Report

Internal Benchmark on Net Interest Income within the Banking industry

Financial Engineering and Management

Version, Date 2.0, 06 June 2023
Project Internal benchmark for W&R Control

Colophon

Version, Date 2.0, 06 June 2023

Our reference M.R. Haring

Author Mark Robert Haring

Document Report v2

Preface

This thesis is written in order to complete the master Industrial Engineering and Management with as track Financial Engineering and Management at the University of Twente. The assignment is performed under supervision of the Rabobank and the University of Twente.

I would like to express my gratitude towards my supervisors of the University of Twente, namely Berend Roorda and Reinoud Joosten, for all the valuable feedback during the execution of my master thesis.

Moreover, I would like to thank Rabobank for the possibility to execute my master thesis for them. The knowledge and expertise shared by my supervisors and colleagues at Rabobank contributed greatly to a good underpinned research. Thus, I would also like to express my gratitude towards my supervisors at Rabobank, namely Liam Eykhout and Frank van Ingen, for their valuable feedback and comments.

Finally, I would like to thank my family and friends which supported me during the execution of the master thesis. They provided me with interesting thoughts and ideas which I could potentially implement in my research.

With kind regards,

Mark Haring

Utrecht, June 6, 2023.

Management summary

Overview

The Rabobank Group (further referred to as Rabobank) was established in the late 19th century as a cooperative agricultural credit union. These days, it is one of the biggest cooperatives in the Netherlands with almost 2 million members. They are active in 40 countries serving approximately 8,5 million customers with over $900 billion in assets (Rabobank, 2022). Rabobank offers services such as retail banking, wholesale banking, private banking, leasing and real estate. Moreover, they are actively involved in leading banking for the Food & Agri sector.

We focus on the Wholesale and Rural (W&R) department within the Rabobank. Wholesale and rural banking offers banking services, such as loans, to large entities who operate in different sectors such as the entire chain of Food & Agriculture. Rabobank is especially interested in knowing more about differences and similarities in Net Interest Income in basepoint (NII bps) between regions or sectors in which they operate. This will give them insights on which regions or sectors are profitable for them and what probable causes are.

Approach

In order to benchmark transaction data between region and sectors, we have developed an suitable Multiple Linear Regression (MLR) model. MLR allows the relationships between multiple variables to be analyzed at different levels of granularity, including individual securities, portfolios, and macroeconomic variables. This helps to identify how variables are related and which variables are most important for our outcomes.

Moreover, we have also build a dashboard tool in which the results are visualized. This enables Rabobank to look in depth into the differences and similarities between regions and sectors. The dashboard can easily be updated with the most recent MLR results.

We have opted to use the tool SPSS to execute our MLR as it can analyze large dataset, provide methods to test for violation of assumptions and is quick in processing output results.

Rabobank has provided access to their database from which a dataset is retrieved. After applying a time series analysis to determine the size of the dataset, transaction data of 2021 and 2022 were used. This dataset is restricted to long-term loans for the departments “Credits & Loans” and “Coverage Large Corporates” within the wholesale department and contains important variables on a monthly basis on contract level. The variables here were month and year of the transaction, Net Interest Income in basepoints (NII bps), location (grouped and ungrouped), region (grouped and ungrouped), sector level 01-03, outstanding, Loss At Default (LAD), Sustainability kpi, Fund Transfer Price in basepoints (FTP bps), North American Industry Classification System (NAICS), Risk ratings (grouped and ungrouped), and facility currency (grouped and ungrouped). Grouped variables are variables containing all the different possibilities, e.g. location grouped is one variable containing all the possible locations. Whereas ungrouped location would be one variable for one specific location.

MLR has several underlying assumptions which need to be met in order for the results to be representative. As these assumptions were not always met, some variables have been excluded from the model. These variables are as follows: nominal, EAD RC, expected loss RC, Regulatory Capital Credit Risk, Net interest income, other net interest income and GII bps.

We have also looked into adding the contract length variable as some contract work with several months of EURIBOR. EURIBOR contracts have floating interest rates agreement and thus different Cost of Fund (CoF). Nevertheless, we were not able to add the above as we were unable to align both databases.

Key findings

We only use standardized beta coefficients in order to analyze the different positive or negative weights of independent variables (IBM, 2016). Using the standardized beta coefficients, we can compare the results of different MLR's and results of independent variables. A strong influence is a standardized beta coefficient higher then 0.2, a moderate influence a value of 0.1, and low to no influence a value of 0 to 0.05 in absolute term.

- ✓ Some currencies have a positive correlation with the NII bps namely US Dollar (USD) and British Pound Sterling (GBP). They have a standardized beta coefficient of 0.222 and 0.265 respectively. This is caused by the CoF for these two currencies. Here the CoF is deemed higher which is calculated towards the client. However, the real CoF is usually lower than what is calculated towards the customer thus increasing the Rabobank's Net Interest Margin.

- ✓ The currency EUR has a negative correlation with the NII bps namely a standardized beta coefficient of -0.259. No currency swaps are needed and thus no additional margin can be made on the spread between currencies. Moreover, no currency swaps are needed for EUR which is part of the FTP and thus, we have a lower CoF and to stay competitive, this is not calculated towards the customer. Finally, as Rabobank is an EUR bank, the fund price of EUR is cheaper compared to other region. Thus, there is no need to search for funds on the interbank market as the cheaper funds imply lower risk and thus lower margins.

- ✓ India, Turkey and Argentina are three locations with high inflation and a strong positive correlation with the NII bps. Their standardized beta coefficients are 0.19, 0.131, and 0.141 respectively. In these countries, higher interest rate base amounts are calculated as these pose higher risk. Due to the higher risk, higher net interest margins are calculated to compensate for the risk. Moreover, due to the high inflation, a higher outstanding amount will have a positive correlation with the NII bps.

- ✓ Our model confirms that some sectors have positive impact on the overall NII bps. Taking Retail Trade (coefficient of 0.184) as an example. This sector has had some difficulties over the past years, starting with the migration towards online. Moreover, this trend has been strengthened by the Covid-lockdowns. Currently, Retail Trade also faces difficulties like hardly no personnel to find, increased energy and rental costs and so on. This all boils down to increased credit risk which Rabobank needs to take into account. Thus, increased credit risk implies higher margins are calculated to compensate for the risk and thus a positive factor on the NII bps. Similar trends can be seen on macro-economic levels for construction and manufacturing.

Recommendations

Although we have found interesting results in our research there is still a lot of untouched potential regarding the benchmarking between regions and sectors. The following are the final recommendations for future research:

1. Apply model to Rural portfolio to observe macro-economic trends.
2. Further analyze why certain variables breach one of the MLR assumptions.
3. Update dataset on monthly basis to observe changes in standardized beta coefficient and thus to be able to notice trend changes.
4. Apply the model per location to spot differences in sectors and or other variables.
5. Migrate the MLR model with current Rabobank's margin models for strengthening predictive and descriptive power
6. Translate the outcome in a more predictive model in order to support strategic decision making.

The implementation of the abovementioned recommendations will offer a more in depth analysis in macro-economic trend changes, analysis of different departments, and offer more predictive power on the factors that influence the NII bps which can be used by Rabobank in order to strengthen their position.

Table of content

List of figures

List of abbreviations

Abbreviation	Meaning
ABON	Advance payment bon
AF	Adjustment Factor
BCBS	Basel Committee on Banking Supervision
Bps	Basis point
BRID	Bridge loan
CALL	Call loan
CoF	Cost of Funds
CROIC	ex-post performance evaluation
CURR	Current Account Overdraft
ECB	European Central Bank
FILE	Financial Lease
FTP	Fund Transfer Price
GII	Gross interest income
GUAG	Guarantee Given
IMPO	Import letter of credit
LAD	Loss at Default
LCST	Standby letter of credit
LGD	Loss given default
LOTA	Long-term advance
NII	Net Interest Income
NIM	Net Interest Margin
PD	Probability of Default
RC	Regulatory Capital
RFR	risk-free rate
RWA	Risk-weighted assets
SHLO	Short-term loans
SME	Small and medium-sized enterprises
SYND	Syndicated loan
TERM	Long-term loan
Wholesale & Rural	Wholesale and Rural

Abbreviation	Meaning
AI	Artificial Intelligence
ARIMA	AutoRegressive Integrated Moving Average
CGCB	Central Governments and Central Banks
EAD	Exposure at default
EL	Expected Loss
IRB	Internal Rating Based
MAE	Mean absolute error
ML	Machine Learning
MLR	Multiple Linear Regression
MSE	Mean squared error
NAICS	North American Industry Classification System
RC	Recovery Rate
RRR	Rabobank Risk Rating
VAR	Vector Autoregression
VECM	Vector Error Correction Model
VIF	Variance inflation factor

1 Introduction

The Rabobank Group (further referred to as Rabobank) was established in the late 19th century as a cooperative agricultural credit union. These days, it is one of the biggest cooperatives in the Netherlands with almost 2 million members. They are active in 40 countries serving approximately 8,5 million customers with over $900 billion in assets (Rabobank, Rabobank, 2023). Rabobank offer services such as retail banking, wholesale banking, private banking, leasing and real estate. Moreover, they are actively involved in leading banking for the Food & Agri sector where, according to their mission statement, they aim to build a better world together.

Banks are under strict regulation from regulators such as the European Central Bank (ECB) and other entities. They are expected to provide external reporting on their assets, risk-weighting and conducts in conform to the Basel system (Bank for international settlements, 2022). The Basel system is a set of international banking regulations established by the Basel Committee on Banking Supervision (BCBS). It prescribes the minimal capital requirements for financial institutions with the goal of mitigating credit risk as experienced in the banking crisis of 2008. Banks that operate internationally are required to maintain a minimum amount of capital based on their risk-weighted assets.

Banks do not only provide external reporting but also internal reporting in order to align their business processes. Thus, it is of importance for the management teams to have quantitative and qualitative analyses of data on a daily basis. In an age where data are one's new gold, one can obtain valuable quantitative and qualitative information from data on ones operations and conducts.

We focus on the Wholesale and Rural (W&R) department within the Rabobank. Wholesale and rural banking offers banking services to large entities who operate in different sectors such as the entire chain of Food & Agriculture.

1.1 Problem context

As explained, W&R department wants to achieve more insights on data available and or the products they offer their customers. They are especially interested in knowing more about differences and similarities between regions or sectors in which they operate and the probable causes.

In order to understand the problem context, one should first understand the basic concept of banking. Some may remember the 3-6-3 rule of bankers back in the day. According to this rule, bankers gathered deposits at 3 percent, lent them at 6 percent, and were on the golf course by 3 o'clock in the afternoon. Of course this rule doesn't hold true these days, but one can grasp the basic concept of how a simple loan would work for retail banking. The idea is that banks lend out money they have at a higher interest rate while ensuring they have enough capital for their risk-weighted assets.

Wholesale banking works a bit differently but has the same basic concepts. The main difference between wholesale and retail banking is that products, in this case loans, are offered to large entities or corporations. These customers can obtain a loan with a nominal amount, which is the maximum amount they may borrow from the bank. However, they cannot decide to use up the entire nominal amount in one go as the ability of the bank to obtain all the repayments or the possibility of default would be too risky. However, customers may use up smaller amounts of the nominal to fund business decisions or processes. Purposes of using funds of the nominal amount can be to build a new factory, buying machines in order to run their business, to increase their inventory or other business decisions where funds are needed. The actual amount that customers have used from the nominal amount is what we call the outstanding amount of the loan. From Figure 1-1, depicted below, we see that the customer has a nominal of 150 million and an outstanding of 100 million

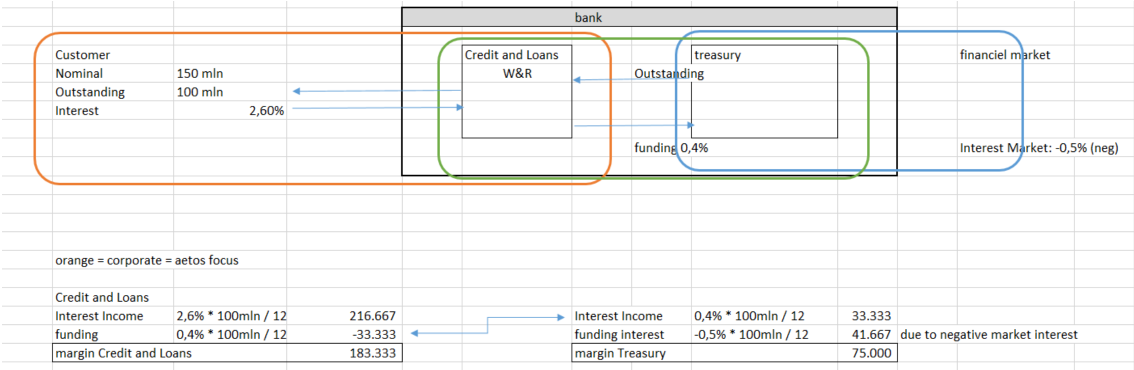


Figure 1-1: Loan structure wholesale banking.

Now we see from Figure 1-1 that W&R department actually does not hold funds for the loans they offer. Wholesale and Rural department ask for a loan at treasury, which is the bank within the bank. Rabobank first checks if the customer is eligible for such loan and if this is the case, treasury offers the loan to the W&R department. Nonetheless, treasury needs to obtain funding for this loan and looks to borrow this money from the Financial markets such as the European Union interbank market. These markets are where banks borrow money from each other and ask a risk-free rate (RFR) plus some margin. In Chapter 2, we explain more on RFRs and on how these are calculated on a daily basis.

Evidently, treasury asks W&R to pay for the interest rate plus an additional markup as treasury has to borrow the money from the open market. Consequently, W&R charges the customer this percentage plus some additional interest rate for the loan. As can be seen from Figure 1-1, this results in Gross Interest Income (GII) for W&R of 216.667. When we subtract the costs W&R incurs, this results in a Net Interest Income (NII) of 183.333.

It must be noted that many different types of loans exist. The situation above describes a Credit & loans long-term loan. However, other products such as short term loans (SHLO), long term advance (LOTA), call loan (CALL), bridge loan (BRID) and many others exist. We will elaborate more on this in Section 2.1.2.

As Rabobank is a large player in the financial markets and especially for Wholesale and Rural banking, they offer large amounts of loans and have a large customer base. Rabobank is interested in a benchmark on the income per transaction between regions and or per sector. This gives them more insights on which regions and or sectors are more profitable and what probable causes are. Moving forward, in our research income per transaction is defined as net interest income as Wholesale control uses this to determine their performance.

1.2 Research objective

Our objective is to investigate the differences in net interest income made by the Rabobank on services provided by W&R department and to get a better understanding on these differences and their possible causes. By shedding light on the differences between region and sectors, we aim to find critical differences in the net interest income which yields focus points for management teams.

Moreover, we will develop a benchmark methodology for a certain loan product within the W&R department which later can be applied to other products yielding valuable additional information.

1.3 Research design

To achieve the research objective, we have formulated research questions to structure the research. The main research question is stated below:
What are the significant differences in Net Interest Income bps (NII bps) between regions and sectors, and how can Rabobank gain a competitive advantage from the factors that influence the NII's bps?

In order to answer this, several sub-research questions are formulated.

1. How does wholesale banking work?
    a. What are the differences between retail and wholesale banking?
    b. What kind of products are offered within Wholesale and Rural banking?
2. What are important variables and definitions within finance and wholesale banking?
    a. What is Gross Interest Income and Net Interest Income?
    b. What is Fund Transfer Price?
    c. What other variables are of importance when conducting an internal benchmark on Net interest Income?
3. What benchmark techniques can be used to compare net interest income?
    a. What are suitable benchmark frameworks?
    b. What is an appropriate model when benchmarking NII

1.4 Scope and demarcations

In order to ensure that the research is feasible within the given time restriction, it is necessary to define the scope and demarcations of the research. We define the scope such that the research can be executed within 26 weeks but is of high complexity and academic level.

From initial data analyses, we see that Wholesale department has the highest nominal and outstanding amounts compared to Rural. Moreover, we note that the departments "Credits and Loans" and "Coverage Large Corporates" amounts for the largest proportion of the portfolio, namely 60.98% of the outstanding for Wholesale and 36.91% for Rural. These two departments are a subset of the Wholesale department.

Furthermore, within the Wholesale portfolio, we see that the groups "Credits and Loans" and "Coverage Large Corporates" constitute 96.83% of the entire outstanding Wholesale portfolio, 65.84% and 28.29% respectively. Figure 1-2 depicts the outstanding amount per super parent (SP) within the wholesale portfolio. It must be noted that the number of outstanding on the y-axis has been hidden due to confidentiality.

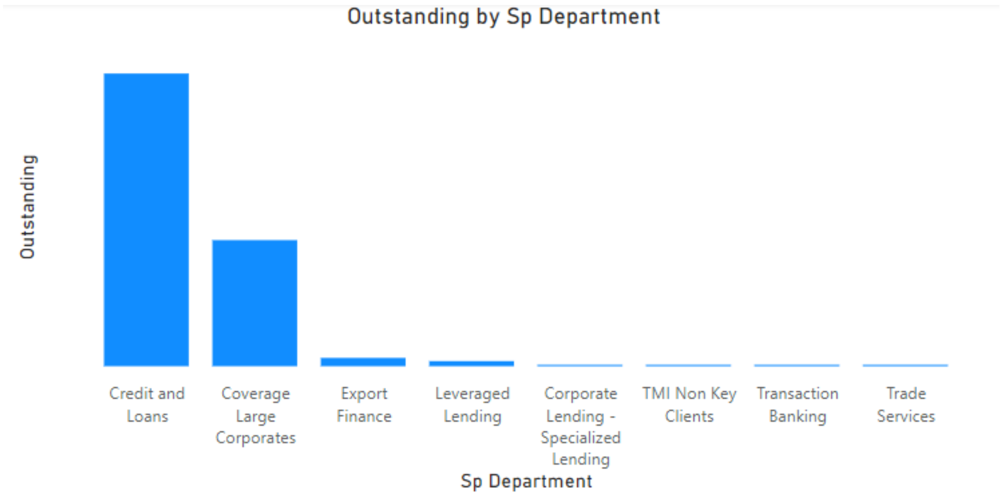


Figure 1-2:Oustanding per Super Parent (SP) department.

Taking a deeper look into the “Credits and Loans” and “Coverage Large Corporates”, we notice that product types labelled “TERM”, which are long-term loans, make up the largest proportion of the portfolio with 69.35%. The second largest proportion are “SHLO”, which are short term loans, and they only account for 6.74% of the portfolio. Thus, we will only focus on the long-term loans which are offered for “Credits and Loans” and “Coverage Large Corporates” within wholesale lending.

Currently, rural lending is only offered in 10 different regions which are four in North America, four in South America, and one in Australia and New Zealand each. Rural lending used to have a region situated in Utrecht however this one was closed in late 2021 and has missing data on GII and NII.

In conclusion, due to time restriction of the research, missing data on GII and NII, and the observation that rural lending has way fewer locations compared to wholesale lending, this research will only focus on long-term loans for “Credits and Loans” and “Coverage Large Corporates” within wholesale lending. Nevertheless, once the methodology is constructed for Wholesale lending, the same model may be applied to Rural lending in the future as 97.4% or the rural lending portfolio constitutes of the product type long-term loans.

We analyze the last 5 years of data for the products described above as it will provide sufficient data for our research. Some assumptions will need to be made concerning certain fees, e.g. commitment fee and penalty fee which are correlated to the maturity date.

1.5 Methodology and thesis outline

We answer the main research question with the help of the sub-research questions. The first two sub-questions aim to gather information by means of a literature study on Wholesale banking, loan structures, Income per transaction, Gross Interest Income, Net Interest Income and FTP. The third sub-research question aim to construct an appropriate model in order to successfully apply a benchmark on NII. Thus, once the model is determined and constructed, we use it to analyse the NII per regions and sectors. In Chapter 4, we elaborate on the results of the proposed benchmark model.

In Chapter 5 we will discuss our findings and results and draw conclusions from them. Moreover, we will elaborate on the limitations of the research. In Chapter 6 we give some recommendations for further research followed by a reference list and appendix.

2 Literature review

2.1 How does wholesale banking work?

This literature review provides an overview of the current research on retail and wholesale banking, including their comparative analysis, the role of wholesale banking in financing small and medium-sized enterprises, and the role of wholesale funding in bank liquidity management. Additionally, a loan structure is an essential aspect of banking and financial services, which can be a complex process that requires a deep understanding of various factors and assumptions.

2.1.1 What are the differences between retail and wholesale banking?

The banking industry has been evolving over the years, with retail and wholesale banking being two major models. These major models are two distinct types of banking services offered by financial institutions to different categories of customers. Retail banking serves individuals and small businesses, while wholesale banking serves large corporations, financial institutions, and governments. The differentiation of the two types of banking services lies in the nature of the services they offer, the risks involved, and the regulatory framework governing them.

Retail banking is a banking service that satisfies the financial needs of individual customers, small businesses, and households. It offers a wide range of products and services, including checking and savings accounts, personal loans, mortgages, and credit cards, among others. On the other hand, wholesale banking is a banking service that caters to the financial needs of large corporations and institutions. Wholesale banking services include trade financing, foreign exchange services, and underwriting services, among others.

One of the significant differences between retail and wholesale banking is the size and type of clients they serve. Retail banks serve smaller clients, while wholesale banks serve larger clients. Another key difference is the nature of services provided by these two types of banks. Retail banks offer a wide range of products and services to their customers, while wholesale banks tend to offer fewer, more specialized complex products and services.

Ahmad & Khan (2016) conducted a comparative analysis of retail and wholesale banking, focusing on their organizational structure, product lines, customer base, and regulatory framework. The authors found that retail banking is highly regulated, offers a wide range of products and services to individual and small business customers, and relies on physical branches to deliver its services. Wholesale banking, on the other hand, is less regulated, offers a narrower range of products and services to large corporate and institutional customers, and relies on sophisticated technology and digital platforms to deliver its services. The authors conclude that while retail banking is more stable and less risky, wholesale banking offers higher returns and growth potential.

Keogh (2015) examined the role of wholesale banking in financing small and medium-sized enterprises (SMEs). The author argued that SMEs face significant challenges in accessing credit from traditional retail banks, and wholesale banking can play a critical role in filling this gap. The author identified several wholesale banking products that can be used to finance SMEs, including asset-based lending, trade finance, securitization, and private placement of debt. The author also highlighted the importance of building relationships between SMEs and wholesale banks to increase access to credit and support long-term growth.

Düllmann & Masschelein (2018) explored the role of wholesale funding in bank liquidity management. The authors argued that wholesale funding plays a crucial role in providing banks with access to low-cost funding sources, but also exposes them to liquidity risks. The authors examined the impact of different types of wholesale funding on bank liquidity, including interbank funding, securitization, and covered bonds. The authors also highlighted the importance of stress testing and contingency planning in managing liquidity risk in wholesale banking.

It must be noted that banks are subject to strict government regulations on the risk they undertake. Although many options, futures and other derivatives are offered within the banking industry, appropriate risk management is demanded and supervised by central banks. This all is done according to the Basel system and regulatory capital needed, as explained in section 0.

In conclusion, retail and wholesale banking are two distinct types of banking services that cater to different customer segments and operate under different regulatory frameworks. While retail banking offers a wider range of products and services to individual and small business customers, wholesale banking offers higher returns and growth potential to large corporate and institutional customers. Wholesale banking can also play a critical role in financing SMEs and providing banks with access to low-cost funding sources, but it also exposes them to liquidity risks. Thus, managing risk is critical in wholesale banking, and stress testing and contingency planning are essential tools for managing liquidity risk. Overall, the above highlights the importance of understanding the differences between retail and wholesale banking and the challenges and opportunities they present for financial institutions.

2.1.2 What kind of products are offered within Wholesale and Rural banking?

Loans are a critical component of the banking industry, and they are provided to both retail and wholesale banking customers. Loan structures differ depending on the type of customer and the amount of capital needed.

In retail banking, loans are usually structured as consumer loans, mortgages, or small business loans. Consumer loans are provided to individuals for personal needs such as purchasing a car or paying for education. Mortgages are provided to individuals for purchasing homes, while small business loans are provided to small businesses for expansion or investment purposes. The loan structure for retail banking customers is typically straightforward and involves a set repayment schedule.

In wholesale banking, loans are usually structured as syndicated loans or project finance. Syndicated loans are large loans provided to corporations that require significant amounts of capital, and they are usually structured as a group of banks sharing the loan risk. Project finance is a type of loan provided to businesses for specific projects such as building infrastructure or investing in renewable energy. The loan structure for wholesale banking customers is typically more complex and involves multiple parties and detailed agreements.

The loan structure offered in both retail and wholesale banking involves a loan agreement between the bank and the borrower. The agreement outlines the terms and conditions of the loan, including the loan amount, interest rate, repayment schedule, and any collateral required. In retail banking, the loan agreement is usually straightforward and standardized, while in wholesale banking, the loan agreement is more complex and tailored to the specific needs of the borrower. We will now elaborate on some of these complex loan agreements (Rabobank, 2018) found within the departments as defined in our scope and demarcation in Section 1.4. Here only the product type “TERM” will be of interest for our research. All other financial product types can be found in Section 8.3.

2.1.2.1 TERM

A term loan is a loan for a period of at least one year, subject to previously agreed conditions. The interest can be fixed or variable. In the case of a fixed rate, the interest can be established for the entire term of the loan or for a shorter period, in which case the period must be at least one year, in accordance to the definition of the Rabobank.
Term loans are used for long term financing needs (fixed assets) and for working capital needs (current assets). A term loan can be paid back in equal monthly, quarterly, semi-annual or annual instalments, or in annuities. Other methods of repayment are repayments with a balloon and bullet repayments. Rabobank policy concerning term loans are given below:

- The maximum tenor for long term loans should not exceed 10 years, with a strong preference for a 5 to 7 years tenors. Exceptions to this policy have to be based on strong motives.
- Repayment should preferably be by the straight line method of repayment. Thus a same amount of interest is allocated in each payment until the debt is repaid in full. Repayment should be dependent on the cash-flow of the business requesting the loan, or the cash flow from the project being financed. The nature of the goods to be financed should also be taken into account.
- Bullet loans may only be granted to primary customers of the bank, whose financial position must also be sound. The maximum term of a bullet loan is limited to 5 years. The credit report should indicate that the bullet loan is a suitable way of financing. The report should also indicate which funds the bullet loan will be repaid from at the date of maturity.

2.2 What are important variables and definitions within finance and wholesale banking?

We firstly elaborate on the difference between Gross Interest Income and Net Interest Income. Moreover, we discuss other important variables within finance. These include variables such as Exposure at Default, Regulatory Capital, Fund Transfer Price, Probability of Default, Rabobank's Risk Rating, Facility currency, and sustainability KPI.

2.2.1 What is Gross Interest Income and Net Interest Income?

Wholesale banking plays a significant role in the global economy, and the interest income generated by these banks is a critical component of their profitability. Gross interest income (GII) and net interest income (NII) are two key indicators of a wholesale bank's performance. GII is the total amount of interest income earned by the bank before any expenses are deducted. In contrast, NII, or also called Net Interest Margin, is the difference between interest earned and interest paid, after deducting any expenses.
Below an overview is given on gross interest income and net interest income in wholesale banking, including their definitions, significance, and factors affecting them.

GII is a measure of a wholesale bank's revenue generated from lending activities, such as corporate loans, trade finance, and project finance as explained in previous section. It is a key indicator of a bank's profitability and reflects the bank's ability to earn income from its assets. However, GII does not account for any expenses, such as funding costs, credit risk, or operational costs. Therefore, it may not provide a comprehensive picture of a bank's performance.

NII is the difference between interest earned and interest paid by a wholesale bank. It is a key indicator of a bank's profitability and reflects the bank's ability to manage its interest rate risk. Wholesale banks typically earn interest income from their assets, such as loans and investments, and pay interest expenses on their liabilities, such as deposits and borrowings. Therefore, a bank's ability to manage its costs of funds (CoF) and interest rate risk is critical to its NII.

GII and NII are important measures of a bank's financial performance as they reflect the bank's ability to generate income from its assets and liabilities. A higher GII indicates that a bank is earning more interest on its assets, which is a positive sign for the bank's financial health. Similarly, a higher NII indicates that a bank is earning more interest on its assets than it is paying on its liabilities, which is also a positive sign for the bank's profitability. However, a low NII may indicate that a bank is paying high interest on its liabilities or is unable to generate sufficient income from its assets.

Several factors affect GII and NII in wholesale banking, including market conditions, interest rate risk, credit risk, CoF, and operational costs. In times of economic downturns or recession, wholesale banks may face higher credit risk and increased default rates, which can lead to lower GII and NII. Similarly, changes in interest rates can affect a bank's NII, as higher interest rates increase CoF, while lower interest rates reduce interest income.

Moreover, it must be noted that it isn't uncommon to talk about GII and NII in basis points (bps). One basis point is equivalent to 0,01% thus 100 bps is equal to 1%. This measure is often used within financial systems. Control Wholesale within Rabobank actually steers based on NII bps.

Research studies have shown the importance of GII and NII in assessing the financial performance of banks. For example, a study by Demirgüc-Kunt & Huizinga (1999) found that GII is positively related to a bank's profitability, while a study by Bikker & Haaf (2002) found that NII is a significant determinant of bank profitability. Hasan & Dridi (2011) found that GII and NII are positively related to bank size, suggesting that larger banks are better able to generate income from their assets and liabilities. Huang & Huang (2018) examine the relationship between interest rate risk and NII for Chinese banks and found that interest rate risk has a significant negative impact on NII. Interest rate risk is the potential for investment losses that can be triggered by a move upward in the prevailing rates for new debt instruments thus resulting in a lower NII (Chen, 2022).

In conclusion, GII and NII are essential indicators of a wholesale bank's performance. GII reflects the bank's ability to generate revenue from lending activities, while NII indicates the bank's ability to manage its interest rate risk and CoF. Factors affecting gross and net interest income include market conditions, credit risk, interest rate risk, CoF, and operational costs. Therefore, wholesale banks must manage their operations effectively, monitor their credit and interest rate risk, and adapt to changing market conditions to maintain their profitability.

2.2.2 What is Funds Transfer Price?

The funds transfer price (FTP) of a transaction or product consists of a cost of funds (CoF) component and an adjustment factor. The CoF component can consist of interest rates, liquidity premium or other product specific FTP add-ons. An overview of the CoF components can be found below.

	Type of FTP Component	FTP Component	Interest/ Liquidity related
2.1	Interest	Interest Rate	I
		Basis swap	I
2.2	Liquidity	Liquidity spread curve for the core currencies	L
		Currency swap spread	L
		Liquidity tenor	L
2.3	FTP add-on – options	Interest floor discount	I
		Prepayment	I/L
		Offer risk	I/L
		Multi-currency	L
		Interest Rate Switch	I
2.4	FTP add-on - collateral cost	Collateral Cost	I
2.5	FTP discounts - Funding benefits	Secured Funding Benefit	L
		Alternative Funding	L
2.6	FTP components related to Liquidity Buffer Cost	Liquidity Buffer Cost:	
		1. Retail liability outflow	L
		2. Undrawn FTP	L
		3. Extension risk	L

Figure 2-1: Overview of Cost of Funds components (Rabobank, 2018).

The CoF is defined as the market price of an asset (liability) if it is funded (invested) in the financial markets with the same interest rate, currency and liquidity risk characteristics.

For each transaction an FTP can be calculated, based on the characteristics of the transaction, prices in the financial markets and a possible Adjustment Factor. Besides steering on the performance metrics as stated in the SF, there is also an element of steering within the FTP called the adjustment factor (”AF”). This is an additional spread (which can be negative or positive) on top of the Cost of Funds. Figure 2-2 shows a schematic overview of the FTP framework:

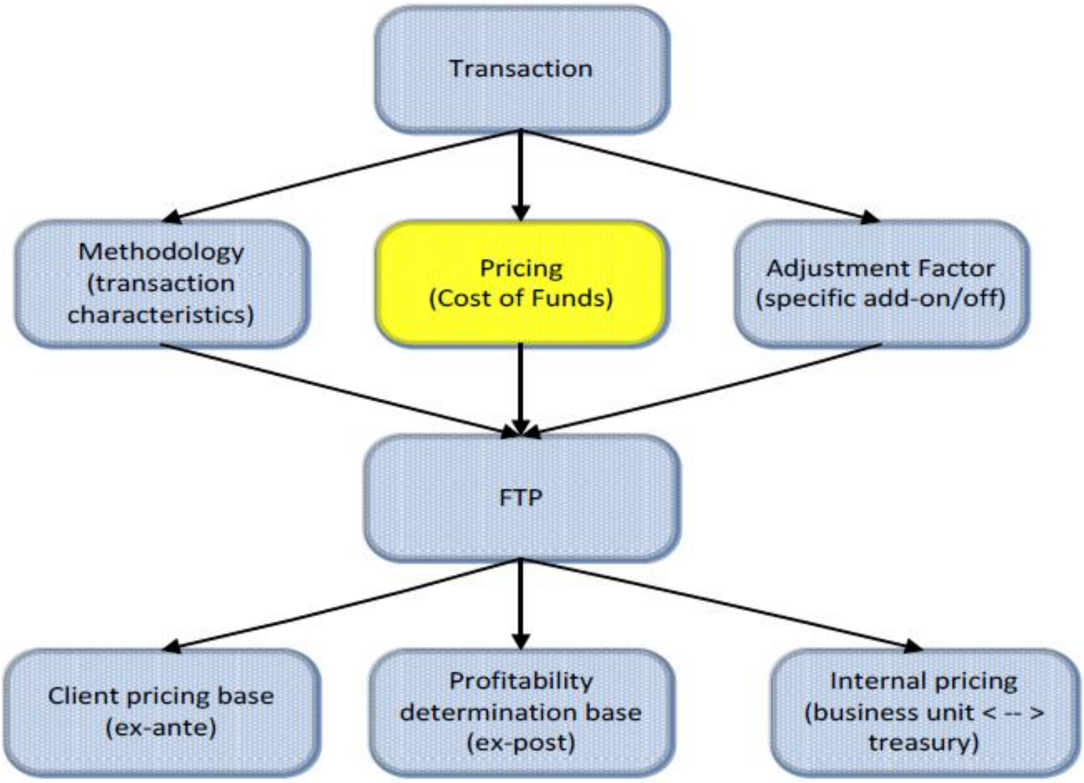


Figure 2-2: Schematic overview of FTP framework (Rabobank, 2018).

2.2.3 What other variables are of importance when conducting an internal benchmark on Net interest Income?

We discuss several variables which are of importance when conducting an internal benchmark on NII bps. These consist of some overall finance related variables which are deemed necessary by the authorities and central banks and other variables which are Rabobank specific.

2.2.3.1 Regulatory Capital

Regulatory Capital (RC) is an important concept in wholesale banking. It refers to the amount of capital that a bank must hold in order to comply with regulatory requirements. In other words, it is the minimum amount of capital that a bank needs to maintain in order to cover potential losses from credit, market, and operational risks.

Regulatory Capital (RC) is a crucial component of bank regulation, and it plays a significant role in ensuring the safety and soundness of the banking system. The concept of RC has evolved over time, and it is currently governed by a set of international standards known as the Basel Accords. The Basel Accords were developed by the Basel Committee on Banking Supervision (BCBS), which is a global regulatory body that sets standards and guidelines for banking supervision (BCBS, 2010).

According to the Basel Accords, banks are required to maintain a minimum level of regulatory capital to cover their risk exposures. Regulatory capital is divided into Tier 1 and Tier 2 capital. Tier 1 capital is composed of equity capital and retained earnings, while Tier 2 capital includes subordinated debt and hybrid capital instruments. The purpose of these two tiers is to ensure that banks have sufficient capital to absorb potential losses.

In addition to the minimum regulatory capital requirements, banks are also subject to capital adequacy ratios. The two most common ratios are the Tier 1 capital ratio and the Total Capital Ratio. The Tier 1 capital ratio is the ratio of Tier 1 capital to risk-weighted assets, while the Total Capital Ratio is the ratio of Tier 1 and Tier 2 capital to risk-weighted assets.

The minimum amount of risk capital needed arises from the materialization of Market Risk, Operational Risk and Credit Risk. The calculation methods of the minimum required amount are determined by legislation/regulators and change over time. The minimum capital requirement is calculated based on the risk-weighted assets are as followed:

$$RC = \frac{RWA}{12{,}5}$$

Regulatory capital has a number of implications for wholesale banking. First, it affects a bank's ability to take on risk. Banks with higher levels of regulatory capital are able to take on more risk than banks with lower levels of capital. Second, regulatory capital can impact a bank's profitability. Banks with higher levels of capital may have lower returns on equity because they are required to hold more capital. Finally, regulatory capital can impact a bank's ability to pay dividends or repurchase shares. Banks that do not meet the minimum regulatory capital requirements may be prohibited from paying dividends or repurchasing shares.

In conclusion, Regulatory Capital (RC) is an important concept in wholesale banking. It is a key component of bank regulation and plays a significant role in ensuring the safety and soundness of the banking system. Banks are required to maintain a minimum level of regulatory capital to cover their risk exposures, and they are subject to capital adequacy ratios. Regulatory capital has implications for a bank's ability to take on risk, its profitability, and its ability to pay dividends or repurchase shares. Overall, a bank's ability to manage its regulatory capital is critical to its success in the wholesale banking sector.

2.2.3.2 Loss at Default

Loss at default (LAD) is a critical aspect of credit risk management in wholesale banking. It refers to the expected loss incurred when a borrower defaults on their loan obligations. Institutions must accurately estimate LAD to determine their overall credit risk and capital requirements.

The estimation of LAD involves several factors, including the probability of default (PD), the loss given default (LGD), and the exposure at default (EAD). These factors help banks to assess the expected loss in the event of a borrower's default and make provisions to cover these losses.

Several studies have explored different models and methods for estimating LAD. For example, Vrins & Verbeek (2007) developed a model that accounts for the correlation between PD and LGD. The model also considers the variability in the timing of the recovery of the outstanding debt after a default.

Another approach to estimating LAD involves using Monte Carlo simulation methods. Garcia et al. (2015) used this method to estimate LAD for a portfolio of commercial loans. The simulation considered the uncertainty in the values of PD, LGD, and EAD and provided a range of possible LAD values.

Several studies have also investigated the impact of different factors on LAD estimates. For example, Krahnen et al. (2011) examined the impact of market volatility and asset correlations on LAD estimates for a portfolio of securitized products. They found that the inclusion of market volatility and asset correlations led to higher estimates of LAD.

The importance of accurate LAD estimates has been highlighted by several studies. For example, Maio & Olsen (2016) found that the underestimation of LAD led to the failure of several European banks during the 2008 financial crisis. Accurate LAD estimates help banks to manage their credit risk and maintain so adequate capital reserves to cover potential losses.

In conclusion, the estimation of LAD is an essential aspect of credit risk management within wholesale banking. Accurate LAD estimates help banks to manage their credit risk and ensure adequate provisions to cover potential losses. Several models and methods are available for estimating LAD, including those that account for correlations between different factors and the use of Monte Carlo simulation. The importance of accurate LAD estimates has been demonstrated by the impact of underestimation on banks during past financial crises.

2.2.3.3 EADRC

Exposure at Default (EAD) and Recovery rate (RC) are two crucial parameters in credit risk modeling, which determine the Expected Loss (EL) of a bank's credit portfolio. Here we focus on EAD and RC in Wholesale Banking, which involves lending to large corporations, financial institutions, and sovereign entities. This review provides an overview of the definition and calculation of EAD and RC, as well as the key factors affecting them.

Exposure at Default (EAD) is defined as the amount of exposure a bank has to a counterparty at the time of default. It is a key parameter in credit risk modeling and is used to calculate Expected Loss (EL), which is the amount that a bank expects to lose due to the default of its counterparties. The calculation of EAD involves the estimation of the amount of funds that a counterparty has drawn down from the credit line, as well as any undrawn amounts that may be available. EAD can be calculated using either the Standardized Approach or the Internal Ratings-Based Approach (IRB). Under the Standardized Approach, EAD is calculated based on a fixed percentage of the total exposure, while under the IRB approach, EAD is calculated using a more sophisticated methodology that takes into account the creditworthiness of the counterparty.

EAD for on balance exposure
For a facility i at point in time t the following formula applies:

$$EAD\ nd_{i,t} = X_m\% * Exposure\ according\ to\ accounting\ principle$$

Where: $X_m = (0, 20, 50)$ based on the maturity

EAD for off balance exposure (undrawn Credit Line)

$$EAD\ nd_{i,t} = CCF_{i,t} * Undrawn\ Exposure\ AMount\ Credit\ Line_{i,t}$$

Where:

$CCF_{i,t}$ equals the (Credit) Conversion Factor that is applicable for the Undrawn Exposure Amount Credit Line l at point in time t.

And where:

$$\begin{aligned} Undrawn\ Exposure\ AMount\ Credit\ Line_{i,t} \\ &= Exposure\ AMount\ Credit\ Line_{i,t} \\ &- \sum_{i=o} Exposure\ according\ to\ accounting\ principles_{i,t,l} \end{aligned}$$

Where:
i,l,t reflects a contract i that is drawn under credit line l at point in time t

Recovery rate (RC) is the percentage of the exposure that a bank can recover in the event of a default. It is another key parameter in credit risk modeling and is used to calculate EL along with EAD and Probability of Default (PD). RC can be estimated based on historical data or expert judgment, and it typically varies depending on the type of collateral and the seniority of the debt. In general, the higher the seniority of the debt and the better the quality of the collateral, the higher the recovery rate.

Several factors affect EAD and RC in wholesale banking. These include the creditworthiness of the counterparty, the type of collateral, the seniority of the debt, the loan structure, and the economic environment. The creditworthiness of the counterparty is a key factor in determining both EAD and RC. A counterparty with a higher credit rating is likely to have a lower EAD and a higher RC than a counterparty with a lower credit rating. The type of collateral and the seniority of the debt also affect RC, with collateralized debt typically having a higher RC than unsecured debt. The loan structure can also affect EAD and RC, with revolving credit lines typically having a higher EAD than term loans. Finally, the economic environment can also affect EAD and RC, with a recession or financial crisis leading to higher EAD and lower RC.

Thus, Exposure at Default (EAD) and Recovery rate (RC) are two important parameters in credit risk modeling. EAD represents the amount of exposure a bank has to a counterparty at the time of default, while RC represents the percentage of the exposure that a bank can recover in the event of a default. Both EAD and RC are affected by various factors, including the creditworthiness of the counterparty, the type of collateral, the seniority of the debt, the loan structure, and the economic environment. Accurate estimation of EAD and RC is essential for effective risk management within wholesale banking.

2.2.3.4 Probability of Default

The probability of default (D1-D4), as depicted in Figure 2-3, of a counterparty over a one-year period. The PD is determined using financial as well as non-financial characteristics of the client in question, and is expressed in basis points. Initially, the Rabobank Risk Rating (RRR) is used to classify clients into groups having a similar PD. To allocate a rating to an enterprise, various internal rating models are used, depending on the segment in which the enterprise operates. For applications other than economic capital (e.g. pricing) a different time horizon than one-year may be considered appropriate.

Rating Scale	Definition
D1	The Client is past due more than 90 days on a Material Contractual Payment and the Technical Default criteria or D2, D3 or D4 are not applicable.
D2	It is unlikely that the Client will pay its debt obligations (principal, interest or fees) in full, without recourse by the bank to actions such as realizing security (if held). This is regardless of the existence of any past-due amount or of the number of days past due. D3 or D4 are not applicable.
D3	A distressed sale or a distressed restructuring has occurred that likely results into a credit-related economic loss, for example involving remission, subordination or postponement of principal, interest or fee (re-)payments, and D4 is not applicable.
D4	The Client has filed for bankruptcy, or a similar order has been granted in respect of the Client.

Figure 2-3: Default rating scales and definitions (Rabobank, Practical Standards for Methodology Deliverables, 2022).

2.2.3.5 Rabobank Risk Rating (RRR)

In line with the guidelines outlined for the 'Internal Rating Based' (IRB) Advanced approach, as set out in the CRD /CRR, Rabobank applies a consistent rating methodology based on internally developed models. The rating system has an obligor rating scale which reflects exclusively quantification of the risk of obligor default.

The non-retail models calculate an obligor probability of default (PD) (with a one year horizon) which is classed to a rating scale for management information; Rabobank's Masterscale is described in the following section. Moreover, the model coverage is provided and the maximum ratings. For retail portfolios the PD's may be assigned at obligor level or at exposure/facility level3. Most rating models require quantitative and qualitative data input. By making use of the override process and group logic, additional information or expert judgement can be included in the rating. In principle, data has to be current, therefore the most recent annual figures available may not be older than 18 months. The rating itself has to be reassessed at least every twelve months for non-retail, extensions of this term are allowed by way of exception. In case outdated ratings are used in capital calculations, several requirements must be met.

For Corporates, Institutions, Central Governments and Central Banks (CGCB), the Masterscale provides twenty-one non-defaulted ratings (R00 – R20) and four default-grades (D01 – D04) as depicted in Figure 2-4 and Figure 2-5.

Obligors with the same rating or grade show a similar level of default risk, irrespective of other aspects such as the industry. Default risk increases as ratings move from R00 to R20 and finally to default (D01 to D04). Associated with each rating or grade (R00 to R20 or D01 to D04) is:

- A range of default probabilities (between lower bound PDL and upper bound PDU)
- A geometric average probability of default (PD)
- A qualitative description

Whilst PD retail pools can always be mapped towards a Rabobank Risk Rating, this is not a requirement. However for (regulatory) reporting, retail pools are mapped towards RRR. By using twenty-one performing ratings and four default grades, the Rabobank's Masterscale:

- Provides sufficient detail to comply with the minimum required seven grades for non-defaulted borrowers and one grade for defaulted borrowers (CRR, article 170, 1b)
- Has a meaningful population of all risk grades without excessive risk concentration in any particular rating (CRR, article 170, 1d)
- Covers the entire default risk spectrum of exposures in the bank

Multiple rebuttable and mandatory default triggers lead to one of the 4 default internal grades. The internal default grades are CRR compliant and the combined four internal default grades comprise the two CRR default indicators:

- 'Past due more than 90 days' in a material way (D01)
- 'Unlikely to pay' (D02 – D04)

The probability of default assigned to obligors/facilities is 100% (CRR article 160, 3).

The Masterscale comprises twenty-one non-defaulted ratings: R00 – R20. Default risk and the probability of default increase as ratings move from R00 to R20 as depicted in Figure 2-4. The twenty-one non-defaulted ratings are grouped into nine risk classes, each with a clear description of the associated degree of default risk. The nine non-defaulted classes range from 'Zero Risk' to 'Weak' and each class represents a distinct and clear range of default risk. Modifiers "+" or "-" can be added to distinguish the relevant individual ratings within each of the nine classes. As such the descriptions are sufficiently detailed to allow for a consistent assignment of obligors with the same risk to the same rating and to allow third parties to understand the ratings-assignment.

The Masterscale ratings (inn bps)

RRR	Geometric	Lowerbound**	Upperbound**	Description risk Classes
R00	0.0000	0.0000	0.0000	**ZERO** risk class for regulatory purposes
R01	1.4000	0.0000	1.6565	**EXTREMELY STRONG** capacity to meet financial commitments. Highest RRR
R02	1.9600	1.6565	2.3191	**VERY STRONG** capacity to meet financial commitments. Counterparties in these grades differ from the highest RRR counterparties only in a small degree
R03	2.7440	2.3191	3.2467	
R04	3.8416	3.2467	4.5454	
R05	5.3782	4.5454	6.3636	**STRONG** capacity to meet financial commitments but somewhat more susceptible to adverse effects of changes in circumstances and economic conditions than higher-rated counterparties.
R06	7.5295	6.3636	8.9091	
R07	10.5414	8.9091	12.4727	
R08	14.7579	12.4727	17.4618	**ADEQUATE** capacity to meet financial commitments. However, major adverse economic conditions or changing sector/ business specific circumstances could lead to a lessened capacity to meet financial commitments.
R09	21.7583	17.4618	27.1120	
R10	33.2053	27.1120	40.6680	
R11	49.8079	40.6680	61.0019	**SATISFACTORY** capacity to meet financial commitments. However, (major) adverse economic conditions or changing sector/ business specific circumstances could lead to a weakened capacity to meet its financial commitments.
R12	74.7118	61.0019	91.5029	
R13	112.0677	91.5029	137.2544	
R14	168.1016	137.2544	205.8816	**MODERATE** capacity to meet its financial commitments. Adverse economic conditions or changing sector/business specific conditions are more likely to weaken capacity to meet its financial commitments.
R15	252.1524	205.8816	308.8224	
R16	378.2286	308.8224	463.2335	
R17	567.3429	463.2335	694.8503	**VULNERABLE**, currently has the capacity to meet its financial obligation but faces major ongoing uncertainties that could impact the capacity to meet its financial commitments.
R18	851.0143	694.8503	1042.2754	
R19	1276.5215	1042.2754	1563.4132	
R20	1914.7823	1563.4132	9999.9999	**WEAK**, currently faces (material) risk of nonpayment and is dependent upon favorable business, financial and economic conditions to meet its financial commitments.

Figure 2-4: Masterscale ratings for R00 to R20 (Rabobank, Global Standard on FTP Pricing Methodology, 2018).

RRR	Geometric	Lowerbound**	Upperbound**	Description risk Classes
D01	10000	10000	10000	**According to the definitions in the Global Standard on Credit Risk Parameters**
D02	10000	10000	10000	
D03	10000	10000	10000	
D04	10000	10000	10000	

* Based on the geometric mean of lower bound and upper bound, except for R01 and R20. The geometric average of R1 is set to 1.4 bps (rounded to 1 decimal).
** The grades exclude the lowerbound and include the upperbound. The exceptions are R00 and the default ratings which include both the lower bounds and upper bounds.

Figure 2-5: Masterscale ratings for D01 to D04 (Rabobank, Global Standard on FTP Pricing Methodology, 2018).

2.2.3.6 Sustainability KPI

Sustainability is the quality of not being harmful to people, communities, environment or depleting natural resources, and thereby supporting long-term social and ecological balance (Rabobank, Rabobank, 2023).

Sustainability related risks and opportunities assessed in relation to the strategy and primary business activities of the client. These can, apart from being client-specific, also be related to the country or sector of the client. Rabobank risk assessment and control looks at the quality of clients' risk assessment and controls, i.e., their mechanisms to identify, prevent, mitigate, remedy and account for potentially adverse impacts (including their risk management of supply chains).

The sustainability KPI incorporates the likelihood of a client causing, contributing to or being linked to adverse social, environmental, or governance impacts. Material and salient sustainability risks are environmental, social and/or governmental risks that, if not properly mitigated:

- Have a negative impact on the environment
- Have a negative impact on the well-being of people or employees
- Have a negative impact on the clients (credit) risk profile caused by direct (e.g. fines) or indirect (e.g. loss of market share, loss of operating license) financial or reputational damage for the client
- Can cause financial or reputational damage to Rabobank

Sustainability performance is a measure of how advanced the sustainability approach is of a client based on its sustainability strategy, governance, transparency and reporting, supply chain management and its business and investment schemes. A clients relative sustainability performance (or client photo) indicates how advanced the sustainability approach of a client is relative to its peers. Global Standard Embedding Sustainability in Credit risk assessment for Wholesale and Rural clients v1.4 7 It identifies whether the client is a frontrunner (A), an average performer (B) or a laggard (C). Furthermore there is a separate client photo category which indicates that a client is policy noncompliant (D) or policy non-compliant with improvement plan (D+) versus Rabobank's Sustainability Policy Framework.

The sustainability assessment is comprised of an assessment of the counterparties' compliance with Rabobank's Sustainability Policy Framework, its sustainability performance (expressed as the numerical sustainability score) and its relative sustainability performance (expressed as the client photo ranging from D+ to A). In the current sustainability version, sustainability automatically translates the numerical sustainability score of Wholesale clients into a client photo based on fixed brackets (1 point: D/D+ (non-compliant), 10 – 20 points: C (Laggard), 21 – 30 points (B (Average performer), 31 – 40 points: A / Frontrunner). It does not take into consideration differentiating factors like company size, the sensitive sectors and region the company is active in and its position in the supply chain. Since the sustainability analysts do take these differentiating factors into consideration when providing an ISA on Potentially High Risk clients, his/her recommendation may differ from the outcome of the sustainability assessment. For trading companies with no investments in activities with large sustainability impacts (e.g. mining or downstream activities like refining or fertilizer production), a separate model has been developed to determine the sustainability performance and relative sustainability performance for this specific category of traders (the Pure Trader Model).

2.2.3.7 NAICS

The North American Industry Classification System (NAICS) is a standardized system used by various industries in North America to classify and categorize businesses into industry sectors based on their primary economic activity. The wholesale banking industry is one such industry that uses the NAICS system to classify its customers based on their economic activity. We aim to provide an overview of the NAICS system, its relevance to wholesale banking, and our research conducted on this topic.

The NAICS system was developed jointly by the statistical agencies of Canada, Mexico, and the United States to provide a common industry classification system to facilitate the collection, analysis, and dissemination of industry data among these countries. The system is updated every five years to reflect changes in the economy and is currently in its sixth revision. The NAICS system uses a hierarchical structure to classify businesses into industry sectors, subsectors, industry groups, and industries based on their economic activity.

Banking is one of the industries that use the NAICS system to classify its customers. Wholesale banks provide financial services to businesses, governments, and other financial institutions rather than individual consumers. Therefore, Rabobank amongst others use the NAICS system to classify their customers based on their primary economic activity to determine their creditworthiness and risk profile. This classification system helps banks to develop products and services that are specific to the needs of different industries and manage risk more effectively.

Research has been conducted to examine the relevance of the NAICS system to the wholesale banking industry. For instance, a study by Li (2018) examined the impact of industry classification on the credit risk of small and medium-sized enterprises (SMEs) in China. The study found that using the NAICS system to classify SMEs into different industries can help banks to better assess their credit risk and develop appropriate risk management strategies. Another study by Serrani-Cinca et al. (2017) examined the impact of the NAICS system on the performance of Spanish banks. The study found that using the NAICS system to classify customers can help banks to develop customized products and services, leading to better customer satisfaction and higher profitability.

The NAICS system is a standardized classification system used by various industries in North America, including wholesale banking but used globally (NAICS, 2022). The system helps wholesale banks to classify their customers based on their primary economic activity to determine their creditworthiness and manage risk more effectively. Research has shown that using the NAICS system to classify customers can help banks to better assess their credit risk, develop customized products and services, and improve their performance. Therefore, the NAICS system is an important tool for wholesale banks in managing risk and providing better services to their customers. Figure 2-6 depicts the depicts the main categories based on sector code in the NAICS.

NAICS

Definition	Sector
Agriculture, Forestry, Fishing and Hunting	11
Mining, Quarrying, and Oil and Gas Extraction	21
Utilities	22
Construction	23
Manufacturing	31
Manufacturing	32
Manufacturing	33
Wholesale Trade	42
Retail Trade	44
Retail Trade	45
Transportation and Warehousing	48
Transportation and Warehousing	49
Information	51
Finance and Insurance	52
Real Estate and Rental and Leasing	53
Professional, Scientific, and Technical Services	54
Management of Companies and Enterprises	55
Administrative and Support and Waste Management and Remediation Services	56
Educational Services	61
Health Care and Social Assistance	62
Arts, Entertainment, and Recreation	71
Accommodation and Food Services	72
Other Services (except Public Administration)	81
Public Administration	92
Unspecified (Rabobank)	99

Figure 2-6: NAICS sector codes.

3 Model

3.1 What benchmark techniques can be used to compare net interest income?

Firstly, we discuss suitable benchmark models. Here we will explain the differences between the models. Hereafter, one model will be chosen which best suits our research goals. Finally, restrictions and adjustments which are made to the model will be explained and presented to the reader.

3.1.1 What are suitable benchmark frameworks?

Machine learning (ML) and artificial intelligence (AI) are rapidly gaining popularity in the finance and wholesale banking sector. These technologies are being used to improve financial forecasting, time series analysis, and multiple linear regression. We briefly discuss the application of machine learning, artificial intelligence, financial forecasting, time series analysis, and multiple linear regression within finance and wholesale banking. Hereafter, an appropriate model is proposed in order to benchmark NII per region and per sector.

3.1.1.1 Financial forecasting

Financial forecasting is the process of predicting future financial performance based on historical data. ML and AI are being used to improve the accuracy of financial forecasts. These technologies are being used to analyze data and make predictions about market trends, stock prices, and other financial indicators.

Financial forecasting is an essential process in the finance and wholesale banking industries. It involves making predictions about future financial performance, which can help businesses make informed decisions about investments, budgeting, and other financial activities. To ensure the accuracy of financial forecasts, various forecasting methods and assumptions are used. Additionally, mathematical calculations are used to quantify the relationship between different financial variables and predict future outcomes.

Financial forecasting methods are divided into two categories: qualitative and quantitative. Qualitative methods are subjective and rely on expert opinions, surveys, and market research. Quantitative methods, on the other hand, use statistical models and mathematical calculations to analyze historical data and make future predictions. Some of the most common quantitative forecasting methods used in finance and wholesale banking include time series analysis, multiple linear regression, and exponential smoothing.

When using financial forecasting methods, we make certain assumptions to ensure the accuracy of predictions. Some of the common assumptions include the stability of relationships between different financial variables, independence of errors, and normal distribution of errors. It is crucial to test these assumptions to ensure that they are valid and reliable.

Mathematical calculations play a crucial role in financial forecasting. For instance, time series analysis involves calculating moving averages, seasonal indices, and trend coefficients to identify patterns and trends in historical data. Multiple linear regression involves calculating standardized beta coefficients and residuals to quantify the relationship between different financial variables. Exponential smoothing involves calculating weighted averages of historical data to predict future outcomes.

Financial forecasting is a critical process that helps businesses make informed decisions about investments and financial activities. Various forecasting methods, assumptions, and mathematical calculations are used to ensure the accuracy of financial forecasts. The use of quantitative methods such as time series analysis, multiple linear regression, and exponential smoothing is prevalent in the finance and wholesale banking industries.

3.1.1.2 Time series

Time series analysis is a statistical technique that is used to analyze data over time. This technique is widely used in finance and wholesale banking for forecasting and risk management. ML and AI are being used to improve the accuracy of time series analysis by analyzing large data sets and identifying patterns.

Forecasting is an important part of financial decision-making, allowing organizations to anticipate future trends and plan accordingly Hyndman et al. (2018). Time series methods have proven to be effective in financial forecasting, as they consider historical data to make predictions about future outcomes. However, there are important assumptions that need to be made when using time series for forecasting, and it is essential to understand how to interpret the results obtained from these methods.

Time series methods for financial forecasting include univariate methods such as ARIMA (AutoRegressive Integrated Moving Average) and exponential smoothing, as well as multivariate methods such as Vector Autoregression (VAR) and Vector Error Correction Model (VECM) Shumway et al. (2017). ARIMA models consider the autocorrelation and seasonality in the data, while exponential smoothing considers the trend and level of the data. VAR and VECM models consider the relationships between multiple time series and can be useful in capturing the interdependence of financial variables.

It is important to make certain assumptions when using time series methods for financial forecasting. For example, we assume that the data are stationary, meaning that its statistical properties do not change over time. Thus, one may assume the mean, variance and autocorrelation structure are constant over time. This assumption is important because time series methods work best on stationary data. In addition, we assume that the relationships between variables do not change over time, which can be tested using cointegration analysis which checks if there is a vector of coefficients to form a stationary linear combination.

Interpretation of results from time series methods is also important. Forecast accuracy measures such as mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE) are commonly used to evaluate the performance of the models Banerjee et al. (1993). Additionally, residual analysis can be used to check for any patterns in the errors, which may indicate that the model needs to be refined.

In conclusion, time series methods have proven to be effective in financial forecasting, and there are several methods available for use, including univariate and multivariate methods. However, it is important to make certain assumptions and to understand how to interpret the results obtained from these methods. The accuracy of the forecasts can be evaluated using various measures, and residual analysis can be used to identify any patterns in the errors. Further research is needed to develop more advanced time series methods for financial forecasting, as well as to improve our understanding of the assumptions and interpretation of results.

3.1.1.3 Machine learning and artificial intelligence

Machine learning (ML) and artificial intelligence (AI) are being used in finance and wholesale banking to automate processes and improve efficiency. These technologies are being used to analyze large data sets and make predictions.

Machine learning is a subfield of artificial intelligence that deals with the design and development of algorithms that can learn patterns and relationships in data without being explicitly programmed. Supervised learning is one of the main categories of machine learning where a model is trained to make predictions based on labelled data. In supervised learning, the model is trained on a labelled dataset, which consists of input features and their corresponding output labels. The model then uses this training data to make predictions on new, unseen data Hastie et al. (2009).

Once a machine learning model has been trained, it's important to evaluate its performance to ensure that it's working as expected. Some common metrics for evaluating the performance of a supervised learning model include accuracy, precision, recall, and F1-score. In addition, confusion matrices, ROC curves, and precision-recall curves can provide additional insight into the model's performance Shalev et al. (2014).

In conclusion, supervised learning is a crucial part of machine learning that involves training a model on labelled data to make predictions on new, unseen data. Understanding the important assumptions and interpretation of results is crucial for developing accurate and effective machine learning models. Despite its popularity, supervised learning is not without its limitations and challenges, and it's important to keep in mind that machine learning models are only as good as the data they are trained on. We will elaborate further on this in Section 0.

3.1.1.4 Multiple Linear Regression

Multiple linear regression (MLR) is a widely used statistical technique in finance and other fields to analyze the relationship between two or more independent variables and a dependent variable. ML and AI are being used to improve the accuracy of multiple linear regression by analyzing large data sets and identifying patterns. The use of MLR assumes certain assumptions to be met, and the resulting output includes standardized and unstandardized coefficients, standardized and unstandardized residuals, predicted values, R, R-squared, adjusted R-squared, and other important components.

The multiple linear regression model is based on several important assumptions, which must be met for the results of the analysis to be valid. These assumptions are similar to ML and include linearity, independence of observations, normality of residuals, homoscedasticity, and absence of multicollinearity James et al. (2013).

- Linearity assumes that the relationship between the dependent variable and the independent variables is linear. This means that the change in the dependent variable is proportional to the change in the independent variable. If the relationship is not linear, nonlinear transformations of the independent variables should be used to make the relationship linear.
- Independence of observations assumes that the observations in the dataset are independent and not influenced by each other. This is important because it affects the accuracy of the regression coefficients.
- Normality of residuals assumes that the residuals, which are the differences between the observed and predicted values, are normally distributed. This is important because it affects the validity of hypothesis tests and the interpretation of confidence intervals.

- Homoscedasticity assumes that the variance of the residuals is constant for all values of the independent variables. This is important because it affects the accuracy of the estimation of the standard errors of the regression coefficients and the hypothesis tests.
- Absence of multicollinearity assumes that the independent variables are not highly correlated with each other. This is important because it affects the accuracy of the estimation of the regression coefficients and the interpretation of the results.

It is important to check that these assumptions are met before interpreting the results of a MLR. If the assumptions are not met, the results may be biased or unreliable. Nevertheless, depending on which assumptions are violated, there are solution which can be applied to the tested dataset. We will elaborate more on these techniques in section 0. Tools such as SPSS, R, Excel, Python, SAS or STATA may be used to apply MLR and test the underlying assumptions.

Interpreting the results of a multiple linear regression analysis involves understanding the coefficient estimates, the overall fit of the model, and the assumptions that were made. The coefficient estimates represent the change in the dependent variable for a one-unit change in the independent variable, holding all other independent variables constant. The overall fit of the model is evaluated by measures such as R-squared, adjusted R-squared, and the residuals plot Hair et al. (2014). It is important to examine the residuals plot to check for the presence of patterns, which could indicate the need for transformations of the independent variables or the presence of outliers. The abovementioned is done in Section 3.1.2.2.

Standardized coefficients (beta coefficients) are a measure of the relative importance of each independent variable in the model. They indicate how much the dependent variable changes in standard deviations when an independent variable changes by one standard deviation, while holding other independent variables constant. The beta coefficients of the independent variables can be used to determine the strength and direction of the relationship between each independent variable and the dependent variable. Moreover, they can be used to analysis different MLR results to each other. Unstandardized coefficients (b coefficients) indicate the change in the dependent variable for a one-unit change in the independent variable.

Standardized residuals are the differences between the actual values of the dependent variable and the predicted values from the MLR model, divided by the standard deviation of the residuals. They are a measure of how well the model fits the data, with smaller residuals indicating a better fit. Unstandardized residuals are the raw differences between the actual and predicted values.

Predicted values are the values of the dependent variable predicted by the MLR model for a given set of values of the independent variables. The dependent variable is the variable we aim to explain or predicts where all other variables are independent variables.

R (correlation coefficient) measures the strength of the linear relationship between the independent variables and the dependent variable. R-squared (coefficient of determination) provides a measure of the proportion of the variance in the dependent variable that is explained by the independent variables in the model. Adjusted R-squared adjusts R-squared for the number of independent variables in the model. The latter gives more conservative prediction on the percentage that the model explains.

Other important components of MLR include the F-test for overall significance of the model, the t-tests for individual coefficients, the Durbin-Watson statistic for detecting autocorrelation in the errors, and the VIF (variance inflation factor) for detecting multicollinearity among the independent variables.

As explained previously, MLR can be used to explain the relationship between variables and/or be used to make predictions in the future. If one would like to use the output of MLR to make predictions about the future, one would use the unstandardized beta coefficient to multiply with its independent variable to predict the dependent variable. Using the unstandardized beta coefficient, one could make a formula as follows:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_1 x_{i2} + .. + \beta_p x_{ip} + \epsilon$$

Where, for $i = n$ observations:

$y_i = dependent\ variable$

$x_i = independent\ (explanatory)\ variables$

$\beta_0 = y\ intercept\ (constant\ term)$

$\beta_h = slope\ coefficients\ for\ eah\ explanatory\ variable \qquad h = 1, \ldots., p$

$\epsilon = the\ model's\ error\ term\ (residuals)$

In conclusion, MLR is a statistical technique for analyzing the relationship between two or more independent variables and a dependent variable. However, certain assumptions need to be met, including linearity, normality, constant variance of the errors, and independence of the errors. MLR also produces important output measures such as standardized and unstandardized coefficients, standardized and unstandardized residuals, predicted values, R, R-squared, and adjusted R-squared, which help to interpret the results and assess the goodness of fit of the model. Understanding the mathematical calculations behind each component is crucial for proper interpretation of the output.

3.1.2 What is an appropriate model when benchmarking NII?

We elaborate which model is best suited in order to find underlying relationships between variables that have an effect on the NII bps. We will first discuss why ML is not suited for our research and the pitfalls when using it. Secondly, we will elaborate on which model is best suited and how we have applied it to our case. Here, we will also discuss adjustments made to our model.

3.1.2.1 Model to be applied

Machine learning (ML) is a tool that has been widely adopted in many industries, including finance and wholesale banking. However, there are also concerns and debates regarding the use of ML in these sectors.

One of the key issues with using machine learning in finance and wholesale banking is that the models can be complex and difficult to interpret. This can make it challenging to understand how the model arrived at its predictions and decisions. The lack of transparency can lead to distrust from stakeholders, regulators, and customers (Merkle, 2019).

Secondly, ML models can be prone to overfitting, where the models perform well on the training data but perform poorly on new, unseen data. This is a significant problem in finance because poor predictions or decisions can lead to significant financial losses.

Moreover, ML models require large amounts of high-quality data to perform well. However, in banking, data can be sparse, incomplete, or of low quality Kim et al. (2019). This can result in inaccurate predictions or decisions that can lead to significant financial losses.

Finally, ML models can be sensitive to changes in the data or the environment, which can make them unstable. This can be a significant concern because the model may not perform well in new or changing market conditions Zhang et al. (2019).

Machine learning is a tool that has been widely adopted in many industries. However, in finance and wholesale banking, there are concerns regarding the use of ML due to issues of transparency and interpretability, overfitting, data quality, and model stability. While ML can be beneficial in some cases, it is important to carefully evaluate the risks and benefits before adopting these technologies in these sectors. Due to the high importance of transparency and regulation within banking, we have opted to not apply a supervised ML model to our internal benchmark.

Thus, which model would suit our defined problem? As we are interested in being able to explain the relationships between variables, financial forecasting would not suit our benchmark model. Time series analysis could help us in our final benchmark model however it could not be used on its own to explain the complex relationships we aim to analyze. Nevertheless, MLR is a statistical technique that allows the relationship between multiple variables to be analyzed. In finance, MLR can be a valuable tool for analyzing variables and predicting outcomes.

MLR allows the relationships between multiple variables to be analyzed, which can help to identify how variables are related and which variables are most important for our outcomes. This can be valuable for analyzing variables such as interest rates, inflation, and exchange rates and thus exactly what we aim for.

Moreover, MLR can be used to predict outcomes based on the relationship between multiple variables. In finance and wholesale banking, this can be valuable for predicting outcomes such as stock prices, loan defaults, and credit ratings Kozak et al. (2014). Note, that we would not want MLR to be used solely to predict the future. Nevertheless, the predictive power could be used for the Control Wholesale business to test their assumptions concerning a company, region, sector or any other variable.

Additionally, MLR can be used to test hypotheses about the relationship between variables. In finance, this can be valuable for testing hypotheses about the effects of policy changes, market trends, and other factors on outcomes such as economic growth, inflation, and interest rates Li & Peng (2017). Here, non-standardized beta coefficient should be used to form the formula of MLR. Nevertheless, for our research we will only use standardized beta coefficients in order to analyze the different positive or negative weights of independent variables (IBM, 2016). Using the standardized beta coefficients, we can compare the results of different MLR's and results of independent variables.

MLR is a flexible technique that can be adapted to a wide range of applications within finance. It can be used to analyze relationships between variables at different levels of granularity, including individual securities, portfolios, and macroeconomic variables.

Thus, our chosen model will be a MLR as we deem it better suited for our problem compared to ML. ML can be complex and difficult to interpret which can lead to distrust of stakeholders and regulators. Moreover, ML is prone to overfitting which could lead to significant financial loses. On the contrary, MLR is a valuable tool for analyzing variables and predicting outcomes in finance and wholesale banking. It allows relationships between multiple variables to be analyzed, which can help to identify how variables are related and which variables are most important for predicting outcomes. It can also be used to predict outcomes, test hypotheses, and is a flexible technique that can be adapted to a wide range of applications in these sectors. We have opted to use the tool SPSS to execute our MLR as it can analyze big dataset, provide methods to test for violation of assumption and is quick in processing output results.

3.1.2.2 Proposed model and important notes and changes

Rabobank has provided access to their database from which a dataset is retrieved. The first dataset retrieved contained monthly transactions within our scope as discussed in section 1.4. Thus, the dataset contained important variables on a monthly basis and on contract level dating back until 2018. The variables here were month and year of the transaction, NII bps, location, region, sector level 01-03, currency, nominal, outstanding, EAD RC, expected loss RC, Regulatory Capital Credit Risk, LAD, Net interest income, other net interest income, Sustainability KPI, GII bps and FTP bps.

While setting the dependent variable as NII bps and using the others as independent variables, we quickly violated one or more of the underlying assumptions of MLR. We found that the variables nominal, EAD RC, expected loss RC, Regulatory Capital Credit Risk, Net interest income, other net interest income and GII bps violated the assumptions of heteroscedasticity, multicollinearity or provided an adjusted R-squared of 1 which is impossible as it would imply that we could predict all the variation of the NII bps. Moreover, while testing the remaining variables, we found that time was having correlation with our results.

Thus, we have applied a time series analysis to determine the size of the dataset. Here we found that using a dataset of the last 2 years (2021 and 2022), time doesn't play a significant role in the NII bps. As is depicted in Figure 3-1, we see that covid has had major impact on the NII bps after which some months are taken into account for the NII bps to stabilize. Here the axis are hidden due to confidentiality however it depicts the NII bps over time.

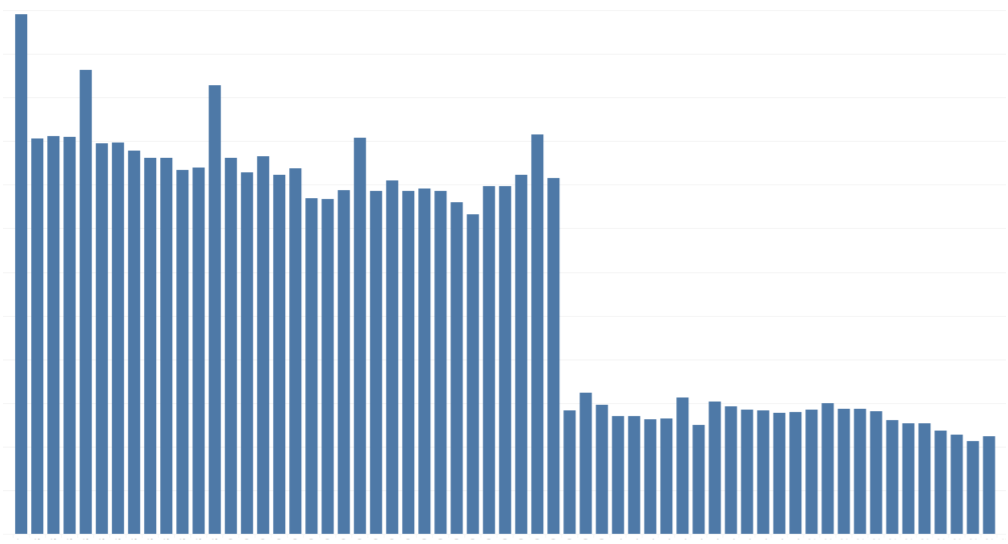


Figure 3-1: NII bps over time.

Moreover, we have expanded our dataset with other variables such as Rabobank Risk Rating, Probability of default and NAICS as these were deemed of importance by business. It must be noted that some variables which do not have scales are coded as dummy variables. This is the case for the following variables: Region, location, RiskRatingGrouped, FacilityCurrency (grouped). A grouped variable imply it can take on all the possible values within that variable. Ungrouped variables is where the possible values, e.g. each location, will become its own variable. We will eventually ungroup the dummy variables and add them as scales to the model. This will give us better use cases for our sensitivity analysis to do further research. A list of the variables found within the dataset and the ones used in SPSS can be found in the table below.

Columns	Meaning and usage
Month	Month (only used for time series analysis)
Contract	Contract corresponding to entity
Contract start date	Start date of contract
Contract end date	End date of contract
Region	Region in which company is situated
Location	Location in which company is situated
Rabobank sector level 01 to 03	Rabobank sector specification
Risk Rating	Rabobank's Risk Rating
Risk Rating PD Average	Probability of default per risk rating
Naics Code (ranging from 2 numbers to 6)	North American Industry Classification System where 2 and 3 numbered NAICS are used
Oustanding	Oustanding amount
EAD RC	EAD RC amount
LAD	Loss at default amount
Sustainability KPI	Sustainability KPI

FTP bps	Fund Transfer Price in basepoints
NII bps	Net Interest Income in basepoints
Location grouped	Locations auto recorded with SPSS as depicted in Section 8.1
Risk Rating grouped	Risk ratings auto recorded with SPSS as depicted in Section 8.1
Currency grouped	Currencies auto recorded with SPSS as depicted in Section 8.1
Location ungrouped	Each location will become its own variable taking on values 1 or 0
Risk Rating ungrouped	Each risk rating will become its own variable taking on values 1 or 0
Currency ungrouped	Each currency will become its own variable taking on values 1 or 0

It must be noted that for grouped variables, the use of auto recoding in SPSS results in categorial variables to be recoded in numerical values in alphabetical order. We have tested other ways to recode these grouped variables with more logic, e.g. according to population size, currency index and so forth, however no noticeable changes were found in the standardized beta coefficients and adjusted R-squared. Nonetheless, the grouped variables will only be used in the first iteration of the MLR after which each variable will be ungrouped in their respectively iteration to see the true influence of a certain risk rating, location or currency.

We have also looked into adding the contract length variable as some contract work with several months of EURIBOR. EURIBOR contracts have floating interest rates agreement and thus different CoF. Nevertheless, we weren't able to add the above as we were unable to align both databases.

Now that we have decided which variables to use, we have added all the datapoints into SPSS. We must note that we have already applied some filtering within excel to remove rows where blank cells are found within the variable columns. Nevertheless, to ensure that each variables (columns) has data, we add an additional filtering within SPSS. Here we also want to get rid of rows containing variables which are defined as "Unspecified". Finally, we will filter out NII bps values bigger then 1000 as these are deemed outliers. The above to mentioned points is achieved by applying the following formula in SPSS:

$$NIIbps$$
$$< 1000\ AND\ NMISS(SpRiskRatingPDAverage, Outstanding, EADRC, LAD, SustainabilityKPI, AvgWeightedP$$
$$< 0.00001\ AND\ Unspecified = 0\ AND\ UnspecifiedCurrency = 0$$

The abovementioned formula ensures that NII bps may not be bigger then 1000, all the used variables may not have a number missing, and no variable may be defined as "Unspecified".

We are now able to run the first iteration of the MLR and investigate if all the assumptions are once again met. The above formula reduces our dataset from 175.000 datapoints to 132.611. The results of the first iteration are as follows:

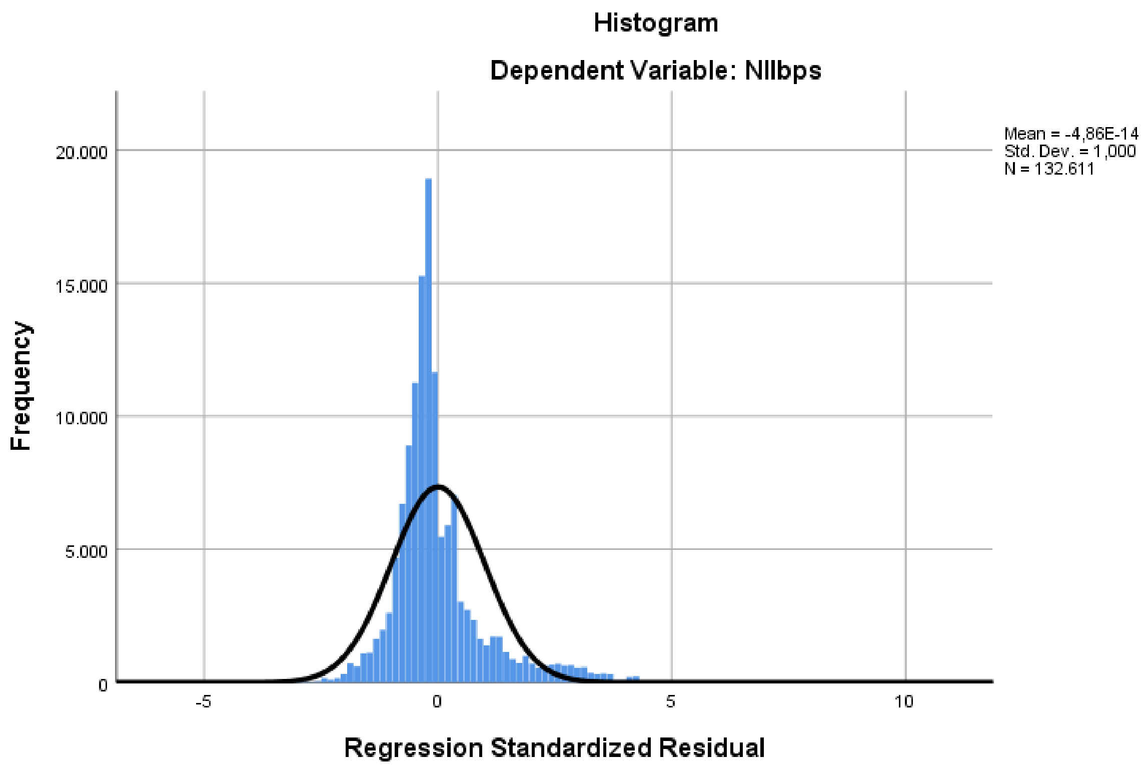


Figure 3-2: Normality plot.

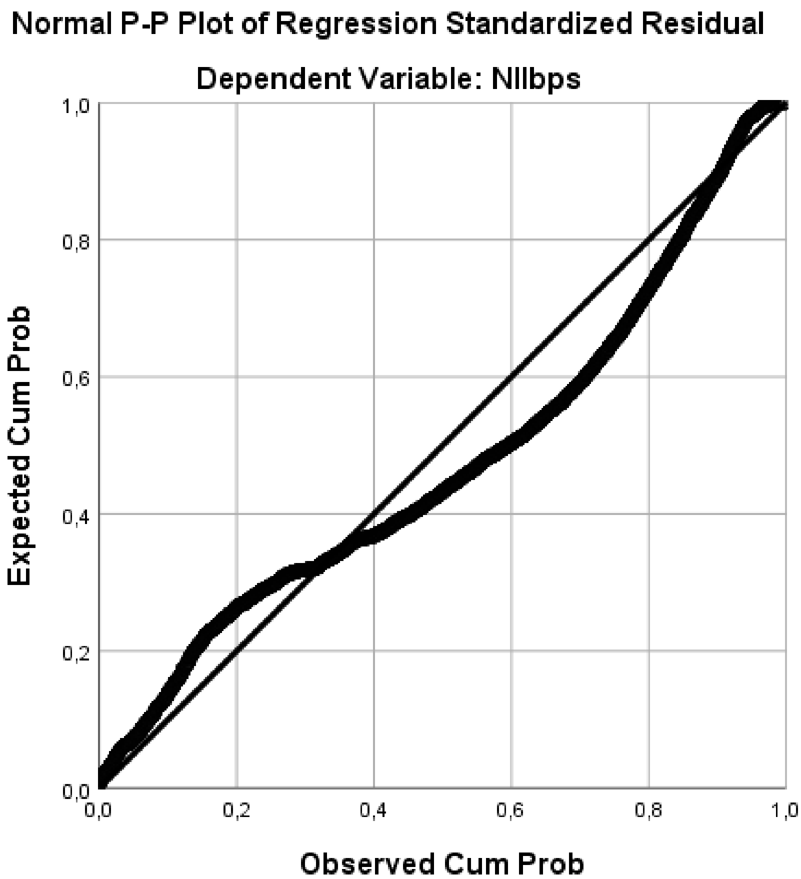


Figure 3-3: Linearity plot.

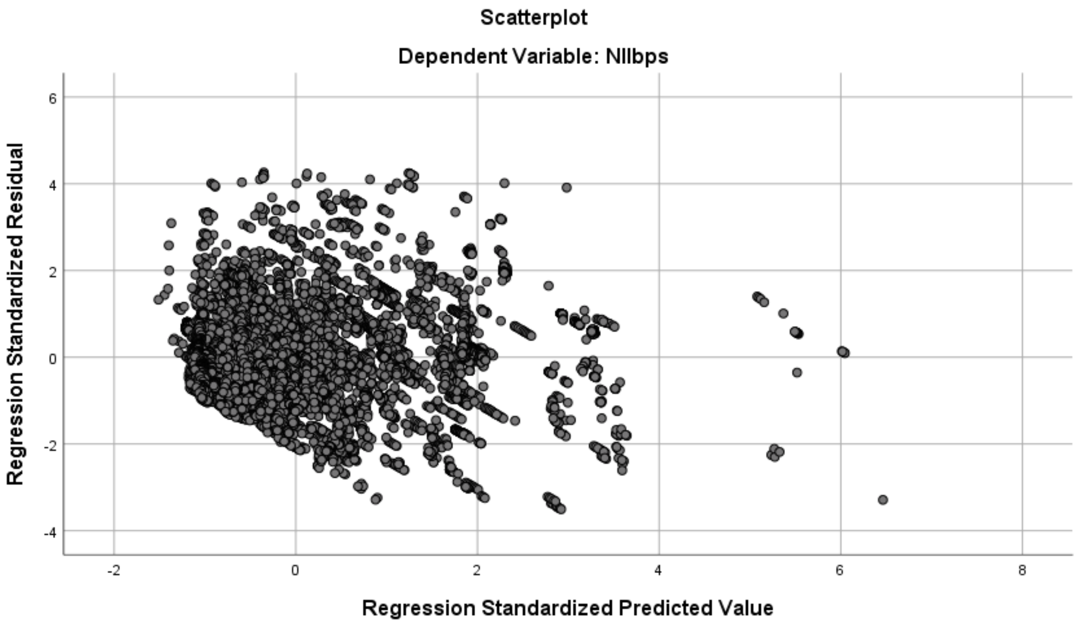


Figure 3-4:Homoscedasticity plot standard residuals versus predicted value.

The standardized residual is found by dividing the difference of the observed and expected values by the square root of the expected value. It is a measure of the strength of the difference between observed and expected values.
Standardized predicted values is a transformation of each predicted value into its standardized form. That is, the mean predicted value is subtracted from the predicted value, and the difference is divided by the standard deviation of the predicted values. Standardized predicted values have a mean of 0 and a standard deviation of 1 (IBM, 2016).

The formula of the first iteration is as follows:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_1 x_{i2} + .. + \beta_p x_{ip} + \epsilon$$

Where, for $i = n$ observations:

$y_i = dependent\ variable$
$x_i = independent\ (explanatory)\ variables$
$\beta_0 = y\ intercept\ (constant\ term)$
$\beta_h = slope\ coefficients\ for\ each\ explanatory\ variable \qquad h = 1, \ldots., p$
$\epsilon = the\ model's\ error\ term\ (residuals)$

And where the x_i's are: Sustainability KPI, Risk Rating PD average, NAICS (two digits), Risk rating grouped, Outstanding, Location, LAD, FTP bps, Facility Currency.

Normality violation will affect the estimates of the standard error (SE) and the confidence interval (CI), and hence the significance of the risk factors. To provide robust estimates of SE, bootstrap techniques or nonparametric regression model can be performed. Nevertheless, these techniques require large sample sizes and are very sensitive to outliers (IOVS, 2019). It is more accurate to check whether the errors in a MLR are normally distributed or the dependent variable has a conditional normal distribution.

By the law of large numbers and the Central Limit Theorem (CLT), the ordinary least squares (OLS) estimators in MLR may assumed to be approximately normal distributed around the true parameters values, thus the CI and estimated parameters remains robust. As our sample size consists of more than 160.000 datapoints, we may assume that Figure 3-3 is still normally distributed as the sample size is large enough, CLT kicks in and justifies that the OLS estimators are well approximated by a multivariate normal distribution regardless of the distributions of the error terms (Zhu, 2022). Moreover, when applying a Jarque Berra test, we find a p-value 0.3921 and thus implying we do not have sufficient evidence to say the dataset is not normally distributed.

Multicollinearity assumptions are met if the VIF values lies between 1-10, thus then there is no multicollinearity. However, if the VIF values are <1 or >10, then there is some form of multicollinearity and thus the assumption is violated. Nevertheless, the above does not hold if one uses dummy variables. In case of more than 3 dummy variables, significance (p) and high level of VIFs do not imply that the multicollinearity assumption is violated (Allison, 2012). As our model uses 40+ dummy variables, some higher amount of VIFs are sometimes detected where there is no true multicollinearity.

For the used dataset, we have tested the dependent variable NII bps versus the standard residuals in order to ensure that the homoscedasticity assumption is met. As the dataset is very large, we cannot easily check this assumption by looking at the residual graph as depicted in Figure 3-4. Figure 3-4 depicts the standard residuals versus the standard predicted values. Moreover, one may spot some linear transformation however these are caused by the coded grouped dummy variables. However, we can test de dependent variable with the standard residuals and figure out if the gamma factor is larger than 0.5. If this is true, it implies our dependent variable contains some form of heteroskedasticity James & Knaub (2019). From the data, we have found that the gamma is near 0 as depicted in Figure 3-6 thus implying the assumption of homoscedasticity is not violated and thus we may apply MLR. Figure 3-6 depicts the absolute difference between NII bps (Y) and standard residuals (Y'), namely $|e| = Y - Y'$.

However, we see some sort of deviation in the test data, as depicted in Figure 3-5. Nonetheless, this can be attributed to time series component. As time progresses, variance of the standard residuals may increase due to unforeseen changes in population size, number of open contracts and other factors. Therefore, we have decided only to look at the last 2 years to decreases the influence of time series and to remove the anomaly/ black swan event caused by the covid crash. A time series analysis in SPSS didn't discover any significance influence of time on the NII bps for the last two years. Thus, it is sound to use the dataset of 2021 and 2022.

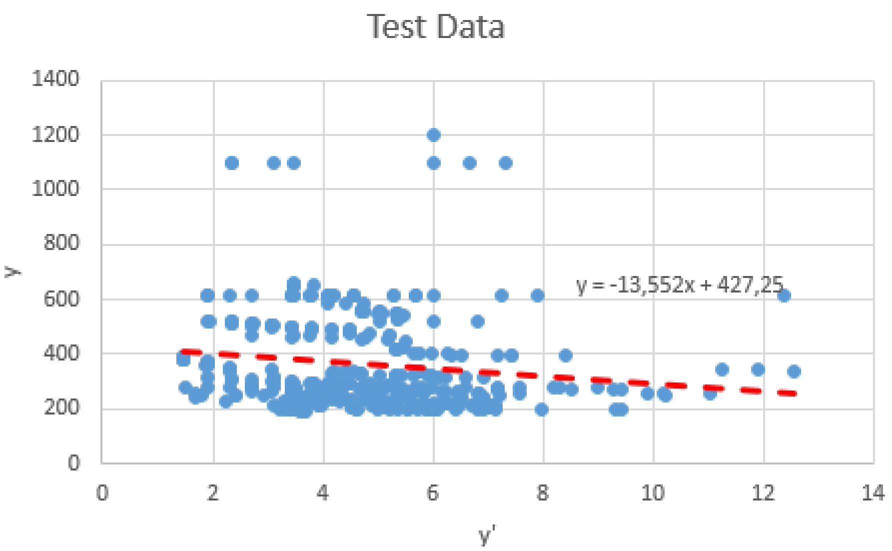


Figure 3-5: NII bps (Y) versus standard residuals (Y').

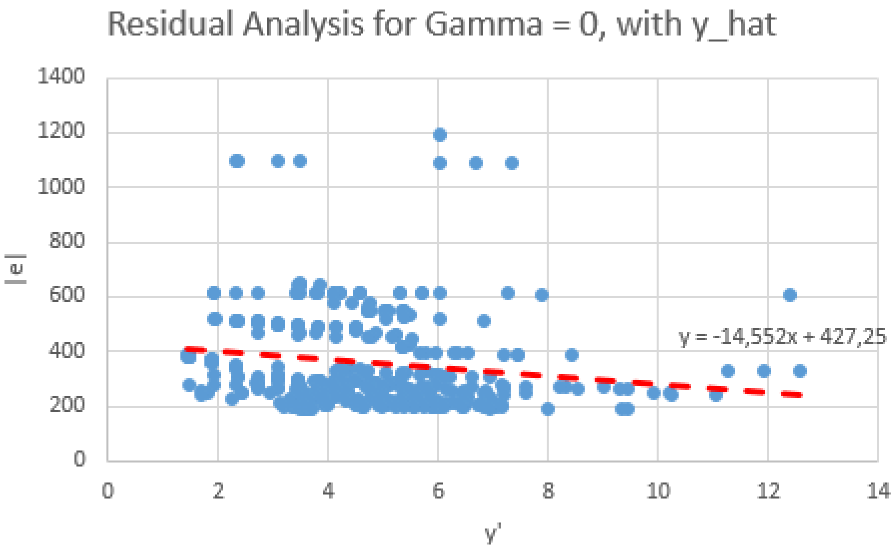


Figure 3-6: Absolute difference between NII bps and standard residuals versus standard residuals.

As described earlier, the OLS estimators are well approximated by a multivariate normal distribution regardless of the distributions of the error terms. It must be noted that we will check the violation of the assumptions after each iteration during the research. Moreover, our model uses 40+ dummy variables, some higher amount of VIFs are sometimes detected where there is no true multicollinearity. Finally, from the data, we have found that the gamma is near 0 as depicted in **Error! Reference source not found.** thus implying the assumption of homoscedasticity is not v iolated and thus we may apply MLR. Thus, all the assumptions of the MLR model are met and thus we may continue applying our model.

For any violations, appropriate action will be taking by removing the corresponding variables. Moreover, it is of importance though to filter out more outliers in order to finetune our MLR output results. Outliers are filtered out when the they have a standard residual value outside of the boundaries of x such as:

$$-3{,}29 < x > 3{,}29 \qquad where\ x = standard\ residual$$

This can be achieved by the following formula in which we change the standardized residuals.

$$NIIbps < 1000\ AND\ NMISS(SpRiskRatingPDAverage, Outstanding, EADRC, LAD, SustainabilityKPI, AvgWeightedP$$
$$< 0.00001\ AND\ Unspecified = 0\ AND\ UnspecifiedCurrency = 0\ AND\ ZRE\_1\ <= 3.29\ AND\ ZRE\_1\ >= -3.29$$

The abovementioned formula ensures that NII bps may not be bigger then 1000, all the used variables may not have a number missing, and no variable may be defined as "Unspecified". Moreover, the standard residuals smaller then -3,29 and bigger then 3,29 are filtered out.

When applying the model per region, we will use the following formula in order to filter them to their respectively region.

$$NIIbps < 1000\ AND\ NMISS(SpRiskRatingPDAverage, Outstanding, EADRC, LAD, SustainabilityKPI, AvgWeightedP$$
$$< 0.00001\ AND\ Unspecified = 0\ AND\ UnspecifiedCurrency = 0\ AND\ Region = 1\ AND\ ZRE\_1\ <= 3.29\ AND\ ZRE\_1\ >= -3.29$$

The abovementioned formula ensures that NII bps may not be bigger then 1000, all the used variables may not have a number missing, and no variable may be defined as “Unspecified”. Moreover, the standard residuals smaller then -3,29 and bigger then 3,29 are filtered out. Finally, the selected region in the formula is equal to 1 which corresponds to Asia.

In Section 8.1, one can find the grouped dummy coded variables. Concerning the dummy coding of the ordinal scale variables, Figure 3-7 depicts a section of how they will look like. Each currency, risk rating and location will be split up and become their own variable with values 0 or 1. Thus a new column will be added for a variable, e.g. R1, where it will take on a value of 1 if the risk rating is equal to R1 and elsewhere it will take on the value 0.

Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
Sector01	Numeric	2	0		{1, 1.1 - Gra...	None	10	Right	Nominal	Input
FacilityCurr...	Numeric	2	0		{1, AUD}...	None	19	Right	Nominal	Input
R1	Numeric	8	2		None	None	10	Right	Nominal	Input
R2	Numeric	8	2		None	None	10	Right	Nominal	Input
R3	Numeric	8	2		None	None	10	Right	Nominal	Input
R4	Numeric	8	2		None	None	10	Right	Nominal	Input
R5	Numeric	8	2		None	None	10	Right	Nominal	Input
R6	Numeric	8	2		None	None	10	Right	Nominal	Input
R7	Numeric	8	2		None	None	10	Right	Nominal	Input
R8	Numeric	8	2		None	None	10	Right	Nominal	Input
R9	Numeric	8	2		None	None	10	Right	Nominal	Input
R10	Numeric	8	2		None	None	10	Right	Nominal	Input
R11	Numeric	8	2		None	None	10	Right	Nominal	Input
R12	Numeric	8	2		None	None	10	Right	Nominal	Input
R13	Numeric	8	2		None	None	10	Right	Nominal	Input
R14	Numeric	8	2		None	None	10	Right	Nominal	Input
R15	Numeric	8	2		None	None	10	Right	Nominal	Input
R16	Numeric	8	2		None	None	10	Right	Nominal	Input
R17	Numeric	8	2		None	None	10	Right	Nominal	Input
R18	Numeric	8	2		None	None	10	Right	Nominal	Input
R19	Numeric	8	2		None	None	10	Right	Nominal	Input
R20	Numeric	8	2		None	None	10	Right	Nominal	Input
D1	Numeric	8	2		None	None	10	Right	Nominal	Input
D2	Numeric	8	2		None	None	10	Right	Nominal	Input
D3	Numeric	8	2		None	None	10	Right	Nominal	Input
D4	Numeric	8	2		None	None	10	Right	Nominal	Input
Unspecified	Numeric	8	2		None	None	13	Right	Nominal	Input
filter_$	Numeric	1	0	NIIbps<1000 A...	{0, Not Sele...	None	10	Right	Nominal	Input
AUD	Numeric	8	2		None	None	10	Right	Nominal	Input
CAD	Numeric	8	2		None	None	10	Right	Nominal	Input
DKK	Numeric	8	2		None	None	10	Right	Nominal	Input
EUR	Numeric	8	2		None	None	10	Right	Nominal	Input
GBP	Numeric	8	2		None	None	10	Right	Nominal	Input
HKD	Numeric	8	2		None	None	10	Right	Nominal	Input
HUF	Numeric	8	2		None	None	10	Right	Nominal	Input
INR	Numeric	8	2		None	None	10	Right	Nominal	Input
NOK	Numeric	8	2		None	None	10	Right	Nominal	Input
NZD	Numeric	8	2		None	None	10	Right	Nominal	Input
PLN	Numeric	8	2		None	None	10	Right	Nominal	Input

Figure 3-7: A selection of dummy variables MLR

The formula of the second iteration is as follows:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_1 x_{i2} + .. + \beta_p x_{ip} + \epsilon$$

Where, for $i = n$ observations:

$y_i = dependent\ variable$

$x_i = independent\ (explanatory)\ variables$

$\beta_0 = y\ intercept\ (constant\ term)$

$\beta_h = slope\ coefficients\ for\ each\ explanatory\ variable \qquad h = 1, \ldots, p$

$\epsilon = the\ model's\ error\ term\ (residuals)$

And where the x_i's are: Sustainability KPI, Risk Rating PD average, NAICS (three numbered), Risk rating grouped, Outstanding, Location, LAD, FTP bps, Facility Currency.

All hereafter iteration's will use the dummy coded variables to find the true influence of the categorial variables. Here the x_i's are:
Sustainability KPI, Risk Rating PD average, two number NAICS (11, 21, 22, 23, 31, 32, 33, 42, 44, 45, 48, 49, 51, 52, 53, 54, 55, 56, 71, 72, 92), R2, R3, R4, R5, R6, R7, R8, R9, R10, R11, R12, R13, R14, R15, R16, R17, R18, R19, R20, D2, D3, D4, Outstanding, Atlanta, Singapore, Brazil, India, Germany, France, Turkey, Hong Kong, Spain, Mexico, Shanghai, Chicago, Canada, Ireland, Australia, Belgium, New Zealand, Italy, London, Argentina, Chile, Kenya, New York, LAD, FTP bps, USD, INR, RMB, MXN, AUD, HKD, GBP, CAD, NZD, SEK, NOK, HUF, PLN, DKK.

4 Results

4.1 Benchmark results

Firstly, we dive into the main results generated by the benchmark model which is applied on the entire dataset for all regions. The benchmark model serves as a standard to which the different regions can be analyzed. We will take a look at the results when all variables are grouped together after which we will further investigate the cause of some standardized beta coefficients. Moving forward, we will refer to standardized beta coefficient as beta coefficient as we do not look into the non-standardized beta coefficients. We only use standardized beta coefficients in order to analyze the different positive or negative weights of independent variables (IBM, 2016). Using the standardized beta coefficients, we can compare the results of different MLR's and results of independent variables.

Moreover, when using the word influence we are talking about correlation between variables and not in the sense of causality. Finally, grouped variables are variables containing all the different possibilities, e.g. location grouped is one variable containing all the possible locations. Whereas ungrouped location would be one variable for one specific location.

4.1.1 All variables grouped

As depicted in Figure 4-1, we see the biggest variable of influence in the sense of correlation is the "Average risk rating Probability of Default" with a positive beta coefficient. Evidently, one would expect to see this, as a higher PD implies a higher risk thus a higher positive return on NII bps. On the contrary, we see that the risk rating grouped have a strong negative correlation with the NII bps. A strong influence is a standardized beta coefficient higher then 0.2, a moderate influence a value of 0.1, and low to no influence a value of 0 to 0.05 in absolute term. Further investigating is needed into the buckets and transaction counts in the different risk ratings. As depicted by Figure 4-2, we see that most companies within the transactions dataset constitute of medium to lower risk rating. In 4.1.2, we will analyze more in depth the standardized beta coefficients of the individual risk rating.

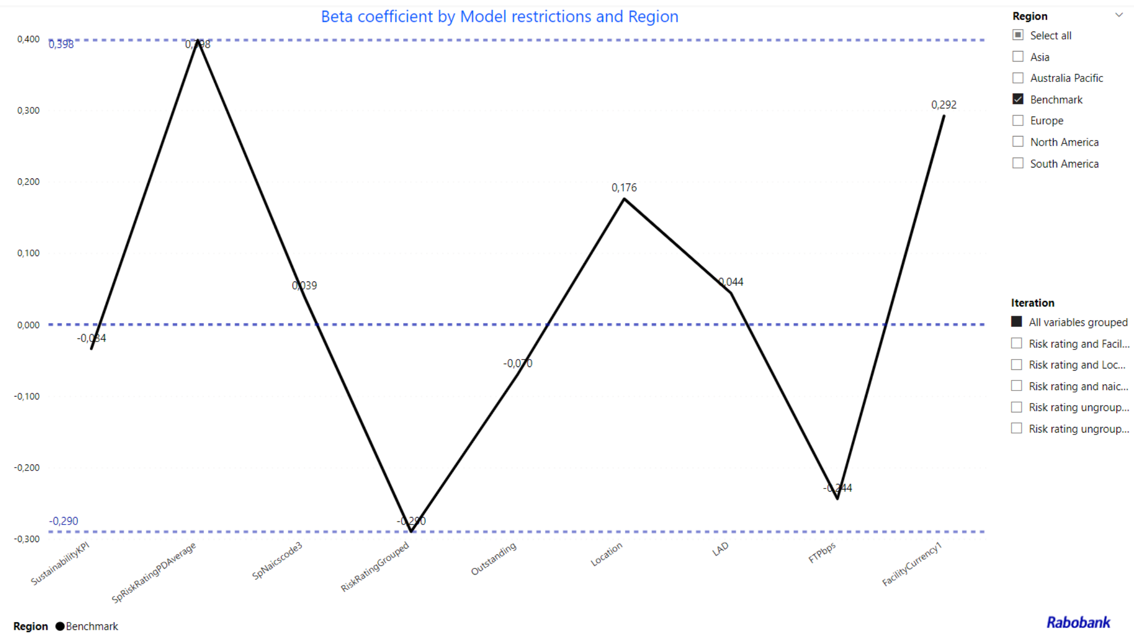


Figure 4-1: Standardized beta coefficient MLR benchmark.

Moreover, we see that FTP bps has a relative high negative influence. Evidently, when the CoF is higher, lower margins are made and thus lower NII bps is expected. With higher CoF, this will partly be calculated onto the customers however here a trade-off must be made between taking lower margins and staying competitive within the banking sector and new clients.

It is interesting to note that both the currency and location have a strong positive correlation with the NII bps. The currency in which transactions are paid can yield positive gains on the NII bps by the spread between different currency as well as the currency index. Additional gains can be made by exchanging one currency to the other and the spread between both. Here, the spread refers to the difference between bid and ask orders on a certain currency. Consequently, the operating location of a company yield some of the benefits for the Rabobank of the abovementioned.

Other variables such as the sustainability KPIs, NAICS code, outstanding, and LAD look to only have a small positive or negligible effect on the NII bps.

Risk Rating	Count
R7	36016
R10	20734
R11	19015
R8	17966
R12	15968
R13	15440
R9	12967
R15	9906
R14	5539
R17	3563
R6	3334
D3	2742
R16	2021
R5	1107
D4	687
R18	562
D2	263
R4	216
R19	126
R3	41
R20	20
R2	7
D1	5

Figure 4-2: Risk rating versus count Benchmark.

Figure 4-3 depicts the standardized beta coefficients which results from our MLR for South America. This is the only region where we see a discrepancy in the influence of Outstanding on the NII bps. Corresponding to the outstanding, we also see a higher negative correlation with NII bps caused by LAD. Higher LAD results in higher amount of capital needed to comply with regulation and thus decrease the NIM. As South America is a region where high inflation is observed, this increases the risks for Rabobank's and the offering of financial products there. Higher risks results in higher margin and thus it has a positive correlation with the NII bps. Moreover, we notice that both the NAICS codes as Location have a significant negative correlation with the NII bps for South America.

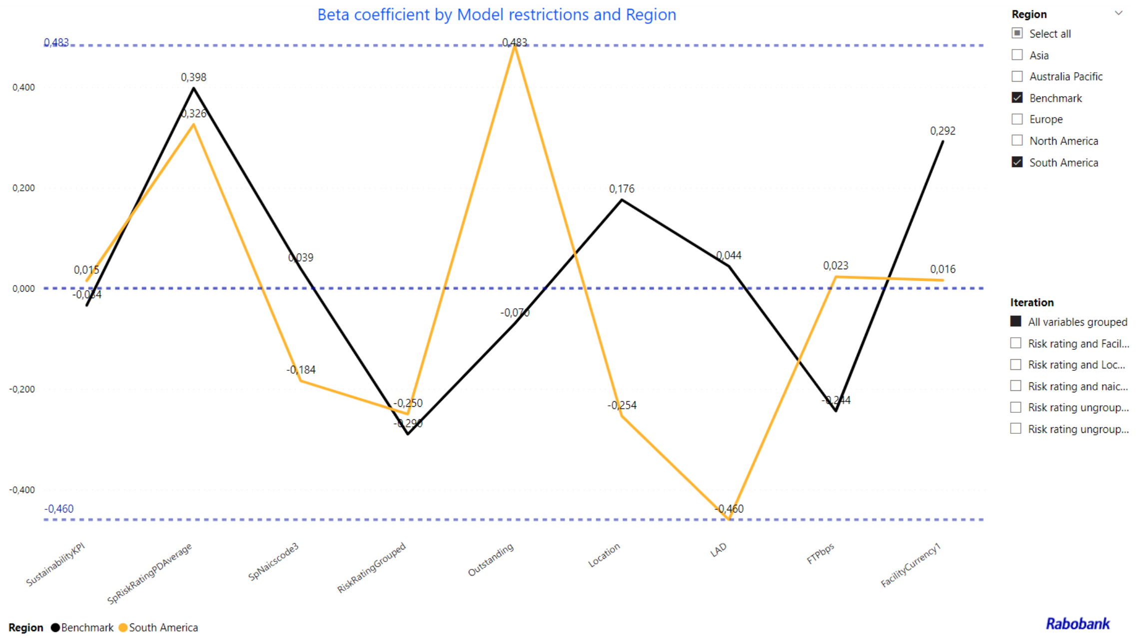


Figure 4-3: Standardized beta coefficient MLR South America.

All the results for region specific in depth results can be found in Section **Error! Reference source not found.** and onwards.

4.1.2 Sensitivity analysis

We further investigate the ungrouped variables such as risk ratings, location and currencies. By analyzing the ungrouped version of these variables, we hope to find additional underlying information which otherwise would have gone unnoticed.

4.1.2.1 Risk rating ungrouped and NAICS2

Ungrouping the risk ratings already reveals a lot on the actual influence of the different risk ratings. As we can see from Figure 4-4, lower risk rating have a negative or negligible correlation with the NII bps whereas higher risk ratings and default rating have a positive correlation with the NII bps. It must be noted that the influence of the risk rating do not perfectly follow the risk to rewards ratio. Some medium risk ratings, e.g. R11-13, still have some noticeable positive correlation with the NII bps. This can be contributed to the higher transaction count in are used dataset, as depicted in Figure 4-2. While ungrouping the risk ratings, we now see a noticeable negative influence of the outstanding amount on the NII bps. Thus, the above observation implies that lending more money does not results in higher profits in terms of NII bps. Note that this observation is found when analyzing all regions and may not hold for a specific region.

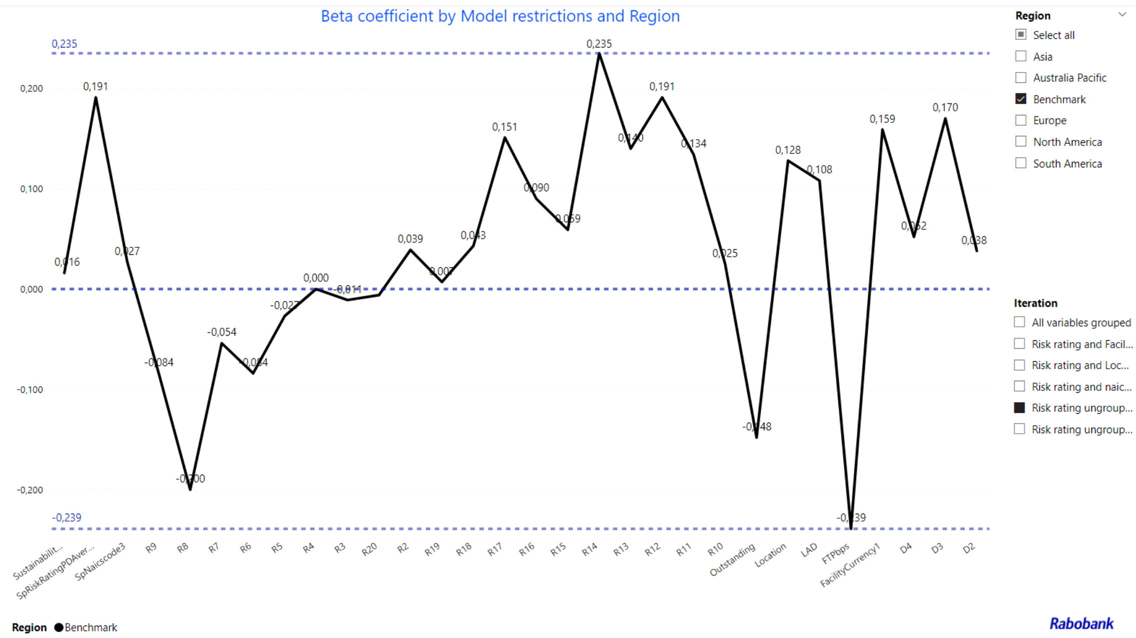


Figure 4-4: Risk rating ungrouped and NAICS2 Benchmark.

We have analyzed less aggregated NAICS code, as depicted in Figure 4-5, but we do not see significant difference in the standardized beta coefficient between a 2 or 3 numbered NAICS code. Evidently, when using NAICS3, more unique NAICS codes are used in the dataset and unless no specific sectors can be found of high correlation with the NII bps, we would expect the standardized beta coefficients to be approximately the same. Additionally, we have used the Rabobank specific sector variable but have not found any significance influence and the standardized beta coefficient were near zero. However, we found that the NAICS code is a better variable to analyze the sectors as this is a standardized method on classifying the different sectors.

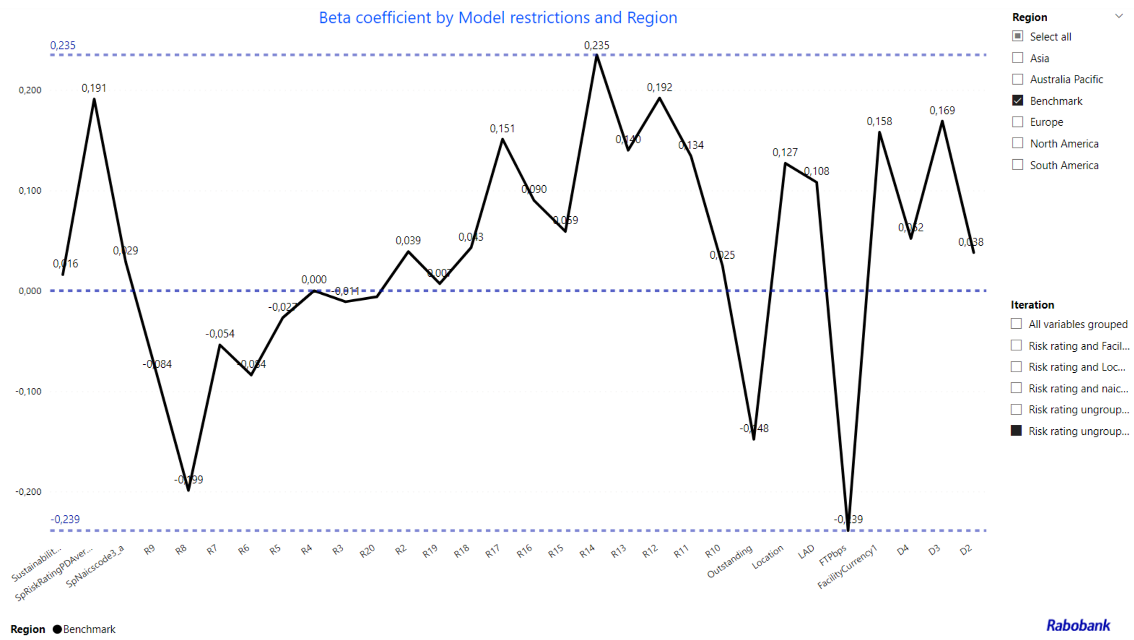


Figure 4-5: Rating ungrouped and NAICS3 Benchmark.

Concluding, we see a logical trend within the risk ratings where higher risk yields higher returns in terms of NII bps. Although the abovementioned trend has some outliers, this can be assigned to the risk rating count in the dataset. Finally, we do not see significant differences in the standardized beta coefficients between a two- and three digits NAICS code.

4.1.2.2 Risk rating and Facility currency ungrouped

In Figure 4-6 we see the influence of different currency on the NII bps. We see that all variables, which are explained in previous sections, stay about the same. Nevertheless, we see some significant positive influence of two different currencies namely US Dollar (USD) and British Pound Sterling (GBP). In consultation with Wholesale Business, they elaborate that this is caused by CoF side for these two currencies. Here the CoF is deemed higher which is calculated towards the client. However, the real CoF is usually lower than what is calculated towards the customer thus increasing the Rabobank's Net Interest Margin. Hence, the positive influence of USD and GBP on the NII bps.

Secondly, from Figure 4-6, we notice a high negative standardized beta coefficient for the EUR. In our dataset, if a different currency then the EUR is used, the price will incorporate the exchange rate and convert the amount to euro's. Thus, one would expect the model to find a negative relationship between EUR and NII bps as compared to the other currencies, it will have a negative influence. Additionally, one could argue as no currency exchanges need to be made, no additional margin can be made on the spread between currencies. Moreover, no currency swaps are needed for EUR which is part of the FTP. Thus, we have a lower CoF and to stay competitive, this is not calculated towards the customer. Finally, as Rabobank is an EUR bank, the fund price of EUR is cheaper compared to other region. Thus, there is no need to search for funds on the interbank market as the cheaper funds imply lower risk and thus lower margins.

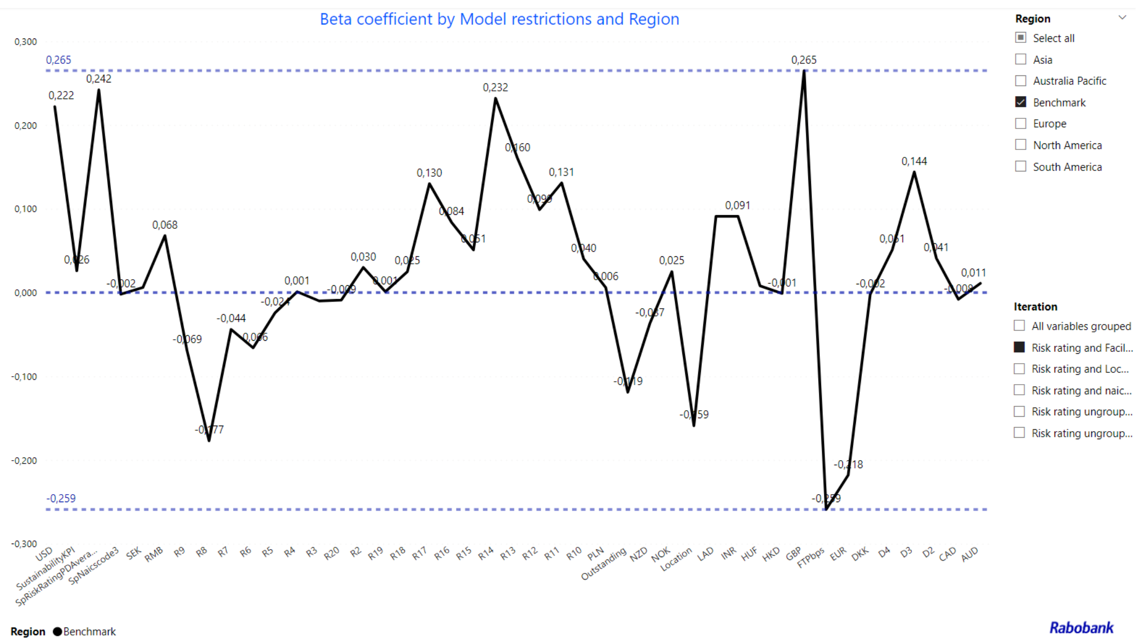


Figure 4-6: Risk rating and Facility currency ungrouped Benchmark.

4.1.2.3 Risk rating and Location ungrouped

While running the MLR again and now with the locations and risk ratings ungrouped, we see some important locations that have correlation with the NII bps. Firstly, we see a high positive correlation with the NII bps caused by London, India, Turkey and Argentina.

India, Turkey and Argentina are three locations with high levels of inflation. In these countries, higher interest rate base amounts are calculated as these pose higher risk. Due to the higher risk, higher net interest margins are calculated to compensate for the risk. Thus, the inflation offers higher risk for business and these locations have a positive effect on the NII bps.

It is interesting to note that both GBP and London have an positive correlation with the NII bps. Thus implying that or the location or the currency is of importance for the NII bps. This is evident in London as many equity financing products are offered there. These are usually paired with higher risk and in order to compensate for that risk, higher margins are charged.

Nevertheless, we see the contrary between USD and the location which usually pay in USD, namely Chicago and Atlanta. Here, the relationship is inversed where the USD is of positive influence however the typical USD location have negative influence. We must however note that companies located in different location then the USA may choose to pay in USD which explains the discrepancy between USD and USA locations. Moreover, more value chain finance products are offered in the US regions. Value chain finance are financial products with lower risk portfolios. Evidently, lower risk also results in lower calculated margins. Hence why both American locations have a negative correlation with the NII bps

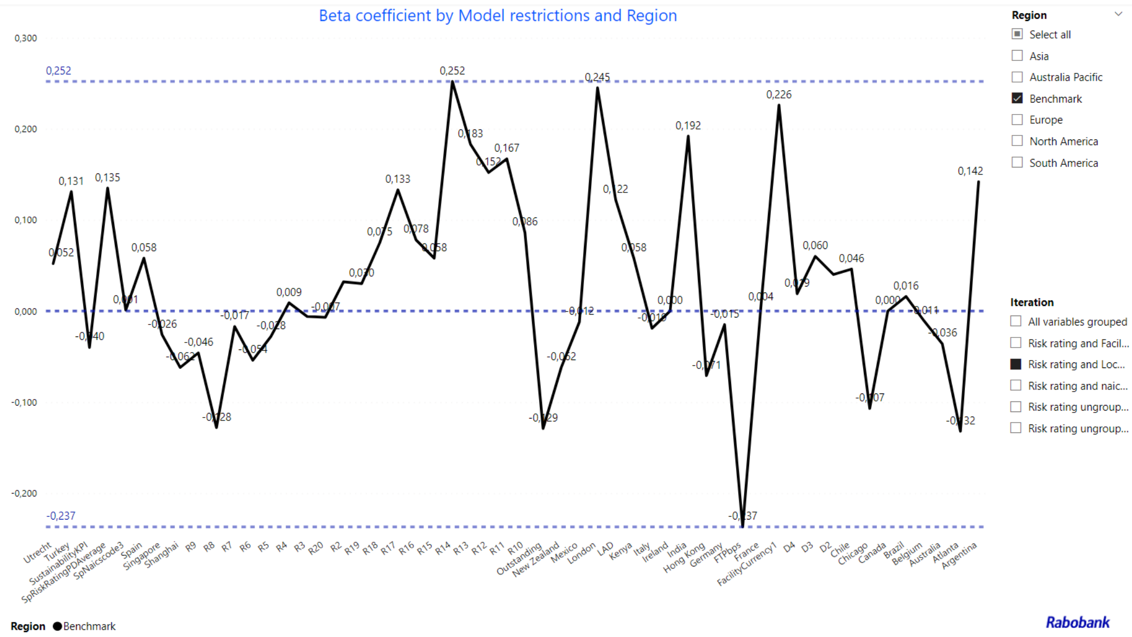


Figure 4-7: Risk rating and location ungrouped Benchmark.

4.1.2.4 Risk rating and NAICS ungrouped

Finally, in Figure 4-8 we see the standardized beta coefficient where risk ratings and NAICS code are ungrouped. The risk ratings again have similar standardized beta coefficient as described in Section 4.1.2.1 where higher risk ratings results in a positive correlation with the NII bps.

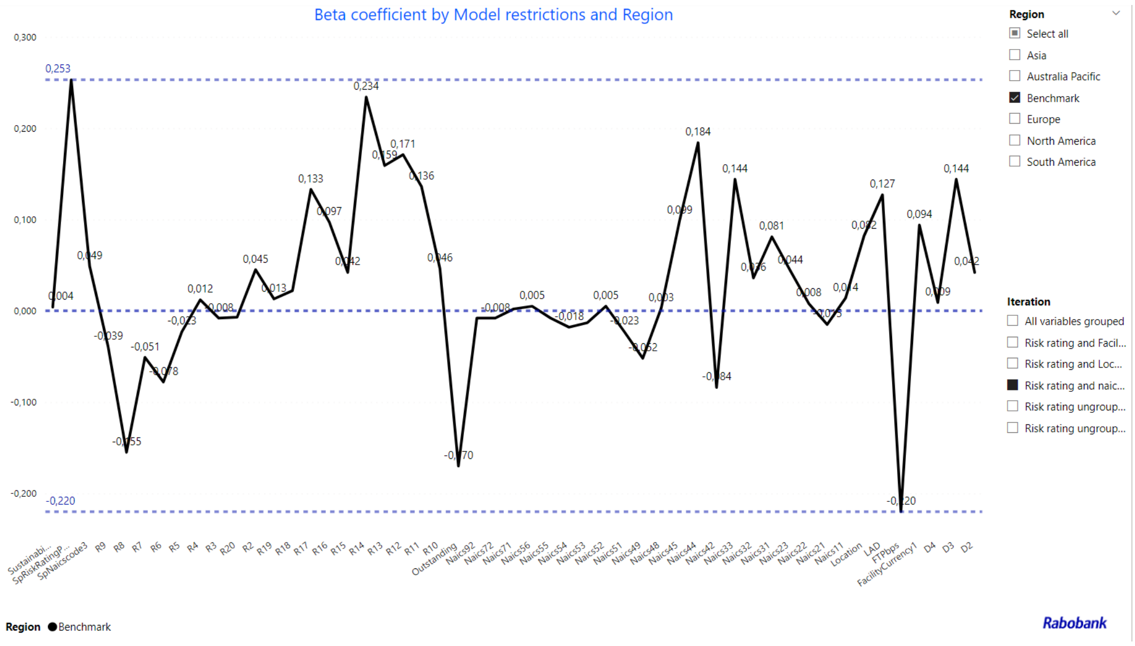


Figure 4-8: Risk rating and NAICS ungrouped benchmark.

Moreover, we now see in more depth the influences of different sectors on the NII bps. As can be seen from Figure 4-8 we see four different sector codes which have a high positive correlation with our dependent variables, namely NII bps. These are sectors 23, 31 to 33, 44, and 45. This are equivalent to the sectors: construction, manufacturing, and retail trade respectively.

Our model confirms that some sectors have positive impact on the overall NII bps. Taking Retail Trade as an example. This sector has had some difficulties over the past years, starting with the migration towards online. Moreover, this trend has been strengthened by amongst others the Covid-lockdowns. Currently, Retail Trade also faces difficulties like hardly no personnel to find, increased energy and rental costs and so on. This all boils down to increased credit risk which Rabobank needs to take into account. Thus, increased credit risk implies higher calculated margins and thus a positive factor on the NII bps. Similar trends can be seen on macro-economic levels for construction and manufacturing.

On the contrary, we also see two sectors which have significant negative correlation with the NII bps. These sectors are numbered as NAICS code 42 and 49, Wholesale trade and Transportation & Warehousing respectively. The abovementioned sectors have been under less stress in last two years. Moreover, they did not incur any demand decreases or other important trend changes. Currently, some macro-economic factors start to play a role in these sectors such as increased fuel costs. To see whether current increased levels of fuel costs have had correlation with these sectors, one should run the model on a smaller dataset. However, this could lead to lower adjusted R-squared thus explaining less of the variance in NII bps. This is something for further research to see if we can spot trend changes on smaller timeframes while ensuring enough predicting power.

Figure 4-9 and Figure 4-10 depict the risk ratings and NAICS codes without any other variables. Note, the reason for filtering of the beta coefficients from high too low for the NAICS was a business decision and does not have any other implications.

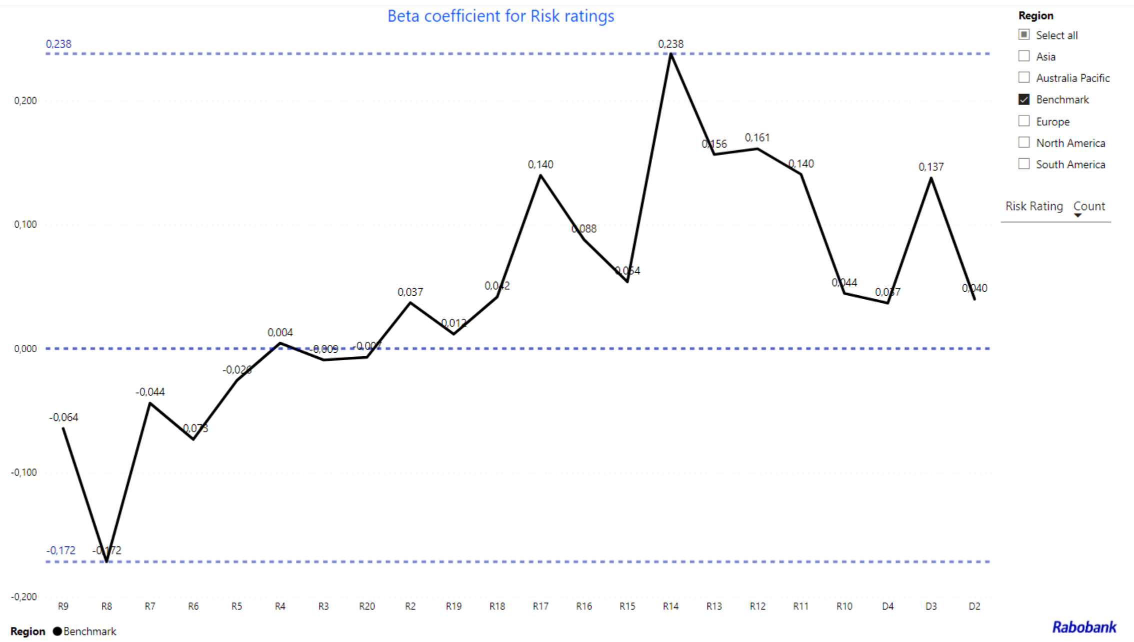


Figure 4-9: Benchmark Risk ratings.

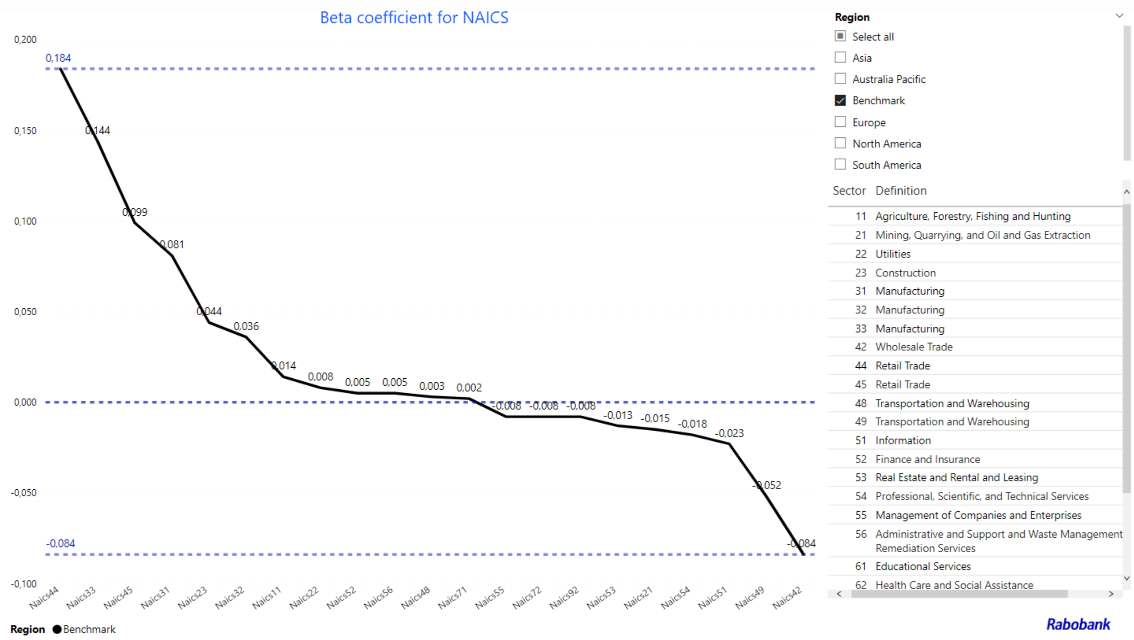


Figure 4-10: Benchmark NAICS.

5 Conclusion

In the final chapter, we conclude the research and answer the research questions set at the beginning of the report. Firstly, we answer the sub research questions before answering the main research question.

How does wholesale banking work?

Retail and wholesale banking are two distinct types of banking services that cater to different customer segments and operate under different regulatory frameworks. While retail banking offers a wider range of products and services to individual and small business customers, wholesale banking offers higher returns and growth potential to large corporate and institutional customers. Wholesale banking can also play a critical role in financing SMEs and providing banks with access to low-cost funding sources, but it also exposes them to liquidity risks. Thus, managing risk is critical in wholesale banking, and stress testing and contingency planning are essential tools for managing liquidity risk.

The loan structure in both retail and wholesale banking involves a loan agreement between the bank and the borrower. The agreement outlines the terms and conditions of the loan, including the loan amount, interest rate, repayment schedule, and any collateral required. In retail banking, the loan agreement is usually straightforward and standardized, while in wholesale banking, the loan agreement is more complex and tailored to the specific needs of the borrower. A wide list of financial products can be found in Section 2.1.2 and Section 8.3.

What are important variables and definitions within finance and wholesale banking?

Many variables are of importance within wholesale banking. Firstly, GII and NII are essential indicators of a wholesale bank's performance. GII reflects the bank's ability to generate revenue from lending activities, while NII indicates the bank's ability to manage its interest rate risk and CoF. Factors affecting gross and net interest income include market conditions, credit risk, interest rate risk, CoF, and operational costs. Therefore, wholesale banks must manage their operations effectively, monitor their credit and interest rate risk, and adapt to changing market conditions to maintain their profitability.

Secondly, the funds transfer price (FTP) of a transaction or product consists of a cost of funds (CoF) component and an adjustment factor. The CoF component can consist of interest rates, liquidity premium or other product specific FTP add-ons. Adjustment Factor in FTP, besides steering on the performance metrics as stated in the SF, there is also an element of steering within the FTP called the adjustment factor (”AF”). This is an additional spread (which can be negative or positive) on top of the Cost of Funds.

Regulatory Capital (RC) is an important concept within wholesale banking. It is a key component of bank regulation and plays a significant role in ensuring the safety and soundness of the banking system. Banks are required to maintain a minimum level of regulatory capital to cover their risk exposures, and they are subject to capital adequate ratios. Regulatory capital has implications for a bank's ability to take on risk, its profitability, and its ability to pay dividends or repurchase shares. Overall, a bank's ability to manage its regulatory capital is critical to its success in the wholesale banking sector.

Exposure at Default (EAD) and Recovery rate (RC) are two important parameters in credit risk modelling within wholesale banking. EAD represents the amount of exposure a bank has to a counterparty at the time of default, while RC represents the percentage of the exposure that a bank can recover in the event of a default. Both EAD and RC are affected by various factors, including the creditworthiness of the counterparty, the type of collateral, the seniority of the debt, the loan structure, and the economic environment. Accurate estimation of EAD and RC is essential for effective risk management within wholesale banking.

What benchmark techniques can be used to compare net interest income?

Several benchmark techniques can be found in literature. Financial forecasting is a critical process that helps businesses make informed decisions about investments and financial activities based on quantitative methods which uses statistical models and mathematical calculations to analyze historical data and make future predictions. Some of the most common quantitative forecasting methods used in finance and wholesale banking include time series analysis, multiple linear regression, and exponential smoothing.

Machine learning is a tool that has been widely adopted in many industries. Machine learning is a subfield of artificial intelligence that deals with the design and development of algorithms that can learn patterns and relationships in data without being explicitly programmed. However, within finance and wholesale banking, there are concerns regarding the use of ML due to issues of transparency and interpretability, overfitting, data quality, and model stability. While ML can be beneficial in some cases, it is important to carefully evaluate the risks and benefits before adopting these technologies in these sectors. Due to the high importance of transparency and regulation within banking, we have opted to not apply a supervised ML model to our internal benchmark.

MLR is a valuable tool for analyzing variables and predicting outcomes in finance and wholesale banking. It allows relationships in the sense of correlation between multiple variables to be analyzed, which can help to identify how variables are related and which variables are most important for predicting outcomes. It can also be used to predict outcomes, test hypotheses, and is a flexible technique that can be adapted to a wide range of applications in these sectors. MLR use assumes certain assumptions to be met, including linearity, normality, constant variance of the errors, and independence of the errors. MLR also produces important output measures such as standardized and unstandardized coefficients, standardized and unstandardized residuals, predicted values, adjusted R-squared, which help to interpret the results and assess the goodness of fit of the model.

What are the significant differences in Net Interest Income bps (NII bps) between regions and sectors, and how can Rabobank gain competitive advantage from the factors that influence the NII's bps?

We see a logical trend within the risk ratings where higher risk yields higher returns in terms of NII bps. It must be noted that the correlation of the risk ratings do not perfectly follow the risk to rewards ratio. Some medium risk rating, e.g. R11-13, still have some noticeable positive correlation with the NII bps. This can be contributed to the higher transaction counts within our used dataset where some ratings are more frequent then others.
Moreover, we see that FTP bps has a relative high negative impact on the NII bps. Evidently, when the CoF is higher, lower margins are made and thus lower NII bps is expected. With higher CoF, this will partly be calculated onto the customers however here a trade-off must be made between taking lower margins and staying competitive within the banking sector and new clients. Moreover, the FTP bps may also change over time as some CoF are secured for the entire term of a loan however shorter CoF agreements are not uncommon such as a three months EURIBOR.

It is interesting to note that both the currency and location have a strong positive correlation with the NII bps. The currency in which transactions are paid can yield positive gains on the NII bps by the spread between different currency as well as the currency index. Additional gains can be made by exchanging one currency to the other and the spread between both. Here, the spread refers to the difference between bid and ask orders on a certain currency in our case. Consequently, the operating location of the company yield some of the benefits of the abovementioned.

Consequently, we see some significant positive influence of two different currencies namely US Dollar (USD) and British Pound Sterling (GBP). In consultation with Wholesale Business, they argument that this is caused by CoF side for these two valuta's. Here the CoF is deemed higher which is calculated towards the client. However, the real CoF is usually lower than what is calculated towards the customer thus increasing the Rabobank's Net Interest Margin. Hence, the positive influence of USD and GBP on the NII bps.

However, we notice a high negative standardized beta coefficient for the EUR. In our dataset, if a different currency then the EUR is used, the price will incorporate the exchange rate and convert the amount to euro's. Thus, one would expect the model to find a negative relationship between EUR and NII bps as compared to the other currencies. Additionally, one could argue as no currency exchanges need to be made, no additional margin can be made on the spread between currencies. Moreover, no currency swaps are needed for EUR which is part of the FTP. Thus, we have a lower CoF and to stay competitive, this is not calculated towards the customer. Finally, as Rabobank is an EUR bank, the fund price of EUR is cheaper compared to other region. Thus, there is no need to search for funds on the interbank market as the cheaper funds imply lower risk and thus lower margins.

India, Turkey and Argentina are three locations with high amounts of inflation. In these countries, higher interest rate base amounts are calculated as these pose higher risk. Due to the higher risk, higher Net interest margins are made. Thus, the inflation offers higher risk for business and these locations have a positive effect on the NII bps. Moreover, due to the high inflation, a higher outstanding amount will have a positive correlation with the NII bps.

It is interesting to note that both GBP and London have an positive correlation with the NII bps. Thus implying that or the location or the currency is of importance for the NII bps. This is evident in London as many equity financing with complex products are offered there. These are usually paired with higher risk and thus resulting in higher margins.

Nevertheless, we see the contrary between USD and the location which usually pay in USD, namely Chicago and Atlanta. Here, the relationship is inversed where the USD is of positive influence however the typical USD location have negative influence. We must however note that companies located in different location then the USA may choose to pay in USD which explains the discrepancy between USD and USA locations. Moreover, more value chain finance products are offered in the US regions. Value chain finance are financial products with lower risk portfolios. Evidently, lower risk also results in lower margin. Hence why both US location have a negative correlation with the NII bps.

Our model confirms that some sectors have positive impact on the overall NII bps. Taking Retail Trade as an example. This sector has had some difficulties over the past years, starting with the migration towards online. Moreover, this trend has been strengthened by the Covid-lockdowns. Currently, Retail Trade also faces difficulties like hardly no personnel to find, increased energy and rental costs and so on. This all boils down to increased credit risk which Rabobank needs to take into account. Thus, increased credit risk implies higher prices and thus a positive factor on the NII bps. Similar trends can be seen on macro-economic levels for construction and manufacturing.

The results indicate that the model can be used to play into macro-economic trends observed as described above in the other to gain competitive advantage and increase their NII bps. Here a trade-off should be made between taking on more risk in order to increase NII bps. Otherwise, one should focus on sectors, locations and currency of interest to increase their NII bps.

6 Recommendations

Although we have found interesting results in our research there is still a lot of untouched potential regarding the benchmarking between regions and sectors. In the following section we will discuss the final recommendations for the proceedings of future research. From our research, proposed model and Section 4.1 a few recommendations arise.

Apply model to Rural portfolio to observe macro-economic trends.
Firstly, the model can be applied to the Rural portfolio of Rabobank. As observed while analyzing the Rabobank's dataset, Rural lending portfolio consists of 97.4% of product type long-term loans. These financial products are the easiest to analyze due to the lack of complexity. Applying the model to the rural portfolio may yield interesting macro-economic trend changes not observed by the wholesale portfolio. Aligning these two might yield additional competitive advantages for business.

Further analyze why certain variables breach one of the MLR assumptions.
Further analysis needs to done on why certain variables breach some of the assumptions of multiple linear regression modelling. Certain variables should not have breached these assumptions according to the theory and thus it is of interest to find the underlying reasoning. More extensive data modification techniques can be applied to be able to incorporate them without any assumption breaches. We have tested some of these techniques however didn't find any significant results. Due to time restrictions, we were not able to apply all techniques which could yield interesting insights.

Update dataset on monthly basis.
The MLR model should be updated on a monthly basis in order to have the latest transaction dataset incorporated and notice trend shifts. Nonetheless, a time series analysis should always be made time to time to ensure that time does not play a significant role in the model. Keeping track of the changes in the standardized beta coefficient within each region and iteration yields the opportunity to notice trend changes. Hereafter, business can play into these trend changes to boost their NII bps.

Apply the model per location to spot differences in sectors and or other variables.
The MLR can be applied per location to spot significant differences in sectors, risk ratings and currency per country. Applying the model per location will result in a smaller dataset used for analysis. Thus, it is of importance to keep a high predictive power concerning the variation of the NII bps namely the adjusted R-squared. A analysis could be made when using a bigger dataset, e.g. 2018-2022, would yield the wanted results. Of course, a time series analysis should be used to support the above.

Link the MLR model to the Rabobank's NIM-tooling.
This is a Rabobank specific recommendation. Without revealing too much information, aligning the model and NIM tooling could yield advice on what to do on short-term. The NIM tooling gives more in depths insights on the interest rate structure and the resulting Net Interest Margin.

Translate the outcome in a more predictive model in order to support strategic decision making.
The model can be used in a more predictive way where the MLR formula can be used for assumption testing for the business. Using the MLR formula, business can fill in their estimates of different independent variables and make a prediction on the expected NII bps. It must be noted this should only be used as a supportive tool and not to make real predictions about the future. This is because one shouldn't want to use predictive information within a bank's balance sheet or make important business decision based on a model which changes over time.

7 References

(BIS), B. f. (2019). *Wholesale Funding of the Global Banking System.* BIS Quarterly review.

Administration, U. E. (2023, 04 01). *PETROLEUM & OTHER LIQUIDS*. Retrieved from U.S. Energy Information Administration: https://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=MCRFPUS2&f=M

Ahmad, A, & Khan. (2016). Retail and Wholesale banking: A comparative analysis. *Journal of Internet Banking and Commerce*, 1-10.

Aliber, R. Z., & Blackwell, D. W. (1971). International banking and finance: A review of some recent trends. *Journal of Finance*, 1029-1053.

Allison, P. (2012, Septemeber 10). *When can you safely ignore multicollinearity?* Retrieved from Statistical Horizons: https://statisticalhorizons.com/multicollinearity/

Alpaydin, E. (2010). *Introduction to Machine Learning.* Cambridge.

Andrew F. Siegel, Michael R. Wagner,. (2022). *Practical Business Statistics.* Science Direct.

Banerjee, A., Dolado, J. J., Galbraith, J. W., & Hendry, D. F. (1993). *Co-intergration, Error Correction, and the Econometric Analysis of Non-stationary Data.* Oxford University Press.

Bank for international settlements. (2022). *Basel Committee on Banking Supervision: Caculation of RWA for credit risk and standardised approach: inidvidual exposures.*

Barboza, J. (2019). Why banks need to be wary of machine learning. *Financial Times*.

BCBS. (2010). *Basel III: A Global Regulatory Framework For More Resilient Banks and Banking Systems.* Bank for International Settlements.

Besley, T., & Ghatak, M. (2017). *Advanced Microeconomic Theory.* Routledge.

Bhattacharya, S., & Thakor, A. V. (1993). Contemporary banking theory. *Journal of financial intermediation*, 2-50.

Bikker, J. A., & Haaf, K. (2002). Competition, concentration and their relationship: An empirical analysis of the banking industry. *Journal of Banking & Finance*, 2191-2214.

Bowerman, B. L., & O'Connell, R. T. (2014). *Essentials of Business-statistics.* Cengage Learning.

Carol M. Woods, Michael C. Edwards,. (2011). *Essential Statistical Methods for Medical Statistics.* Science Direct.

Chatfield, C., & Prothero, D. L. (2020). *The Analysis of Time Series: An Introduction.* CRC Press.

Chen, J. (2022). *Interest Rate Risk Definition and Impact on Bond Prices*. Retrieved from Investopedia: https://www.investopedia.com/terms/i/interestraterisk.asp

Chen, J., & Hsieh, J. J. . (2018). Machine learning in finance. *Journal of Financial Data Science*, 3-23.

Chirwa, E. W., & Mlachila, M. (2019). Determinants of net interest income and gross income in Africans banks. *Journal of African Business*, 469-490.

Demirgüc-Kunt, A., & Huizinga, H. (1999). Determinants of commercial bamk interest margins and profitability: Some international evidence. *The World Bank Economic Review*, 379-408.

Düllmann, K., & Masschelein, N. (2018). The Role of Wholesale Funding in Bank Liquidity Management. *Financial Intermediation*.

Enders, W. (2010). *Applied Econometric Time Series.* John Wiley & Sons.

Field, A. P. (2013). *Discovering Statistics using SPSS: and Sex and Drugs and Rock 'n' Roll.* London: Sage.

Fox, J. (2016). *Applied Regression Analysis and Generalized Linear Models.* Sage Publication.

Garcia-Feijoo, L., Moreno, J., & Rubio, G. (2015). Estimating loss given default under a Bayesian approach for commercial loans portfolio. *Journal of Risk Finance*, 407-423.

Gelman, A., & Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models.* Cambridge University Press.

Gujarati, D. N. (2003). *Basic Econometrics.* McGraw-Hill Education.

Hair Jr, J. F., Black, W. C., Babin, B. J. & Anderson, R. E. (2014). *Multivariate Data Analysis.* Pearson Education.

Hair Jr, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate Data Analysis.* Pearson.

Haldane, A. G. (2018). *The Race Against the Machine: How the Combination of AI and Big Data are Transforming the Financial Landscape.* Bank of England.
Hamilton, J. D. (1994). *Time Series Analysis.* Princeton university press.
Hasan, I., & Dridi, J. (2011). The effects of the global crisis on Islamic and conventional banks: A comparative study. *International Monetary Fund*, 11-16.
Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Interference, and Prediction.* New York: Springer.
Hill, R. C., Griffiths, W. E., & Liam, G. C. (2018). *Principles of Econometrics.* John Wiley & Sons.
Huang, Y., & Huang, Y. (2018). The impact of interest rate risk on net interest income of Chinese commercial banks. *Journal of Chinese Economic and Business Studies*, 131-148.
Huitema, B. E. (2020). *Democratic Experiments in Africa: Regime Transitions in Comparative Perspective.* Cambridge University Press.
Hull, J. C. (2018). *Options, Futures, and Other Derivatives.* Pearson Education Limited.
Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and Practice.* OTexts.
IBM. (2016). *Statistics and Econometrics for Finance Quantitative Analysis and IBM® SPSS® Statistics A Guide for Business and Finance.* Springer.
IOVS. (2019). Are linear regression techniques appropriate for analysis when the dependent (outcome) variable is not normally distributed. *IOVS*.
James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning.* Springer.
James, R., & Knaub, Jr. (2019). *Estimating the Coefficient of Heteroscedasticity.* ResearchGate.
Keenan, J., & Delhaise, P. (2010). *Corporate Finance Handbook: A Practical Approach to Financials Management.* John Wiley & Sons.
Keogh, J. W. (2015). The role of wholesale banking in financing small and medium-sized enterprises. *Small Business Management*, 669-681.
Kim, Y., & Lee, J. . (2019). Can machine learning improve risk management in banking? *Journal of Risk Management in Financial Institutions*, 292-301.
Kozak, S., & Yildiz, B. (2014). *Multivariate Linear Regression Models in Banking Sector: A Case Study From Turkey.* Procedia-Social and Behavioral Sciences.
Krahnen, J. P., Wilde, C. J., & Gürtler, M. (2011). The impact of asset correlations on the estimation of LAD for securitizations. *Journal of Banking & Finance*, 1577-1588.
Kruschke, J. K. (2015). *Doing Bayesian Data analysis.* Science Direct.
Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2004). *Applied Linear Statistical Models.* McGraw-Hill.
Li, J., & Peng, Y. . (2017). Exploring the impact of interest rate changes on bank profitability in China; A Multiple linear regression analysis. *International Journal of Economics, Commerce and Management*, 142-155.
Li, Y., & Li, Z. (2018). Industry Classification and Credit Risk Assessment for Small and Medium-Sized Enterprises. *Journal of Risk and Financial Management*, 27.
Maio, P., & Olsen, K. B. (2016). Loss given default: A comprehensive study. *Journal of Banking & Finance*, 45-66.
Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (2008). *Forecasting: Methods and Applications.* John Wiley & Sons.
Merkle, C. (2019). *Machine Learning in Finance: Why Transparency is Key.* Medium.
Mingo, J. J. (2018). *Credit Risk Modeling using Excel and VBA.* John Wiley & Sons.
Mishkin, F. S. (1991). *Asymmetric Information and Financial Crises: A Historical Perspective.* American economic review.
Mishkin, F. S. (2006). *The Economics of Money, Banking, and Financial Markets.* National Bureau of Economic Research.
Molyneux, P., & Nguyen, L. . (2019). Artificial intelligence in finance. *Journal of Banking & Finance*, 1-7.

NAICS. (2022). *North American Industry Classification System.* North American Industry Classification System.

Neter, J., Wasserman, W., & Kutner, M. H. (1989). *Applied Linear Statistical Models.* Irwin.

Oviatt, B. M., & McDougall, P. P. (1994). Understanding sustained competitive advantage. *Journal of management*, 697-727.

Pindyck, R. S., & Rubinfield, D. L. (2014). *Econ.* Econ.

Rabobank. (2018). *CRDW Portfolio Data, System Interface Specfication (SIS).*

Rabobank. (2018). *Global Standard on FTP Pricing Methodology.* Rabobank Gobal Policy.

Rabobank. (2022). Basel IV Business Information: Refresher on Capital. *Rabobank*.

Rabobank. (2022). *Global Standard on Loss Given Default and Exposure at Default.* Rabobank CLR - Enterprise Risk.

Rabobank. (2022). *Global Standard on Sustainable Development.* The Cooperative Rabobank.

Rabobank. (2022). *Loan Quality Classification & Rabobank Risk Rating (Credit Risk Parameters).* Country banking Rabobank Standard Credit.

Rabobank. (2022). *Practical Standards for Methodology Deliverables.* Rabobank Compliance Legal Risk.

Rabobank. (2023). *Rabobank*. Retrieved from About Rabobank: https://rabobank.jobs/en/about-rabobank/

Rime, D., & Schrimpf, A. (2013). *The Anatomy of the Global FX Market Through the Lens of the 2013 Triennial Survey.* Bank for International Settlements Quarterly Review.

Saunders, A., & Cornett, M. M. (2014). *FInancial Institutions Management: A Risk Management Approach.* McGraw-Hill Education.

Serrani-Cinca, C., Gutiérrez-Nieto, B., & Mar-Molinero, C. (2017). NAICS classification and bank performance. *Journal of Business Research*, 1-10.

Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding Machine Learning: From Theory to Algorithms.* New York: Cambridge University Press.

Shumway, R. H., & Stoffer, D. S. (2017). *Time Series Analysis and its Applications: With R Examples.* Springer.

Sironi, A. (2033). Testing for market discipline in the European banking industry: Evidence from subordinated dept spreads. *Journal of Money, Credit and Banking*, 443-472.

Supervision, B. C. (2015). *International Convergence of Capital Measurement and Capital Standards: A Revised Framework.* Basel Committee on Banking Supervision.

Van Roy, P., & Van Roy, P. (2014). *Bank Regulation, Risk and Return: Evidence.*

Vrealey, R. A., & Myers, S. C. (2011). *Principles of Corporate Finance.* McGraw-Hill.

Vrins, F. A., & Verbeek, M. J. (2007). A PD-LGD frameork for retail loans with correlated factors. *Journal of Banking & Finance*, 785-804.

Wooldridge, J. M. (2010). *Econometirc Analysis of Cross Section and Panel Data.* MIT Press.

Zhang, X., Lai, K. K., & Zhang, Y. . (2019). Can machine learning be used to forecast stock prices? Evidence from China. *Journal of Forecasting*, 730-744.

Zhu, A. (2022). *Are the OLS Estimators Normally Distributed in a Linear Regression Model?* Retrieved from towardsdatascience: https://towardsdatascience.com/are-ols-estimators-normally-distributed-in-a-linear-regression-model-89b688fa8dc3

8 Appendix

Below one can find all appendices which were of importance of the research. We have decided to leave out the SPSS results of each iteration as it would take 20 plus pages.

8.1 Dummy variables auto recoded

SpLocationRegion into Region

Old Value	New Value	Value Label
Asia	1	Asia
Australia Pacific	2	Australia Pacific
Europe	3	Europe
North America	4	North America
South America	5	South America

Figure 8-1: Values for regions MLR

SpLocation into Location

Old Value	New Value	Value Label
Argentina	1	Argentina
Atlanta	2	Atlanta
Australia	3	Australia
Belgium	4	Belgium
Brazil	5	Brazil
Canada	6	Canada
Chicago	7	Chicago
Chile	8	Chile
France	9	France
Germany	10	Germany
Hong Kong	11	Hong Kong
India	12	India
Ireland	13	Ireland
Italy	14	Italy
Kenya	15	Kenya
London	16	London
Mexico	17	Mexico
New Zealand	18	New Zealand
Shanghai	19	Shanghai
Singapore	20	Singapore
Spain	21	Spain
Turkey	22	Turkey
Utrecht	23	Utrecht

Figure 8-2: Values for locations MLR

```
SpRabobankSectorLevel03 into Sector03
Old Value  New Value  Value Label

F&A                  1  F&A
Non F&A              2  Non F&A
```

Figure 8-3: Values for sectors MLR

```
SpRabobankSectorLevel02 into Sector02
Old Value                          New Value  Value Label

01 - Commodities                           1  01 - Commodities
02 - Animal Protein                        2  02 - Animal Protein
03 - Beverages                             3  03 - Beverages
04 - Consumer Foods                        4  04 - Consumer Foods
05 - Dairy                                 5  05 - Dairy
06 - Farm Inputs                           6  06 - Farm Inputs
07 - Fresh Produce                         7  07 - Fresh Produce
08 - Other F&A                             8  08 - Other F&A
09 - Manufacturing (Non-F&A)               9  09 - Manufacturing (Non-F&A)
10 - Energy                               10  10 - Energy
11 - Construction                         11  11 - Construction
12 - Wholesale and retail trade           12  12 - Wholesale and retail trade
(excluding F&A and energy)                12  (excluding F&A and energy)
13 - Transportation and storage           13  13 - Transportation and storage
14 - Accommodation and food               14  14 - Accommodation and food
service activities                        14  service activities
15 - Information and                      15  15 - Information and
communication                             15  communication
16 - Financial and insurance              16  16 - Financial and insurance
activities                                16  activities
17 - Real estate activities               17  17 - Real estate activities
18 - Professional, scientific             18  18 - Professional, scientific
and technical activities                  18  and technical activities
19 - Administrative and support           19  19 - Administrative and support
service activities                        19  service activities
22 - Arts, entertainment and              20  22 - Arts, entertainment and
recreation                                20  recreation
23 - Other Non F&A                        21  23 - Other Non F&A
```

Figure 8-4: Values for sectors level 02 MLR

FacilityCurrency into FacilityCurrency1

Old Value	New Value	Value Label
AUD	1	AUD
CAD	2	CAD
DKK	3	DKK
EUR	4	EUR
GBP	5	GBP
HKD	6	HKD
HUF	7	HUF
INR	8	INR
NOK	9	NOK
NZD	10	NZD
PLN	11	PLN
RMB	12	RMB
SEK	13	SEK
Unspecified Facility Currency	14	Unspecified Facility Currency
USD	15	USD

Figure 8-5: Values Facility currency MLR

SubRiskRating into RiskRatingGrouped

Old Value	New Value	Value Label
D2	1	D2
D3	2	D3
D4	3	D4
R10	4	R10
R11	5	R11
R12	6	R12
R13	7	R13
R14	8	R14
R15	9	R15
R16	10	R16
R17	11	R17
R18	12	R18
R19	13	R19
R2	14	R2
R20	15	R20
R3	16	R3
R4	17	R4
R5	18	R5
R6	19	R6
R7	20	R7
R8	21	R8
R9	22	R9
Unspecified Sub Risk Rating	23	Unspecified Sub Risk Rating

Figure 8-6: Values Risk ratings MLR

8.2 Product type TERM

Product	Description	Asset or Liability	Applicable for Exposure Type
TERM	A term loan is a loan for a period of at least one year, subject to previously agreed conditions. The interest can be fixed or variable. In the case of a fixed rate, the interest can be established for the entire term of the loan or for a shorter period, in which case the period must be at least one year. Purpose Term loans are used for long term financing needs (fixed assets) and for working capital needs (current assets). Repayment A term loan can be paid back in equal monthly, quarterly, semi-annual or annual instal-ments, or in annuities. Other methods of repayment are repayments with a balloon and bullet repayments. Rabobank policy - The maximum **tenor** for long term loans should not exceed 10 years, with a strong preference for 5 to 7 years tenor. Exceptions to this policy will have to be based on strong motives. - **Repayment** should preferably be by the straight line method of repayment. Repayment should be dependent on the cash-flow of the business requesting the loan, or the cash flow from the project being financed. The nature of the goods to be financed should also be taken into account. - **Bullet loans** may only be granted to primary customers of the bank, whose financial position must also be sound. The maximum term of a bullet loan is limited to 5 years; for an exception see par. 5.3. The credit report should indicate that the bullet loan is a suitable way of financing. The report should also indicate which funds the bullet loan will be repaid from at the date of maturity.	Asset or Liability	

Figure 8-7: Product type TERM (Rabobank, Rabobank, 2023)

8.3 Other financial products

8.3.1 SYND

SYND are Syndicated loans. A loan of Rabobank together with a group of bankers, insurers, etcetera. The registered syndicated loan here is only the participation of the Rabobank within the syndicated loan. SYND belongs to the list of product types applicable to facilities with exposure type Trade Related Bank Exposure (TRBE).

8.3.2 SHLO

A short term advance is a loan granted for a short period, at previously agreed interest rates and terms. The term of short term advances is usually a round number of months (1, 2, 3, 6 or 12) or 1 or more days.
A short term advance can be made under a revolving line of credit and under short term facilities; availability is restored as the short term advances are repaid. When short term advances are taken up within a facility, the advances may not exceed the final date agreed for the facility.
Short term advances are used for short term financing (current assets). Repayment is made in a single payment at the end of the term of the short term advance.

8.3.3 ABON

An advance payment bond (= prepayment bond) is a bank guarantee to ensure repayment by the contracting company of advance payments paid out to this contracting company, in case the delivery is not correct

With large contracts, the preparations themselves are often very expensive. Large numbers of workers and machines often have to be moved into position (an advance payment is also referred to as a mobilisation payment) or work has to be undertaken on pre-fabricated units. To enable all this to happen the contracting company receives an advance payment.
It is normally an amount of between 10% and 20% of the total contract and is reduced against certificates from some relatively neutral body, such as a firm of consulting engineers or architects. It is evident that the advance payment bond comes into force only, when the advance payment has actually been paid into the account of the applicant.

To guarantee the repayment of the prepayment in the case the contract does not come into force or the applicant does not fulfil his contractual obligations. ABON is recognized as a valid product type for exposure type TRBE. The following definition are applicable within Rabobank.

- Applicant: the contractor or exporter
- Guarantor: the bank of the contractor/exporter as issuing bank
- Intermediary: the bank of buyer/importer;
- Beneficiary: the buyer/importer or main contractor.

8.3.4 GUAG

A GUAG is a Guarantee Given. This is a transaction-related bank guarantee, being a unilateral commitment of the bank to pay, on behalf of our customer, a specific amount of money to a beneficiary, if the client does not meet certain predetermined obligations from the transaction towards the beneficiary, e.g. observance of a purchase agreement.

8.3.5 CURR

A CURR is a Current Account Overdraft. An account overdraft (facility) is an arranged between the bank and its client whereby the bank permits the client to make payments in varying amounts against its current account up to a previously agreed overdraft limit and subject to previously agreed conditions.

8.3.6 LCST

A standby letter of credit (LCST) is furnished as security to a beneficiary and not by way of payment. The security is in the form of a promise by the issuing bank only when the applicant defaults on its obligations. Payment under the standby letter of credit is against a certificate (prescribed in the standby letter of credit) presented by the beneficiary, stating that default has occurred.
The standby letter of credit can be issued as a performance letter of credit, in which case it is identical to a performance bond.
Standby letters of credit can be used to guarantee the execution of an obligation by the applicant. Parties involved are: Applicant, issuing bank, advising/negotiating bank, and beneficiary.

Rabobank policy on LCST are as follows:

- Automatic extension (evergreens) of multi-year standby letters of credit is not allowed;
- Standby letters of credit must be irrevocable;
- Standby letters of credit must refer to the UCP rules;
- A facility under which standby letters of credit can be issued should be in place;
- The application has to clarify within what tenor possible drawings have to be repaid.

8.3.7 BILF & BILT

The risk that the bill is not paid on the expiry date is taken over by the bank of the presenter. The bank now waives its right of recourse against the presenter.

In fact the bank buys the bill from the presenter and so accepts the risk against a foreign country. The bank is only willing to do this if payment of the bill of exchange on the expiry date is guaranteed in one way or another (aval of the bank of the drawee or a bank guarantee).
If the ‘drawee’ and the guarantor are in default, the bank cannot have recourse against the presenter of the bill of exchange. BILF and BILT is recognized as a valid product type for exposure type TRBE.

8.3.8 CALL

A call loan is a loan granted for an indefinite period, which can be cancelled every day before 12.00 a.m. by either the lender or the borrower, in which case it must be repaid within two working days of its cancellation. Call loans are not the same as day-to-day deposits. Call loans are used for temporary financing needs and it is repaid in a single payment after the loan has been cancelled by either of the parties.

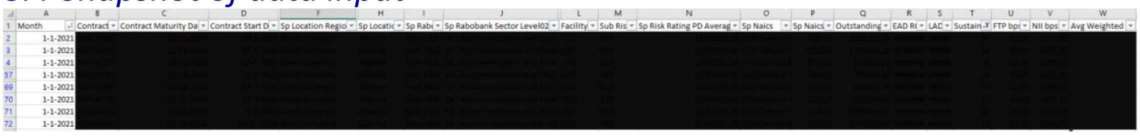
8.3.9 IMPO

An import letter of credit is a L/C concerning an import transaction. The importer has an account with a domestic bank, the exporter has an account with a foreign bank. The importer orders the bank to issue a L/C. In the L/C the domestic bank commits itself to pay the exporter when the latter submits the correct documents.
The issuing bank issues an import L/C. Usually a foreign bank (mostly the bank of the exporter) is asked to advise the L/C to the exporter.

8.3.10 FILE

A FILE is a Financial Lease. A lease considered to have the economic characteristic of asset ownership. Lessor has legal ownership. Lessee has option to buy.

8.4 Snapshot of data input



	A	B	C	D	F	H	I	J	L	M	N	O	P	Q	R	S	T	U	V	W
1	Month	Contract	Contract Maturity Da	Contract Start D	Sp Location Regio	Sp Locatic	Sp Rab	Sp Rabobank Sector Level02	Facility	Sub Ris	Sp Risk Rating PD Averag	Sp Naics	Sp Naics	Outstanding	EAD R	LAD	Sustain	FTP bp	NII bps	Avg Weighted
2	1-1-2021	[redacted]																		
3	1-1-2021	[redacted]																		
4	1-1-2021	[redacted]																		
57	1-1-2021	[redacted]																		
69	1-1-2021	[redacted]																		
70	1-1-2021	[redacted]																		
71	1-1-2021	[redacted]																		
72	1-1-2021	[redacted]																		

Figure 8-8: Small snapshot of dataset (hidden due to confidentiality)

8.5 Asia results

We will elaborate on the results of standardized beta coefficients within the region Asia. We firstly provide the adjusted R-squared, which is the amount the model explains, after which we will take a deeper dive into the different grouped and ungrouped variables. As can been seen from Figure 8-10, the model has a high predictive power. While all variables are grouped, we can explain 36% of the variation in the NII bps. However, ungrouping the variables, we are able to explain up to 76% of the variation in the NII bps. Thus, our model can explain a significant amount of the NII bps with the used independent variables.

Region	Iteration	Dataset count	R	R Square	Adjusted R Square
Asia	Risk rating and NAICS ungrouped	7573	0,87	0,76	0,76
Asia	Risk rating and facility currency ungrouped	7426	0,86	0,73	0,73
Asia	Risk rating and location ungrouped	7582	0,86	0,73	0,73
Asia	Risk rating ungrouped & NAICS 2	7571	0,85	0,73	0,73
Asia	Risk rating ungrouped & NAICS 3	7571	0,85	0,73	0,73
Asia	All variables grouped	7514	0,60	0,36	0,36

Figure 8-10: Adjusted R-squared for each iteration region Asia.

8.5.1 All variables grouped

In Figure 8-11, we see similar results for region Asia compared to the benchmark model. For the variables close to the benchmark, similar reasoning can be applied as explained in section 4.1.1. We see some small differences in the risk ratings grouped and location. Moreover, we note that LAD has a negative correlation with the NII bps compared to a small positive influence while using the benchmark model. Secondly, we note that FTP bps is of no influence in Asia specific. Nevertheless, we observe a significant difference in the facility currency between the benchmark model and Asia, thus the currency which is used to pay in the transactions dataset. In our sensitivity analysis, we take a deeper look into the currencies to find underlying relationships.

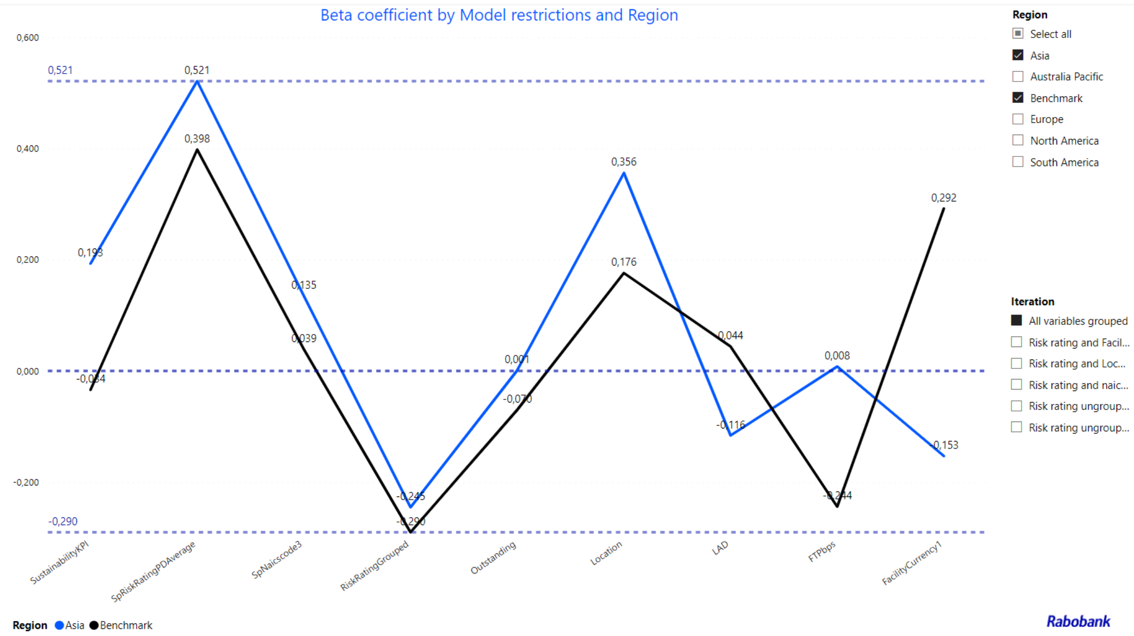


Figure 8-11: Standardized beta coefficient MLR Asia.

8.5.2 Sensitivity analysis

We further investigate the ungrouped variables such as risk ratings, location and currencies for Asia region specific. By analyzing the ungrouped version of theses variables, we hope to find additional underlying information which otherwise would have gone unnoticed. Note that observed missing data points in the figures of regions specific charts are caused due to the benchmark model is applied on the entire dataset. Region specific MLR's is run on the dataset corresponding to the respectively region and thus some variables may not be applicable for the region. E.g. a region may contain only a few risk ratings, sectors or currencies.

8.5.2.1 Risk rating ungrouped and NAICS2

Ungrouping the risk ratings already explains some of the observed results found when all variables are grouped. As can be seen from Figure 8-12, we observe a high positive influence of the D3 and D4 rating in Asia. Obviously, this deviates from the benchmark concerning the risk rating but nevertheless it explains the significant negative influence of the LAD. Higher risk ratings, and especially default ratings, result in a higher LAD amount. This will impact the NII bps as the probability of a real defaults is extremely high in D3 and D4. Companies in D3 and D4 go into a special process to help them repay their debts. While in this process, the companies pay what they can and thus lower NII bps can observed.

Some companies may move from D4 back to D3 which is usually paired by significant payables made on the loan and thus explaining the high positive correlation with the NII bps.

Additionally, we see that NAICS codes have a higher positive influence in Asia compared to the benchmark. This will be further analyzed in the following sensitivity analysis.

In conclusion, due to the high risk ratings, e.g. D3 and D4, the Rabobank does not obtain the agreed Interest Income, a higher LAD is booked and thus LAD has a negative correlation with the NII bps. Companies moving from D4 to D3 ratings have an increased amount in debts repaid and thus explains the positive correlation with the NII bps. Finally, as mentioned a few times, higher credit risk results in higher NIM.

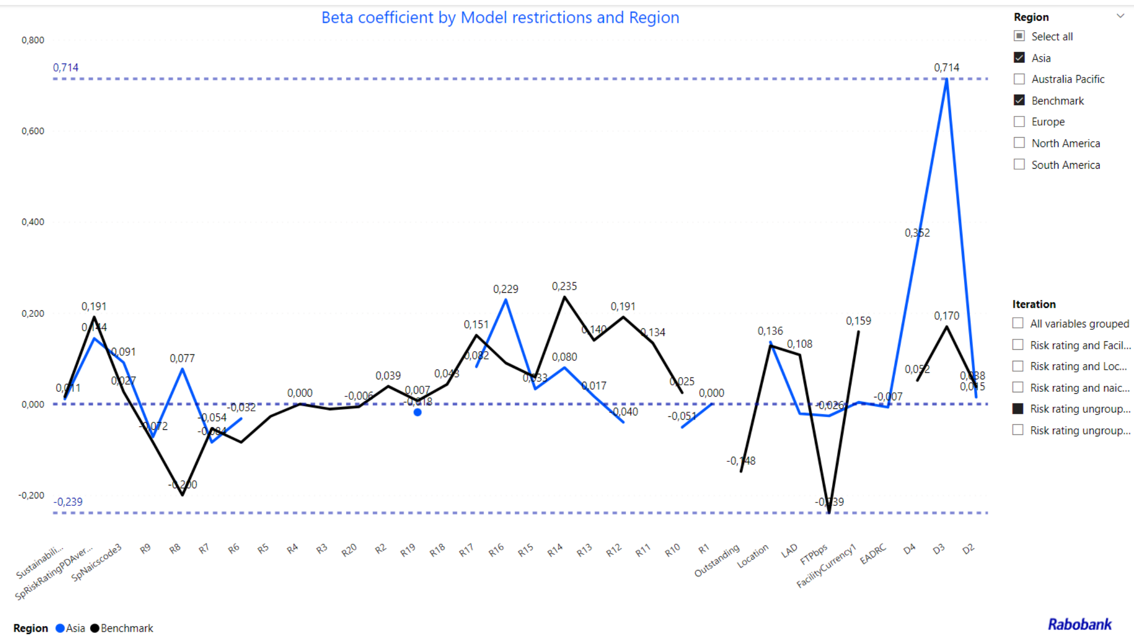


Figure 8-12: Risk rating ungrouped and NAICS2 Asia.

8.5.2.2 Risk rating and Facility currency ungrouped

In Figure 8-13, we see the difference in standardized beta coefficient between the Benchmark model and the MLR for Asia specific. As we noted a significant difference between the facility currency in Section 8.5.1 and thus we are interested by which currencies this is caused. As one may notice, all but one currency have a small negative correlation with the NII bps. These are NZD, HKD, AUD and RMB with a small negative influence. Only the INR has a small positive correlation with the NII bps. As explained in Section 4.1.2.2, this is due to the high inflation in the country and thus higher inflation, implies higher risk which results in higher NIM.

We however now also observe a high positive influence of the locations on the NII bps. We will dive deeper into the location in the following section.

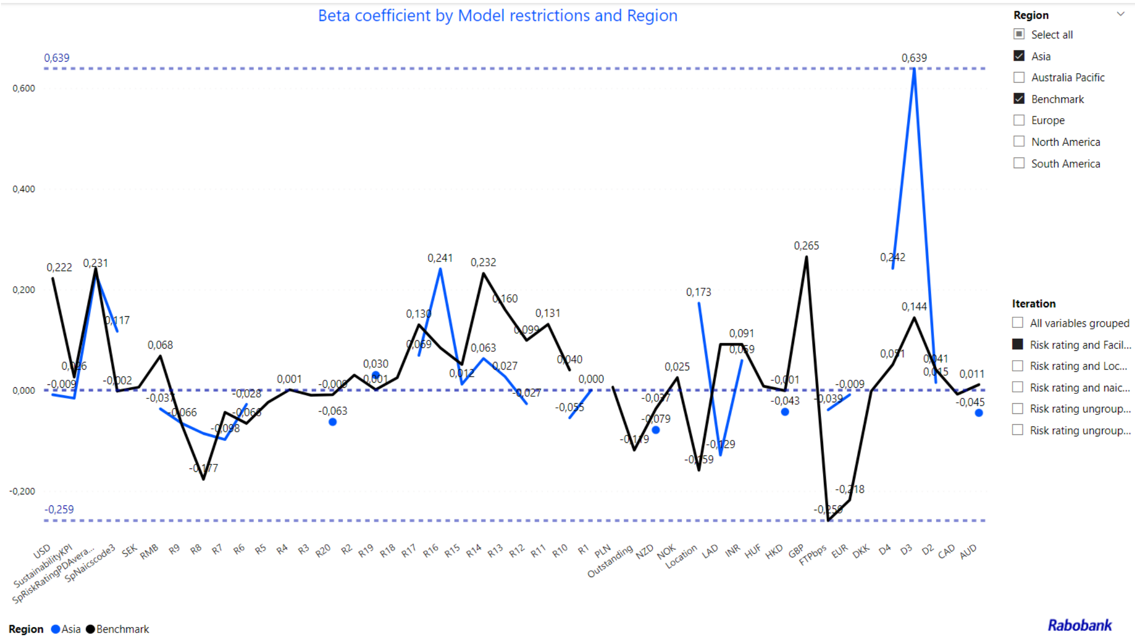


Figure 8-13: Risk rating and Facility currency ungrouped Asia.

8.5.2.3 Risk rating and Location ungrouped

While running the MLR with locations and risk rating ungrouped we notice a significant difference between the two Asia specific region, as depicted in Figure 8-14: Risk rating and Location ungrouped Asia. We notice that Singapore has a high positive correlation with the NII bps whereas Hong Kong has small negative influence. Thus, for Wholesale business, if they have the opportunity to choose between these two location, one should give the preference to Singapore. Singapore is a city where many development opportunities arise and thus explains the high positive correlation with the NII bps. Hong Kong has been subject to many diplomatic problems with tension between China and the rest of the world. Although this is at play, we do not see a significant positive or negative correlation with the NII bps.

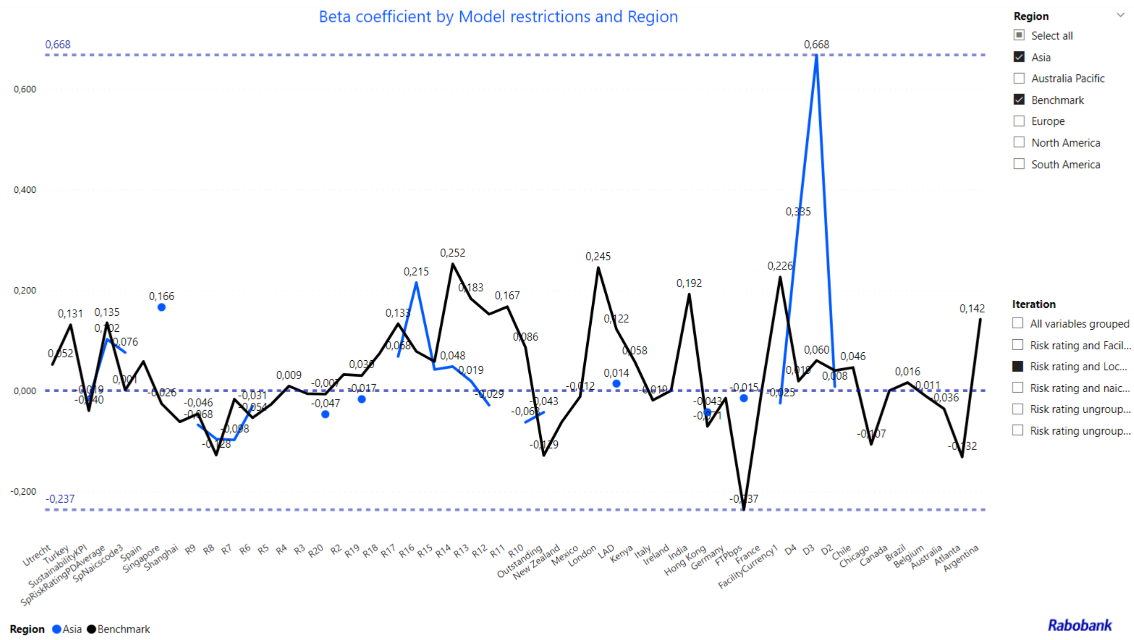


Figure 8-14: Risk rating and Location ungrouped Asia.

8.5.2.4 Risk rating and NAICS ungrouped

Figure 8-15 depicts the risk ratings and NAICS code ungrouped for the Benchmark and Asia MLR models. Here, the other variables are included as well. Figure 8-16 only depicts the risk ratings for both the Benchmark as the Asia model. Here again we see the obvious relationship that higher risks results in higher returns and thus a higher NII bps. Nevertheless, we see a higher positive significant relationship caused by the D3 and D4. Moreover, from Figure 8-16, we see that risk rating D3 has a high count within the dataset. In is interesting to note that D4 ratings only have a few but still have a significant positive correlation with the NII bps.

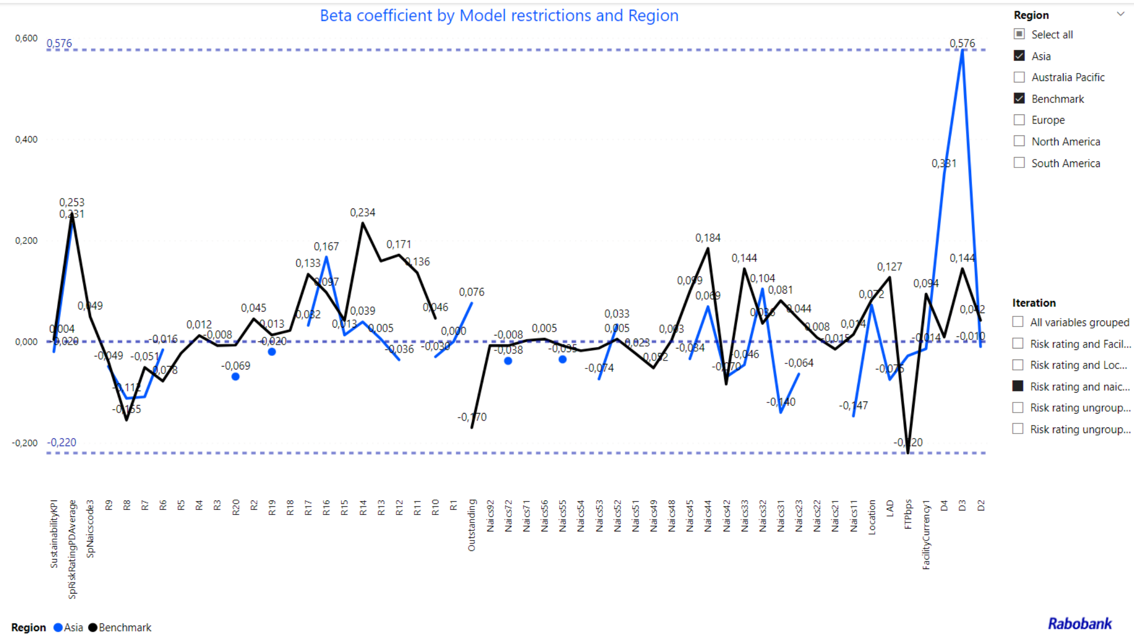


Figure 8-15: Risk rating and NAICS ungrouped Asia.

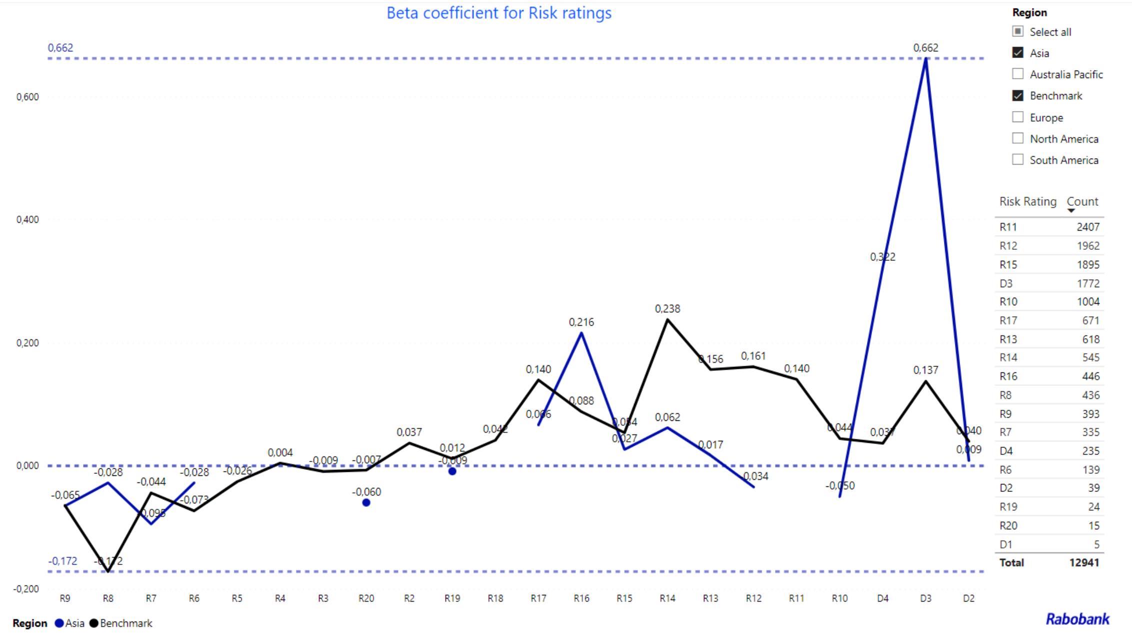


Figure 8-16: Asia Risk ratings.

Figure 8-17 gives us a better insight into the influences of sectors on the NII bps for region Asia versus the benchmark. Where in Section 4.1.2.4 we noted a positive influence of the manufacturing sector, we see that these sectors have a negative correlation with the NII bps in Asia. Thus, one should restrict the long-term loans offered for these sectors in Asia. This also aligns with a macro-economic trend we have seen where businesses move their manufacturing back to western countries due to vocalized human rights issues.

Moreover, we note that the sector Agriculture, Forestry, Fishing and Hunting, which corresponds to a NAICS code of 11, have a high negative correlation with the NII bps. In the benchmark we observed that this sector had no significant impact on the NII bps however it is of interest for Asia. Thus, this is also a sector to be avoided in Asia.

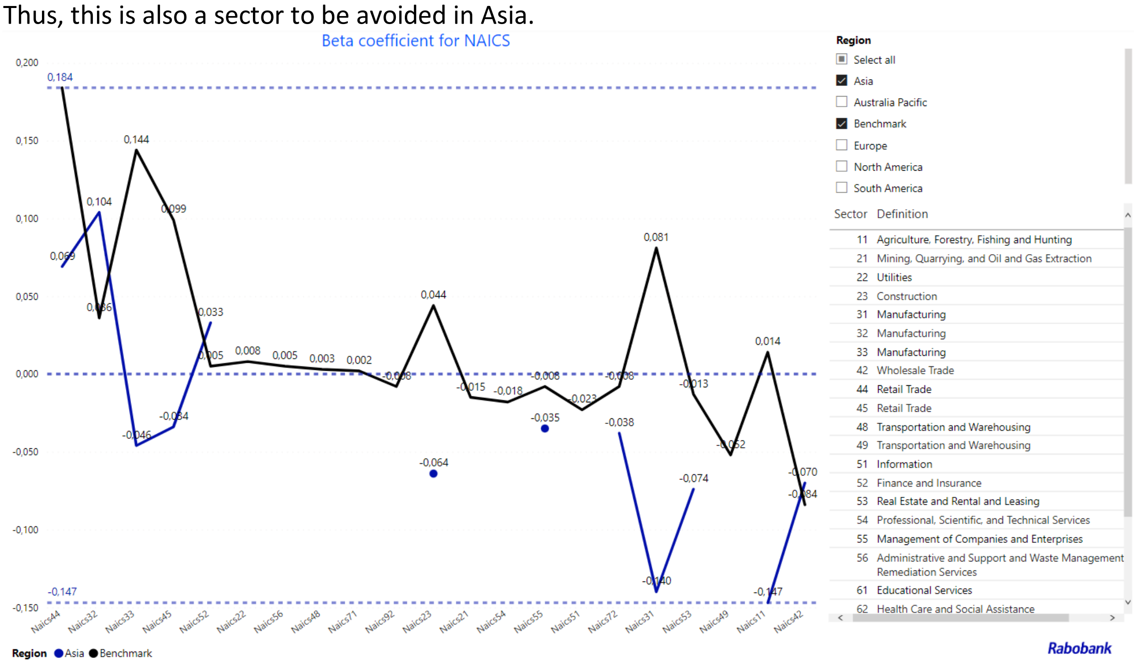


Figure 8-17: Asia NAICS.

8.6 Australia Pacific results

Sadly, the data and the applied model did not explain any significant proportion of the variation in the NII bps. As can be seen from Figure 8-18, each iteration applied to region Australia Pacific resulted in an adjusted R-squared of 0.11 to 0.16. Thus, we were only able to explain 16% of the variation in the NII bps for Australia Pacific. As one may note, there is a lower count of transaction for this region. Moreover, there is a lack of difference within the dataset which makes it difficult for our model to explain underlying relationships.

Region	Iteration	Dataset count	R	R Square	Adjusted R Square
Australia Pacific	Risk rating and NAICS ungrouped	5785	0,41	0,17	0,16
Australia Pacific	Risk rating ungrouped & NAICS 2	5784	0,40	0,16	0,16
Australia Pacific	Risk rating ungrouped & NAICS 3	5784	0,40	0,16	0,16
Australia Pacific	Risk rating and facility currency ungrouped	5784	0,40	0,16	0,16
Australia Pacific	Risk rating and location ungrouped	5773	0,36	0,13	0,13
Australia Pacific	All variables grouped	5769	0,34	0,11	0,11

Figure 8-18: Adjusted R-squared for each iteration region Australia Pacific.

As can be seen from Figure 8-19, in 2021 and 2022 most risk rating of the transaction were of risk rating R10 and hereafter a few other risk rating.

Risk Rating	Count
R10	2374
R12	958
R9	862
R11	777
R7	434
R8	403
R16	260
R13	174
R14	87
R15	45
D2	20
D4	14
Total	**6408**

Figure 8-19: Risk rating count Australia Pacific.

In conclusion, we decided to ignore the results for the region Australia Pacific. Figure 8-20, Figure 8-21, Figure 8-22, Figure 8-23, and Figure 8-24 give a visualization the standardized beta coefficient in each iteration. In the future, as the dataset expands over time, another analysis can be applied on this region to see if a higher adjusted R-squared is achieved and thus a higher variation of the NII bps can be explained.

8.6.1 All variables grouped

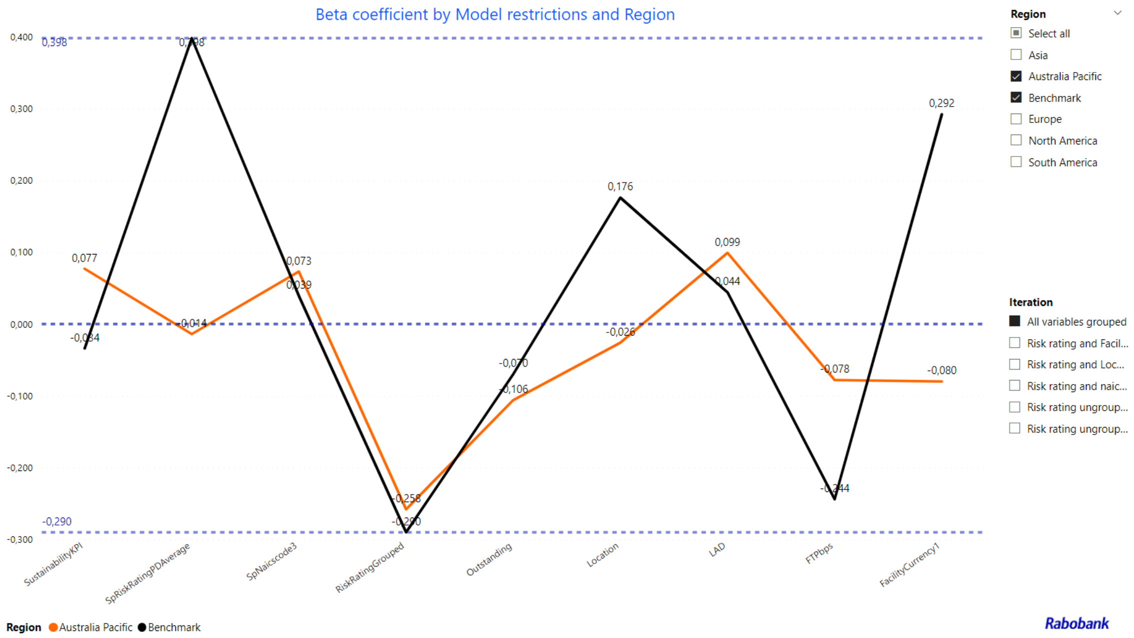


Figure 8-20:Standardized beta coefficient MLR Australia Pacific.

8.6.2 Sensitivity analysis

Again, we will skip the following sensitivity analysis for Australia Pacific due to the low observed adjusted R-squared. Note that observed missing data points in the figures of regions specific charts are caused due to the benchmark model is applied on the entire dataset. Region specific MLRs are run on the dataset corresponding to the respectively region and thus some variables may not be applicable for the region. Exempli gratia, a region may contain only a few risk ratings, sectors or currencies.

8.6.2.1 Risk rating ungrouped and NAICS2

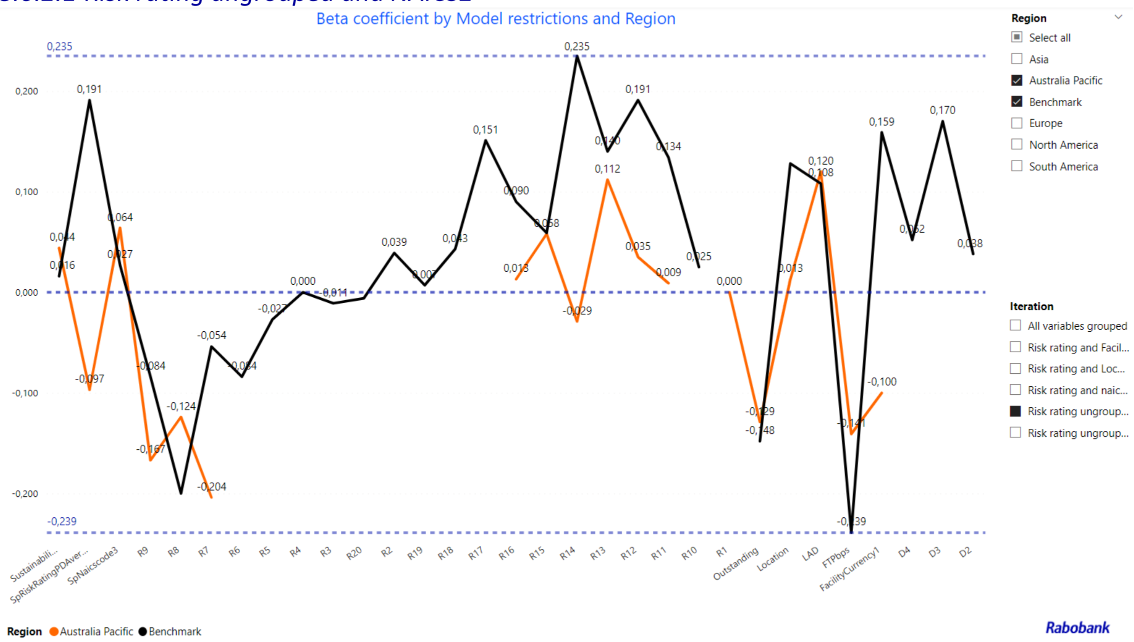


Figure 8-21: Risk rating ungrouped and NAICS2 Australia Pacific.

8.6.2.2 Risk rating and Facility currency ungrouped

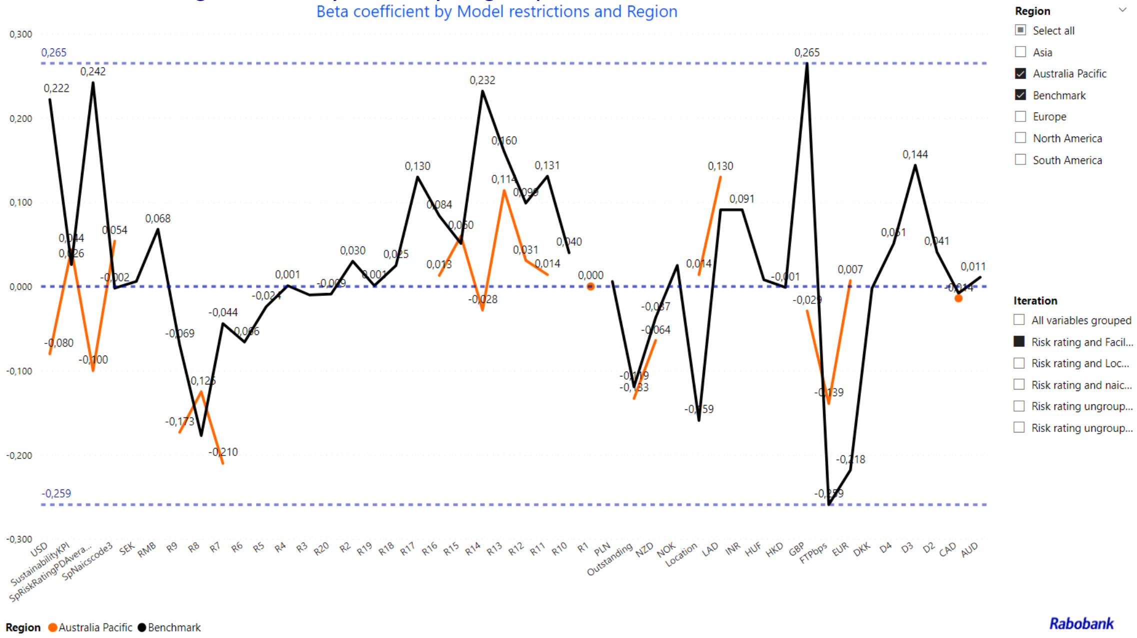


Figure 8-22: Risk rating and Facility currency ungrouped Australia Pacific.

8.6.2.3 Risk rating and Location ungrouped

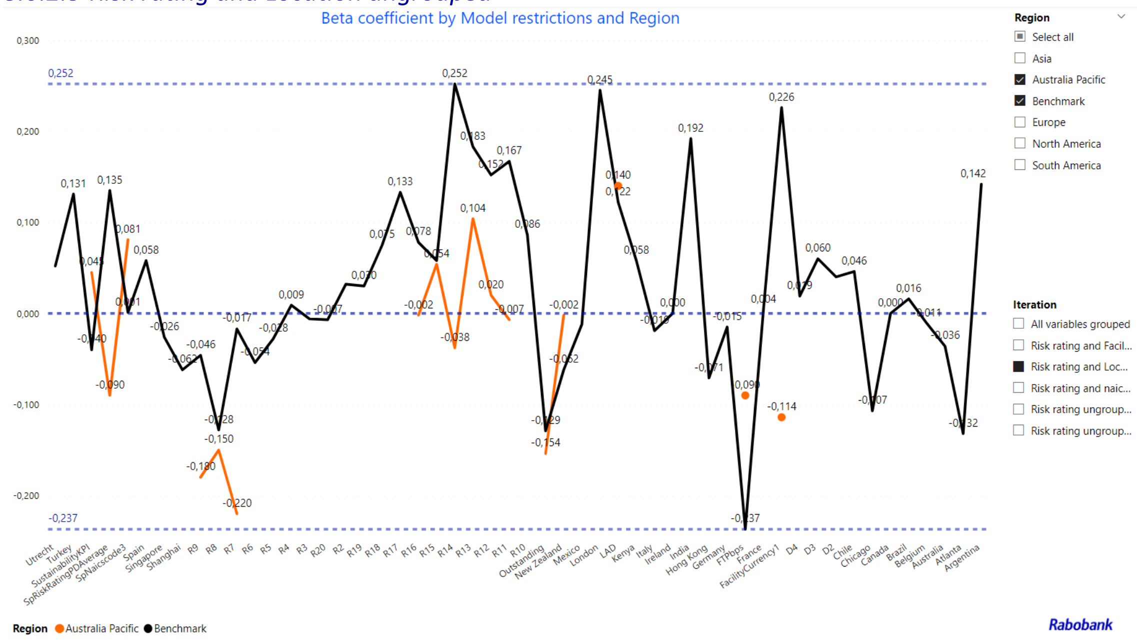


Figure 8-23: Risk rating and Location ungrouped Australia Pacific.

8.6.2.4 Risk rating and NAICS ungrouped

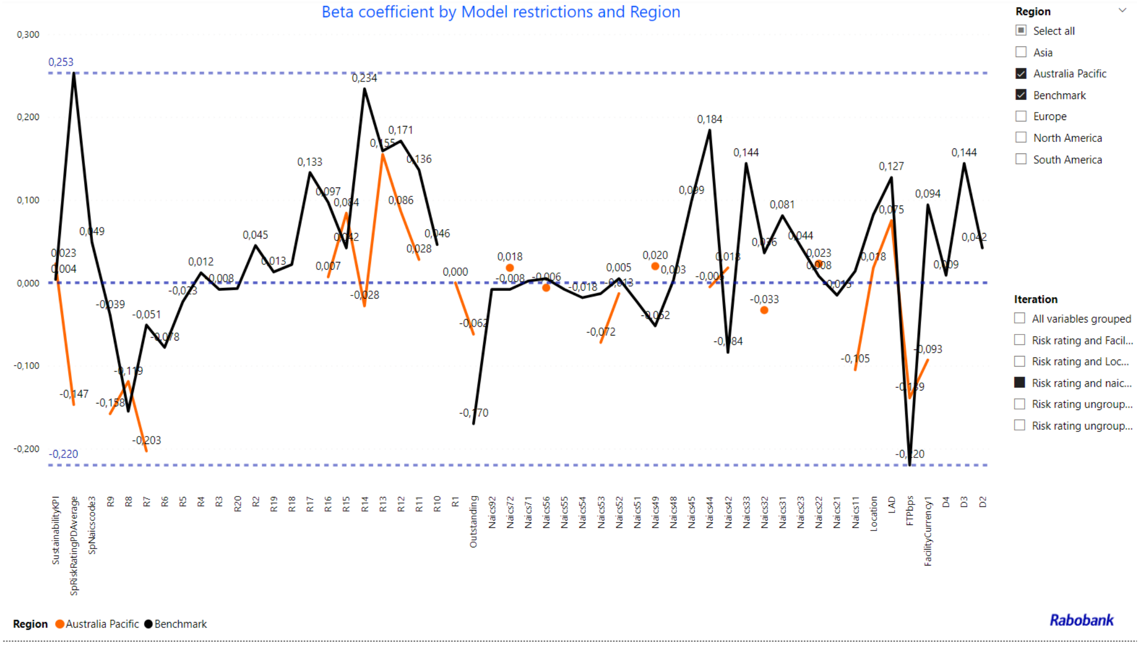


Figure 8-24: Risk rating and NAICS ungrouped Australia Pacific.

8.7 Europe results

We elaborate on the results of standardized beta coefficients within the region Europe. We firstly provide the adjusted R-squared, which is the amount the model explains, after which we take a deeper dive into the different grouped and ungrouped variables. As can be seen from Figure 8-25, the model has a high predictive power. While all variables are grouped, we can explain 46% of the variation in the NII bps. However, ungrouping the variables, we are able to explain up to 62% of the variation in the NII bps. Thus, our model can explain a high significant amount of the NII bps with the used independent variables.

Region	Iteration	Dataset count	R	R Square	Adjusted R Square
Europe	Risk rating and NAICS ungrouped	93055	0,79	0,62	0,62
Europe	Risk rating and facility currency ungrouped	93034	0,79	0,62	0,62
Europe	Risk rating and location ungrouped	92943	0,78	0,61	0,61
Europe	Risk rating ungrouped & NAICS 2	93194	0,76	0,58	0,58
Europe	Risk rating ungrouped & NAICS 3	93194	0,76	0,58	0,58
Europe	All variables grouped	93357	0,68	0,46	0,46

Figure 8-25: Adjusted R-squared for each iteration region Europe.

8.7.1 All variables grouped

In Figure 8-11, we see similar results for region Europe compared to the benchmark model. For the variables close to the benchmark, similar reasoning can be applied as explained in Section 4.1.1. We however see a significant difference between the variable outstanding in Europe compared to the benchmark. As Europe is a less risky region, lending out more money does not result in higher NII.

Secondly, we note that the LAD has a high positive correlation with the NII bps compared to the benchmark model. A higher LAD amount implies higher risk taken on by the Rabobank and thus higher risk results in higher NII. Moreover, we see observe a higher negative correlation with the NII bps caused by FTP bps, thus the CoF.

Finally, we note a significant difference between the standardized beta coefficient for the currency. In our sensitivity analysis, we will take a deeper look into the currencies to find underlying relationships.

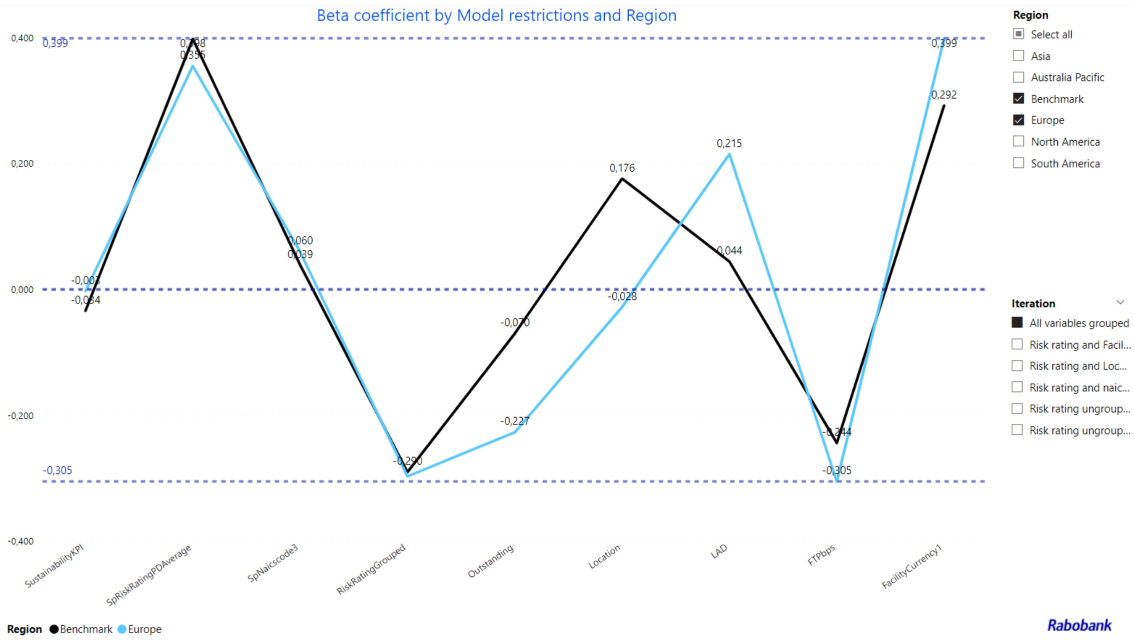


Figure 8-26: Standardized beta coefficient MLR Europe.

8.7.2 Sensitivity analysis

In the following section, we will further investigate the ungrouped variables such as risk ratings, location and currencies for Europe region specific. By analyzing the ungrouped version of theses variables, we hope to find additional underlying information which otherwise would have gone unnoticed. Note that observed missing data points in the figures of regions specific charts are caused due to the benchmark model is applied on the entire dataset. Region specific MLR's is run on the dataset corresponding to the respectively region and thus some variables may not be applicable for the region. Exempli gratia, a region may contain only a few risk ratings, sectors or currencies.

8.7.2.1 Risk rating ungrouped and NAICS2

From Figure 8-27, we see that Europe region follows the benchmark model closely. However, we see some small deviations within the risk ratings which will be analyzed in section 8.7.2.4. Once again, we see a significant difference in the variables FTP bps and FacilityCurrency in Europe. Both have a higher observed standardized beta coefficient compared to the benchmark model. As CoF are cheaper in Europe and less money is needed from interbank market, lower risk is observed and thus a negative correlation with the NII bps. The influence of different currencies will be analyzed in the following section.

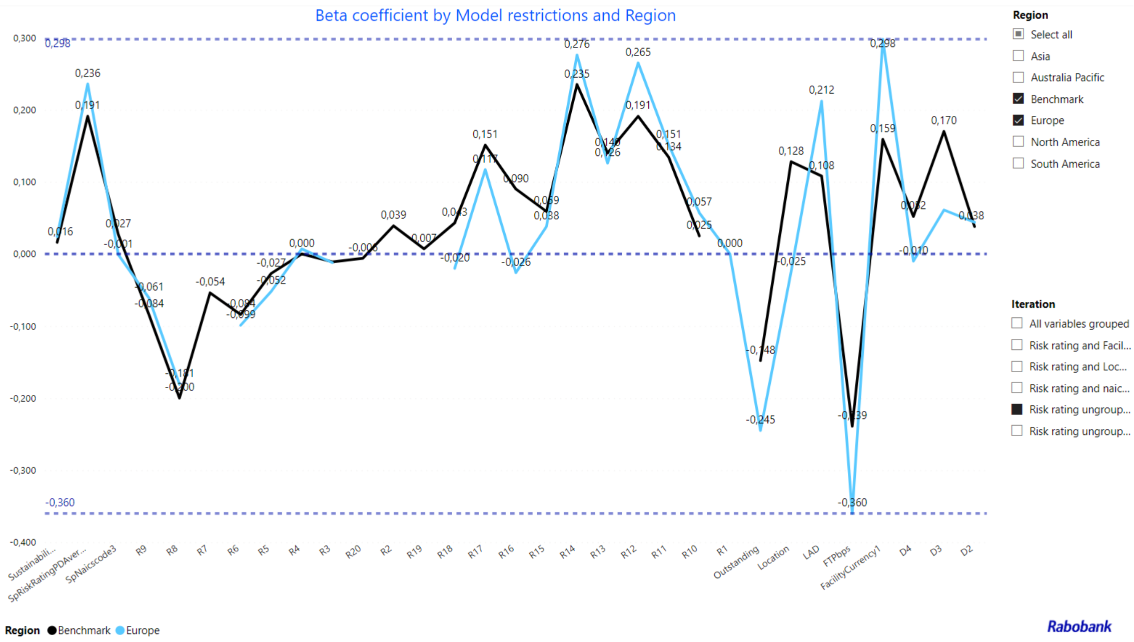


Figure 8-27: Risk rating ungrouped and NAICS2 Europe.

8.7.2.2 Risk rating and Facility currency ungrouped

From Figure 8-28, we see the influence of different currencies on the NII bps. Here we again note that the USD has high significant correlation with the NII bps as explained in section 4.1.2.2. However, we note that the RMB, the Chinese Yuan, also has a positive correlation with the NII bps. This can be explained by the spread between the EUR and the RMB where significant margins can be made.

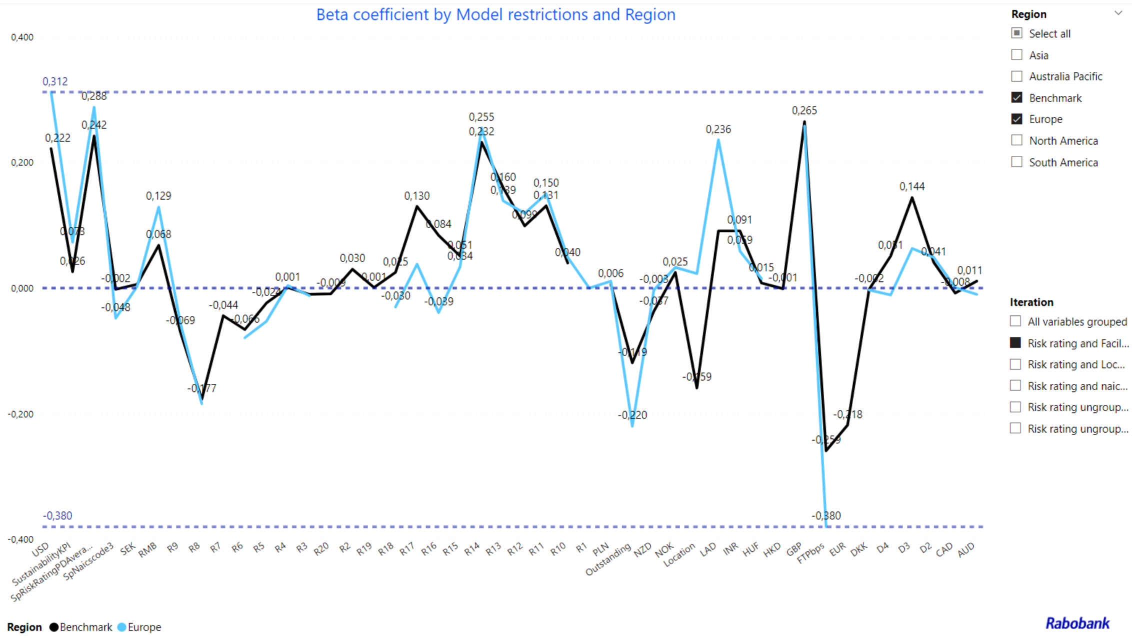


Figure 8-28: Risk rating and Facility currency ungrouped Europe.

8.7.2.3 Risk rating and Location ungrouped

From Figure 8-29, we do can only spot one location with significant correlation with the NII bps in Europe. This concerns the region Utrecht and is caused by a lot of facilities being booked in Utrecht. Moreover, after analyzing the dataset, we also see that these booking usually tend to have lower risk ratings and thus resulting in lower margins. One may also spot that Kenya is attributed to the region Europe. As this is the only country in Africa where wholesale lending is offered for now, it is currently still attributed to the region Europe.

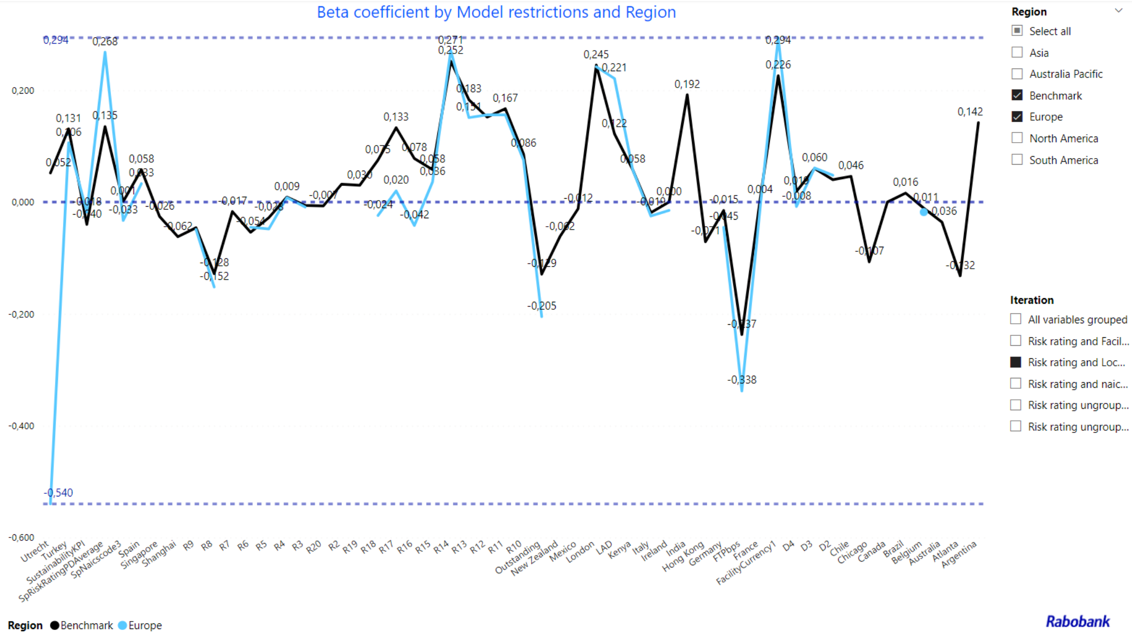


Figure 8-29: Risk rating and Location ungrouped Europe.

8.7.2.4 Risk rating and NAICS ungrouped

Figure 8-30 depicts all the variables with the risk ratings and NAICS codes ungrouped for the MLR model applied in Europe. One may find it hard to see obvious difference between the benchmark model and Europe. Overall, Europe tends to follow the benchmark model closely however some small differences can be spotted.

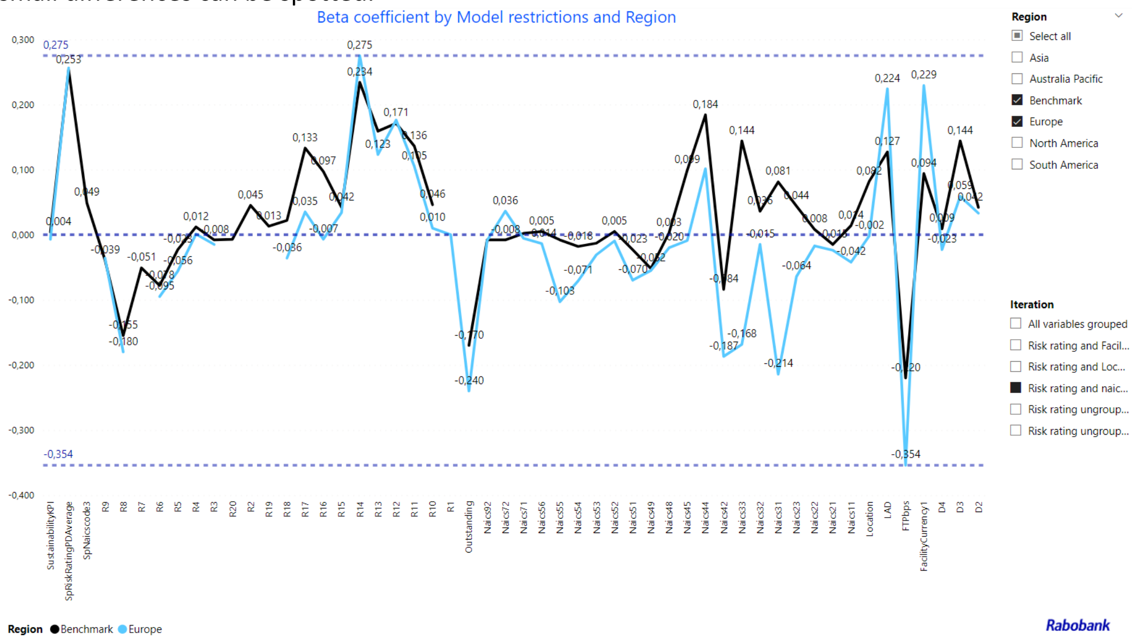


Figure 8-30: Risk rating and NAICS ungrouped Europe.

From Figure 8-31, we see the different risk rating for Europe specific versus the benchmark model. Firstly, we note that Europe overall has a lot of different risk ratings compared to other regions. We see some small deviation in risk ratings R16-R18 however this is caused by the lack of data for these regions. Overall, Europe follows the benchmark model where higher risk ratings results in higher margins and thus a higher NII bps can be observed.

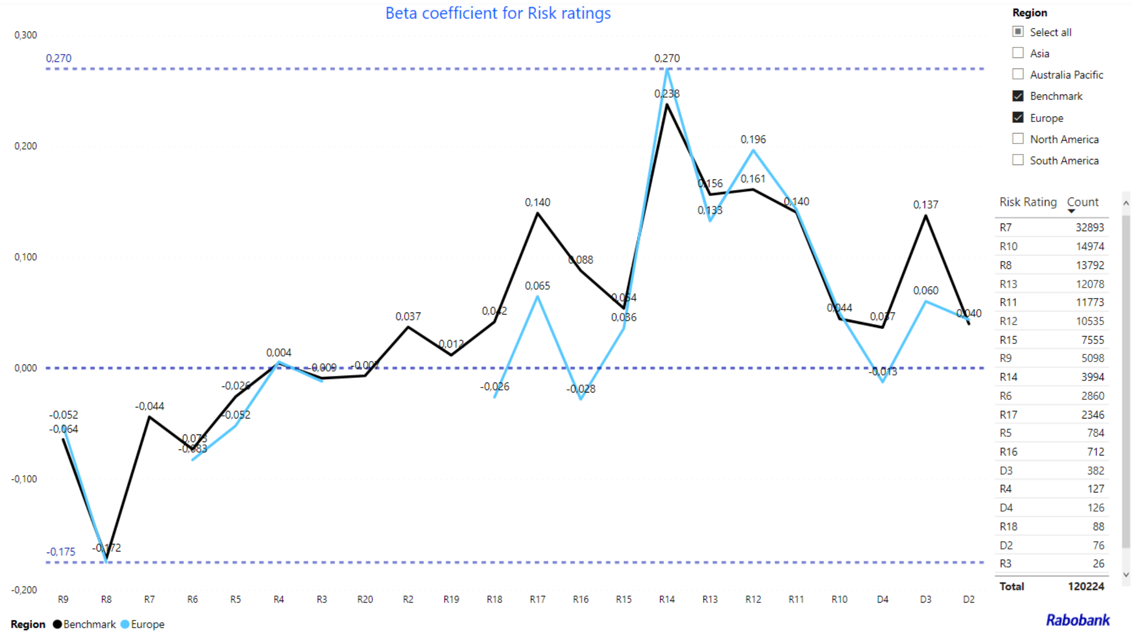


Figure 8-31: Europe Risk ratings.

From Figure 8-32, we see a more in depth analysis on the NAICS for Europe and the benchmark model. The figure gives us a clear visualization of the important sectors to be in and especially not to be in in Europe. As explained in Section 4.1.2.4, retail trade is an important sector to operate within as it results in a higher NII bps. However, in contrary to what is observed in the benchmark model, sectors Construction, manufacturing and Wholesale trade should be avoided in Europe. These sectors have been under less stress in Europe. Although some CO2 emission restrictions and higher labor costs are at play in Europe, this does not significantly put these sectors at higher risk in Europe. Moreover, in Europe, the Covid restrictions got lifted more gradually and quicker compared to other regions and thus enabling business to run a full capacity again. Risk rating and NAICS ungrouped

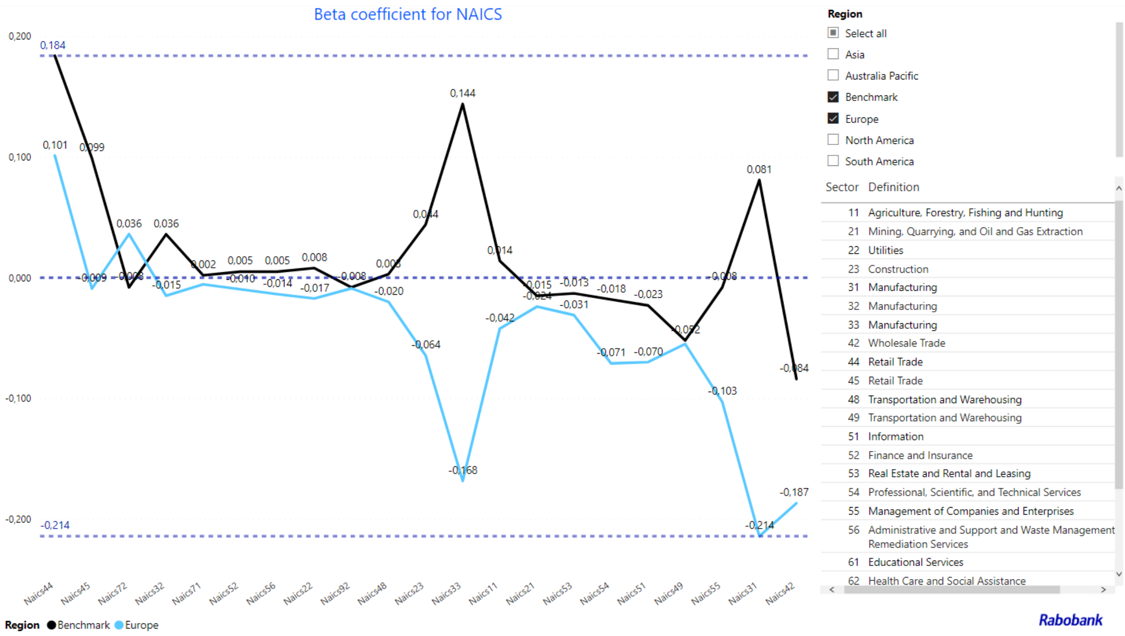


Figure 8-32: Europe NAICS.

8.8 North America results

In the following section, we will elaborate on the results of standardized beta coefficients within the region North America. We will firstly provide the adjusted R-squared, which is the amount the model explains, after which we will take a deeper dive into the different grouped and ungrouped variables. As can been seen from Figure 8-33, the model has a high predictive power. While all variables are grouped, we can explain 49% of the variation in the NII bps. However, ungrouping the variables, we are able to explain up to 60% of the variation in the NII bps. Thus, our model can explain a high significant amount of the NII bps with the used independent variables.

Region	Iteration	Dataset count	R	R Square	Adjusted R Square
North America	Risk rating and NAICS ungrouped	14819	0,78	0,60	0,60
North America	Risk rating and facility currency ungrouped	14830	0,75	0,57	0,57
North America	Risk rating and location ungrouped	14832	0,75	0,57	0,57
North America	Risk rating ungrouped & NAICS 3	14828	0,65	0,57	0,57
North America	Risk rating ungrouped & NAICS 2	14830	0,75	0,57	0,56
North America	All variables grouped	14833	0,70	0,49	0,49

Figure 8-33: Adjusted R-squared for each iteration region North America.

8.8.1 All variables grouped

In Figure 8-11, we see similar results for region North America compared to the benchmark model. For the variables close to the benchmark, similar reasoning can be applied as explained in Section 4.1.1. We however see a significant difference between the variable currency in North America compared to the benchmark. We will need to ungroup the currencies in order to find the underlying variables which have a negative correlation with the NII bps.

Finally, we note that PD has a higher positive correlation with the NII bps and the risk rating grouped a small lower negative influence. We will take a closer look to these variables in the following sensitivity analysis.

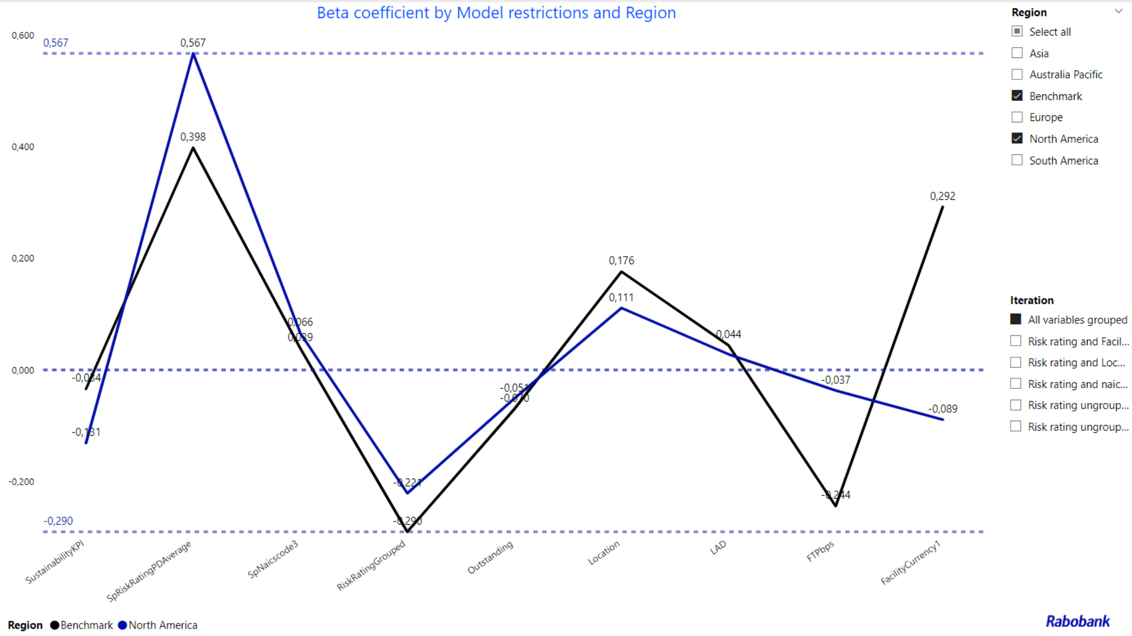


Figure 8-34: Standardized beta coefficient MLR North America.

8.8.2 Sensitivity analysis

We further investigate the ungrouped variables such as risk ratings, location and currencies for North America region specific. By analyzing the ungrouped version of theses variables, we hope to find additional underlying information which otherwise would have gone unnoticed. Note that observed missing data points in the figures of regions specific charts are due to the benchmark model applied on the entire dataset. Region specific MLRs are run on the dataset corresponding to the respectively region and thus some variables may not be applicable for the region. E.g. a region may contain only a few risk ratings, sectors or currencies.

8.8.2.1 Risk rating ungrouped and NAICS2

Figure 8-35 depicts the standardized beta coefficients results from the MLR model for North America and the Benchmark. As one may notice, we see significant difference in the risk ratings R17 to R19 in North America compared to the benchmark model. These risk rating are close to the default rating and evidently higher credit risk results in higher NII for the Rabobank.

It is more interesting to note that once ungrouping the risk rating, the standardized beta coefficient for the average PD has flipped from a significant positive standardized beta coefficient to a significant negative one. From Figure 8-39, we see that most risk ratings within the transaction dataset are of lower risk ratings in North America. As these are lower risks, and thus the PD is lower, both yield a negative correlation with the NII bps.

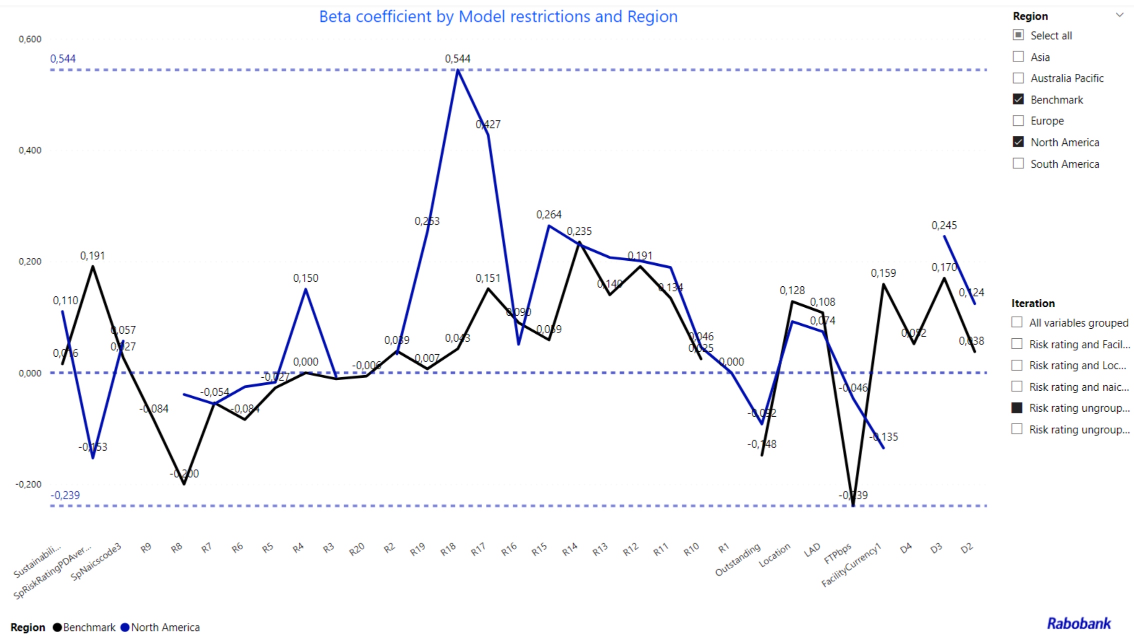


Figure 8-35: Risk rating ungrouped and NAICS2 North America.

8.8.2.2 Risk rating and Facility currency ungrouped

From Figure 8-36 we only notice a small difference for CAD between the benchmark and North America MLR. We note that CAD has a higher positive influence in North America whereas in the benchmark model no significant influence can be observed. As can be seen from Figure 8-37 in the following section, we observe the location Canada to be of positive correlation with the NII bps. Thus, one can expect that the currency of that location is also of interest.

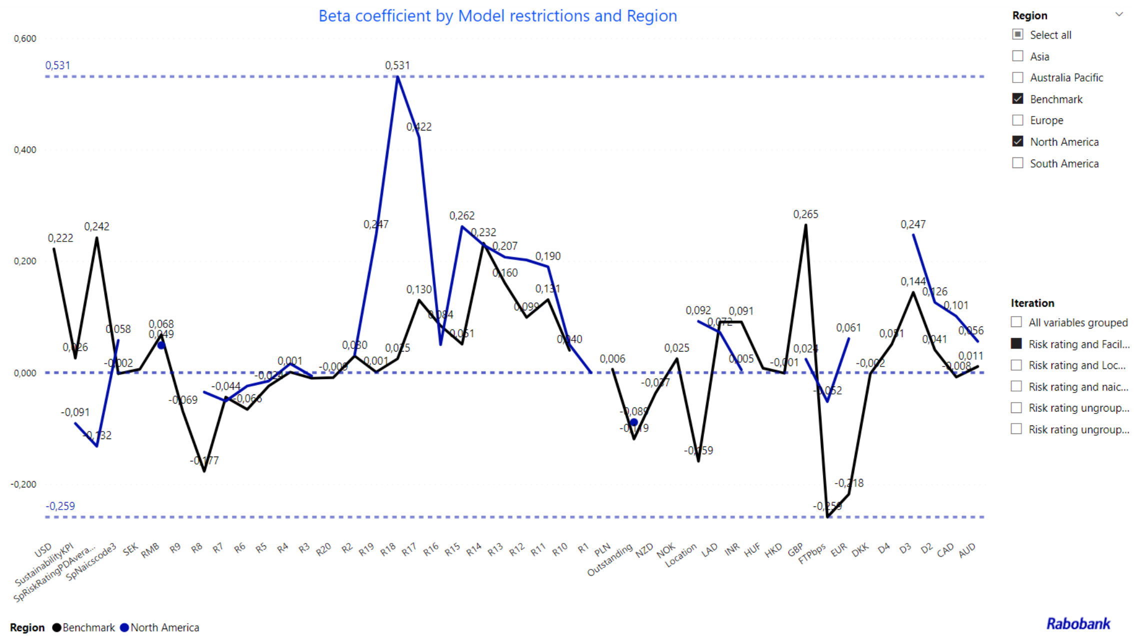


Figure 8-36: Risk rating and Facility currency ungrouped North America.

8.8.2.3 Risk rating and Location ungrouped

From Figure 8-37 we notice that both Canada and Mexico have a positive correlation with the NII bps. Both these locations have seen a significant increase in inflation during the selected period. Evidently, higher inflation demands for a higher base rate where CoF is not always that expensive and thus results in higher margins.

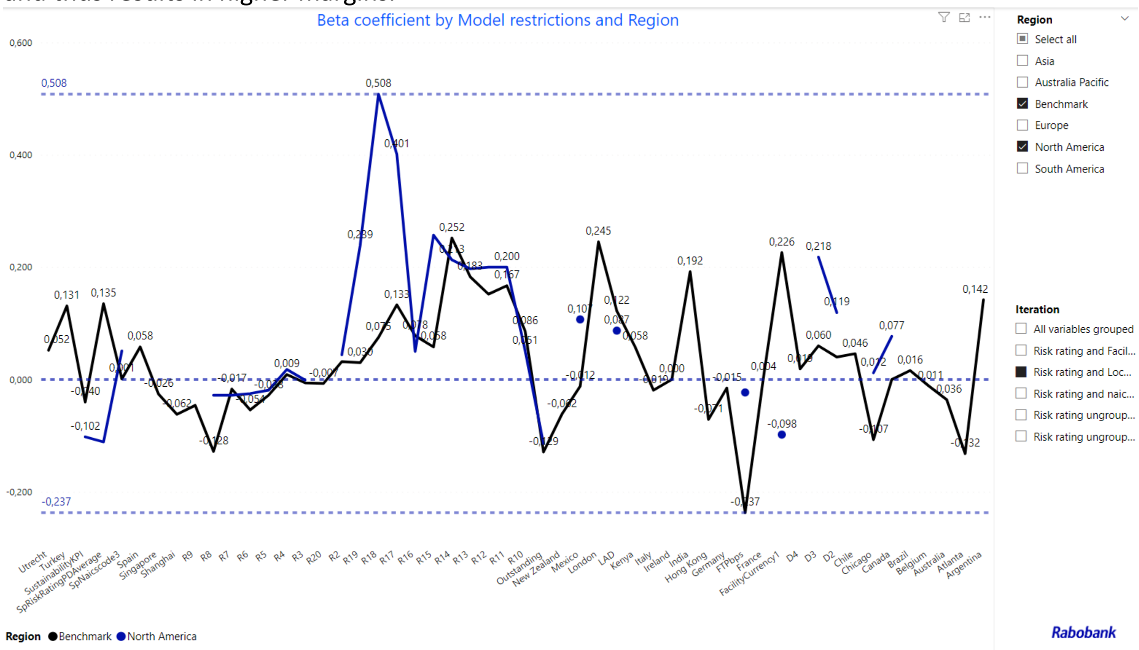


Figure 8-37: Risk rating and Location ungrouped North America.

8.8.2.4 Risk rating and NAICS ungrouped

As explained in Section 8.8.2.1, we see that most risk ratings within the transaction dataset are of lower risk ratings in North America as depicted in Figure 8-39. As these are lower risk, and thus the PD is lower, both yield a negative correlation with the NII bps. However, we see that the risk rating R17 to R19 yield a significant positive correlation with the NII bps. Although close to the default ratings, these companies seem to be able to pay the interest on their loan and as they have higher credit risks, this yields higher NII bps for the Rabobank

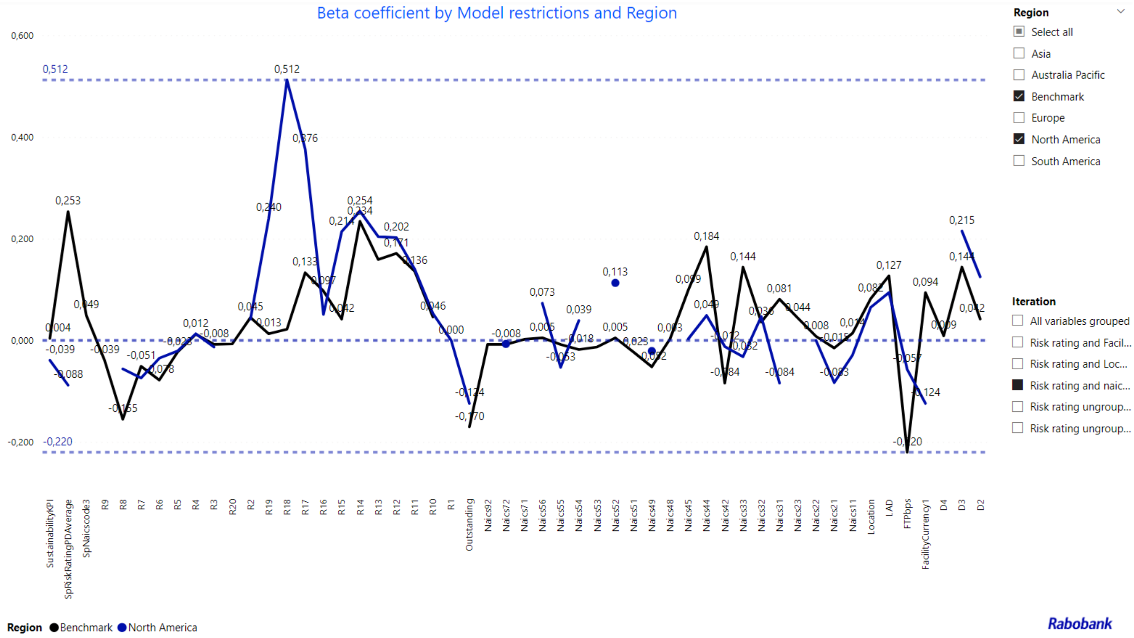


Figure 8-38: Risk rating and NAICS ungrouped North America.

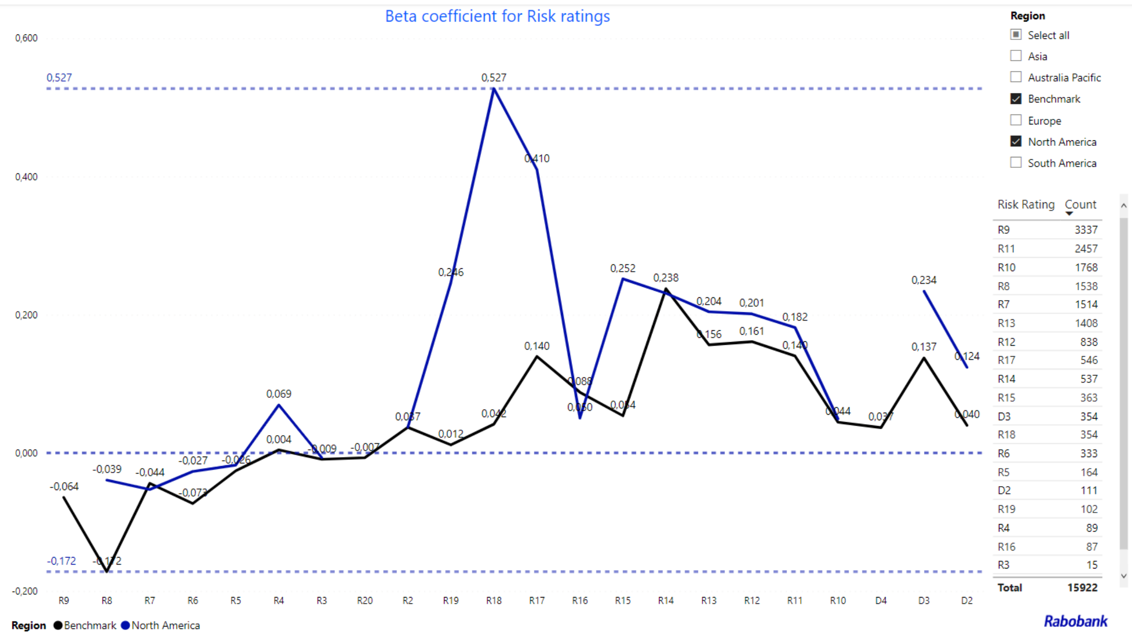


Figure 8-39: North America Risk ratings.

From Figure 8-40, we notice that some sectors should be avoided, similar to Europe. Firstly, manufacturing should be avoided in North America as this has a significant negative correlation

with the NII bps. Moreover, the sector “Mining, Quarrying, and Oil and Gas Extraction” should be avoided as well due to their negative correlation with the NII bps. This sector is on high demand due to the war in Ukraine and the restriction posed on Russian oil amongst others. We see an increase in crude oil production in the North America (Administration, 2023).
Additionally, we note that the sector manufacturing does not have any significance influence in North America on the NII bps compared to the benchmark. Thus this sector is not of interest when focusing on NII bps.
However, sectors with NAICS code 52 and 56 are of interest as they have a positive significant correlation with the NII bps. These are sectors “Finance and Insurance” and “Administrative and Support and Waste Management Remediation Services” respectively. After Covid lockdown, we have seen a lot of quantative easing in America which resulted in a lot of money inflow to certain sectors. These sectors have taken on higher risk and higher risk results in higher margin for the Rabobank.

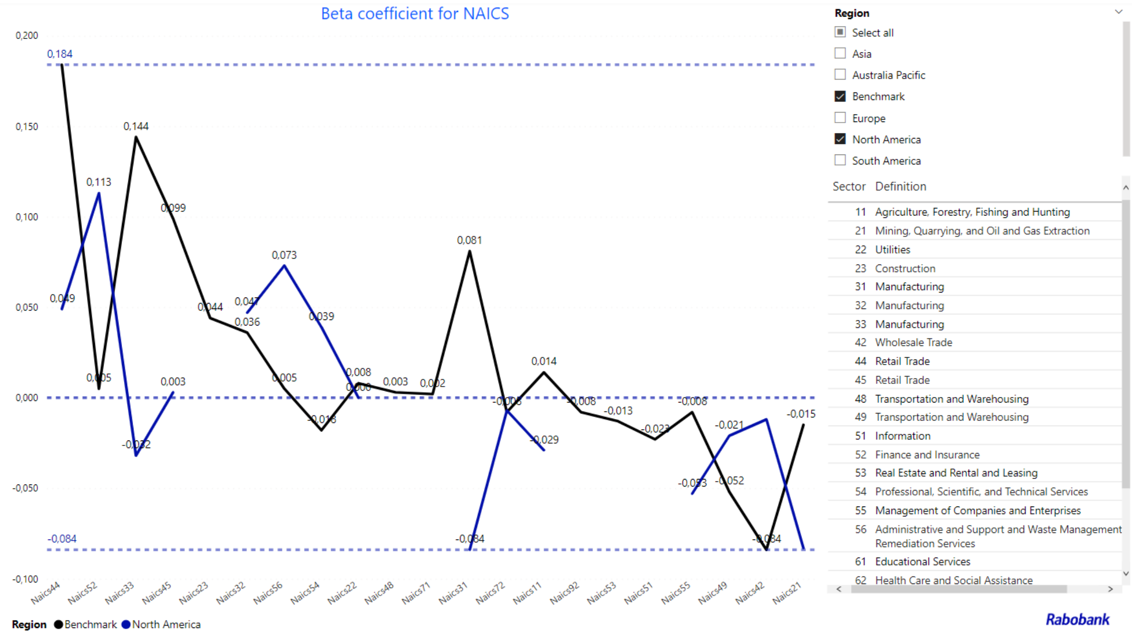


Figure 8-40: North America NAICS.

8.9 South America results

We elaborate on the results of standardized beta coefficients within the region South America. We firstly provide the adjusted R-squared, which is the amount the model explains, after which we take a deeper dive into the different grouped and ungrouped variables. As can been seen from Figure 8-41, the model has a high predictive power. While all variables are grouped, we can explain 37% of the variation in the NII bps. However, ungrouping the variables, we are able to explain up to 54% of the variation in the NII bps. Thus, our model can explain a high significant amount of the NII bps with the used independent variables.

Region	Iteration	Dataset count	R	R Square	Adjusted R Square
South America	Risk rating and NAICS ungrouped	9206	0,74	0,54	0,54
South America	Risk rating and location ungrouped	9274	0,68	0,46	0,46
South America	Risk rating and facility currency ungrouped	9268	0,64	0,42	0,41
South America	Risk rating ungrouped & NAICS 2	9268	0,64	0,42	0,41
South America	Risk rating ungrouped & NAICS 3	9268	0,64	0,42	0,41
South America	All variables grouped	9246	0,61	0,37	0,37

Figure 8-41: Adjusted R-squared for each iteration region South America.

8.9.1 All variables grouped

Figure 8-42 depicts the standardized beta coefficients which results from our MLR for South America. As one may note, this is the only region where we see a discrepancy in the influence of Outstanding on the NII bps. Corresponding to the outstanding, we also see a higher negative correlation with NII bps caused by LAD. Higher LAD results in higher amount of capital needed to comply with regulation and thus decrease the NIM. As South America is a region where high inflation is observed, this increases the risks for Rabobank's and the offering of financial products there. Higher risks results in higher margin and thus it has a positive correlation with the NII bps. Moreover, we notice that both the NAICS codes as Location have a significant negative correlation with the NII bps for South America. We will analyze these ungrouped variables in Sections 8.9.2.4 and 8.9.2.3.

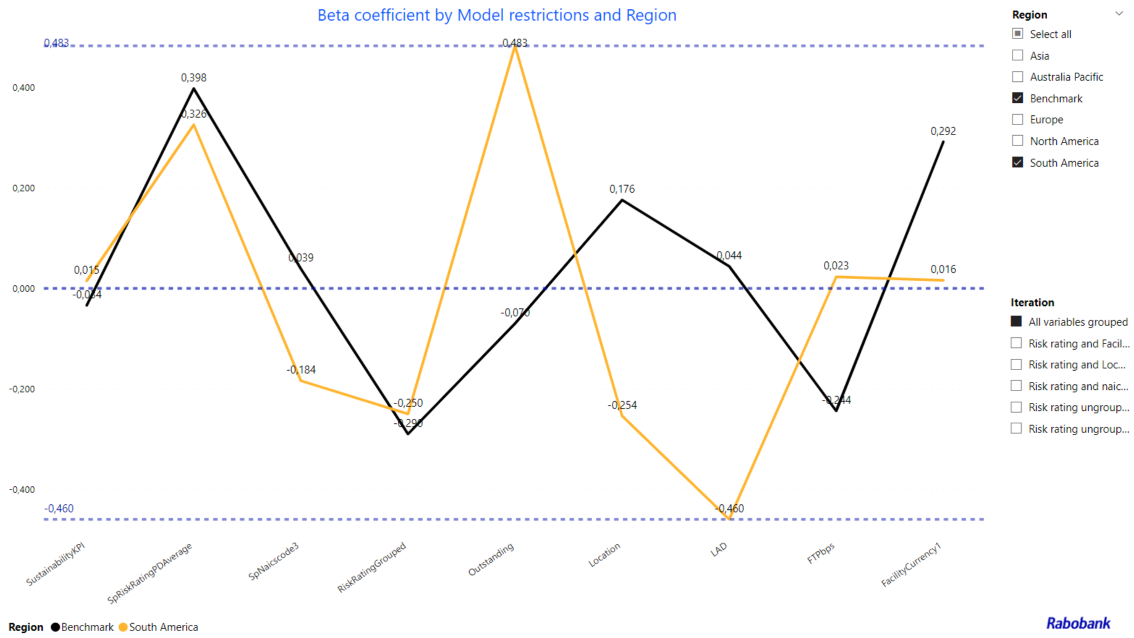


Figure 8-42: Standardized beta coefficient MLR South America.

8.9.2 Sensitivity analysis

We further investigate the ungrouped variables such as risk ratings, location and currencies for South America region specific. By analyzing the ungrouped version of theses variables, we hope to find additional underlying information which otherwise would have gone unnoticed. Note that observed missing data points in the figures of regions specific charts are caused due to the benchmark model is applied on the entire dataset. Region specific MLR's is run on the dataset corresponding to the respectively region and thus some variables may not be applicable for the region. E.g. a region may contain only a few risk ratings, sectors or currencies.

8.9.2.1 Risk rating ungrouped and NAICS2

Figure 8-43, gives us more insights in the ungrouped risk ratings and the grouped NAICS code. We see some risk ratings deviating from the benchmark model however no big significance differences besides the observed variables in the previous section.

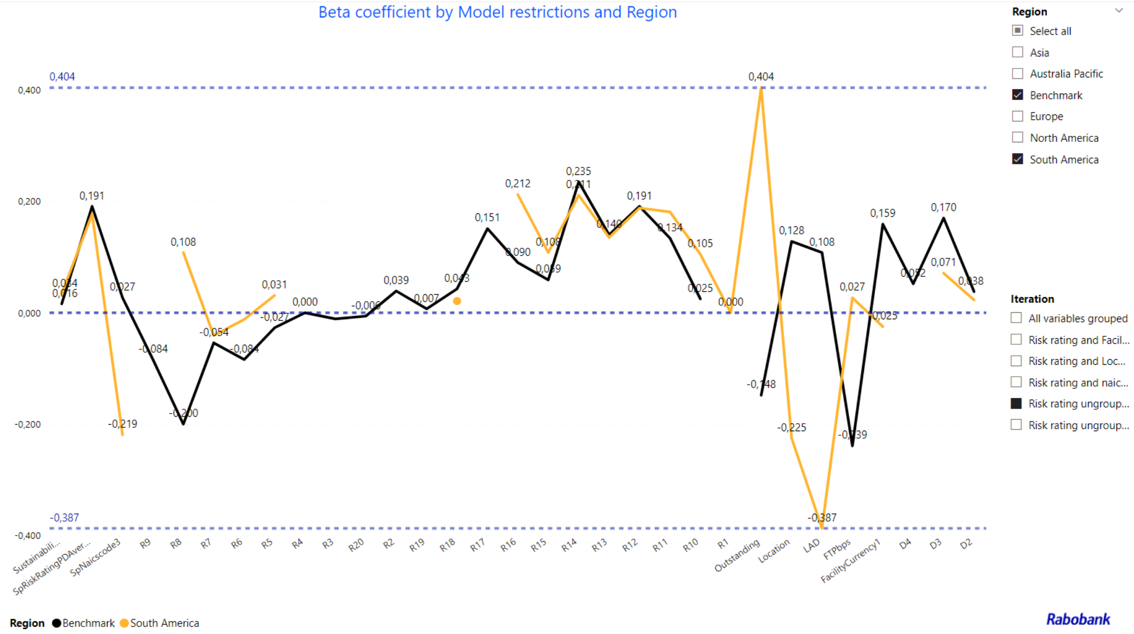


Figure 8-43: Risk rating ungrouped and NAICS2 South America.

8.9.2.2 Risk rating and Facility currency ungrouped

While ungrouping the currencies for South America, we notice only one currency is used to pay for the long-term loans namely the USD as depicted in Figure 8-44. However we do not see any significance importance of this currency for South America specific.

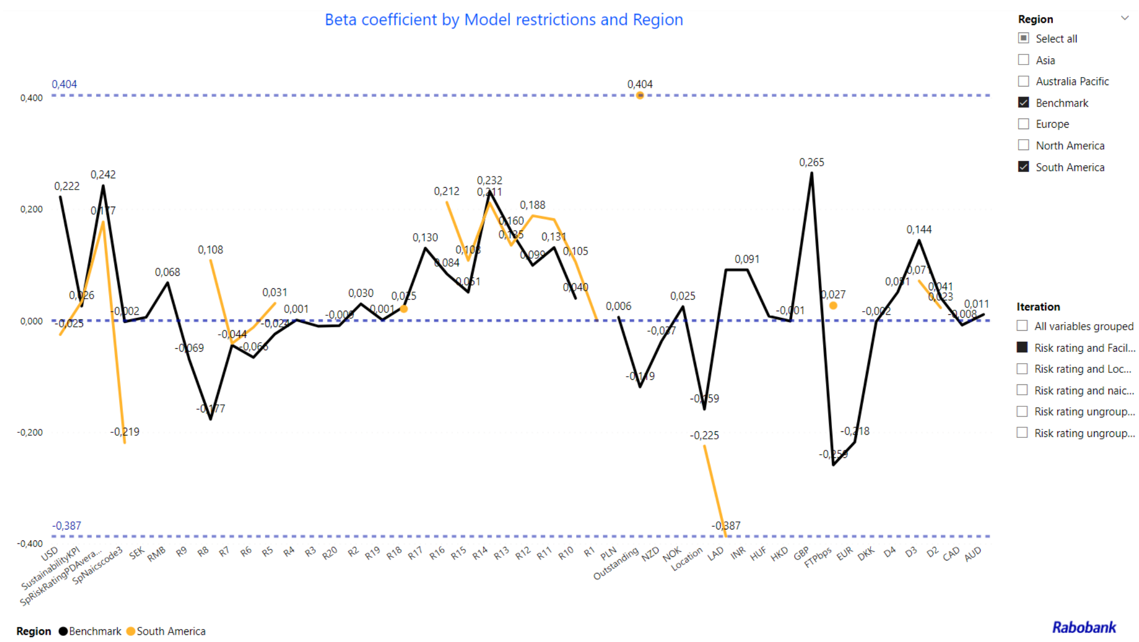


Figure 8-44: Risk rating and Facility currency ungrouped South America.

8.9.2.3 Risk rating and Location ungrouped

From Figure 8-45, we see two locations causing a high negative correlation with the NII bps namely Brazil and Chile. However, we once again notice the positive influence the location Argentina has on the NII bps. Argentina has currently very high inflation levels and one could argue it is a defaulting state. Due to the inflation and higher credit risk, higher margin can be made and thus explains the positive influence.

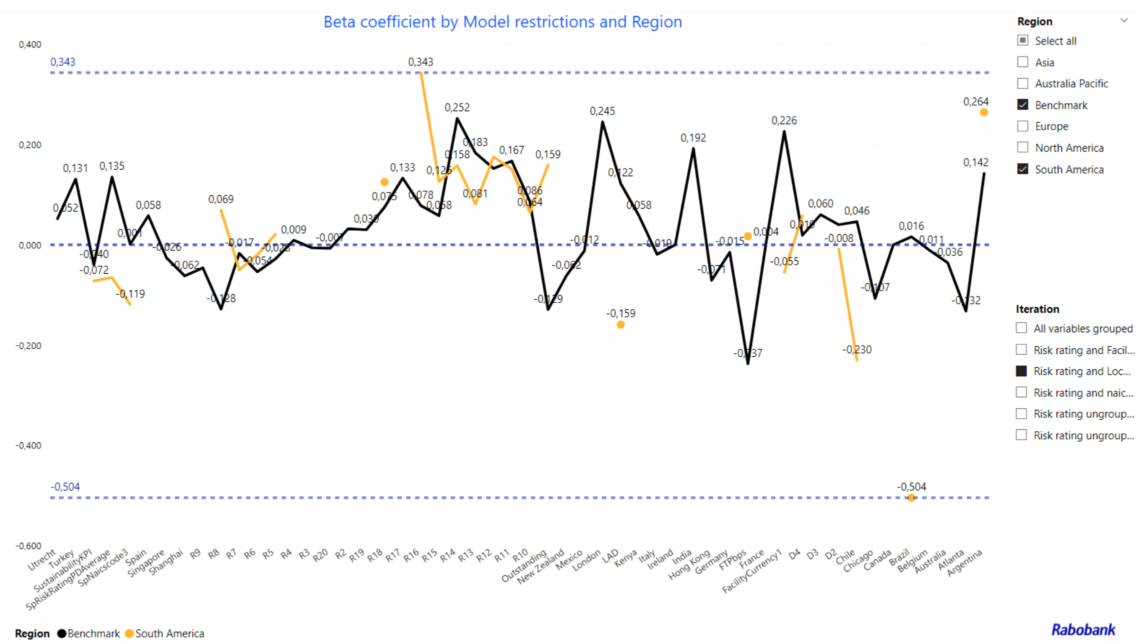


Figure 8-45: Risk rating and Location ungrouped South America.

8.9.2.4 Risk rating and NAICS ungrouped

From Figure 8-46, we see that risk ratings and NAICS codes are of importance in South America. Thus, we will analyze the figures where only risk ratings and NACIS are used.

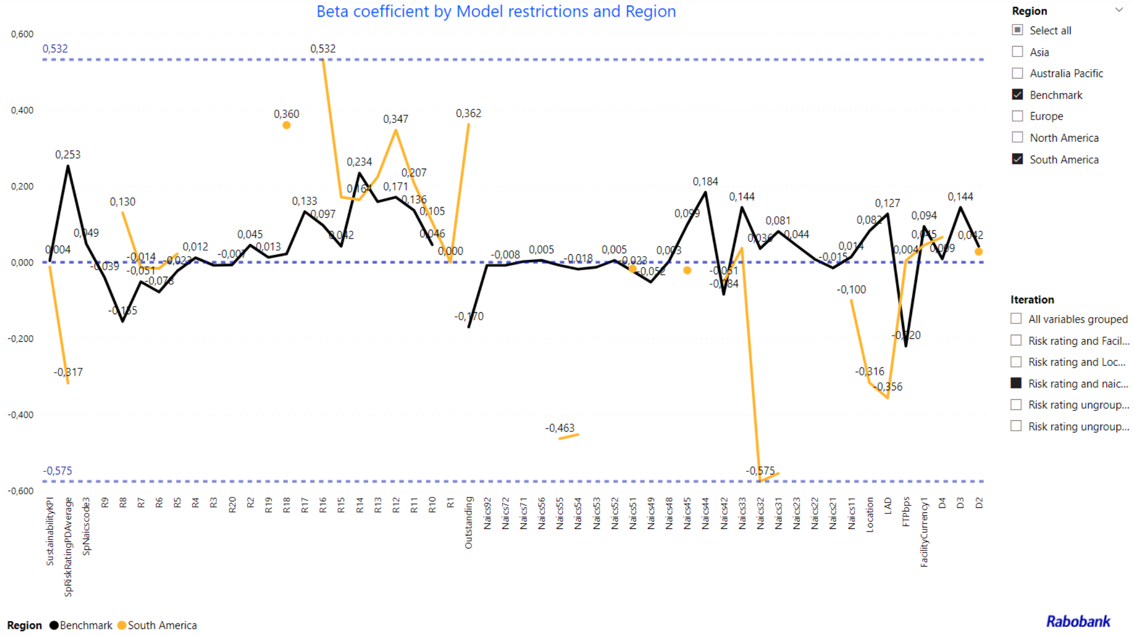
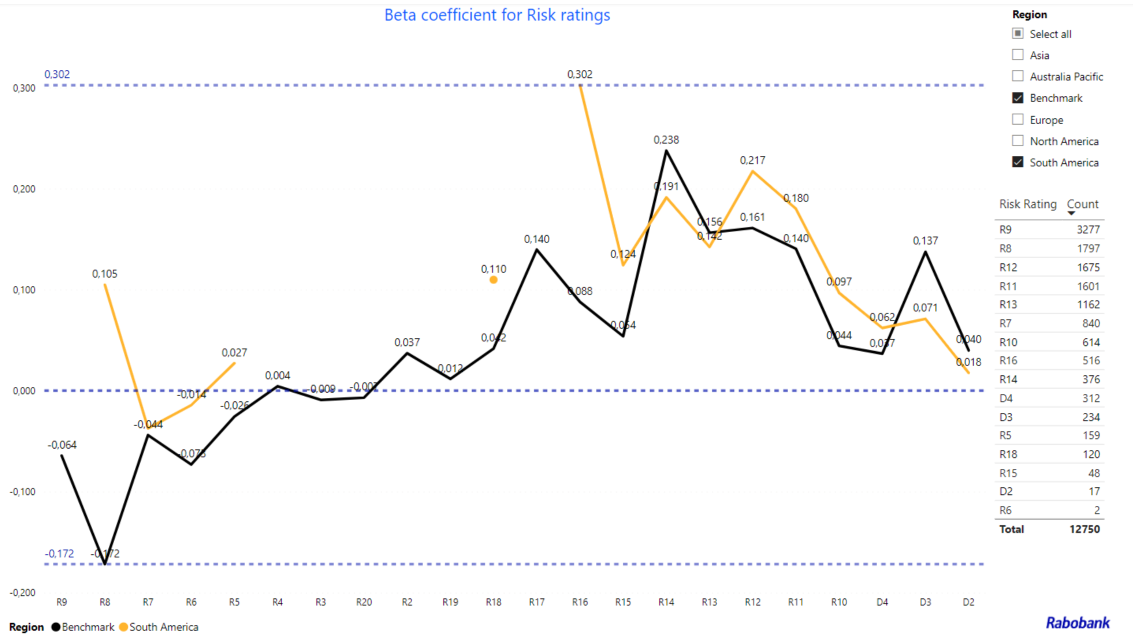


Figure 8-46: Risk rating and NAICS ungrouped South America.

Overall, we see that the risk ratings follow the benchmark closely in South America, as depicted in Figure 8-47. However, we notice some deviation for risk rating R8 where it has a positive influence in South America. As can be seen from the same figure, a high number of companies are situated in the R8 risk bucket where higher NII bps is observed. We also see R16 have a higher standardized beta coefficient compared to the benchmark model. These transaction in South America are 25% of the total R16 risk rating observed in the benchmark model. Thus, the region specific MLR gives better insights in the important risk ratings.



Risk Rating	Count
R9	3277
R8	1797
R12	1675
R11	1601
R13	1162
R7	840
R10	614
R16	516
R14	376
D4	312
D3	234
R5	159
R18	120
R15	48
D2	17
R6	2
Total	12750

Figure 8-47: South America Risk ratings.

Finally, from Figure 8-48, we notice certain sectors which should be avoided in South America as these sectors have a negative correlation with the NII bps. The sector NAICS codes are 54, 55, 31 and 32. These correspond to “Professional, Scientific, and Technical Services”, “Management of Companies and Enterprises”, and “Manufacturing” respectively.

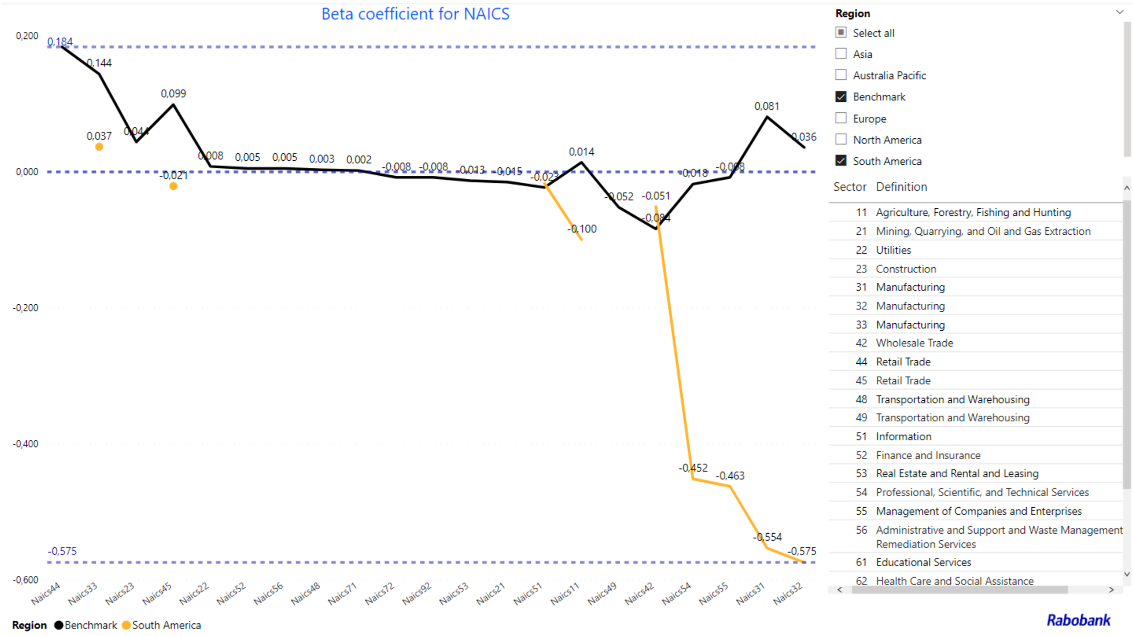


Figure 8-48: South America NAICS.