U.PORTO

MASTER
IN ECONOMICS OF BUSINESS AND STRATEGY

Risk Premium and Bank Spread: Financial Determinants in the Context of Portuguese SMEs

Tomás Pontes Torrado

M

2025

FACULDADE DE **ECONOMIA**

U.PORTO

FACULDADE DE ECONOMIA
UNIVERSIDADE DO PORTO

Risk Premium and Bank Spread: Financial Determinants in the Context of Portuguese SMEs

Tomás Pontes Torrado

Dissertation

Master in Economics of Business and Strategy

Supervised by
Professor José Manuel Peres Jorge

2025

Acknowledgements

First, I would like to thank my family for their support throughout my academic journey.

A very special thank you to my girlfriend, whose presence and understanding have been a constant source of strength.

I am also grateful to my advisor, Professor José Manuel Peres Jorge, for his guidance throughout this process.

Abstract

One of the most anticipating challenges for small and medium-sized enterprises (SMEs) is access to credit, and how interest rates are calculated. This dissertation investigates the financial determinants of bank spreads applied to Portuguese small and medium-sized enterprises (SMEs). The study was conducted using firm-level data from over 10,000 companies with active bank loans in 2023 and estimates a multiple linear regression model to identify how internal financial indicators influence the bank spread and therefore the risk premium that a company must pay in their interest rate. The analysis incorporates variables capturing profitability, leverage, operational efficiency, debt servicing capacity, and labor cost intensity, along with sectoral and regional fixed effects. Results reveal that higher profit margins and stronger interest coverage are associated with significantly lower spreads, while greater operational intensity, measured by EBITDA/Assets and Personnel Expenses/Assets, tend to increase financing costs. These findings provide empirical insight into credit risk pricing for SMEs and offer practical implications for firms and financial institutions aiming to improve access to credit and risk-adjusted lending practices.

Key Words: Bank Spread; SMEs; Risk Premium; Credit Risk; Financial Determinants; Portugal; Credit Pricing

Table of Contents

Table Index

1. Introduction

Access to credit remains one of the most difficult challenges for SMEs, especially in economies where bank lending represents the main source of external financing.

In the Portuguese context, SMEs account for a big share of economic activity and employment yet often face significant barriers in obtaining financing on favorable terms. The bank spread, defined as the difference between the interest rate charged on loans and a market reference rate such as the Euribor, reflects the cost of capital and the bank's perception of borrower risk, sectoral dynamics, and broader macroeconomic conditions.

Although access to bank financing has been widely studied in the context of Portuguese SMEs, particularly following the financial and sovereign debt crises, existing literature tends to focus either on credit volumes or on the effects of policy interventions. For instance, while Farinha and Félix (2014) analyze supply and demand dynamics using a disequilibrium framework, and Custódio et al. (2024) explore the effects of public credit guarantees, neither study models the bank spread as a function of firm-level characteristics. As such, a significant gap remains regarding the microeconomic determinants of lending rates applied to SMEs, which this thesis seeks to address.

Developing a clear perception on the determinants of the bank spread is essencial for policymakers and business leaders, as it provides insights into how financial institutions price credit risk and allocate capital. Since the spread over a risk free or benchmark rate, like Euribor, inherently reflects the risk premium charged by banks, studying the firm-level financial variables that influence this spread is, by extension, an empirical analysis of how banks assess and price borrower risk. In other words, identifying the financial determinants of the spread is equivalent to understanding the composition and variation of the risk premium applied to heterogeneous SME profiles.

Regardless of the growing availability of firm-level data and the increased sophistication of credit models, the way lending terms are priced remains only partly understood in empirical studies.

This dissertation examines the main determinants influencing the bank spread applied to Portuguese SMEs, using detailed financial and structural information at the firm level. The econometric model developed includes a set of independent variables, covering dimensions such as profitability, capital structure, operational efficiency, debts servicing capacity, and firm size, along with sectoral and regional fixed effects.

By providing empirical support on the relationship between firm characteristics and credit pricing, this dissertation aims to contribute for a clearer understanding of risk assessments and financial decision processes in the banking sector.

2. Literature Review

2.1. Characterization of SMEs in Portugal

In the context of European and Portuguese economic environment, the European Commission (Recommendation 2003/361/EC) defines small and medium-sized companies as companies that employ less than 250 people and have an annual revenue of less than 50 million euros or whose yearly total balance does not exceed 43 million euros. These companies are essential in Portugal´s economic landscape, representing 99,9% of the total Portuguese companies and are responsible for 76% of the employment in the private sector, representing 68,3% of Portugal's Gross Value Added (INE, 2023).

SMEs operate in various sectors, ranging from commerce and services to industry and tourism, according to Koporcic et al. (2025) they are also distinguished by their flexibility, innovation, and ability to adapt to market changes, qualities that make them resilient in times of crisis.

2.2. Empirical Evidence on Bank Credit and Spread Determinants for Portuguese SMEs

Access to bank financing has been widely studied in the context of SMEs, given their strong dependence on traditional credit and limited access to capital markets. In Portugal, several studies have addressed the financial constraints faced by these firms, particularly during and after the financial and sovereign debt crises.

One of the most relevant contributions in this area is the work of Farinha and Félix (2014), who apply a disequilibrium model estimated with firm-level microdata to jointly analyze the determinants of credit supply and demand for Portuguese SMEs between 2010 and 2012. The authors demonstrate that cash flow generation capacity, debt levels, and availability of collateral mainly drive credit supply, at the same time, they demonstrated that demand is negatively affected by interest rates and positively associated with working capital needs.

The disequilibrium model developed in this study identifies that the detected amount of credit in the market may not be a result of a market equilibrium between supply and demand. As an alternative, it assumes that the amount provided reflects the minimum between demanded and supplied credit ($Q = \min(Q_d, Q_s)$), thereby reflecting the existence of credit rationing. This approach is especially relevant in the presence of market imperfections, such as bank credit, where the interest rate does not fully adjust to eliminate excess demand or supply.

Although it uncovers fundamental factors influencing credit terms, this study does not explicitly model the bank spread as a dependent variable, treating it only as an explanatory factor included on the demand side.

More recently, Custódio et al. (2024) examined the impact of a public credit guarantee program, The PME Líder Guarantee Program, on the behavior of Portuguese SMEs, using a regression discontinuity design. The authors stated that credit guarantees have a significant positive effect on improving financing conditions, including the applied interest rates: "The maximum spread that banks could apply to credit lines granted to *PME* Líder companies in 2015 ranged from 2.7 to 3 percentage points above the 6-month Euribor" (Custódio et al., 2024, p. 6). While contributing to the understanding of imposed spread limits, the study's focus is on eligible firms, making it impossible to generalize to the broader SME population.

In another study Pinto et al. (2023) analyze the relationship between bank financing and trade credit in the context of Portuguese SMEs during the period from 2010 to 2019. The evidence suggests that operational profitability, measured by ROA, is negatively linked with the use of bank credit, indicating that firms with higher profitability make less use of external financing. In contrast, SMEs exhibiting lower levels of liquidity and profitability are more likely to rely on trade credit, indicating that suppliers often act as alternative lenders when access to bank financing is constrained. The study also finds that companies with greater bankruptcy risk have a higher probability to use bank credit and are less likely to access trade credit, as suppliers often display higher levels of risk aversion relative to banks. Overall, the findings indicate that trade credit and bank credit act as substitute financing sources, particularly during periods of more restrictive credit conditions, highlighting the importance of how financial characteristics at the firm level influence access to different types of credit.

2.3. Relationship Between Banks and SMEs

Probably the main obstacle to business investments is liquidity constraints, particularly in the case of small and medium-sized enterprises (SMEs). According to Fazzari et al. (1988), the conventional premise of capital market perfection, in which internal and external financing are treated as perfect substitutes, proves inadequate when confronted with the realities of information asymmetries and the costs associated with external financing. These authors demonstrate that many companies, especially those retaining a higher proportion of their earnings, display a significantly greater dependence of investment on internal cash flow. This empirical evidence indicates that, in contexts of restricted financial access, investment decisions are no longer driven solely by the expected profitability of projects but become increasingly limited by the availability of internal capital. The study also identifies a financing hierarchy in which firms rely first on retained earnings, then on debt, and only as a last resort on the issuance of new equity often avoided due to dilution effects and associated costs. This behavior is particularly pronounced among SMEs, which face greater difficulties in accessing external funding due to their lower transparency and visibility in the eyes of investors. Therefore, liquidity constraints not only restrict investment decisions at the microeconomic level, but also contribute to the amplification of business cycles, given that firms' cash flows tend to correlate strongly with aggregate economic fluctuations.

Many researchers like Beck et al (2008), found that SMEs fund a minor portion of their investment through external funds, particularly bank credit. This limitation stems largely from the reduced ability of these firms to overcome problems of asymmetric information, which discourages banks and limits the availability of financing. Although some SMEs resort to alternative sources such as informal financing or trade credit, these alternatives are insufficient to compensate for the lack of access to traditional credit, especially in countries with weak legal and financial systems.

Petersen and Rajan (1994) have focused on the effects of customer relationships, providing empirical proof that lending connections with banks facilitates credit access for small companies. On their analysis it is demonstrated that close ties with a single lender do not always results in lower interest rates, they do ease financial constraints, especially for younger and smaller firms facing information asymmetries.

On the same study, firms that borrow from multiple banks tend to pay higher interest rates, suggesting that strong and concentrated relationships with a primary bank are more beneficial than diversified but superficial ones. Overall, the study concludes that the main benefit of bank-firm relationships lies in the increased quantity of credit available rather than its price. This form of relationship lending facilitates banks collecting soft information over time, which is not easily captured through financial statements alone. Petersen and Rajan (1994) argue that these close ties allow lenders to mitigate adverse selection and moral hazard problems, leading to greater loan availability. Nevertheless, the study also points out that this dependency may present disadvantages, after a strong relationship is established, firms may become dependent on a single lender and lose bargaining power, potentially leading to non-competitive pricing if the main bank makes use of its informational advantage.

On the other hand, Ioannidou and Ongena (2010) offer empirical support that long-term banking relationships with firms can lead to rising loan costs over time. Their study, relying on comprehensive data from the Bolivian credit registry, finds that when a firm changes to a new bank, it receives a considerably lower loan rate, about 87 basis points lower, compared to loans from its previous banks, however, after around three years, the interest rates charged by the new bank gradually increase, eventually converging back to the level observed prior to the switch. This cyclical pattern in loan pricing suggests the presence of informational hold-up, where banks use the informational advantage obtained during the lending relationship to extract rents from firms. In other words, while new relationships may start with favorable terms, over time, banks may use their informational advantage to raise interest rates, even if the firm's credit quality remains stable.

2.4. Bank Spreads and Risk Premium

Risk premium as demonstrated by Mishkin (2019), represents the additional compensation required by banks when taking on riskier assets. This compensation is reflected in the interest rates charged to clients and is therefore an integral part of the banking spread so the greater the default risk, the larger the risk premium will be, as consequence, the higher will be the interest rate that borrowers must pay.

The theoretical model proposed by Ho and Saunders (1981) formalizes this concept, in their model the bank is treated as a financial intermediary operating under market uncertainty, including credit risk. For determining interest margins, the authors demonstrate that the existence of positive loan risk premiums, that is, an interest rate above the risk-free or deposit rate to compensate for credit risk is a necessary condition for financial intermediation to occur, so because of that the bank spread is composed, among other factors, by a component that represents compensation for the risk taken in lending, being sensitive to the variance of expected returns and to the risk aversion of bank management.

From an empirical point of view, Brock and Rojas-Suarez (2000) reinforce this interpretation by analyzing the behavior of spreads in Latin America, they find that high levels of impaired loans and macroeconomic uncertainty substantially increase bank spreads. Their analysis shows that spreads are amplified whenever the banking system faces increased credit risk and instability, leading banks to incorporate higher risk premiums into the interest rates charged to borrowers. The evidence clearly shows that higher spreads are directly linked to the perception of risk by banks and to the instability of returns associated with credit operations.

To conclude, existing studies consistently show that the risk premium is a central determinant of the bank spread, working as an adaptation mechanism in response to uncertainty and credit market risk. This relationship is particularly clear in borrower segments with limited collateral capacity, such as SMEs, where information asymmetry leads banks to charge higher premiums, which inevitably translates into higher spreads.

2.5. Risk Valuation

2.5.1. Altman Z-Score

Throughout the years, an important area of study in corporate finance and risk management has been the development of financial distress and bankruptcy prediction models, one of the most famous and academically recognized is the one developed by Altman (1968), the Altman Z-Score. This model is a linear discriminant analysis that combines five key financial ratios into a single score to measure the likelihood of bankruptcy. The original model, formulated for publicly traded manufacturing firms, is expressed as follows:

$$Z = 1.2 \cdot \left(\frac{\text{Working Capital}}{\text{Total Assets}}\right) + 1.4 \cdot \left(\frac{\text{Retained Earnings}}{\text{Total Assets}}\right) + 3.3 \cdot \left(\frac{\text{EBIT}}{\text{Total Assets}}\right) + 0.6 \cdot \left(\frac{\text{Market Value of Equity}}{\text{Book Value of Total Liabilities}}\right) + 1.0 \cdot \left(\frac{\text{Sales}}{\text{Total Assets}}\right)$$

Relying on the calculated Z-Score, firms are classified as financially sound with $Z > 2.99$, in a grey area with $1.81 < Z < 2.99$ or distressed with $Z < 1.81$.

The model achieved notable precision in forecasting outcomes in Altman's original study, properly identifying approximately 95% of the firms one year prior to bankruptcy, with Type I errors (failing to predict bankruptcy) as low as 5% but when applied to data from two years before the event, the accuracy declined moderately to 83%, demonstrating the model's robustness over short forecasting periods.

However, several researchers such as Grice and Ingram (2001) conclude that Altman's Z-Score model, despite being widely used, demonstrates a substantial decrease in forecasting performance when applied to more recent data and non-manufacturing firms, highlighting the temporal instability of its coefficients. The authors advocate for regular re-calibration of the model with updated, proportionally representative datasets, alerting against its uncritical application across different contexts.

2.5.2. Credit Scoring Model

The credit scoring model is an automated system that measures a client's credit risk by attributing weighted values to specific characteristics and estimating the probability of default based on past behavioral trends.

According to Carvalho (2009), these models rely on historical relationships between client characteristics and their payment practices, assuming the continuation of these trends in the future. By assessing current customer data, the model predicts the probability of payments default and is used as a risk evaluation tool.

Abdou and Pointon (2011) demonstrate that scoring systems replace subjective evaluations with quantitative, data-driven approach. The score results from the sum of points assigned to each one of the predictor variables, such as number of credit accounts, previously correlated with payment behavior. The use of these models has advantages because they are built using large historical samples, which allows for objective and fast decision making.

However, credit scoring models have limitations, Abdou and Pointon (2011) highlight the fact that as the models rely on historical data, their output could not be accurate if the client's behavior shifts over time. There is also the risk of indirect discrimination and classification errors, especially when models don't take into account relevant variables or are not updated regularly.

2.5.3. Expected Loss Model

The Expected Loss (EL) model is a fundamental element of credit risk management, especially considering the regulatory environment established by the Basel Committee on Banking Supervision. This model estimates potential losses that can result from borrower's default and is formally integrated into the Basel III standards as part of the Internal Ratings-Based (IRB) approach to credit risk.

The Expected Loss is computed as:

$EL = EAD \times LGD \times PD$

where Exposure at Default (EAD) measures the total value at risk, Loss Given Default (LGD) represents the proportion of loss in the event of default and Probability of Default (PD) estimates the likelihood of borrower default.

According to Basel III (BCBS, 2017), this model is fundamental to the calculation of minimum capital requirements, supporting sufficient reserves and adjusting loan rates with risk levels. The framework requires banks to estimate EL to ensure sufficient capital buffers and to promote stability within the financial system.

2.5.4. Risk-Adjusted Return on Capital (RAROC)

RAROC is a performance measurement tool, widely used in the USA, that assesses the return on capital and adjusts it for the risk undertaken, it is calculated as:

RAROC= (Expected Loan Revenue−Expected Loss) / Economic Capital

Where:

Expected Loan Revenue: Anticipated income from the loan, including interest payments.

Economic Capital: Capital set aside to incorporate potential losses, reflecting the risk of the loan.

Federal Deposit Insurance Corporation (FDIC, 2004) demonstrates that the use of RAROC enables financial institutions to evaluate the risk-adjusted performance of different loan portfolios. In their illustrative case, two portfolios with similar structures display different outcomes: although one generates higher gross income, the other with a higher RAROC demonstrates superior risk-return efficiency, requiring less economic capital and delivering greater economic profit. According to Bauer and Zanjani (2021), RAROC serves as a benchmark for evaluating the marginal profitability of risk exposures and remains valid even in dynamic, multi-period models, provided that appropriate adjustments are made to reflect the firm's effective risk aversion and capitalization strategy.

3. Data

3.1. Dataset

The dataset used in this study was extracted from SABI (System for the Analysis of Iberian Balance Sheets) platform, through a structured search to identify Portuguese SMEs with active bank financing and complete financial statements for the fiscal year **2023**.

The sample selection was based on several economic, legal, and organizational criteria to ensure internal consistency and homogeneity. Following the SME definition established by the European Commission (Recommendation 2003/361/EC), the sample was restricted to firms with operating revenues of up to €50 million and between 0 and 250 employees, based on the latest available fiscal year.

To build this dataset, the degree of corporate independence was also assessed using the Bureau van Dijk (BvD) Independence Indicator, this indicator categorizes firms based on their ownership structure and level of external control. Only companies rated A+, A, A–, or B+ were included in the dataset, corresponding to autonomous or associated enterprises not majority-controlled by larger corporate groups or multinational holdings, ensuring that the financing conditions reflected in the data set are determined primarily by the firm's characteristics and not by those of a parent company.

As a result of this selection process, the final dataset comprised 12,787 valid observations, which were subsequently subjected to data cleaning, transformation, and statistical validation, as described in the following section.

3.2. Dataset Cleaning

As part of the dataset cleaning process, several exclusion criteria were applied to ensure the consistency and reliability of the final sample. First, all firms reporting a zero balance for the sum of current and non-current financial liabilities were excluded, since these companies had no outstanding financial debt, they could not be used to compute the dependent variable, the bank spread, which requires the presence of interest obligations.

Next, firms that reported financial debt, but no interest expenses were also removed, this mismatch may stem from accounting anomalies, grace periods, or non-remunerated credit agreements. In any case, the absence of interest charges makes it impossible to estimate an effective interest rate, and retaining this kind of observations would compromise the analysis.

Observations with an estimated effective interest rate exceeding 15% were excluded to improve data quality, these extreme values are considered statistical outliers and inconsistent with standard corporate lending conditions in Portugal, including them could distort the spread distribution and the robustness of the econometric estimates.

After applying these exclusion rules and completing the necessary variable transformations, the final dataset comprised 10,368 firms, all presenting consistent financial data, active bank debt exposure, and valid interest rate information suitable for the analysis.

Data Processing Step	No. of Observations	Observations Removed
Initial dataset (SABI – 2023)	12.787	0
Removal of firms without bank financing	11.416	1.371
Exclusion of firms with debt but no interest payments	10.938	478
Elimination of observations with interest rates above 15%	10.368	570
Final number of valid observations for analysis	10.368	2.419

Table 1- Evolution of the Sample Throughout the Data Cleaning Process - Sabi 2023

3.3. Variables

3.3.1. Construction of the Dependent Variable: Spread

In this study the central dependent variable is the bank spread, defined as the difference between the effective interest rate borne by the company and the benchmark interest rate in the interbank market. This variable plays a central role in the research, as it represents the risk perception attributed by financial institutions to each company, ultimately reflecting the cost of bank financing based on the economic, financial, and institutional profile of each borrower.

As it is not possible to know the details of the credit contracts, and therefore, to know the spread and the interest rate of each contract, the operationalization of the spread was based on the construction of a proxy for the effective interest rate, using the ratio between interest expenses on financial debt and the total amount of financial debt obtained by the company (including both current and non-current financing obtained).

$$Effective\ Interest\ Rate = \frac{Interest\ Expenses}{Current\ Financing\ Obtained + Non - Current\ Financing\ Obtained}$$

This ratio directly estimates the average effective interest rate the company bears. The construction of proxies is widely used in literature as a valid approximation in contexts where direct access to credit contracts is limited or unavailable, as highlighted by Faulkender & Petersen (2006). Indirect proxies for the cost of financing are frequently used when data on the terms of debt contracts are not available.

The effective interest rate was adjusted to obtain the spread by subtracting the benchmark interest rate prevailing in the European interbank market, specifically the 12-month Euribor. The reference rate used was retrieved from Euribor Rates and corresponds to the average observed over 2023 and 2022, set at 2.089%. This two-year average was chosen to smooth out short-term fluctuations in interest rates and to align with the temporal structure of financial debt, which in many cases was contracted in previous years and not necessarily within the year of analysis.

This choice is based on the typical structure of corporate credit agreements in Portugal, where loans, particularly those with maturities longer than one year, are generally indexed to the 12-month Euribor, with an additional premium reflecting the borrower's credit risk.

Accordingly, the formula adopted for the construction of the spread was:

$$\text{Spread} = \text{Effective Interest Rate} - \overline{\text{Euribor}}_{2022-2023}$$

$$\overline{\text{Euribor}}_{2022-2023} = 2.089\%$$

This approach, although based on a methodological simplification, the assumption of homogeneity in the debt structure, without explicitly distinguishing between short- and long-term debt, or between different types of financing, allowed the construction of an objective, and comparable metric across firms.

3.3.2. Independent Variables and Theoretical Justification for Their Inclusion in the Model

Considering theoretical and empirical literature, econometric modelling of the determinants of the bank spread requires the inclusion of a multidimensional set of explanatory variables that capture firms' structural, operational, and financial characteristics. The selection of variables in this study follows criteria of theoretical relevance, empirical significance in prior research, and suitability to the institutional context of the Portuguese credit market.

Each variable was designed to reflect specific dimensions of the firm's risk profile, which, according to financial intermediation theory and bank risk assessment models, are direct determinants of the cost of financing applied by lending institutions.

Profitability

Profitability is one of the core dimensions in credit risk assessment, as it reflects the firm's ability to generate sustainable cash flows and returns for its stakeholders. Financial literature, particularly fundamental analysis models such as those proposed by Altman (1968) and Beaver (1966), argues that firms with strong profitability indicators tend to exhibit a lower probability of default, benefiting from a lower cost of external capital. To represent this dimension the following two indicators were chosen:

Return on Equity (ROE) represents the ratio between net income and shareholders' equity. This indicator summarizes management's effectiveness in generating returns on shareholders' resources and is often interpreted as a reflection of value creation. However, as suggested by Damodaran (2015), in the presence of high financial leverage or sectoral volatility, an unusually high ROE may signal latent risks, such as increased exposure to the economic cycle or aggressive financing practices, which may adversely affect lenders' perception of the firm's risk profile.

$$\text{Return on Equity (ROE)} = \frac{\text{Net Income}}{\text{Shareholders' Equity}}$$

Profit Margin is the ratio between net income and total revenue; this variable reflects the company's capability to produce net earnings from sales. Traditional financial analysis considers high margins as an indicator of operational efficiency and competitive advantage. While Fama and French (2002) do not address net profit margin directly, their findings reinforce the general concept that profitability metrics are inversely related to leverage, consistent with the pecking order theory. This demonstrates that more profitable firms may rely less on external debt, possibly contributing to more favorable credit risk assessments.

$$\text{Profit margin} = \frac{\text{Net income}}{\text{Turnover}}$$

Capital Structure

The configuration of a firm's capital structure constitutes a major factor in defining the cost of credit, as it influences financial risk exposure and long-term solvency. Capital structure theories like Modigliani & Miller (1958) or Myers (1977) demonstrated that the level of indebtedness affects the expected return and the associated risk, directly impacting external capital pricing.

Leverage according to the SABI calculation is the ratio between non-current liabilities plus loans and shareholders' funds and reflects the extent to which a firm relies on external financing. While it can amplify returns for shareholders, it also increases financial risk, according to Modigliani & Miller (1958), higher leverage raises the expected return on equity due to increased risk exposure. Myers (1977) further argues that excessive leverage may lead to suboptimal investment behavior, reducing firm value and increasing credit risk, factors that are often penalized by lenders through higher credit spreads.

$$\text{Leverage} = \frac{\text{Non-current liabilities} + \text{Loans}}{\text{Shareholders' funds}}$$

Operating Efficiency

Operational performance is a central factor in a firm's ability to generate sustainable results from its resources.

EBITDA is widely recognized as an important indicator of a firm's operational performance, since it reflects the earnings generated from core business activities before the impact of financing and accounting decisions. It reflects the firm's ability to produce positive cash flows from operations, which is crucial for servicing debt and sustaining long-term financial health. Damodaran (2007), uses EBITDA as a proxy for cash flow and is especially relevant in credit risk assessment because it isolates the profitability of the business from capital structure effects and non-cash charges.

Turnover measures the scale and activity level of a firm. While Berger and Udell (1998) do not focus explicitly on turnover volume, they emphasize the importance of financial statements and continuous information, including cash flows and transactions, in monitoring borrower firms. In the case of SMEs, where standard financial transparency is often lacking, turnover data can play a crucial role in assessing liquidity and repayment potential.

Debt Servicing Capacity

A firm's financial health is also measured by its ability to generate sufficient earnings to cover its financial obligations, making this one of the primary factors in credit risk assessment.

Interest Coverage Ratio expresses the relationship between EBIT and interest expenses and is one of the most direct indicators of current financial solvency. A high value suggests a greater margin of safety and a lower probability of default, leading to a reduction in the perceived risk by financial institutions. Damodaran (2015) shows that it plays a central role in estimating synthetic credit ratings and the corresponding default spreads applied by lenders, higher values of this ratio indicate a more substantial capacity to meet interest obligations and are typically associated with lower credit spreads.

$$\text{Interest Coverage Ratio} = \frac{\text{EBIT}}{\text{Interest Expenses}}$$

Human Resources Structure

The composition of a firm's operating cost structure, particularly regarding the labor factor, represents an additional dimension of analysis, especially relevant in labor intensive sectors.

Personnel Expenses is supported by empirical evidence from Pontuch (2011), who demonstrates that labor intensity is positively associated with expected returns, particularly among small and medium-sized firms. By isolating the effect of labor intensity relative to the industry median, the author shows that SMEs have a positive association between returns and labor intensity, even if when controlling for other traditional firm-level determinants of returns.

Sectoral and Regional Dummies

Beyond firm-specific financial indicators, other factors such as industry sector and geographic location can also play an important role in shaping credit conditions and can influence lenders' risk perception.

To control these effects, a sector fixed dummy (based on NACE Rev. 2 sections) was included to capture differences in business models, market stability, and sector-specific risk. Firms in more volatile or capital-intensive sectors may face higher credit spreads due to increased default risk.

Similarly, regional dummy, based on the firm's district, considers unobserved geographic influences, such as banking competition, economic development, and infrastructure access. Both variables were encoded as categorical factors using Stata's encode function and included as dummies, allowing for a more precise estimation of firm-level financial impacts on the observed bank spread.

3.3.3. Variables Transformation

The proposed econometric model in this dissertation was developed through the transformation and normalization of some selected explanatory variables. The initial dataset included financial ratios and accounting figures expressed in absolute terms. To ensure consistency and comparability, all financial statement variables were converted into financial ratios expressed relative to total assets. This approach ensured a uniform scale across variables, allowing for a consistent analytical framework and reducing the influence of absolute differences in magnitude in the estimation process. Additionally, this normalization aligns conceptually with the dependent variable, bank spread, which is itself a relative measure.

This approach is accepted by financial literature and widely adopted in credit risk, scoring, and bankruptcy prediction models such as the previously mentioned Altman Z Score. Ratios such as EBITDA/Assets, Sales/Assets, or Personnel Expenses/Assets are commonly used in empirical research as they capture efficiency, profitability, and cost structure in a way adjusted for firm size.

EBITDA/Assets is the ratio between EBITDA and total assets is commonly employed as a proxy for operational performance, as it abstracts from accounting choices related to depreciation and amortization. While its use is widespread in corporate valuation practice (Damodaran, 2015), caution is advised when interpreting it in isolation, particularly for firms with significant reinvestment needs. Nonetheless, higher values of this indicator may signal stronger operating performance and, under certain conditions, be associated with lower perceived credit risk, potentially leading to narrower credit spreads demanded by lenders.

$$\frac{\text{EBITDA}}{\text{Total Assets}}$$

Turnover/Assets, also known as asset turnover, relates total turnover to total assets. It is a structural productivity metric that reflects the firm's efficiency in utilizing its economic resources to generate sales. Higher values are generally associated with more dynamic management models, greater capacity to generate cash flows, and stronger resilience to demand shocks. Altman (1968) suggests that firms with higher asset turnover tend to have greater operational efficiency and lower insolvency risk, which logically supports the inference that such firms are likely to benefit from more favorable credit conditions and lower credit spreads.

$$\frac{\text{Turnover}}{\text{Total Assets}}$$

Personnel Expenses/Assets relate personnel expenses to total assets, allowing for assessing the degree of labor intensity in the company's operational costs, its relevance could vary across industries, as it may reflect either legitimate structural characteristics or productivity weaknesses, high ratios can indicate structural rigidity or low mechanization, depending on the sectoral context. For this reason, it is included in the model as a variable for structural cost intensity.

$$\frac{\text{Personnel Expenses}}{\text{Total Assets}}$$

To further reduce the influence of extreme values without compromising the sample's integrity, bilateral winsorization was applied at the 1st and 99th percentiles of the financial ratios to preserve the central distribution while limiting the distortion from outliers.

Variable	Observations	Mean	Median	Standard Deviation	Minimum	Maximum
Spread	10368	2.14	2.10	2.49	-2.09	12.88
Return on Equity (%)	10368	16.33	11.48	29.89	-104.94	166.85
Profit Margin (%)	10368	6.58	3.35	13.37	-19.944	98.21
Leverage (%)	10368	113.07	51.98	230.02	-444.39	1545.76
EBITDA/Assets	10368	12.82	10.76	10.51	-14.18	51.48
Turnover/Assets	10368	163.56	128.56	138.49	3.11	888.05
Interest Coverage	10368	635.04	9.17	4252.94	-40.94	38174.52
Personnel Expenses/Assets	10368	30.50	21.47	31.54	0	190.36

Table 2 - Descriptive Statistics of the Variables

4. Methodology

The methodology adopted in this study is based on the estimation of a multiple linear regression, with the aim of identifying the determinants of the banking spread applied to Portuguese companies.

The empirical analysis focuses on a set of financial variables representing firms' performance, capital structure, operational efficiency, and credit risk. As mentioned above, literature widely uses these variables to explain the cost of corporate financing and reflect the firm's ability to generate profits, meet financial obligations, and operate efficiently.

Although the dataset is cross-sectional , to control unobserved heterogeneity associated with structural factors external to the firm, such as the economic sector or the region in which it operates, the model also includes categorical variables corresponding to the industry classification and the district where the company is headquartered, this inclusion allows for the isolation of fixed effects associated with structural differences between sectors and geographic locations, which could otherwise bias the estimates. In this way, it is possible to capture systematic variations in the spread that do not result directly from the individual characteristics of the firm, but rather from the sectoral and regional context in which it operates.

$$\begin{aligned}\text{Spread}_i = \alpha + \beta_1 \text{ROE}_i + \beta_2 \text{Profit Margin}_i + \beta_3 \left(\frac{\text{EBITDA}}{\text{Total Assets}}\right)_i + \beta_4 \text{Leverage}_i \\ + \beta_5 \text{Interest Coverage Ratio}_i + \beta_6 \left(\frac{\text{Turnover}}{\text{Total Assets}}\right)_i + \beta_7 \left(\frac{\text{Personnel Expenses}}{\text{Total Assets}}\right)_i \\ + \delta_j \text{Sector}_j + \theta_k \text{District}_k + \varepsilon_i\end{aligned}$$

4.1. Empirical Results

This section presents and discusses the results of the multiple linear regression model estimated in the previous section. Based on the financial and categorical variables included in the model, it aims to identify the main determinants of the banking spread applied to Portuguese companies.

Model

	Spread
Return on Equity (ROE)	-0.0016
	(0.0011)
Profit Margin	-0.0062**
	(0.0025)
EBITDA/Assets	0.0185***
	(0.0033)
Leverage	-0.0037***
	(0.00010)
Interest Coverage Ratio	-0.00013***
	(0.000004)
Turnover/Assets	0.0009***
	(0.0002)
Personnel Expenses/Assets	0.0036***
	(0.0010)
Sectoral Dummy	Yes
Regional Dummy	Yes
Constant	1.8395***
	(0.0550)
R-squared	0.0696
Observations	10368

Standard errors in parentheses

Includes sector and district fixed effects

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3 - Regression Model

Table 3 presents the estimates for the central regression equation in this study. The results obtained with this regression show that several financial variables significantly affect the spread. Profit Margin displays a negative and highly significant coefficient (-0.0062; p<0.05), indicating that companies with higher profitability benefit from lower financing costs, which is in line with the assumption that financially stable companies are perceived by banks as less risky. On the other hand, the EBITDA/Assets have a positive and statistically significant coefficient (0.0185; $p<0.01$), possibly indicating that, despite being profitable, these companies operate in more volatile sectors or lack sufficient collateral, leading to a higher perceived risk.

The variable Leverage is negatively associated with the spread (-0.0037; $p<0.01$), indicating that higher debt levels may reflect efficient financial management or a more experienced use of bank financing in specific contexts. The variable with the most substantial absolute effect is the Interest Coverage Ratio (-0.00013; $p<0.01$), confirming that companies with greater solvency are strongly rewarded with significantly lower spreads, highlighting the fundamental role of financial health and the ability to meet interest obligations in credit risk assessments.

Additionally, operational efficiency indicators such as Turnover/Assets (0.0009; $p<0.01$) and Personnel Expenses/ Assets (0.0036; $p<0.01$) are also statistically significant with positive coefficients. These effects may suggest that banks view more operationally intensive structures or higher fixed costs more cautiously, resulting in higher spreads as a form of risk compensation.

The estimated model yields an R^2 of 0.0696, indicating that the included variables explain approximately 7% of the variation in bank spreads. While relatively small, this level of explanatory power is consistent with findings in the literature, where R^2 values often range between 5% and 20%, particularly in studies using firm-level or cross-country data. Even in Demirgüç-Kunt et al. (1999) extensive analysis using data from 80 countries, the highest R^2 values reported range between 0.50 and 0.60, leaving a substantial portion of variation unexplained, underscoring the fact that modest R^2 values are common in banking spread regressions, because of the complex and partially unobservable determinants of credit pricing decisions.

4.2. Robustness

To evaluate the regression's robustness, an alternative specification of the model was tested to verify whether the main results remain stable with the exclusion of the variable Interest Coverage Ratio, which demonstrated a high level of statistical significance and scale. The goal of this specification was to verify whether the main results remain consistent in the absence of this dominant variable. Additionally, another robustness check was conducted by removing fixed effects for sector and region, this allowed for the assessment of whether the results are sensitive to unobserved heterogeneity related to industry and geographic location.

Comparison of Models

	Model 1	Model 2	Model 3
	Spread	Spread	Spread
Return on Equity (ROE)	-0.0016	0.0019	-0.0016
	(0.0011)	(0.0010)	(0.001)
Profit Margin	-0.0062**	-0.0097***	-0.006**
	(0.0025)	(0.00264)	(0.0024)
EBITDA/Assets	0.0185***	0.0155***	0.015***
	(0.0033)	(0.0034)	(0.0032)
Leverage	-0.0037***	-0.000338***	-0.00043***
	(0.00010)	(0.00010)	(0.0001)
Interest Coverage Ratio	-0.00013***	-	-0.000129***
	(0.000004)	-	(0.0000043)
Turnover/Assets	0.0009***	0.00086***	0.0012***
	(0.0002)	(0.00025)	(0.00023)
Personnel Expenses/Assets	0.0036***	0.0037***	0.0026***
	(0.0010)	(0.0011)	(0.00094)
Sectoral Dummy	Yes	Yes	No
Regional Dummy	Yes	Yes	No
Constant	1.8395***	1.8192***	1.868***
	(0.0550)	(0.0563)	(0.0528)
R-squared	0.0696	0.0209	0.0622
Observations	10368.0000	10368.0000	10368.0000

Standard errors in parentheses

Model 1: Baseline Model. Model 2: Model without Interest Coverage Ratio. Model 3: Model without Sectoral and Regional Fixed Effects

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4 - Robustness Test

Table 4 presents the results for the two alternative regressions developed to evaluate the robustness of the baseline model. The regression results demonstrate a strong consistency observed across all model specifications. Variables such as Profit Margin, EBITDA/Assets, Leverage, Turnover/Assets, and Personnel Expenses/Assets consistently demonstrate statistically significant relationships with the banking spread. In general terms these variables maintain their expected signs and levels of significance across all three models tested, confirming the robustness of the core findings.

Model 2 shows minor variations in the coefficients, being the Profit Margin variable, the variable exhibiting the biggest change, increasing in magnitude and significance, suggesting that, in the absence of the Interest Coverage Ratio, this variable captures a larger portion of the explanatory power.

The results demonstrate strong stability across the three different specifications:

1. The baseline model;
2. A model estimated without the Interest Coverage Ratio variable;
3. A model estimated without sectoral or regional fixed effects.

Across these, coefficient signs, magnitudes, and significance levels remain largely unchanged. Even when excluding fixed effects, the explanatory variables retain their influence, demonstrating that the results are not driven by unobserved heterogeneity.

However, in Model 2, where Interest Coverage Ratio is excluded, there is a considerable drop in the explanatory power, with R-squared decreasing from 0.0696 to 0.0209. This indicates the important role that Interest Coverage Ratio plays in capturing variance in the bank spread, reinforcing its dominant predictor in the baseline model.

These findings confirm that the estimated relationships are not only statistically significant but also empirically robust, supporting the credibility of the analysis.

4.3. Out-of-Sample Validation

In addition to the robustness checks, an out-of-sample validation test was implemented. Out-of-sample validation refers to the process of testing a model's predictive performance on a subset of data that was excluded from the estimation phase, with the goal of assessing how well the model can generalize to unseen observations.

The procedure involved dividing the full firm sample into two subsets: a training set (70%), used to estimate the model, and a test set (30%), reserved exclusively for out-of-sample validation, the observations were randomly assigned to each subset, ensuring representativeness across sectors, firm sizes, and geographic locations.

The model was first estimated using the training set, and the resulting coefficients were then applied to the test set to generate predicted values of the bank spread., then, the model's predictive performance was assessed using standard forecast error metrics, namely the Mean

Absolute Error (MAE) and the Root Mean Squared Error (RMSE), which quantify the accuracy of the predictions on data not used in estimation.

4.3.1. Out-of-Sample Estimation - 70%

Out-of-the Sample Estimation

	spread
Return on Equity (ROE)	-0.00096
	(0.00122)
Profit Margin	-0.00834**
	(0.00306)
EBITDA/Assets	0.01811***
	(0.00468)
Leverage	-0.00059***
	(0.00011)
Interest Coverage Ratio	-0.00013***
	(0.000005)
Turnover/Assets	0.00090***
	(0.0003)
Personnel Expenses/Assets	0.00346**
	(0.00128)
Sectoral Dummy	Yes
Regional Dummy	Yes
Constant	1.898***
	(0.0659)
R-squared	0.0746
Observations	7243.0000

Standard errors in parentheses

Model 4: 70%* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5 - Regression Test (70%)

The regression results confirm the robustness of the main relationships, showing that Profit Margin remains negatively and significantly associated with the credit spread, while EBITDA/Assets show a positive and significant effect, suggesting that higher operating returns may carry perceived risk.

Both Leverage and the Interest Coverage Ratio are negatively associated with the spread, in line with the previous expectations. Additionally, Turnover/Assets and Personnel Expenses/Assets are positively and significantly related to the spread, possibly reflecting operational or structural risk.

4.3.2. Out-of-the Sample Validation - 30%

Variable	Obs	Mean	Std. dev.	Min	Max
Absolute error	3125	-0.0802835	2.394954	-4.98957	10.24073
Squared error	3125	5.740412	10.88949	1.26e-08	104.8725

Table 6 - Out-of-the Sample Validation

To evaluate the predictive accuracy of the model, the absolute and squared errors were computed for the test group, comprising 3,125 observations. The results reveal a nuanced picture of the model's out-of-sample performance.

The MAE is –0.080, which indicates that the model slightly overestimates the spread on average, but this bias is minimal and close to zero indicating the absence of a systematic tendency to consistently over or underestimate the dependent variable, which is a positive indicator of model balance.

The standard deviation of the absolute error is 2.40, reflecting a moderate dispersion of prediction errors around the mean. While most errors remain within a reasonable range, the minimum (–4.99) and maximum (10.24) values show that, in extreme cases, the model can diverge significantly from the true values, these extreme values can result from unusual observations or patterns not fully captured by the model's independent variables.

The mean MSE is approximately 5.74, representing the average squared differences between the model's output and the actual values in the out-of-sample data. Since the errors are squared, this value is sensitive to large deviations providing a conservative estimate of average error magnitude by giving greater weight to large errors. The standard deviation of the squared error is relatively high (10.89), indicating a wide variability in the errors, suggesting that while many predictions can be close to the real values, there are also many predictions with much larger errors, resulting in a certain volatility in the error sizes. Regarding the maximum squared error, this metric reaches 104.87, confirming the presence of extreme prediction errors in the sample, having a strong influence on MSE, pulling the average error upward.

These results indicate that the model performs well on average, with no significant bias and generally moderate error levels. However, the presence of high variance and extreme outliers suggests that in some instances, the model may struggle to capture more volatile or atypical spread behaviors.

5. Conclusion

The main objective of this study was to identify the determinants of the bank spread applied to Portuguese firms by estimating a multiple linear regression model using firm-level financial data. The spread was defined as the difference between the effective interest rate paid by firms and the 12-month Euribor, therefore representing the additional cost of financing relative to the market benchmark.

This study demonstrates that several financial variables have a statistically significant impact on the spread. Profit Margin and Interest Coverage Ratio demonstrated negative and highly significant effects, confirming that more profitable and solvent firms benefit from better financing conditions, in contrast with these results, the EBITDA/Assets ratio showed a positive and robust association, which may reflect banks' perception of higher risk.

Leverage was also negatively associated with the spread, suggesting that higher debt levels may signal financial maturity and greater ability to manage debt. Finally, the operational efficiency indicators, Turnover/Assets and Personnel Expenses/Assets were statistically significant with positive effects, possibly indicating that lenders perceive more complex or cost-intensive structures as riskier.

The model incorporates sectoral and regional fixed effects to control unobserved heterogeneity related to industry and geographic context. These fixed effects were included to isolate the influence of firm-level financial characteristics on the borrowing spread.

In summary, this study aims to contribute to a deeper understanding of the factors influencing corporate borrowing costs in Portugal, highlighting the importance of SME's financial health.

6. Bibliography

Abdou, H. A., & Pointon, J. (2011). Credit scoring, statistical techniques and evaluation criteria: A review of the literature. *Intelligent Systems in Accounting, Finance and Management, 18*(2–3), 59–88. https://doi.org/10.1002/isaf.325

Altman, E. I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance, 23*(4), 589–609. https://doi.org/10.2307/2326758

Basel Committee on Banking Supervision. (2017). *Basel III: Finalising post-crisis reforms*. Bank for International Settlements. https://www.bis.org/bcbs/publ/d424.pdf

Bauer, D., & Zanjani, G. (2021). Economic capital and RAROC in a dynamic model. *Journal of Banking and Finance, 125*, 106071. https://doi.org/10.1016/j.jbankfin.2021.106071

Beaver, W. H. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research, 4*(Supplement), 71–111. https://doi.org/10.2307/2490171

Beck, T., Demirgüç-Kunt, A., & Maksimovic, V. (2008). Financing patterns around the world: Are small firms different? *Journal of Financial Economics, 89*(3), 467–487. https://doi.org/10.1016/j.jfineco.2007.10.005

Brock, P. L., & Rojas-Suarez, L. (2000). Understanding the Behavior of Bank Spreads in Latin America. *Journal of Development Economics, 63*(1), 113–134.

Carvalho, P.V. (2009). Fundamentos da Gestão de Crédito: uma contribuição para o valor das organizações. *Edições Sílabo.*

Custódio, C., Hansman, C., & Mendes, B. (2024). Credit access and market access: Evidence from a Portuguese credit guarantee scheme (Working paper). Imperial College London, Emory University, & London Business School.

Damodaran, A. (2015). *Applied corporate finance* (4th ed.). John Wiley & Sons.

Demirgüç-Kunt, A., & Huizinga, H. (1999). Determinants of commercial bank interest margins and profitability: Some international evidence. *The World Bank Economic Review, 13*(2), 379–408.

European Commission. (2003). *Commission Recommendation of 6 May 2003 concerning the definition of micro, small and medium-sized enterprises (2003/361/EC). Official Journal of the European Union*. https://single-market-economy.ec.europa.eu/smes/sme-fundamentals/sme-definition_en

European Commission. (2023). *Annual Report on European SMEs 2022/2023*. Directorate-General for Internal Market, Industry, Entrepreneurship and SMEs. https://single-market-economy.ec.europa.eu/system/files/2023-08/Annual%20Report%20on%20European%20SMEs%202023_FINAL.pdf

Fama, E. F., & French, K. R. (2002). Testing trade-off and pecking order predictions about dividends and debt. *The Review of Financial Studies, 15*(1), 1–33. https://doi.org/10.1093/rfs/15.1.1

Farinha, L., & Félix, S. (2014). Credit rationing for Portuguese SMEs. *Economic Bulletin, Autumn*, 109–123. Banco de Portugal. https://www.bportugal.pt/sites/default/files/anexos/pdf-boletim/box3_en_aut14.pdf

Fazzari, S. M., Hubbard, R. G., & Petersen, B. C. (1988). Financing constraints and corporate investment. *Brookings Papers on Economic Activity, 1988*(1), 141–195. https://doi.org/10.2307/2534426

Federal Deposit Insurance Corporation. (2004). Economic capital and the assessment of capital adequacy. *Supervisory Insights, Winter 2004*, 5–16. https://www.fdic.gov/bank-examinations/economic-capital-and-assessment-capital-adequacy

Ho, T. S. Y., & Saunders, A. (1981). The Determinants of Bank Interest Margins: Theory and Empirical Evidence. *Journal of Financial and Quantitative Analysis, 16*(4), 581–600.

INE. (2023). *Estatísticas das empresas – Análise por dimensão da empresa*. Instituto Nacional de Estatística. https://www.ine.pt

Ioannidou, V., & Ongena, S. (2010). Time for a change: Loan conditions and bank behavior when firms switch banks. *The Journal of Finance, 65*(5), 1847–1877. https://doi.org/10.1111/j.1540-6261.2010.01596

Koporcic, N., Kukkamalla, P. K., Markovic, S., & Maran, T. (forthcoming). Resilience of small and medium-sized enterprises in times of crisis: an umbrella review. *Review of Managerial Science*. https://doi.org/10.1007/s11846-025-00883-0

Mishkin, F. S. (2019). *The Economics of Money, Banking and Financial Markets* (13th ed., Global Edition). Harlow: Pearson Education.

Modigliani, F., & Miller, M. H. (1958). The Cost of Capital, Corporation Finance and the Theory of Investment. *The American Economic Review, 48*(3), 261–297.

Myers, S. C. (1977). Determinants of Corporate Borrowing. *Journal of Financial Economics, 5*(2), 147–175.

Petersen, M. A., & Rajan, R. G. (1994). The benefits of lending relationships: Evidence from small business data. *The Journal of Finance, 49*(1), 3–37. https://doi.org/10.1111/j.1540-6261.1994.tb04418.x

Pinto, A. P. S., Henriques, C. M. R., Cardoso, C. E. O. S., & Neves, M. E. D. (2023). Bank Credit and Trade Credit: The Case of Portuguese SMEs from 2010 to 2019. *Journal of Risk and Financial Management, 16*(3), 170. https://doi.org/10.3390/jrfm16030170

Pontuch, P. (2011). *Labor intensity and expected stock returns*. Université Paris Dauphine - DRM Finance. https://ssrn.com/abstract=1967489

7. Appendixes

Correlation Matrix

A correlation matrix was done to access the correlation between the explainable variables used. This analysis identifies strong linear relations between the variables that could result in redundancy and overlap information.

Levels of correlation above 0.80 can be problematic as they can impair the precision of coefficient estimates in the regression. In this study, no excessively high correlations were observed among the variables, which provides an initial indication of the absence of severe multicollinearity as it shows in the matrix below.

	ROE	Profit Margin	EBITDA/Assets	Leverage	Interest Coverage Ratio	Sales/Assets	Personnel Expenses/Assets
ROE	1						
Profit Margin	0.2344	1					
EBITDA/Assets	0.4681	0.4369	1				
Leverage	0.0297	-0.1002	-0.1226	1			
Interest Coverage Ratio	0.0581	0.0980	0.0924	-0.0290	1		
Turnover/Assets	0.1288	-0.2466	0.1140	-0.0099	0.0022	1	
Personnel Expenses/Assets	0.0567	-0.2088	0.1049	0.0360	-0.0145	0.4407	1

Table 7 - Correlation Matrix

Stata Code

```
drop if (Financiamentosobtidoscorrentes + Financiamentosobtidosnãocorre ) == 0

drop if (Financiamentosobtidoscorrentes + Financiamentosobtidosnãocorre ) > 0 & Jurosdefinanciamentosobtidos == 0

drop if TaxadeJuro > 15

winsor2 Retornosobrecapitalpróprio, cut(1 99) replace

winsor2 Margemdelucro, cut(1 99) replace

winsor2 Alavancagem, cut(1 99) replace
```

```
winsor2 EBITDAATIVO, cut(1 99) replace

winsor2 VendasAtivo, cut(1 99) replace

winsor2 Capacidadedecobrirjuros, cut(1 99) replace

winsor2 GastosPessoalAtivo, cut(1 99) replace

gen spread = TaxadeJuro - 2.089

encode SecçãoCAE, gen(sec_cae_cat)

encode Distrito, gen(distrito_cat)

reghdfe spread Retornosobrecapitalpróprio Margemdelucro EBITDAATIVO Alavancagem Capacidadedecobrirjuros VendasAtivo GastosPessoalAtivo, absorb(sec_cae_cat distrito_cat) vce(robust)


// set seed ensures that the randomization will always be the same when running the code.
set seed 12345


// runiform() generates continuous random values between 0 and 1.

// Here, we use these values to split the sample: 70% training, 30% testing.

gen rand = runiform()

gen grupo_treino = (rand <= 0.7)

gen grupo_teste  = (rand > 0.7)


reghdfe spread Retornosobrecapitalpróprio Margemdelucro EBITDAATIVO Alavancagem Capacidadedecobrirjuros VendasAtivo GastosPessoalAtivo, absorb(sec_cae_cat distrito_cat) == 1 vce(robust)


predict spread_estimado if grupo_teste == 1
```

```
gen erro = spread - spread_estimado if grupo_teste == 1

gen erro_quad = erro^2 if grupo_teste == 1

summarize erro erro_quad if grupo_teste == 1

// End Script
```