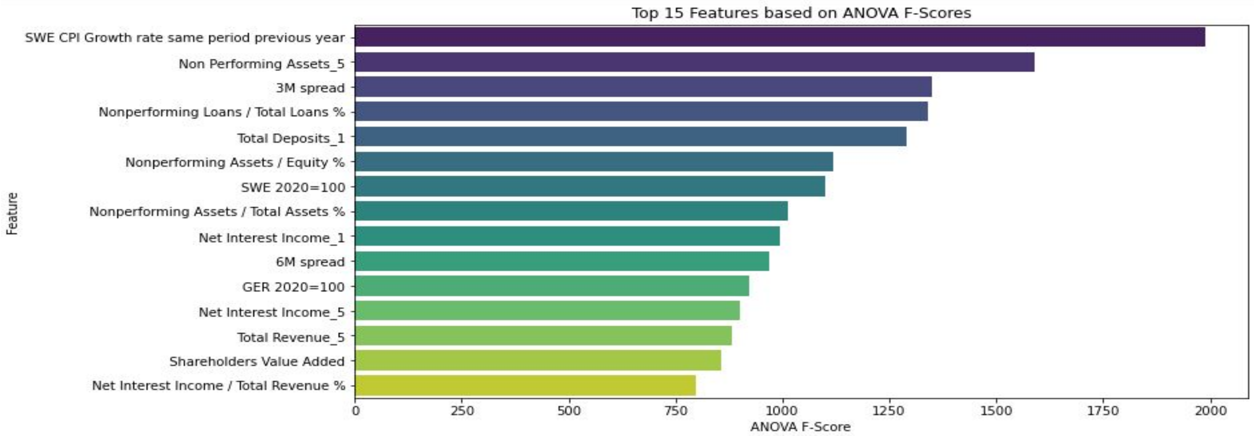

Illustration 11. 1-F ANOVA for Financial companies

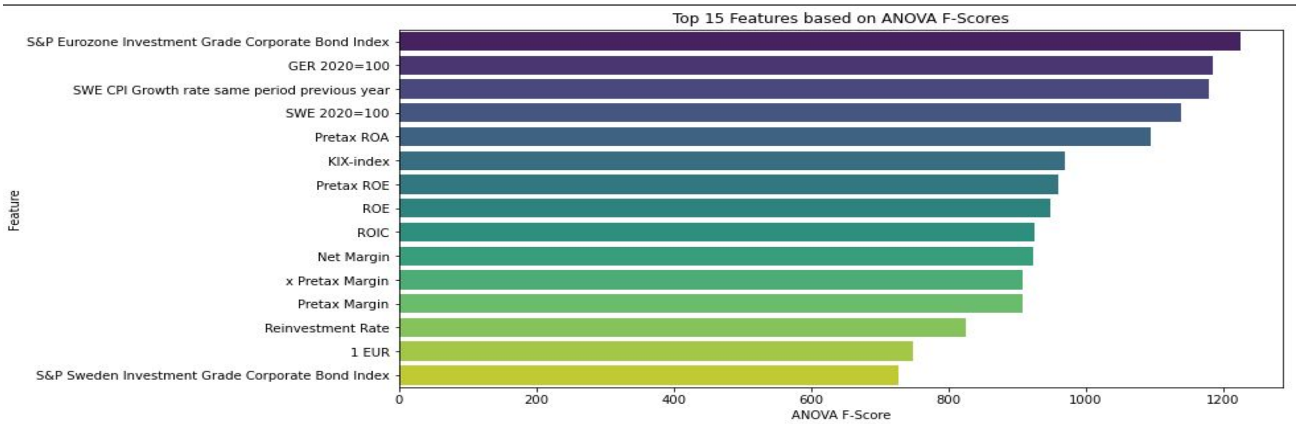

Illustration 12. 1-F ANOVA for Real Estate companies

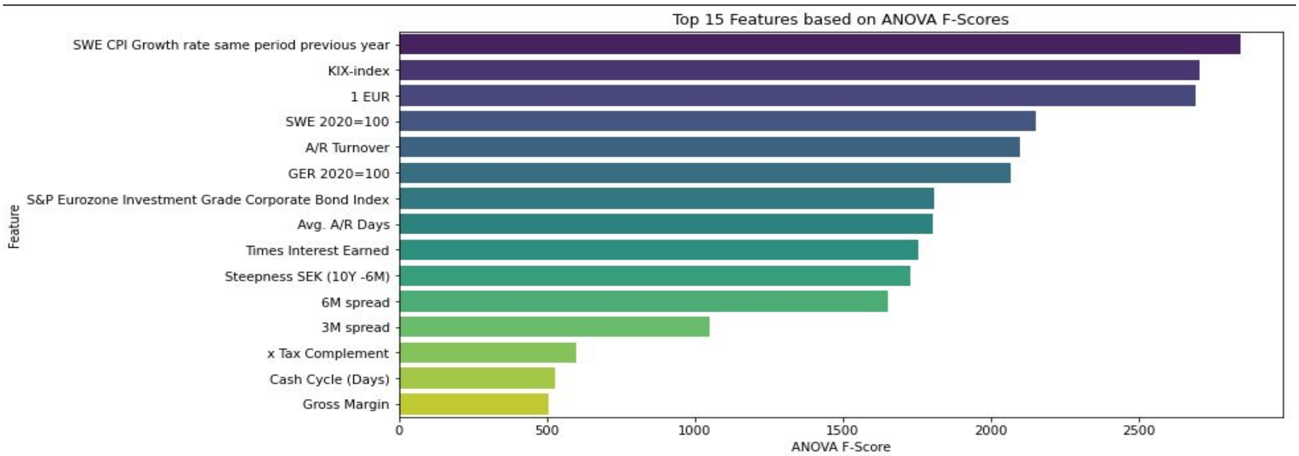

Illustration 13. 1-F ANOVA for Industrial companies

Due to the fact that ANOVA test is agnostic regarding the direction of the relationship and cares only about its magnitude (it is how it orders them as well) this method provides the results that are a mixture of variables from the edges of the correlation spectrum (top 15 most important variables as per ANOVA is a list including elements exclusively from the top 10 positive and top 10 negatively correlated features of Pearson).

1-F ANOVA results are consistent with previously discussed Pearson's correlation results and independently confirm them, thus, increasing the credibility of the research.

4. EMPIRICAL RESULTS - Variable Selection - Sequential Selectors

As mentioned in the methodology, alongside Pearson correlation and ANOVA, we employ sequential selection methods such as SFS, SBS, and SFFS.

Identical to the ANOVA these methods are myopic about the direction of the relation and are based only on the overall performance improvement that addition and keeping of each variable is bringing to the model.

For Financial firms, the results coincide with the correlation results such that among macroeconomic factors the inflation level remains significant.

With respect to company specific factors, the value of total assets, revenue, deposits, and non-performing assets are significant. The provision for loan losses as a percentage of net charge-offs was a distinct factor selected by these variable selection methods.

For both real estate and non-financial companies, the selected company specific factors are similar to the correlation results. One notable distinction is that the non-financial companies appear to be affected by Swedish inflation growth while the real estate companies are significantly explained by the German inflation growth.

Table 5. Top 10 significant explanatory factors based on SFS (accuracy by R^2)

Financial	Real Estate	Nonfinancial
SWE CPI Growth Prov. for Loan Losses/Net Charge-offs % Avg. Total Equity / Avg. Assets % Net Loans / Total Deposits % Dividend per Share_1 Total Assets_1 Total Revenue_5 Non Performing Assets_5 Total Deposits_5 Total Equity_5	GER CPI Growth GER 2020=100 DAX S&P Sweden IG Corporate Bond Index S&P Eurozone IG Corporate Bond Index Operating Margin Net Margin Asset Turnover ROE Quick Ratio	SWE CPI Growth GER 2020=100 3M spread KIX-index Current Ratio Times Interest Earned Cash Cycle (Days) Debt/Equity A/R Turnover ROIC

Table 6. Top 10 significant explanatory factors based on SBS (accuracy by R^2)

Financial	Real Estate	Nonfinancial
SWE CPI Growth Return on Equity % Net Income Margin% Net Income_1 Non Performing Assets_1 Net Interest Income_5 Non Interest Income_5 Total Revenue_5 Net Income_5 Total Equity_5	SWE 2020=100 1 EUR KIX-index 'EBITDA Margin Operating Margin Effective Tax Rate Net Margin x Tax Complement ROE Inv Turnover	Steepness SEK (10Y -6M) KIX-index S&P Sweden IG Corporate Bond Index S&P Eurozone IG Corporate Bond Index Asset Turnover Pretax ROA Quick Ratio Times Interest Earned % LT Debt to Total Capital (Total Debt - Cash) / EBITDA

Table 7. Top 10 significant explanatory factors based on SFFS (accuracy by R^2)

Financial	Real Estate	Nonfinancial
SWE CPI Growth Prov. for Loan Losses/Net Charge-offs % Avg. Total Equity / Avg. Assets % Net Loans / Total Deposits % Dividend per Share_1 Total Assets_1 Total Revenue_5 Non Performing Assets_5 Total Deposits_5 Total Equity_5	GER CPI Growth GER 2020=100 DAX S&P Sweden IG Corporate Bond Index S&P Eurozone IG Corporate Bond Index Operating Margin Net Margin Asset Turnover ROE Quick Ratio	SWE CPI Growth GER 2020=100 3M spread KIX-index Current Ratio Times Interest Earned Cash Cycle (Days) Debt/Equity A/R Turnover ROIC

4. EMPIRICAL RESULTS - Variable Selection - Information Criteria (Akaike and Bayes)

Table 8. Top 10 significant explanatory factors based on AIC

Financial	Real Estate	Nonfinancial
SWE CPI Growth rate same period previous year Non Performing Assets_5 3M spread Nonperforming Loans / Total Loans % Total Deposits_1 SWE 2020=100 Nonperforming Assets / Equity % Nonperforming Assets / Total Assets % Net Interest Income_1 GER 2020=100	S&P Eurozone IG Corporate Bond Index SWE CPI Growth GER 2020=100 SWE 2020=100 Pretax ROA KIX-index Pretax ROE ROE 'Net Margin Pretax Margin	S&P Eurozone IG Corporate Bond Index SWE CPI Growth GER 2020=100 SWE 2020=100 Pretax ROA 'KIX-index Pretax ROE ROE Net Margin Pretax Margin

Table 9. Top 10 significant explanatory factors based on BIC

Financial	Real Estate	Nonfinancial
SWE CPI Growth Non Performing Assets_5 3M spread Nonperforming Loans / Total Loans % Total Deposits_1 SWE 2020=100 Nonperforming Assets / Equity % Nonperforming Assets / Total Assets % Net Interest Income_1 GER 2020=100	S&P Eurozone IG Corporate Bond Index SWE CPI Growth GER 2020=100 SWE 2020=100 Pretax ROA KIX-index Pretax ROE ROE 'Net Margin Pretax Margin	S&P Eurozone IG Corporate Bond Index SWE CPI Growth GER 2020=100 SWE 2020=100 Pretax ROA 'KIX-index Pretax ROE ROE Net Margin Pretax Margin

The results are identical between both information criteria utilized and is extremely close to all previous methods as well.

4. EMPIRICAL RESULTS - Prediction Model - Logistic regression

After the variables playing increasingly important role have been selected the next step is to attempt to build the model capable of predicting the outcome of the Bernoulli variable of interest. Logistic regression is the workhorse of the paper and has been selected to play a major role due to it following advantages:

- Efficiency with Binary Outcome - LR specifically designed for binary outcome variables.
- Assumption of Linearity not Required
- Robustness to Outliers (no outliers removed)
- Providing Probabilistic Predictions - output - likelihood of the outcome variable
- Less Prone to Overfitting (pivotal thing for in-and-out of sample model performance issue

Based on the accuracy - ability to accurately classify credit spread size as either positive or negative in all cases general model performance can be considered to be at satisfactory strongly above average with values ranging between 70 and 90 per cent.

Precision - proportion of correctly predicted positive cases among all cases predicted as positive by the model is sizably similar to accuracy, but lower for all sectors except for real estate where it is actually slightly (by 2%) higher with overall values between 65 and 85%

Table 10. Logistic regression performance measures for all sectors

	Financial	**Real Estate**	**Non-Financial**
Accuracy	0.7169	0.7870	0.8718
Precision Score	0.6570	0.8021	0.8561

Table 11. Confusion matrix for Financial sector

Financial		
	Predicted Negative (0)	Predicted Positive (1)
Actual Negative	385	271
Actual Positive	86	519

Table 12. Specialized performance measures for Financial sector

	Precision	Recall	F-1-score	Support
0	0.90	0.79	0.84	667
1	0.86	0.93	0.89	893
macro avg	0.88	0.86	0.87	1560
weig. avg	0.87	0.87	0.87	1560

Table 13. Confusion matrix for Real Estate sector

Real Estate		
	Predicted Negative (0)	Predicted Positive (1)
Actual Negative	196	248
Actual Positive	77	1005

Table 14. Specialized performance measures for Real Estate sector

	Precision	Recall	F-1-score	Support
0	0.72	0.44	0.55	444
1	0.80	0.93	0.86	1082
macro avg	0.76	0.69	0.70	1526
weig. avg	0.78	0.79	0.77	1526

4. EMPIRICAL RESULTS - Prediction Model - Logistic regression

Table 15. Confusion matrix for Industrial sector

Industrial		
	Predicted Negative (0)	Predicted Positive (1)
Actual Negative (0)	527	140
Actual Positive (1)	60	833

For the all sectors the logistic regression model exhibits above average accuracy in predicting outcomes within the sector. This is better seen from the confusion matrices (below), meaning that model is more prone of committing Type I error:

$$\text{Type I Error (False Positive Rate)} = \frac{FP}{FP+TN}$$

than Type II error:

$$\text{Type II Error (False Negative Rate)} = \frac{FN}{FN+TP}$$

The idea is getting off-main-diagonal elements as close to zero as possible while concentrating all the values in the main.

Recall - measure of the proportion of correctly predicted positive cases (true positives) among all actual positive cases - model's ability to capture or "recall" positive instances from the dataset is higher for class 1 that for class o indicating that model is better “seeing” higher spreads that “seeing” bad ones.

$$\text{Recall} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)+False Negatives (FN)}}$$

F-1 score - metric unifying Precision and Recall together as a harmonic mean is high as well:

$$F1\ score = 2 \times \frac{precision \times recall}{precision + recall}$$

Although having macro average in the output as well we will focus on the weighted one as in classification tasks with imbalanced datasets. Because we are having unequal support - number of observation of each class, the WA is a method to compute the average performance across all classes, while considering the class imbalance.

$$\text{Weighted Average} = \frac{1}{N}\sum_{i=1}^{N} \text{Metric}_i \times \text{Support}_i$$

It calculates the metric of interest (e.g., precision, recall, F1-score) for each class individually and then averages the results across all classes, weighted by the support.

Table 16. Specialized performance measures for Industrial sector

	Precision	Recall	F-1-score	Support
0	0.82	0.59	0.68	656
1	0.86	0.93	0.89	605
macro avg	0.74	0.72	0.71	1261
weig. avg	0.74	0.72	0.71	1261

In general, model results for all sectors are at the high level, however, as visible from the confusion matrices lowest performance is attributable for Financial sector dataset, Real estate and Industrial sectors show better values.

4. EMPIRICAL RESULTS - Prediction Model - Random Forest

For financial companies, the random forest prediction model achieves a satisfactory accuracy of 73.83% and an Area under the Curve- PR of 0.6570 indicating the model's balance in precision and recall. Precision for class 0 (81%) and class 1 (86%) is notable, suggesting good discrimination between positive and negative instances, while the recall for both classes is strong, especially for class 1 (93%).

For the Real estate sector companies, the random forest model achieves an accuracy of 76.07%. A high AUC-PR, such as 0.9089, implies that the model has achieved a strong balance between precision and recall across different thresholds. Precision for class 0 (61%) and class 1 (81%) indicates room for improvement, especially for correctly predicting negative instances, while the recall for class 1 (87%) is strong, but class 0 (49%) could be enhanced.

For Non-financial sector companies, the random forest model achieves an accuracy of 92.24%. Similar to the real estate companies, the Area under the Curve- Precision Recall of 0.9823 indicates strong performance in distinguishing positive and negative instances. The uniformly high precision, recall, and F-1 scores reflects robust predictive capacity of the model.

Table 17. Random Forest performance measures for all sectors

	Financial	Real Estate	Non-Financial
Accuracy	0.7383	0.7602	0.9224
Area Under the Precision-Recall Curve (AUC-PR):	0.6570	0.9089	0.9823

Table 18. Confusion matrix for Financial sector

Financial		
	Predicted Negative (0)	Predicted Positive (1)
Actual Negative	426	230
Actual Positive	100	505

Table 19. Specialized performance measures for Financial sector

	Precision	Recall	F-1-score	Support
0	0.81	0.65	0.72	656
1	0.86	0.93	0.75	605
macro avg	0.75	0.74	0.74	1261
weighted avg	0.75	0.74	0.74	1261

Table 20. Confusion matrix for Real Estate sector

Real Estate		
	Predicted Negative (0)	Predicted Positive (1)
Actual Negative	217	227
Actual Positive	139	943

Table 21. Specialized performance measures for Real Estate sector

	Precision	Recall	F-1-score	Support
0	0.61	0.49	0.54	444
1	0.81	0.87	0.84	1082
macro avg	0.71	0.68	0.69	1526
weighted avg	0.75	0.76	0.75	1526

Table 22. Confusion matrix for Industrial sector

Non-Financial		
	Predicted Negative (0)	Predicted Positive (1)
Actual Negative	607	60
Actual Positive	61	832

Table 23. Specialized performance measures for Industrial sector

	Precision	Recall	F-1-score	Support
0	0.91	0.91	0.91	667
1	0.93	0.93	0.93	893
macro avg	0.92	0.92	0.92	1560
weighted avg	0.92	0.92	0.92	1560

4. EMPIRICAL RESULTS - Prediction Model - Gaussian Naive Bayes Classifier

For Financial sector companies, the model achieves a balanced performance with good precision and recall for both classes (0 and 1). The precision for class 0 (83%) indicates accurate identification of actual negatives, while recall for class 1 (87%) suggests effective capturing of actual positives.

For real estate companies, the model performs best in identifying actual positives with high precision (87%) and recall (73%). Here, the precision for actual negatives (class 0) is lower (52%), indicating potential false positives.

For Non-financial sector companies, the model shows strength in capturing actual negatives with high precision (83%) and recall (91%).

Table 24. Gaussian Naive Bayesian Classifier performance measures for all sectors

	Financial	Real Estate	Non-Financial
Accuracy	0.7264	0.7254	0.6635
Area Under the Precision-Recall Curve (AUC-PR):	0.7162	0.8897	0.9028

Table 25. Confusion matrix for Financial sector

Financial		
	Predicted Negative (0)	Predicted Positive (1)
Actual Negative	392	264
Actual Positive	81	524

Table 26. Specialized performance measures for Financial sector

	Precision	Recall	F-1-score	Support
0	0.83	0.60	0.69	656
1	0.66	0.87	0.75	605
macro avg	0.75	0.73	0.72	1261
weighted avg	0.75	0.73	0.72	1261

Table 27. Confusion matrix for Real Estate sector

Real Estate		
	Predicted Negative (0)	Predicted Positive (1)
Actual Negative	217	227
Actual Positive	139	943

Table 28. Specialized performance measures for Real Estate sector

	Precision	Recall	F-1-score	Support
0	0.52	0.73	0.61	444
1	0.87	0.73	0.79	1082
macro avg	0.69	0.73	0.70	1526
weighted avg	0.77	0.73	0.74	1526

Table 29. Confusion matrix for Industrial sector

Non-Financial		
	Predicted Negative (0)	Predicted Positive (1)
Actual Negative	607	60
Actual Positive	61	832

Table 30. Specialized performance measures for Financial sector

	Precision	Recall	F-1-score	Support
0	0.56	0.97	0.71	667
1	0.96	0.43	0.60	893
macro avg	0.76	0.70	0.65	1560
weighted avg	0.79	0.66	0.65	1560

LUND UNIVERSITY FINANCE SOCIETY EST 1991

4. EMPIRICAL RESULTS - Prediction Model - Logistic Regression

Interpretation and analysis of ROC curves for LR

Perfect Receiver Operating CHaracteristic (ROC) curve would be a right angle at the top left corner, where sensitivity is 1 (100%) and specificity is also 1 (100%), forming a perfect square.

Only for 1 sector (Ind.) curve shape approaches perfect case - steep vertical rise from (0,0) to (o,1) indicating that almost all true positives are identified without any false positives.

Then, the curve extends horizontally from (0,1) to (1, 1) indicating that all positives are correctly classified for both positive and negative instances without any false positives or false negatives.

Other sectors (Fin. and RE.) cannot boast with the same precision due to their more “bow-shaped” form. Overall observation is that our model has more troubles with separating true positives from false positives (less steep vertical part), but is much better at classifying both true positives and negatives (almost all have “good” horizontal line.

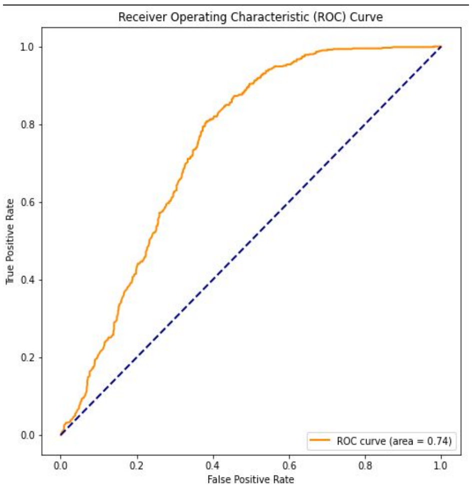

Illustration 14 . ROC curve for Financial Sector (LR)

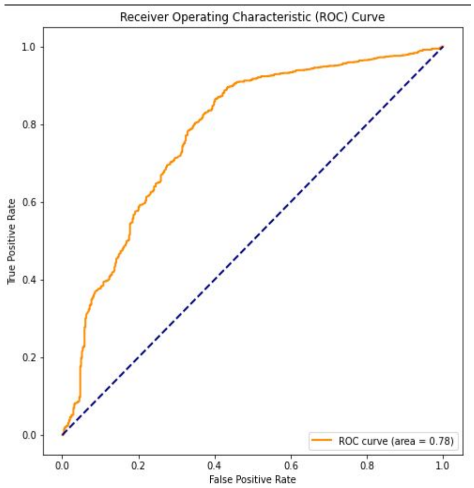

Illustration 15. ROC curve for Real Estate Sector (LR)

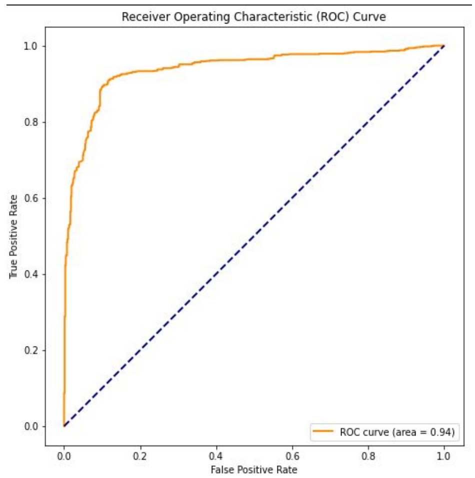

Illustration 16. ROC curve for Industrial Sector (LR)

Interpretation and analysis of PR curves for LR

Perfect precision-recall curve would be a straight line from the top left corner to the top right corner, with precision and recall both being equal to 1 (or 100%) for all threshold values.

It is clearly visible that the model is not able to produce perfect precision from the beginning.

(all graphs start from below 1 on the Y axis), but deals very with maintaining the precision (not falling below the starting point).

It is peculiar to observe that the last case (Ind.) while starting almost at 1 exactly is abruptly falling towards the end of the sample, meaning that the bonds at the end of the sample are more complicated to process that ones at the beginning.

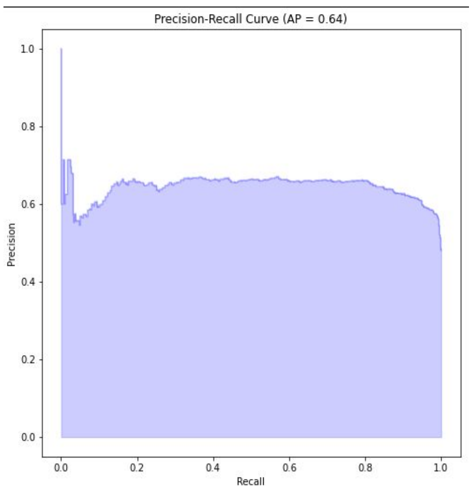

Illustration 17. PR curve for Financial Sector (LR)

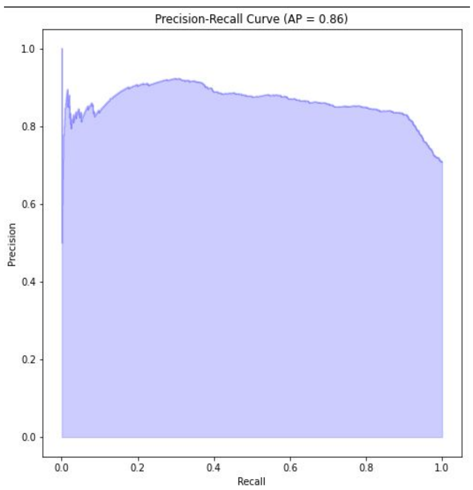

Illustration 18. PR curve for Real Estate Sector (LR)

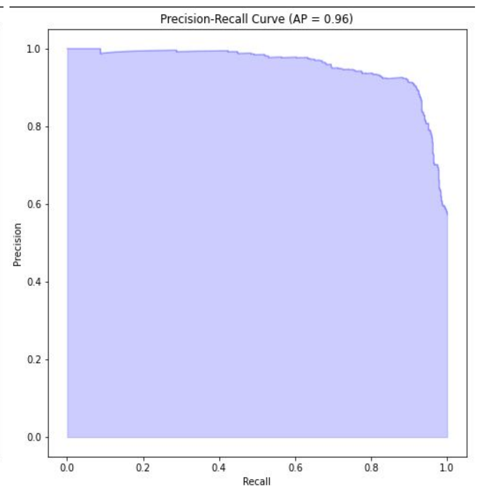

Illustration 19. ROC curve for Industrial Sector (LR)

4. EMPIRICAL RESULTS - Prediction Model - Logistic Regression

Interpretation and analysis of PR curves for LR

Interpretation of the PR curves for the random forest and Gaussian naive Bayes classifier is identical to one mentioned above.

As the characteristic feature of all machine learning approaches they all start from the exact precision and then generally fail to maintain more or less the same precision throughout the sample quickly falling down.

There is no much difference between RF and GNB, both methods behave worse in the beginning of the sample (sudden fall between 0 and 0.2-0.3 on the horizontal axis, then they both plato (except RF for RE) and fall very steeply after around 0.9).

It seems that logistic regression performs much better at maintaining the precision although its starting level might be lower than one of RF and GNB.

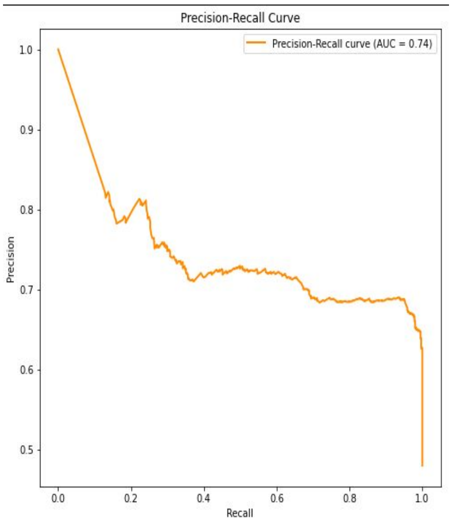

Illustration 20. PR curve for Financial Sector (RF)

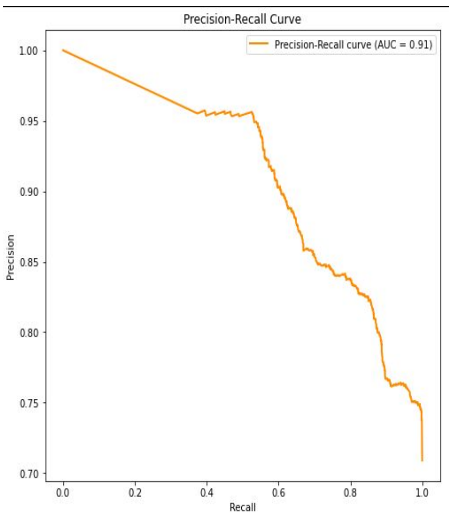

Illustration 21. PR curve for Real Estate Sector (RF)

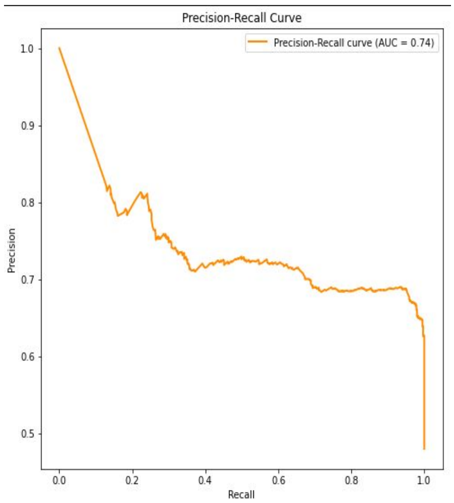

Illustration 22. ROC curve for Industrial Sector (RF)

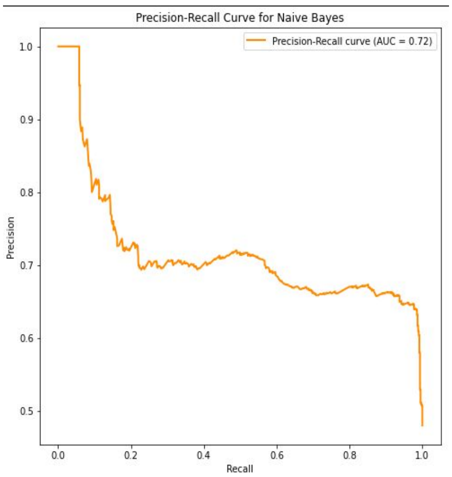

Illustration 23. PR curve for Financial Sector (GNB)

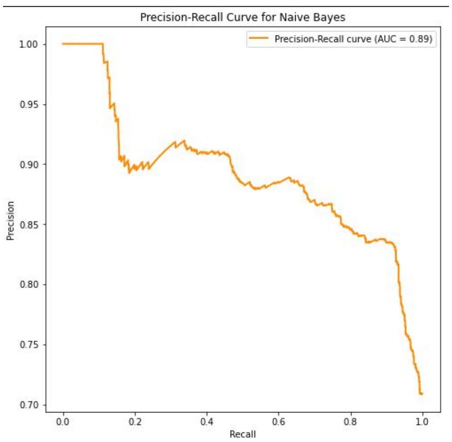

Illustration 24. PR curve for Real Estate Sector (GNB)

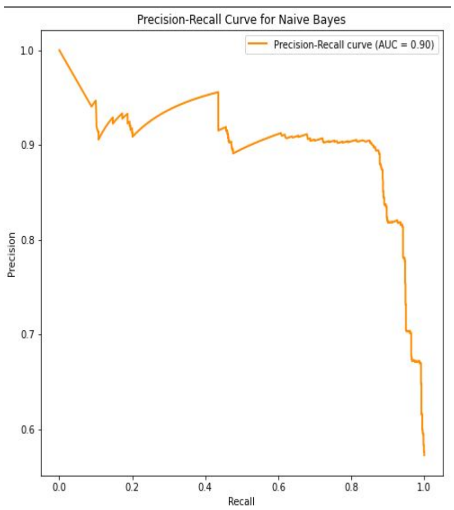

Illustration 5. ROC curve for Industrial Sector (GNB)

5. CONCLUSIONS

N.B. All of the aforementioned methods are inherently incapable of explaining the causal relationship mechanism between the currency-credit spread and the explanatory factors (they cannot answer why this is related in this way). Attributing particular scores and ranking inputs is a mere technical connection between them that can be, actually, completely spurious. Equality of the way how these connections has been captured by all our selection mechanisms means that all of them work not differently from each other (which is good for research consistency).

The empirical results from the variable selection show that across the three industrial classifications in the study, several factors such as Sweden CPI growth rate, yield curve spread of maturities 3 and 6-month, and GER 2020 influence the credit spread differences all the sectors. While the factors affecting the non-financial and real estate companies are similar, credit spreads of bonds issued by financial firms are additionally affected by the net interest income, shareholder value added and the total revenue. Conforming to the existing literature, Eurozone IG Corporate Bond index affects the choice of SEK bonds negatively, as a strong performance in the index suggests increased investor confidence in eurozone market.

Aligning with the theory, higher inflation may erode the purchasing power of bond returns affecting the creditworthiness of the corporate bonds. The 3-month spread, representing the spread between short-term interest rates affects the cost of financing for the companies, across the sectors and can affect the profitability and creditworthiness of the firms thereby influencing the spread of bonds. The level of non-performing assets of the firm is considered a crucial indicator of the asset quality, where a high level of NPA can suggest financial distress and can increase the risk associated with the bonds issued by the firm. The total deposit level reflects stability and liquidity of the firm, while the revenues of the firm indicate robustness and diversification which impacts its creditworthiness.

While economic indicators like inflation growth rate are common across sectors, there are sector-specific financial metrics that play a significant role. The financial sectors rely heavily on interest rate spreads, while real estate and non-financial sector companies focus on metrics indicating stability and risk. However, each sector has unique determinants reflecting its specific dynamics and risks. It is noticed that for financial sector companies the provision of loan losses as a percent of net charge-offs is significant given that risk management is crucial in this sector. In addition, importance of variables like level of deposits and level of loans as a percent of deposit reflect the lending and borrowing dynamics of a financial institution, which can ultimately impact the credit spread. Solvency and capital adequacy are also key factors in the financial sector.

For the real estate sector, the total scale of assets, as an indicator of financial health, becomes significant in influencing credit spread variations. In addition, the level of revenue and non-performing assets influences credit risk and thereby affects spread differences.

For non-financial companies, economic indicators like CPI index can influence credit spreads due to their sensitivity to economic conditions. In addition, profitability metrics like operating margin and performance metrics such as return on equity and return on assets are important micro-economic factors that affect the spread.

The differences in significant explanatory factors highlight the unique drivers and risk factors specific to each sectors. Understanding these sector-specific factors is crucial for accurate credit spread prediction and assessment.

In our analysis of predictive models across various sectors, we observe distinct strengths in all three models. Random Forest consistently outperforms in terms of accuracy and AUC-PR, showcasing its robustness in handling complexities and achieving a balanced precision-recall trade-off. On the other hand, the Gaussian Naive Bayes Classifier demonstrates comparable performance due to its simplicity, effective in scenarios where interpretability is required. The logistic model exhibits best predictive capacity for non-financial sector data and moderate performance for real-estate and financial sector data. The former achieves the highest accuracy among the three sectors maintains a balance between precision and recall for positive instances. The random forest model consistently performs well across all sectors, while Gaussian Naive Bayes model performs especially well in the real estate sector.

APPENDIX

Table A1. Research biases and their mitigation techniques

Bias	Description	Impact on Model Selection if not addressed	Mitigation Technique
Sample Size Bias	Model performance metrics are sensitive to the size of the dataset. Smaller datasets - not representative view of a model's true performance.	Model chosen based on metrics that are unreliable due to the limited sample size.	Ensuring sufficiently large dataset for reliable model evaluation with representative stratification (bootstrapping not considered)
Selection Bias and Data Leakage	Selection process favors certain types of models over others, leading to an inaccurate estimation of model performance.	Models are chosen based on criteria that favor a specific type of algorithm or hyperparameter setting, leading to suboptimal performance on the out-of-sample data.	❖ Introduction of different groups of independent and uncorrelated factors ❖ Introduction of the sample splitting into train and test subsamples.
Overfitting	Model learns the training data too well, capturing noise and irrelevant patterns that do not generalize to new, unseen data	Exceptionally good performance on the training set, low prediction power on the out-of-sample data.	Utilization of feature selection mechanisms
Hyperparameter Tuning Bias	Repeatedly tuning hyperparameters on the same validation set can lead to overfitting to that specific set.	Chosen hyperparameters may not generalize well to new data, as they are tailored to the idiosyncrasies of the train set.	Utilization of a separate validation set for hyperparameter tuning. Hyperparameters tuned on the separate validation set not connected to the set on which the performance of the model is evaluated.
Look-back Bias	Information that was not available at a particular point in time is used in the analysis or decision-making process as if it were available.	Overstated performance, and failure to learn from mistakes as bad observations are not included	Usage of historical information available at the time of observation.

APPENDIX

Full list of financial sector's variables "fed" to the data selection mechanisms

SEK-EUR, GER CPI Growth rate same period previous year, SWE CPI Growth rate same period previous year, GER CPI level 2020=100, SWE CPI level 2020=100, Steepness SEK (10Y -6M), Steepness EUR (10Y-6M), 3M spread, 6M spread, 5Y spread, 10Y spread, 1 EUR Exc. Rate, KIX-index, OMX level, OMX_Return (d), DAX level, DAX_Return (d), EURO_STOXX_50 level, EURO_STOXX_50_Return, S&P Sweden Investment Grade Corporate Bond Index level, S&P Eurozone Investment Grade Corporate Bond Index level, S&P Sweden Investment Grade Corporate Bond Index return (d), S&P Eurozone Investment Grade Corporate Bond Index return (d), Return on Assets %, Return on Equity %, Shareholders Value Added, SG&A Margin %, Net Interest Income / Total Revenue %, EBT Margin % Net Income Margin%, Normalized Net Income Margin %, Nonperforming Loans / Total Loans %, Nonperforming Assets / Total Assets %, Nonperforming Assets / Equity %, Prov. for Loan Losses / Net Charge-offs %, Avg. Total Equity / Avg. Assets %, Total Equity + Allowance for Loan Losses / Total Loans %, Net Loans / Total Deposits %, Tier 1 Capital Ratio %, Total Capital Ratio %, Core Tier 1 Capital Ratio %, Tier 2 Capital Ratio %, EBT + Int. Exp. / Int. Exp., Net Interest Income_1, Non Interest Income_1, Provision for Loan Losses_1, Total Revenue_1, Net Income_1, Normalized Net Income_1, Diluted EPS before Extra_1, Dividend per Share_1, Net Loans_1, Non Performing Loans_1, Non Performing Assets_1 Total Assets_1, Total Deposits_1, Total Equity_1, Net Interest Income_5, Non Interest Income_5, Provision for Loan Losses_5, Total Revenue_5, Net Income_5, Normalized Net Income_5, Diluted EPS before Extra_5, Net Loans_5, Non Performing Loans_5, Non Performing Assets_5, Total Assets_5, Total Deposits_5, Tangible Book Value_5, Total Equity_5.

Total = 72

Full list of industrial and real estate sectors' variables "fed" to the data selection mechanisms

SEK-EUR, GER CPI Growth rate same period previous year, SWE CPI Growth rate same period previous year, GER CPI level 2020=100, SWE CPI level 2020=100, Steepness SEK (10Y -6M), Steepness EUR (10Y-6M), 3M spread, 6M spread, 5Y spread, 10Y spread, 1 EUR Exc. Rate, KIX-index, OMX level, OMX_Return (d), DAX level, DAX_Return (d), EURO_STOXX_50 level, EURO_STOXX_50_Return, S&P Sweden Investment Grade Corporate Bond Index level, S&P Eurozone Investment Grade Corporate Bond Index level, S&P Sweden Investment Grade Corporate Bond Index return (d), S&P Eurozone Investment Grade Corporate Bond Index return (d), Gross Margin, EBITDA Margin, Operating Margin, Pretax Margin, Effective Tax Rate, Net Margin, Asset Turnover, x Pretax Margin Pretax ROA, x Leverage (Assets/Equity) Pretax ROE, x Tax Complement ROE, x Earnings Retention, Reinvestment Rate, Quick Ratio, Current Ratio, Times Interest Earned, Cash Cycle (Days), Assets/Equity, Debt/Equity, % LT Debt to Total Capital, (Total Debt - Cash) / EBITDA, A/R Turnover, Avg. A/R Days, Inv Turnover, Avg. Inventory Days, Fixed Asset Turnover, WC / Sales Growth, ROIC.

Total = 54

APPENDIX

Table A2. Descriptive Statistics on systemic factors used for all companies during 2018 - 2024 time period

Facor	*count*	*mean*	*std*	*min*	*max*
SEK-EUR	843	58.955	48.8	-41.2	144.495
GER CPI Growth rate same period previous year	843	0.275	0.341	-0.424	1.172
SWE CPI Growth rate same period previous year	843	5.416	3.868	0.291	11.363
GER 2020=100	843	109.372	6.284	99.7	117.8
SWE 2020=100	843	111.497	8.17	101.247	123.295
Steepness SEK (10Y -6M)	842	-0.016	0.835	-2.104	1.762
Steepness EUR (10Y-6M)	841	3.481	2.505	-1.98	10.07
3M spread	836	-0.262	0.521	-1.07	1.006
6M spread	842	-0.199	0.314	-0.886	0.614
5Y spread	841	4.51	2.657	-1.065	8.909
10Y spread	841	3.299	2.412	-0.969	9.069
1 EUR	842	10.726	0.579	9.889	11.941
KIX-index	842	120.586	5.774	111.814	132.241
OMX	842	2148.906	163.116	1716.03	2456.17
OMX_Return	840	0	0.011	-0.039	0.042
DAX	843	14655.93	1157.651	11556.48	16469.75
DAX_Return	842	0	0.011	-0.076	0.045
EURO_STOXX_50	843	3970.195	343.038	2958.21	4564.11
EURO_STOXX_50_Return	842	0	0.012	-0.072	0.051
S&P Sweden Investment Grade Corporate Bond Index	843	229.295	4.791	220.09	236.91
S&P Eurozone Investment Grade Corporate Bond Index	843	226.979	13.06	205.12	244.59
S&P Sweden Return	843	0	0.001	-0.006	0.004
S&P Eurozone Return	843	0	0.002	-0.013	0.012

APPENDIX

Table A3. Positive Correlation Results

Top 10 positively correlated significant explanatory factors (all sig. at at least alpha=0.05)					
Financial		**Real Estate**		**Nonfinancial**	
SWE CPI Growth	0.53	SWE CPI Growth	0.40	SWE CPI Growth	0.57
3M spread	0.46	GER 2020=100	0.40	1 EUR	0.55
SWE 2020=100	0.43	SWE 2020=100	0.40	KIX-index	0.55
GER 2020=100	0.4	KIX-index	0.37	SWE 2020=100	0.51
Net Interest Income_1	0.4	1 EUR	0.33	A/R Turnover	0.51
6M spread	0.39	3M spread	0.32	GER 2020=100	0.5
Net Interest Income / Total Revenue %	0.38	6M spread	0.26	6M spread	0.46
Net Interest Income_5	0.38	A/R Turnover	0.16	3M spread	0.39
Shareholders Value Added	0.37	EURO_STOXX_50	0.14	x Tax Complement	0.29
Total Revenue_5	0.37	Effective Tax Rate	0.12	Gross Margin	0.28

Table A4. Negative Correlation Results

Top 10 negatively correlated significant explanatory factors					
Financial		Real Estate		Nonfinancial	
Core Tier 1 Capital Ratio %	-0.19	S&P Sweden IG Corporate Bond Index	-0.32	Steepness EUR (10Y-6M)	-0.08
SG&A Margin %	-0.2	Reinvestment Rate	-0.34	(Total Debt - Cash) / EBITDA	-0.08
S&P Eurozone IG Corporate Bond Index	-0.33	ROIC	-0.36	Pretax ROA	-0.09
Steepness SEK (10Y -6M)	-0.33	Pretax Margin	-0.36	Effective Tax Rate	-0.24
Total Assets_1	-0.34	x Pretax Margin	-0.36	S&P Sweden IG Corporate Bond Index	-0.27
Nonperforming Assets / Total Assets %	-0.41	Net Margin	-0.36	Cash Cycle (Days)	-0.29
Nonperforming Assets / Equity %	-0.42	ROE	-0.37	Steepness SEK (10Y -6M)	-0.46
Total Deposits_1	-0.44	Pretax ROE	-0.37	Times Interest Earned	-0.47
Nonperforming Loans / Total Loans %	-0.46	Pretax ROA	-0.39	Avg. A/R Days	-0.48
Non Performing Assets_5	-0.49	S&P Eurozone IG Corporate Bond Index	-0.41	S&P Eurozone IG Corporate Bond Index	-0.48

APPENDIX

Table A5. 1-Factor ANOVA Results

Top 10 significant explanatory factors based on F-score (all sig. at at least alpha=0.05)					
Financial		Real Estate		Nonfinancial	
SWE CPI Growth	1988.30	S&P Eurozone IG Corporate Bond Index	1226.66	SWE CPI Growthr	2844.78
Non Performing Assets_5	1588.93	GER 2020=100	1185.35	KIX-index	2703.55
3M spread	1350.45	SWE CPI Growth	1180.01	1 EUR	2692.53
Nonperforming Loans / Total Loans %	1341.08	SWE 2020=100	1138.47	SWE 2020=100	2149.81
Total Deposits_1	1289.30	Pretax ROA	1095.76	A/R Turnover	2096.74
Nonperforming Assets / Equity %	1120.11	KIX-index	969.55	GER 2020=100	2068.54
SWE 2020=100	1101.06	Pretax ROE	960.31	S&P Eurozone IG Corporate Bond Index	1806.13
Nonperforming Assets / Total Assets %	1013.94	ROE	949.81	Avg. A/R Days	1802.12
Net Interest Income_1	993.80	ROIC	925.53	Times Interest Earned	1756.33
6M spread	969.80	Net Margin	923.60	Steepness SEK (10Y -6M)	1728.01

Disclaimer

Disclaimer

These analyses, documents and any other information originating from LINC Research & Analysis (Henceforth "LINC R&A") are created for information purposes only, for general dissipation and are not intended to be advisory. The information in the analysis is based on sources, data and persons which LINC R&A believes to be reliable. LINC R&A can never guarantee the accuracy of the information. The forward-looking information found in this analysis are based on assumptions about the future, and are therefore uncertain by nature and using information found in the analysis should therefore be done with care. Furthermore, LINC R&A can never guarantee that the projections and forward-looking statements will be fulfilled to any extent. This means that any investment decisions based on information from LINC R&A, any employee or person related to LINC R&A are to be regarded to be made independently by the investor. These analyses, documents and any other information derived from LINC R&A is intended to be one of several tools involved in investment decisions regarding all forms of investments regardless of the type of investment involved. Investors are urged to supplement with additional relevant data and information, as well as consulting a financial adviser prior to any investment decision. LINC R&A disclaims all liability for any loss or damage of any kind that may be based on the use of analyzes, documents and any other information derived from LINC R&A.

Conflicts of interest and impartiality

To ensure LINC R&A's independence, LINC R&A has established compliance rules for analysts. In addition, all analysts have signed an agreement in which they are required to report any and all conflicts of interest. These terms have been designed to ensure that COMMISSION DELEGATED REGULATION (EU) 2016/958 of 9 March 2016, supplementing Regulation (EU) No 596/2014 of the European Parliament and of the Council with regard to regulatory technical standards for the technical arrangements for objective presentation of investment recommendations or other information recommending or suggesting an investment strategy and for disclosure of particular interests or indications of conflicts of interest.

Other