KAUNAS UNIVERSITY OF TECHNOLOGY

AIDAS MALAKAUSKAS

THE EVALUATION OF ACCESS TO CREDIT FOR SMALL AND MEDIUM ENTERPRISES

Doctoral dissertation
Social Sciences, Economics (S 004)

2023, Kaunas

This doctoral dissertation was prepared at Kaunas University of Technology, School of Economics and Business, Academic Centre of Economics, Business and Management, Sustainable Economics research group during the period of 2019 to 2023.
The doctoral right has been granted to Kaunas University of Technology together with Klaipėda University and Lithuanian Energy Institute.

Scientific Supervisor:
Assoc. Prof. Dr. Aušrinė LAKŠTUTIENĖ (Kaunas University of Technology, Social Sciences, Economics, S 004).

Edited by: English language editor Dr. Armandas Rumšas (Publishing House *Technologija*), Lithuanian language editor Aurelija Gražina Rukšaitė (Publishing House *Technologija*).

Dissertation Defense Board of Economics Science Field:
Prof. Dr. Rytis KRUŠINSKAS (Kaunas University of Technology, Social Sciences, Economics, S 004) – **chairperson**;
Assoc. Prof. Dr. Asta BALIUTĖ (Kaunas University of Technology, Social Sciences, Economics, S 004);
Prof. Dr. Andrzej BUSZKO (University of Warmia and Mazury in Olsztyn, Poland, Social Sciences, Economics, S 004);
Prof. Dr. Astrida MICEIKIENĖ (Vytautas Magnus University, Social Sciences, Economics, S 004);
Prof. Dr. Dalia ŠTREIMIKIENĖ (Lithuanian Energy Institute, Social Sciences, Economics, S 004).

The public defense of the dissertation will be held at 1 p.m. on 10 November, 2023 at the public meeting of Dissertation Defense Board of Economics Science Field in the meeting room A228 at Santaka Valley of Kaunas University of Technology.

Address: K. Baršausko 59-A228, Kaunas, LT-51423, Lithuania
Phone: (+370) 608 28 527; e-mail doktorantura@ktu.lt

Doctoral dissertation was sent out on 10 of October, 2023.
The doctoral dissertation is available on the internet http://ktu.edu and at the libraries of Kaunas University of Technology (Gedimino 50, Kaunas, LT-44239, Lithuania), Klaipėda University (H. Manto 84, Klaipėda, LT-92294, Lithuania) and Lithuanian Energy Institute (Breslaujos 3, Kaunas, LT-44403, Lithuania).

KAUNO TECHNOLOGIJOS UNIVERSITETAS

AIDAS MALAKAUSKAS

KREDITO PRIEINAMUMO VERTINIMAS MAŽOMS IR VIDUTINĖMS ĮMONĖMS

Daktaro disertacija
Socialiniai mokslai, ekonomika (S 004)

2023, Kaunas

Disertacija rengta 2019–2023 metais Kauno technologijos universiteto Ekonomikos ir verslo fakultete, Ekonomikos, verslo ir vadybos akademiniame centre, Tvarios ekonomikos mokslo grupėje.
Doktorantūros teisė Kauno technologijos universitetui suteikta kartu su Klaipėdos universitetu ir Lietuvos energetikos institutu.

Mokslinis vadovas:
doc. dr. Aušrinė LAKŠTUTIENĖ (Kauno technologijos universitetas, socialiniai mokslai, ekonomika, S 004).

Redagavo: anglų kalbos redaktorius dr. Armandas Rumšas (leidykla „Technologija"), lietuvių kalbos redaktorė Aurelija Gražina Rukšaitė (leidykla „Technologija")

Ekonomikos mokslo krypties disertacijos gynimo taryba:
prof. dr. Rytis KRUŠINSKAS (Kauno technologijos universitetas, socialiniai mokslai, ekonomika, S 004) – **pirmininkas**;
doc. dr. Asta BALIUTĖ (Kauno technologijos universitetas, socialiniai mokslai, ekonomika, S 004);
prof. dr. Andrzej BUSZKO (Varmijos ir Mozūrijos universitetas Olštyne, Lenkija, socialiniai mokslai, ekonomika, S 004);
prof. dr. Astrida MICEIKIENĖ (Vytauto Didžiojo universitetas, socialiniai mokslai, ekonomika, S 004);
prof. dr. Dalia ŠTREIMIKIENĖ (Lietuvos energetikos institutas, socialiniai mokslai, ekonomika, S 004).

Disertacija bus ginama viešame Ekonomikos mokslo krypties disertacijos gynimo tarybos posėdyje 2023 m. lapkričio 10 d. 13 val. Kauno technologijos universiteto „Santakos" slėnyje, Posėdžių kambaryje A228.

Adresas: K. Baršausko g. 59-A228, Kaunas, LT- 51423, Lietuva.
Tel: (+370) 608 28 527; el. paštas doktorantura@ktu.lt

Disertacija išsiųsta 2023 m. spalio 10 d.
Su disertacija galima susipažinti interneto svetainėje http://ktu.edu ir Kauno technologijos universiteto bibliotekoje (Gedimino g. 50, Kaunas, LT-44239, Lietuva), Klaipėdos universiteto bibliotekoje (H. Manto g. 84, Klaipėda, LT-92294, Lietuva) ir Lietuvos energetikos instituto bibliotekoje (Breslaujos g. 3, Kaunas, LT-44403, Lietuva).

Table of Contents

List of Figures

List of Tables

INTRODUCTION

Relevance of the topic. Small and medium-sized enterprises (SMEs) play a crucial role in the world economy. One of the main challenges they are facing is the limited ability in obtaining funding from external credit providers (Muller et al., 2022). Approximately one-third of SMEs cite a lack of access to affordable funding as a major obstacle to their growth, resilience, and survival (ECB, 2022b). Limited ability to access external credit leads to reduced sales, liquidity constraints, and supply chain shocks, which adversely impacts growth opportunities and can force SMEs to lay off employees or shut down (Khan, 2022). Compared to large companies, SMEs are more likely to be credit rationed and face worse financing conditions, such as shorter credit facility maturities, higher collateral requirement, and higher interest rates (Chodorow-Reich et al., 2022). Such adverse conditions are attributed to the inherent information asymmetry, which leads to the inability of financing providers to appropriately assess a company's creditworthiness. The limited SME ability to access credit is related to a range of individual factors, including limited collateral availability, weaker financial health, and higher susceptibility to market and industry-specific risks (Angori et al., 2019). Macro-specific factors, such as the Lending Infrastructure and Financial Institution Structure, determine the underlying credit market conditions which have direct impact on SME ability to access credit. Competitive markets tend to have a significant positive effect on the access to credit through lower interest rates and higher loan amounts for SMEs (Kärnä and Stephan, 2022). Improving access to credit for SMEs is crucial for supporting their growth and economic well-being. A potential way of achieving it is by reducing the informational opaqueness of SME entities by evaluating the underlying SME access to credit and determining the underlying conditions that impact it.

This dissertation proposes a model for evaluating access to credit for small and medium-sized enterprises, which enables to identify the underlying conditions that define a company's ability to access credit and are important for small and medium-sized enterprises, as well as for state legal regulatory authorities. A comprehensive evaluation of the access to credit and the underlying conditions for small and medium-sized enterprises would allow to reduce SME informational opaqueness and create further opportunities to obtain external credit.

Research problem. The scientific literature analysis suggests that SME access to credit can be studied through a range of proxies and underlying factors impacting it. This dissertation defines *access to credit* as the ability of an SME to access external credit from the traditional commercial banks. Studies on access to credit can be broadly classified into two categories: credit demand and credit supply (Maier, 2016; Angori et al., 2019; Altavilla et al., 2021). Credit demand studies focus on the factors influencing the borrower's decisions to apply and the reasons for borrower discouragement (Mac An Bhaird et al., 2016; Nguyen et al., 2021; Altavilla et al., 2021). Credit supply studies tend to focus on either the bank loan portfolio and macro-specific conditions

(Bolton et al., 2016; Peón and Guntín, 2021; Altavilla et al., 2021), or the application outcomes and factors affecting a company's ability to receive approval or be rationed (Kirschenmann, 2016; Chodorow-Reich et al., 2021). The selection of the actual proxy depends on the research problem and data availability (Lee et al., 2015). While studying SME access to credit through the bank loan portfolio is useful in capturing the macro-level impact on credit supply, it may not consider company-specific characteristics that could be crucial in understanding the factors influencing access to credit. Conversely, the latter, which examines the outcomes of individual financing applications and decisions, has the potential for greater data granularity, but it is dependent on the openness and robustness of the data sources (Kirschenmann, 2016). By studying access to credit through application or decision outcomes, researchers can understand specific reasons for a lower or higher access to credit, as well as the cases where conditional approvals are issued. Negative financing application outcomes are usually categorized as first-degree rationing, where financing applications are completely rejected (Jiménez et al., 2012), or second-degree rationing, where financing applications are approved but with adjustments in product conditions, such as the amount of financing, the cost of borrowing, the maturity term, and/or the requested collateral (Berger et al., 2022). Worse product conditions can lead to the company's inability to access credit, thus underscoring the importance of studying access to credit through the application or decision outcomes. The relevance of the topic is highlighted by its comparability to estimating a company's financial distress, as banks base their credit decision-making processes on evaluating a borrower's ability to make future loan installments (Molina and Preve, 2012).

The ability of SMEs to access credit is influenced by various factors which can be classified into macro-specific and individual application factors (Berger and Udell, 2006). The macro-specific factors include the lending infrastructure and the financial institution structure, which set the underlying market conditions and are beyond the control of individual entities. These factors are essential in determining the overall financial health of the market, and, consequently, the underlying access to credit that SMEs are exposed to (Dobbie et al., 2020; Angori et al., 2020). A market with strong accounting standards (Florou and Kosi, 2015; Deno et al., 2020), marked-to-market balance sheets (Adrian and Shin, 2010), and active rating agencies is likely to have lower financial constraints (Bosch and Steffen, 2011) and a higher access to credit for SMEs. Moreover, competitive markets have a significant positive impact on the access to credit, which leads to lower interest rates and higher loan amounts for SMEs (Love and Pería, 2015; Wang et al., 2020). Individual application factors refer to the underlying characteristics of a potential borrower and can be grouped into Lending Technology, Firm Characteristics, and Product Characteristics factor groups (Berger and Udell, 2006). These factors are specific to individual entities and play a crucial role in determining SME access to credit. The term *Lending Technology factors* refers to the technological tools and set-ups that lenders use to evaluate SME creditworthiness. It is generally split into two groups: *Transaction Lending* and *Relationship*

Lending. Transaction Lending considers financial statement data and credit history to evaluate borrowers (Motta and Sharma, 2020). Therefore, it is typically employed for larger, more transparent borrowers who provide audited and comprehensive financial statements (Palazuelos et al., 2018; Ferri et al., 2019). The availability of standardized and verifiable information is critical for credit screening and monitoring processes, whereas a good credit history increases the likelihood of obtaining a credit (Cassar et al., 2015). In contrast, Relationship Lending is the preferred approach when the information about the company is limited, and creditworthiness can be assessed based on past relationships (Durguner, 2017; Rabetti, 2022). These technological differences have significant implications for the SME access to credit, as an individual borrower may be able to secure funding based on factors belonging to one lending technology but not the other (Angori et al., 2019; Chodorow-Reich et al., 2022). Though the difference between credit supply and credit demand is distinct, it is evident that the relative fuzziness between the two groups of factors actually exists, which suggests that variables belonging to both underlying factor groups should be employed when assessing access to credit (Ferri et al., 2019). The *Firm Characteristics* factor group is unique to each entity and may include such factors as the company's size, age, sector (Mina et al., 2013), or ownership structure (Aterido and Hallward-Driemeier, 2011; Sikochi, 2020; de Andrés et al., 2021). Younger SMEs may have a more challenging time accessing credit than the more established ones due to the lack of a proven track record (Mac An Bhaird et al., 2016). Finally, the *Product Characteristics* factor group relates to the features of the financing product itself, such as the amount of collateral required (Gurara et al., 2020; Berger et al., 2022), the interest rate (Xu et al., 2020; Kärnä and Stephan, 2022), and the contract maturity (Minnis and Sutherland, 2017; Aoki, 2021). The choice of the right financing product can have a significant impact on the credit access an SME ultimately receives (Adam and Streitz, 2016; Gurara et al., 2020). Scientific literature analysis has shown that research, which would comprehensively evaluate SME access to credit by utilizing factors considering both macro-specific and individual application factor groups, is scarce. Therefore, it is necessary to develop and apply the SME access to credit evaluation methodology which would utilize factors from both factor groups and would define factors which are important for SMEs when accessing credit.

Scientific studies have explored various access to credit modelling techniques with a focus on the perceived creditworthiness of the applying company (Molina and Preve, 2012; Kruppa et al., 2013; Pal et al., 2016). However, estimating access to credit is challenging due to several factors, such as the multicollinearity of independent variables, data availability, and human biases (Dastile et al., 2020). To accurately evaluate the SME access to credit and the importance of the underlying factors, it is essential to consider both macro- and individual application factors (Berger and Udell, 2006). Some studies estimate country specific models without utilizing macro-specific factors which help to mitigate the possibility of the omitted variable bias which would be present if a cross-country model were developed (Angori et al., 2020; Calabrese et al., 2022; Kärnä and Stephan, 2022). A range of modelling techniques are utilized to study SME access to credit ranging from the traditional techniques, such as Discriminant Analysis (Barboza et al., 2017) and Logistic Regression (Wang et al., 2020; Malakauskas and Lakštutienė, 2021; Medianovskyi et al., 2023), to state-of-the-art machine learning techniques, such as decision trees (Trivedi, 2020), random forest (Medianovskyi et al., 2023), artificial neural networks (Hadji Misheva et al., 2021), support vector machines (Silva et al., 2020), and k-nearest neighbor (Hussin Adam Khatir and Bee, 2022). While the traditional techniques like Logistic Regression hold an advantage in terms of variable interpretability and stability, they are not well-suited for managing larger datasets and variable interdependencies (Correa Bahnsen et al., 2016). On the other hand, machine learning techniques have become increasingly popular in recent years, as they offer the potential to improve accuracy and reduce bias in credit scoring models. Several studies have explored various modelling techniques in estimating SME access to credit and shown promising results for a number of modelling techniques, such as Support Vector Machines, Random Forest, Multi-layer Perceptron, and Gradient Boosting (Barboza et al., 2017; Malakauskas and Lakštutienė, 2021; Medianovskyi et al., 2023). However, each modelling technique is denoted by its own strengths and limitations, and the choice of the most appropriate technique depends on the specific problem at hand (Preece et al., 2018). The inherent black-box nature of these models means that the interpretability may be limited, and the use of explainability methods, such as Shapley Additive Explanations, would be required (Arya et al., 2019; Arrieta et al., 2020). Overall, the use of machine learning techniques can improve SME access to the credit modelling accuracy and determine non-linear dependencies and variable interactions.

Although there has been considerable research on separate factor groups and individual factors, the importance of each factor group, individual factors and their interactions remains unclear. It is evident that the evaluation of SME access to credit is a multidimensional problem which requires a systemic approach in defining SME access to credit measurement, along with the selection of the underlying factors that would be utilized in modelling and choosing the appropriate modelling techniques which would estimate an accurate and interpretable model. The currently available scientific literature is limited in providing a comprehensive methodology for evaluating SME access to credit. Since there is a lack of research on SME access to credit evaluation in terms of using machine learning techniques, this dissertation comprehensively evaluates access to credit for Small and Medium Enterprises for the first time by using the example of the Baltic States. This dissertation fits into a growing body of literature which uses SME financing application outcomes to evaluate the underlying SME access to credit. The relevance of the topic and the research problem is based on the need to create a model that would help to comprehensively evaluate the availability of credit and the factors which are important in accessing credit for small and medium-sized enterprises. This dissertation studies the research problem regarding the ways how to evaluate SME access to credit.

The object of the research: underlying factors, impacting access to credit of small and medium enterprises.

Research aim: to create an SME access to credit evaluation model and apply it empirically.

Research objectives:

1. To analyze the importance of access to credit for Small and Medium Enterprises.
2. To identify proxies used in measuring SME access to credit and the underlying factors which determine SME access to credit.
3. To analyze SME access to credit modelling and explainability techniques.
4. To create a comprehensive SME access to credit evaluation model.
5. To empirically apply the SME access to credit evaluation model based on the example of the Baltic States.

Research methods. To determine the significance of SME access to credit, identify the factors impacting it and determine modelling techniques, analytical research is employed, which involves systematic organization, comparison, generalization, analysis and synthesis of scientific literature. To empirically evaluate SME access to credit, the following methods are used: descriptive statistic data analysis to carry out the comparative analysis for the determining the underlying SME access to credit, dimensionality reduction process which utilizes correlation heat-maps and Euclidean distance clustering for creating the representative feature vectors, machine learning techniques (Logistic Regression, Random Forest, Multi-layer Perceptron, and Gradient Boosting) for estimating the SME access to credit model, Receiver Operating Characteristic Area and the Curve and Average Precision for evaluating the performance of the estimated models, mean absolute SHAPley additive explanations and Permutation Feature Importance metrics to evaluate the global importance of variables, SHAP and partial dependence plots to evaluate the local importance of variables and their interactions.

Research information sources and dataset. To conduct the analysis and create an SME access to credit model, the dissertation uses studies presented in scientific publications and included in the following database: Elsevier, CA Web of Science, Scopus, EBSCO, Emerald Management, Springer, Google Scholar. Financing applications were retrieved from a credit institution operating in the Baltic States. The data includes SME credit applications received throughout the period of 2018-2022.

Research limitations:

1. The empirical model does not consider formal and informal connections between applying SMEs and other larger group companies due to treating them as a single

category. This homogenization can have implications on the research findings, particularly for SMEs that are part of a large group, which have an advantage in accessing credit over independent SMEs in terms of availability of the resources, such as collateral and guarantees.

2. The empirical application of the SME access to credit model utilizes only financing application approvals and 1st degree rationing outcomes as a proxy to evaluate the SME access to credit, which effectively does not account for conditional approvals (2nd degree rationing) and cases when the potential borrowers were entirely discouraged from applying. The inclusion of other proxies would enable for a more complete SME access to credit evaluation.

Scientific novelty and significance of the research findings. The scientific novelty of this dissertation lies in the creation of a comprehensive model for evaluating SME access to credit, which not only considers the underlying factor groups but also considers individual factors and their interactions. This dissertation utilizes the cross-disciplinary approach to conducting research by combining economic and mathematical sciences. Through the development of a conceptual model and the selection of state-of-the-art modelling techniques, the dissertation proposes a three-stage SME access to credit evaluation model which is empirically applied in a country specific context. The created model combines the latest SME access to credit studies, proposes factor selection procedures, and utilizes state-of-the-art modelling techniques and explainability methods. This model is novel, and it has been applied for the first time in an empirical setting. It enables for a more comprehensive understanding of the SME access to credit and the driving forces behind it. The purpose of developing an SME access to credit evaluation model is to provide a tool for assessing the SME access to credit along with the underlying factors and the issue of making better informed decisions by borrowers – when evaluating their ability to access credit, lenders – when evaluating received applications or internal policies, and regulators – when considering legislation.

The empirical research findings reveal that SME access to credit is not uniform across countries, and the importance and the impact of the underlying factor groups and individual factors also vary significantly. The dissertation findings suggest that the underlying conditions which define the SME access to credit differ between countries, and therefore the policymakers and financial institutions need to consider country-specific factors when designing policies and products to support SME access to credit. Additionally, the research provides insight into the underlying factors which are crucial when SMEs try to access credit, which is of relevance to SME companies, governments, and financial institutions. The research findings can aid in further scientific investigations related to access to credit and could help in developing policies and financial products to improve SME access to credit. This methodology is country agnostic, and therefore it can be applied in different country settings.

Logical structure of the dissertation. This dissertation consists of an introduction, 3 parts, conclusions and references. The dissertation was prepared by utilizing literature sources. The first section analyzes the importance of the access to credit for Small and Medium Enterprises, identifies proxies for measuring SME access to credit and the underlying factors which determine SME access to credit and analyzes the SME access to credit modelling and explainability techniques. In the second section, a methodology for evaluating SME access to credit is created. In the third section, SME access to credit is empirically evaluated in a country-specific setting. The summarized logical structure of the thesis is shown in the Figure 1. In addition, the dissertation contains Annexes, a Summary of the Dissertation, a List of References, a List of Author's Scientific Publications on the Topic of the Dissertation, a List of Scientific Conferences where the Results of the Research were Presented, Copies of the Published Articles, and Author's Curriculum Vitae.

The volume of the dissertation is: 184 pages excluding the Annexes. The thesis contains 45 tables, 26 figures, and 10 annexes. The List of References contains 273 references.

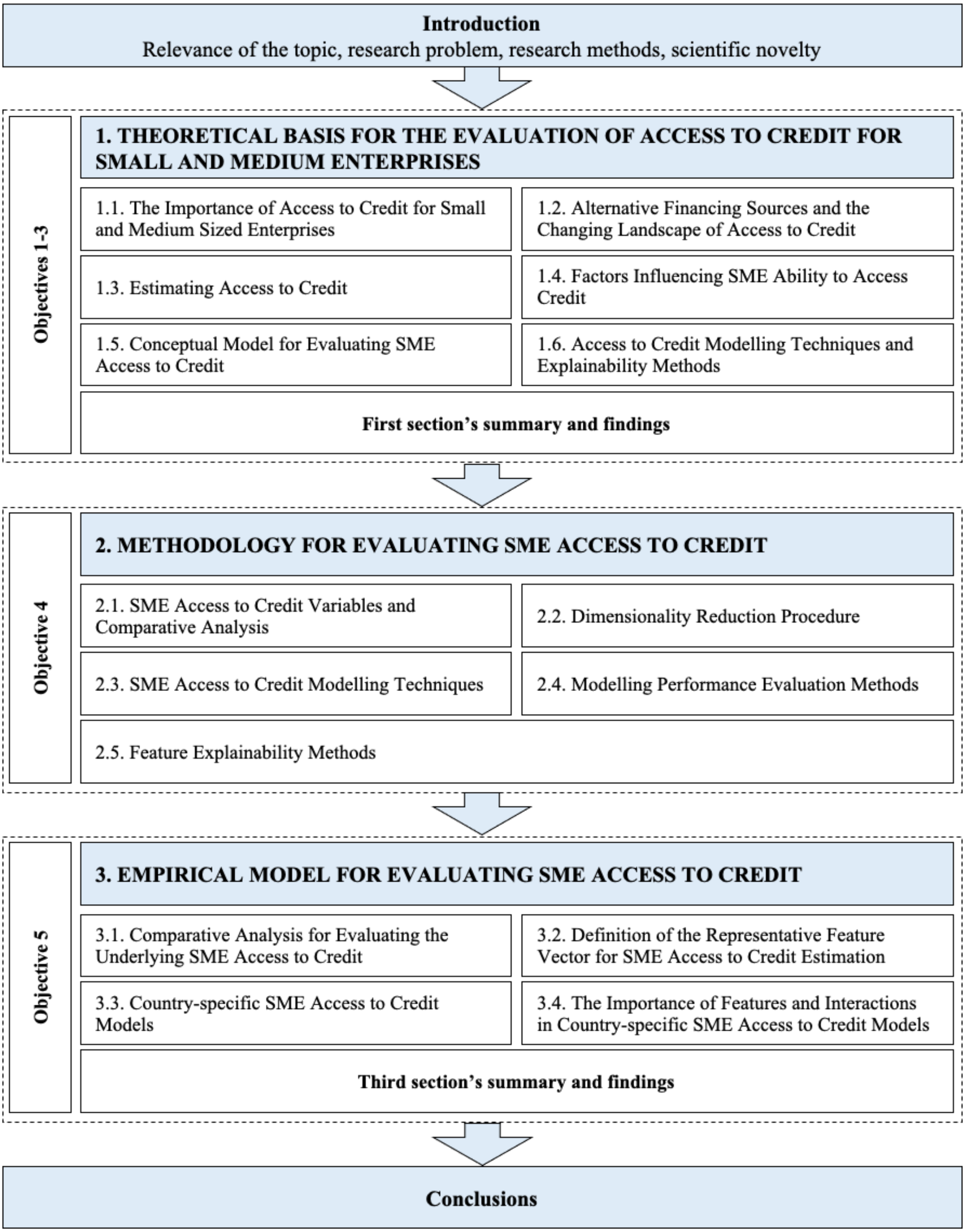


Figure 1. The logical structure of the dissertation. Created by the author.

1. THEORETICAL BASIS FOR THE EVALUATION OF ACCESS TO CREDIT FOR SMALL AND MEDIUM ENTERPRISES

In the first section, the dissertation solves objectives 1, 2 and 3. First, it examines the significance of the access to credit for SMEs. It discusses proxies that can be used to measure SMEs access to credit and identifies the key underlying factors. Additionally, the section analyzes the different models, techniques and explainability methods used to evaluate SME access to credit and the underlying factors. The findings and the summary are presented at the end of the section.

1.1. The Importance of Access to Credit for Small and Medium Sized Enterprises

The definition of small and medium enterprises (SMEs) varies across different regions due to differences in the economic, legal and cultural contexts. In the United States, the definition of SMEs depends on the industry the company is operating in and the agency that is evaluating the company. The Small Business Administration (SBA) is one of the main agencies responsible for defining and supporting small businesses in the US. According to the SBA, a small business is defined as an independent business having fewer than 500 employees. The SBA also provides definitions for small businesses in specific industries, based on the *North American Industry Classification System* (NAICS) code, which defines the size standard in terms of the average annual receipts. The size standards range from 7.5 to 38.5 million USD in average annual receipts, depending on the industry. As defined by OECD (2022), SMEs are non-subsidiary, independent firms which employ fewer than a given number of employees and are not exceeding individual or a group of financial asset rules. The actual thresholds vary across countries, and the most frequently accepted ones are set by the European Union and the United States. In the European Union, SMEs are defined by a combination of four criteria: the number of employees, the annual turnover, the balance sheet total, and the ownership structure. These criteria are used to determine whether a company is considered a micro, small, or medium-sized enterprise. For a business to be considered an SME in the EU, the upper limit for employees is 250, the sales revenue should not exceed EUR 50 million, or the balance sheet should not exceed EUR 43 million (EC, 2003).

Table 1. The definition of SMEs by company thresholds. Based on OECD (2022).

Country	Headcount	Other measures
USA	<500	Depending on the industry, the total amount of annual receipts up to US$7.5 million to US$38.5 million.
European Union	<250	The annual turnover <EUR 50 million or the balance sheet total <EUR 43 million. The company must also be autonomous in terms of ownership and resource availability.
China	from <200 to <3000 (depending on the industry)	Depending on the industry, the annual turnover <RMB 300 million or the balance sheet total <RMB 400 million.
Australia	<200	The Australian Bureau of Statistics uses only headcount, while the Australian Tax Office uses only annual turnover <AUS$250 million.

Based on the European Union definition, the size is not the only factor which is considered when determining whether an entity is an SME. In fact, a business can be fulfilling all size requirements to be considered an SME, but might still have access to significant resources due to being owned, linked, or partnered with a different large company. The European Commission defines three categories of enterprises based on the resource availability: autonomous, if an enterprise is completely independent or has one or more minority (<25%) partnerships; partner if holdings are more or equal to 25% but not more than 50%; and a linked enterprise if the holdings exceed 50%. Depending on the category, when determining if an enterprise is an SME, different inputs should be used. If a company has a 30% stake in a different company, 30% of the partner's headcount should be added to the original business. Such stipulations are necessary, as companies which have linkages to large businesses have advantages in comparison to stand-alone ones, notably from the angle of financial resource availability. Notably, such resource criteria are not used in the United States, where the SME definition depends on specific industries (USITC, 2010).

Table 2. The European Commission definition of SMEs by company size thresholds. Based on EC (2003).

Enterprise Category	Headcount	Annual turnover	or	Annual balance sheet
Micro	<10	<EUR 2 million	or	<EUR 2 million
Small	<50	<EUR 10 million	or	<EUR 10 million
Medium-sized	<250	<EUR 50 million	or	<EUR 43 million

As pointed out by Senderovitz (2009), the term 'SMEs' is commonly used in both academic and regulatory literature, it is often not clearly defined and unambiguous. For comparability and consistency, this dissertation uses the definition and size segmentation of SMEs as defined by the European Commission without accounting for the resource availability, which is a common limitation arising due to inherent informational opaqueness of SME entities (see Table 2). Voiding resource availability condition from the SME definition homogenizes companies, which are subsidiaries (or are dependents) of larger (non-SME-size) businesses, into a common category group with independent SMEs. Given that SMEs with substantial resource availability can potentially be subject to different banking standards than independent constituents, studies should consider any potential implications.

SMEs are crucial in shaping a nation's economy and are viewed as a vast and dynamic source of innovation. With their socially and economically positive impact, this sector is deemed of strategic importance amongst national and international regulators (Manzoor et al., 2021). Across the world, SMEs account for 99% of all companies, while creating between 50-60% of value added. Two out of three persons are employed by an SME, while one in three works in a micro company employing less than 10 individuals. They are the key drivers in ensuring a sustainable economic growth as well as successful adaptation to changes coming from globalization, ageing population, and digitalization (González et al., 2019). As SMEs are the key providers of employment, they are keeping the industrial fabric in many regions as well as the social identities of both urban and rural communities. Even though the role of SMEs and their importance have been recognized by policy makers, barriers to operate efficiently are still evident. Many SMEs are struggling with the unnecessary regulatory and administrative burden, the ability to access strategic resources such as skills, knowledge, and finance. Barriers which, if not mitigated, risk to trap SMEs in a low-productivity, low-innovation cycle, would leave SMEs in an endless low-growth, low-wages, and low-employment cycle (Khan, 2022). Though growth barriers for SMEs are relatively similar across different countries, the proportion of SMEs in countries differs. The overwhelming majority of SME employment (around 25%) across the OECD countries is concentrated to the wholesale and retail sector, which has relatively low barriers to operate in terms of skills or investment. On the other hand, the manufacturing sector by virtue of being prone to high capital intensity, accounts for approximately 20% of SME employment. Though general SME specialization trends do exist, there are significant differences between countries concerning specific industries (Muller et al., 2022). Namely, in Greece, there are almost 6 times more SMEs operating in the Accommodation and Food sector than there are in Poland, while the United Kingdom has three times more SMEs in Information and Communication than Canada. Such a variation could be explained with the policy and framework differences which are driving the country's specialization. South Korea is a good illustration of the policy impact on the employment structure, through the targets supported by policy actions, Korea promoted SME linkages with large Chaebols, which led to the SME employment proportion of the vehicle and transport sector becoming much higher than in other economies. Though some changes in employment structures are evident, there have not been any significant changes in the general sector groups, rather, most changes are occurring in specific sub-sectors (such as ICT). Though there are significant differences between SMEs operating in different countries – a common challenge exists – which is the limited ability to access finance (Mazanai and Fatoki, 2012; OECD, 2022).

Access to credit presents a significant challenge for most businesses, but it constitutes an especially formidable barrier for SMEs. Due to small size, SMEs are particularly susceptible to encountering obstacles that impede their growth prospects and hinder overall economic development. In their research, Byiers and et al. (2010) examined the constraints related to credit and the demand for manufacturing companies. They found that the factors responsible for heightened credit constraints were also the ones reducing credit demand, even in cases where a company faced financial constraints. Larger companies displayed a greater propensity to seek credit, yet they were less likely to encounter such constraints. Researchers have categorized credit constraints into two distinct types: 'internal' constraints (e.g., a lack of competences or time) and 'external' constraints (e.g., high market competitiveness and the availability of resources) (Buckley and Prescott, 1989). As Beck et al. (2004) demonstrated, obstacles associated with finance, legal issues, and corruption have adverse effects on firm growth. However, not all obstacles carry equal weight, as individual companies identify the need for an established bank relationship and accessibility to financing as key growth constraints. There exists a positive feedback loop linking access to financing and SME performance (Rajan and Zingales, 1998; Giovannini and Moran, 2013). Three main literature streams that define the impact of credit constraints on business operations can be identified.

The first stream of research examines the intricate relationship between financial constraints and a firm's international endeavors, particularly the strategy of expanding through exporting. However, this growth approach is not devoid of challenges. On one hand, financing constraints may curtail exports through mechanisms tied to high fixed costs (Bellone et al., 2010; Manova et al., 2015) and variable expenses (Manova, 2013). Operating in multiple markets demands substantial resources to sustain foreign distribution channels, navigate extended time-to-sale periods, and cover the expenses associated with varying regulations and customs charges. As most of these investments are upfront and can be considered sunk costs, businesses engaged in exporting often grapple with a heightened need for additional liquidity compared to those operating solely within their domestic borders. Notably, export-oriented firms can face credit constraints due to elevated relative interest costs stemming from sovereign risk and disparities in financial sector development, as demonstrated by Kletzer and Bardhan (1987). Additionally, higher working capital intensity and limited access to relief measures steer companies, particularly in less affluent regions, toward manufacturing less capital-intensive products, which can have detrimental long-term implications for economic development. Building on the framework outlined by Kletzer and Bardhan (1987), Beck (2002) revealed that countries with more developed and efficient financial markets and lower financing constraints tend to boast higher export market shares and more favorable trade balances. Nevertheless, as highlighted by Minetti and Zhu (2011), limited access to bank debt negatively impacts both a firm's likelihood of exporting and the total sales derived from exports. Pietrovito and Pozzolo (2021) expand on this by showing that financially constrained companies not only exhibit reduced export capabilities but, if they do manage to export, they typically have a lower overall proportion of exports. Specifically, the probability of successfully accessing foreign markets for credit-constrained companies decreases by approximately 3%, while the potential share of exported goods dwindles by around 17%. Moreover, Brooks and Dovis (2020) illustrate that credit constraints curtail the responsiveness of trade volumes to trade liberalization. Furthermore, Miao and Wang (2012) link credit constraints to total factor productivity, disrupting the efficient allocation of capital between less and more efficient firms. On the contrary, exporting can mitigate financing constraints, as it allows companies to tap into international financial markets, broadening their credit supply options. Manova et al. (2015) and Chaney (2016) underscore, at the firm level, that businesses facing lower liquidity constraints, such as multinational corporations with access to capital markets and support from parent companies, generally exhibit stronger export performance compared to private domestic firms. Furthermore, exporting companies often experience more stable sales due to international diversification, which enhances their liquidity position and overall financial health (Greenaway et al., 2007). In a similar vein, a reverse causality argument posits that higher participation in international trade is associated with enhanced productivity. Consequently, a diversified sales portfolio can serve as a certification of quality, diminishing information asymmetry and bolstering credit availability (Manole and Spatareanu, 2010). The productivity of small and medium-sized enterprises (SMEs), relative to large firms, exhibits considerable variation across countries. However, when considering sectoral disparities, SMEs often lag behind larger counterparts in the manufacturing sector, characterized by its substantial capital intensity and susceptibility to economies of scale—challenges that pose particular difficulties for smaller businesses to overcome. Conversely, in many nations, SMEs demonstrate robust productivity in the service sector, which tends to be more diverse and populated with businesses emphasizing brands or intellectual property.

The second strand of literature delves into the intricate relationship between financial constraints and innovation. Historically, a straightforward assumption prevailed that financial constraints had a detrimental impact on R&D investments, driven by factors such as underlying information asymmetries (Myers and Majluf, 1984), the absence of suitable collateral, and relatively high borrowing costs (Arrow, 1962). Recent empirical findings, however, have shed light on the complexity of this relationship. While it holds true that R&D investment is influenced by underlying financial constraints, some researchers present evidence suggesting that, concerning R&D investments, the availability of internal financing sources bears greater significance than external ones (Himmelberg and Petersen, 1994; Czarnitzki and Hottenrott, 2011). Conversely, others provide evidence that R&D investments are just as sensitive to financial constraints as regular investments (Mulkay et al., 2001). A comparative study by Brown et al. (2011) involving Western European firms concluded that credit-constrained firms are less inclined to invest in R&D and introduce new products. Furthermore, Altomonte and Békés (2016) and Ferrando and Ruggieri (2018) found that financing constraints faced by SMEs exert a significantly negative impact on their productivity by curbing investment opportunities, with this effect being most pronounced in companies operating in innovative, R&D-intensive industries. In contrast, Bond et al. (2005) did not find compelling evidence of R&D investments being adversely affected by financial constraints, although they did acknowledge variations between countries. The disparities in empirical findings can be attributed to the inherent nature of R&D investments, as well as significant variations in the measurement methods employed to assess credit constraints. Investing in R&D entails allocating substantial funds to pay programmers, scientists, or engineers, and these investments often result in intangible assets that could potentially be deemed worthless. The considerable variability in these characteristics poses challenges when selecting appropriate methods and metrics for quantifying financing constraints, prompting researchers to use proxies. Fazzari et al. (1988) employed investment sensitivity to cash flows as an indirect proxy linking credit constraints to investments, but this approach was heavily criticized by Kaplan and Zingales (1997) for its disconnection from actual credit availability and apparent endogeneity issues associated with the selected measures. Some authors have employed direct measures to provide empirical evidence suggesting that credit constraints have a detrimental impact on R&D investments (Aghion et al., 2012; Brown et al., 2013; Mancusi and Vezzulli, 2014), as well as innovation (Savignac, 2007; Gorodnichenko and Schnitzer, 2013). On a related note, findings by Caggese (2019) suggest that credit constraints indirectly affect innovation through self-selection mechanisms, such as entry barriers. Furthermore, Brown et al. (2013) highlighted that banking institutions remain the primary external funding source for SMEs seeking to finance R&D investments. Meanwhile, Bougheas (2004) demonstrated that banks' willingness to finance R&D projects is largely influenced by their willingness to monitor their customers' investment activities through the use of funds.

The third and final research stream focuses on the self-reinforcing cycle between exporting and innovation, with productivity at its core. Innovation serves as the primary driver of productivity, thereby facilitating and enabling exporting. Operating in a foreign environment grants access to a broader array of knowledge sources, which can be harnessed to enhance products and processes. This, in turn, feeds back into the innovation process, further boosting productivity (Castellani and Zanfei, 2007). Van Beveren and Vandenbussche (2010) demonstrated that the anticipation of expanding into foreign markets encourages firms to innovate, a finding corroborated by Bustos (2011), who illustrated that trade liberalization spurs firms to engage in both innovation and exporting, resulting in heightened productivity. The link between productivity and access to credit was highlighted by Gatti and Love (2008), who concluded that access to financing is strongly and positively associated with a firm's total factor productivity. Additionally, Cassiman and Golovko (2011) demonstrated that innovation directly influences productivity, prompting businesses to embark on exporting endeavors. Conversely, Keller (2004) were unable to establish a clear connection between exporting and increased innovation through information diffusion. Instead, they argued that information diffusion occurs organically through a deliberate commitment to learning and alignment with international standards.

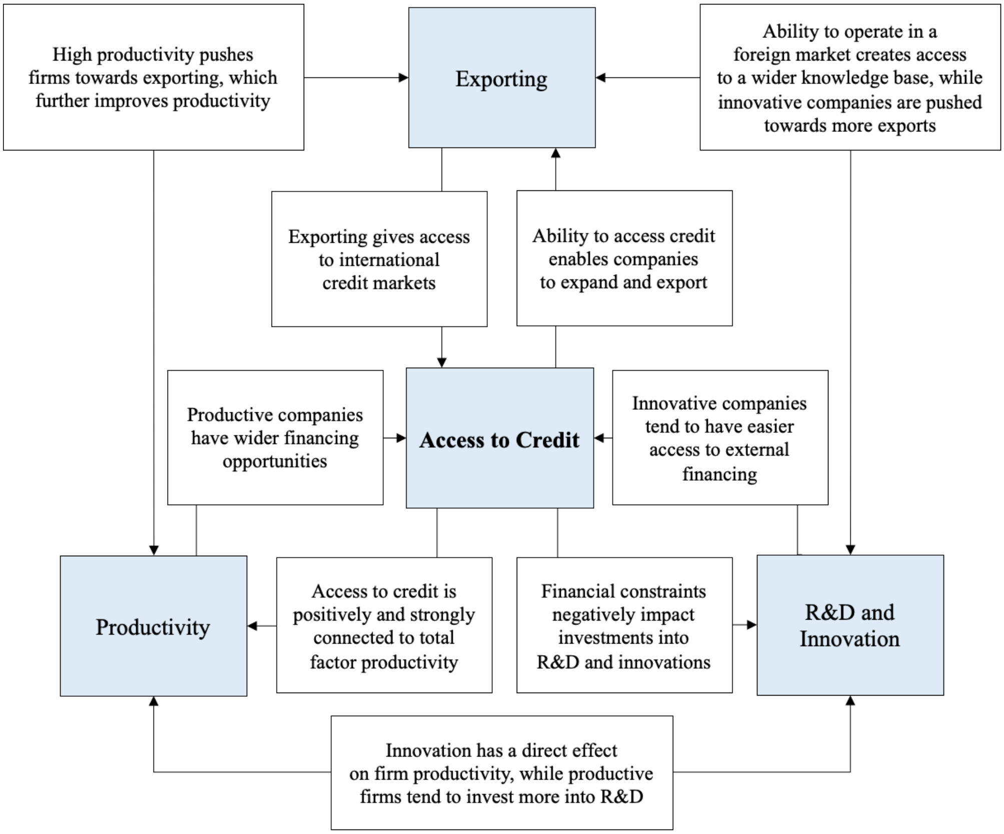


Figure 2. The importance of access to credit for SMEs. Created by the author.

Figure 2 summarizes the relationships among access to credit, exporting, productivity, and R&D and innovation. The significance of access to credit for SMEs extends across multiple dimensions, including investments in business development, company productivity, and the ability to engage in international trade. Financing constraints can hinder a firm's entry into foreign markets through a self-selection mechanism, primarily driven by high fixed and variable costs (Kletzer and Bardhan, 1987; Bellone et al., 2010; Manova, 2013; Manova et al., 2015). Additionally, limited access to bank debt has detrimental effects on both a company's probability of exporting and the total sales derived from exports (Minetti and Zhu, 2011; Pietrovito and Pozzolo, 2021). Conversely, successful international operations can alleviate credit constraints by providing access to international capital markets (Manova et al., 2015; Chaney, 2016). An inverse relationship exists between access to credit and a company's investments in R&D and innovations (Arrow, 1962; Myers and Majluf, 1984; Mulkay et al., 2001; Aghion et al., 2012; Brown et al., 2013; Mancusi and Vezzulli, 2014; Caggese, 2019). Similarly, businesses with a high capacity to access financing tend to exhibit higher productivity levels (Gatti and Love, 2008). Given the interconnectedness of these dimensions and their mutual influence, the adverse effects of financing constraints are compounded through a self-reinforcing loop, potentially resulting in reduced competitiveness and hindered economic growth. A constrained company not only suffers the direct effects of limited financial resources but also faces linked repercussions, such as diminished productivity and competitiveness on the international stage. Ultimately, as highlighted by Gross et al. (2020), sufficient access to credit plays a pivotal role in fostering increased investment and consumption.

It has been determined that limited access to credit for SMEs can lead to reduced sales, liquidity restraints, and supply chain disruptions, which may result in SMEs laying off employees or ceasing operations. Furthermore, a self-enforcing loop exists between the SME access to credit and business competitiveness and growth. It was also found that country- and academia-wide differences exist in defining what constitutes an SME and what the term ‘access to credit’ means. Therefore, this dissertation establishes the definition of an SME as defined by the European Commission.

1.2. Alternative Financing Sources and the Changing Landscape of Access to Credit

During the past decade, fund raising sources from external lenders have been changing - as specialized, niche banks and fintechs which employ innovative financing technologies are successfully challenging the traditional banks. New financing technologies are emerging not only due to changes in the supply factors like new regulations or technological progress, but also due to the shifting credit demand towards growing disintermediation (Block et al., 2018). Platform-based marketplaces enable entrepreneurs to attract financing from a larger pool of potential investors, which was previously limited only to high-net-worth individuals or local financing providers. As more capital is available for businesses to employ, alternative financing sources are successfully contributing to closing the funding gap (Coakley et al., 2018). For an alternative lender platform to operate successfully, activity from the supply side of the borrower requests and the demand side of the investment availability should match (Maier, 2016). As Maier’s (2016) study suggests, the key drivers for attracting new borrowers are the process transparency and convenience. Notably, borrowers in alternative lending providers are not looking for a complete replacement for their banking relationships, due to dissatisfaction or other reasons, but they rather want a flexible alternative. Therefore, to induce switching and attract borrowers, alternative lending platforms should work towards exceeding banks in borrower convenience. From the demand perspective, to match the growth of supply of borrowers, alternative lending platforms should focus on the investment opportunity presentation. Furthermore, borrower assessment is carried out by hundreds of potential investors, who carry out screening individually. As shown by Ralcheva and Roosenboom (2020), the alternative lending market has matured, and recent campaigns are being led by larger, older companies with more external financing options. One of the most popular forms of alternative lending - peer-to-peer (p2p) lending - is quite different from the traditional lending as lenders never meet borrowers - since a platform works as an intermediary providing the marketplace and its interface. Finally, and of utmost importance, alternative lending platforms act as non-experts in making credit decisions, which can impact the appropriate due diligence. As alternative lending opportunities enable smaller investors to access the investment market, subsequently, it creates possibility for different types of businesses to secure funding. As demonstrated by Kgoroeadira et al. (2019), in contrast to the traditional financing, alternative lenders tend to ignore business specific characteristics and prefer to look at the SME owner's characteristics, such as the credit score, assets and income. A self-employed, non-homeowner is less likely to receive funding and pay a higher interest rate, while an employed, homeowner will receive financing much more easily with a lower interest rate. These findings are also confirmed by Nisar et al. (2020) who add that married, high-income borrowers are not only more likely to get funding but also are more creditworthy. Such results tend to disfavor start-up financing as financing is limited to pilot or small-scale projects, where the underlying entrepreneur can maintain a job. Therefore, alternative funding mainly acts as an alternative to financing for riskier ventures which are ready to pay higher interest spreads.

It is important to note the dissociation between the decisions being made and the actual credit risk when comparing the traditional and alternative lenders. The traditional ones, depending on the lending technology, tend to rely on the market logic, borrower creditworthiness and past relationships, while alternative investors are motivated to invest based on their aesthetic perception, emotional value, and the novelty of the project. When deciding to invest, a higher weight is put on non-financial factors like clean tech, impact on the society or environment over the traditionally accepted risk-return metrics. Yet, the non-financial factor importance dissipates the larger the investment is (Bento et al., 2019). Wasiuzzaman et al. (2021) further investigate alternative lender profiles and show that investors participating in alternative lending are less experienced, younger individuals who tend to overlook possible returns in exchange for a good cause or presentation. As investments happen on the micro-level, a higher risk and lower returns are acceptable if the projects that are backed are in line with environmental and societal goals. The successfully accessed alternative lending credit depends not only on the underlying strength of a business or a project but also on the founder's social influence. As demonstrated by Liu et al. (2021), the founder's digital reputation and post-sharing cascades are positively associated with successfully accessing a credit. Results reveal that, through founder-funder relationship, information influence could work as a certification effect on the potential project leading to a decision to fund a project. The influence of project-sharing cascades suggests a normative influence effect – the founders lead one to conform with the others and invest. It is worth noting that the social status should not be mixed with the personal traits. Butticè and Rovelli (2020) provide evidence that owners with narcissistic characteristics are less likely to successfully access credit from alternative lending sources. Nonetheless, the entrepreneurial context matters as, for certain project funding (particularly in the creative industry), the personal traits could significantly improve the access to credit. Some studies argue that, under certain economic conditions, credit rationing is not observed, on the contrary, excess of credit is prominent. Bonnet et al. (2016) discuss that, for banks, over-lending does not necessarily mean that, in the case of default, the credit will be lost, as banks would be amongst the first ones to be paid back through collateral and secured creditors. Factors like the non-resident status, being jobless, which are related to the underlying business, appear to be detrimental for higher credit rationing. Meanwhile, having public financial aid, or a higher invested capital corresponds to the factors which are related to the project, thus influencing over-lending.

The definition of the access to credit differs between literature. As understood by Claessens and Tzioumis (2006) from the World Bank: ‘‘access to credit refers to the availability of supply of quality financial services at a reasonable cost”. This definition allows for a wide range of interpretations, particularly when it comes to defining ‘quality financial services’ and establishing what can be deemed ‘reasonable’ in terms of costs. Beck et al. (2009) use the term ‘access to finance’ together with the term ‘access to financial services’, which refers to the general access to all financial services. The term is also common when referring to the general company ability to access capital, both internally and externally, and both in terms of equity and debt. Frank et al. (2020) refer to financing as a general term to study the Pecking Order Theory which explains the company selection of different sources of funding. In this form, the term includes three possible sources of financing - internal funds, external debt, and external equity. As the theory suggests, the selection of the financing source depends on the availability and cost of funds (whether these are direct through interest or indirect through the transfer of equity).

The most commonly used term ‘access to credit’ across different studies such as Claessens and Tzioumis (2006); Ayalew and Xianzhi (2019); Ademosu and Morakinyo (2021); Akande et al. (2021); Amadasun and Mutezo (2022) refer to only one part of the capital structure - the external debt. In other studies, such as (Gatti and Love, 2008; Angori et al., 2019, 2020; de Andrés et al., 2021; Basiglio et al., 2022), it is referred specifically to the access to bank credit. In line with the aforementioned studies, this dissertation uses the terms ‘access to finance’, ‘access to credit’ and ‘credit access’ interchangeably to refer to the availability of external debt in a form of bank credit. This dissertation focuses specifically on the SME access to external credit from the traditional commercial banks.

1.3. Estimating Access to Credit

SMEs’ access to credit is a critical area of interest for both scholars and global policymakers. The assessment of this access varies across studies, with proxies selected depending on data availability and country-specific contexts, making it challenging to identify a single ’best’ measurement method and thereby limiting comparability and research replicability. Access to credit is commonly approached through the supply side, bifurcated into two primary streams: modeling the entire bank credit portfolio (Miao and Wang, 2012; Molina and Preve, 2012; Deyoung et al., 2015; Bolton et al., 2016; Peón and Guntín, 2021; Altavilla et al., 2021) and modeling financing application outcomes, often termed credit rationing (Jiménez et al., 2012; Kirschenmann, 2016; Berger et al., 2022; Chodorow-Reich et al., 2021). The former, utilized for macro-level credit supply insights, offers ease of quantification and tracking but may fall short in capturing individual-specific factors crucial for understanding access to credit. In contrast, the latter method estimates access to credit by examining the outcomes of individual financing applications and decisions, potentially providing rich data granularity. However, it is contingent on data source accessibility and dataset robustness (Lee et al., 2015). Analyzing access to credit through application or decision outcomes allows for a nuanced understanding of the specific factors contributing to variations in access levels and the cases where conditional approvals are granted. Negative financing application outcomes are typically categorized into two types: 1st degree rationing, where a financing application is wholly rejected (Jiménez et al., 2012), and 2nd degree rationing (Berger et al., 2022), wherein a financing application is approved but subject to adjustments in product conditions, including the amount of financing, borrowing costs, maturity term, and requested collateral. In certain instances, these adjusted product conditions, often less favorable, can prevent a company from accessing credit (Kirschenmann, 2016; Durguner, 2017). Key studies evaluating access to credit by employing credit supply as a proxy are summarized in Table 3. As noted by Molina and Preve (2012), modeling access to credit is akin to estimating a company's financial distress, as banks' credit decision-making processes revolve around evaluating a borrower's ability to meet future loan obligations.

The second approach to studying access to credit focuses on the demand side, which involves estimating companies' communicated need for financing and understanding whether these companies that require financing actually submit loan applications (Brown et al., 2011; Berger et al., 2017; Nguyen et al., 2021; Altavilla et al., 2021). Unlike modeling credit supply, assessing access to credit through the demand side helps uncover entity-specific barriers based on self-perception, often gathered through external questionnaires (Angori et al., 2019). Evaluating access to credit should not be limited to cases where potential borrowers attempted to secure financing; it should also encompass situations where companies were discouraged from applying (Mac An Bhaird et al., 2016; Nguyen et al., 2021). The demand-driven approach allows for the consideration of not only companies with financing needs but also those deterred from applying. When assessing overall credit availability, a common mistake is to focus solely on approvals or rejections while ignoring discouraged borrowers who do not even approach potential lenders. Discouragement can be beneficial in a well-functioning financing market when it deters non-creditworthy firms from applying. However, if creditworthy firms are reluctant to apply, it can lead to suboptimal levels of investment and, consequently, suboptimal market growth. For example, as

Table 3. Credit supply indicators used as proxies for access to credit evaluation. Created by the author.

Dependent variable and description	Study
Loan portfolio	
Studied how macroeconomic effects, specifically stock-price bubbles, have a positive effect on credit supply.	Miao and Wang (2012)
Studied a theoretical loan supply function from the loan portfolio optimization model with market imperfections.	Deyoung et al. (2015)
Studied a credit supply model for differentiating the effect of relationship- and transaction-lending bank responses to the crisis.	Bolton et al. (2016)
Studied the impact of the relationship lending on the credit access by estimating the total credit issued and shifts in conditional approvals.	Banerjee et al. (2021)
Studied the effects of the credit scoring risk evaluation technique on a small loan portfolio size.	Berger et al. (2022)
Studied the impact of bank credit restrictions over trade credit.	Peón and Guntín (2021)
Loan portfolio growth rate	
Studied the importance of the credit supply and the credit demand on loan growth.	Altavilla et al. (2021)
Studied the impact of raising bank capital requirements on the lending portfolio growth.	Fang et al. (2022)
Application outcome	
Studied the effect of the changes in bank credit supply on loan applications and their subsequent approval rates.	Jiménez et al. (2012)
Studied the impact of the accounting quality on decision outcomes.	Cassar et al. (2015)
Studied the effect of innovation on SME access to finance.	Lee et al. (2015)
Studied the impact of macro-prudential policies on SME access to bank credit.	Ćehajić and Košak (2022)
Conditional approval	
Studied the effect of the lending relationship and demographics on contract conditions.	Neuberger and Räthke-Döppner (2015)
Studied the effect of informational opaqueness on credit rationing in terms of conditional approvals.	Kirschenmann (2016)
Studied whether financial lending technologies have any impact on the size of approved funding.	Motta and Sharma (2020)
Studied loan contract terms by investigating whether relationship-lending borrowers fare better or worse in comparison to others.	Berger et al. (2022)
Studied the conditional approval model for SMEs to evaluate the impact of size on the obtained credit conditions.	Chodorow-Reich et al. (2021, 2022)

demonstrated by Brown et al. (2011), the presence of foreign banks in Eastern Europe compared to Western Europe has led creditworthy firms to be discouraged from applying for financing. Smaller, informationally opaque companies, in particular, exhibit high levels of discouragement, indicating that perceived creditworthiness relies more on 'hard information.' A study by Mac An Bhaird et al. (2016) examines SMEs that choose not to apply for financing due to their belief that their applications would likely be rejected. Significant country variations in non-application rates were evident, with discouragement rates reaching 44% in Ireland and a relatively modest 5% in Finland. These findings shed further light on the nature and impact of information asymmetry, a common issue in SME lending. Discouraged borrowers are often young, small companies with recent turnover declines and higher debt-to-assets ratios. The financial system and the economic environment, acting as macro-level transition mechanisms, significantly influence borrower discouragement. Building on Mac An Bhaird et al. (2016), Nguyen et al. (2021) argue that the fear of rejection is not the sole reason for borrower discouragement and analyze entrepreneur education as a factor in debt aversion. In SMEs, credit decisions are often based on the personal characteristics of owners or managers, unlike large corporations where credit decisions are thoroughly considered. Entrepreneur education has substantial implications not only for borrower discouragement due to debt aversion or a cumbersome application process but also for how SMEs secure financing. Entrepreneurs with higher education tend to be less discouraged from applying for financing. Policies aimed at alleviating financial constraints may not yield targeted results if SMEs managed by entrepreneurs with lower education levels are unwilling to apply for financing. Altavilla et al. (2021) demonstrated that following a monetary policy shock, credit demand is influenced by bank strength, suggesting that discouragement diminishes during positive economic times. Table 4 provides an overview of key studies evaluating access to credit using credit demand proxies.

Table 4. Credit demand indicators used as proxies for access to credit evaluation. Created by the author.

Dependent variable and description	Study
Credit request	
Studied how relaxed credit conditions increased demand for loans, which contributed to credit boom and crisis.	Dell'Ariccia et al. (2012)
Studied how concentration in local banking market affects credit demand.	Chong et al. (2013)
Studied the relationship between the requested loan amount and the underlying small bank-firm relationship lending factors.	Kirschenmann (2016)
Studied whether small banks have a comparative advantage in fulfilling customer financing needs.	Berger et al. (2017)
Studied how banking variables affected the credit demand and the credit request's outcome.	Angori et al. (2019)
Studied credit rationing for perceived financing needs based on transaction lending factors and soft information.	Ferri et al. (2019)
Studied the role of relationship lending and transaction lending in firms' access to credit in a form of credit requests.	Angori et al. (2020)
Discouragement	
Studied the access to bank credit by identifying the determinants which are impacting business propensity to apply.	Brown et al. (2011)
Studied the underlying factors for SME owner's discouragement.	Mac An Bhaird et al. (2016)
Studied the impact of education on borrower discouragement through self-credit-rationing.	Nguyen et al. (2021)
Studied factors impacting demand and rationing and connected them to firm dependence on the state support.	Aristei and Angori (2022)
Studied how the level of company innovativeness impacts their likelihood to apply for credit.	Brown et al. (2022)

As noted by Jiménez et al. (2012); Mina et al. (2013); Maier (2016); Angori et al. (2019); Altavilla et al. (2021), the credit supply and demand are closely related and often influenced by similar forces. Given the adverse economic conditions as the credit supply may contract, due to increasing agency costs for banks, contemporaneously, the credit demand may fall, due to lower growth expectations and a higher cost of borrowing. At the same time, firms negatively affected by economic conditions may borrow more. Such a connection implies that any study based solely on micro-level or macro-level data and ignoring country specifics might suffer from the omitted-variables problem.

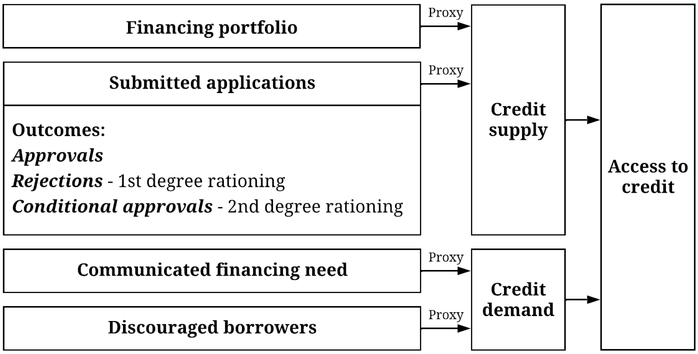


Figure 3. Options for evaluating access to credit. Created by the author.

It has been determined that the access to credit can be evaluated by the use of the credit supply or demand proxies which are estimated with the use of the underlying factors. The selection of supply- or demand- proxies depends on the specific research problem, the availability of data, and the focus on either macro- or micro-effects. As demonstrated in Figure 3, credit supply specific proxies are the loan portfolio numbers and the submitted application outcomes (approvals, rejections, and conditional approvals), while credit demand specific proxies are both the perceived need for financing and the borrower discouragement. It is important to note that distinguishing between credit supply and credit demand proxies can be difficult as the underlying access to credit is when both meet. Finally, regardless whether credit supply or demand proxies are used, it is necessary to account for both macro and micro specific factors, which reduces the potential omitted variable-bias risk. The absolute majority of academic literature is focused not on which proxy is the most appropriate one, but rather what the underlying factors that impact access to credit actually are.

1.4. Factors Influencing SME Ability to Access Credit

1.4.1. Lending technology factors

The ultimate decision whether a financing application will be approved or rejected depends on the lender and the factors related to internal and external policies, the risk appetite and, in some cases, personal preferences. Early work by Petersen and Rajan (1995) set up the framework which defines the key factor groups that are impacting the access to credit:

- The debtor's willingness to make debt repayments based on historic experience (or simply – the entity's credit history)
- The debtor's ability to make debt repayments, which is purely connected to the entity's financial strength
- Macro-environment conditions, which could impact the debtor's ability to repay on national or local levels

- Risk mitigating measures, which underline what collateral the debtor is inclined to lose if the debt is not repaid properly.

While novel at the time, the initial framework did not account for factors such as the underlying market structure and lending technology, both of which have significant implications for credit access. Berger and Udell (2006) enhanced the access to credit framework by explicitly recognizing key factor groups, including lending infrastructure factors, financial institution structure, and lending technology, as essential dimensions for assessing credit access for both individuals and countries. An oversimplification of lending technologies typically categorizes them into two types: transaction lending, which relies on quantitative data, and relationship lending, which revolves around 'soft' or qualitative data. This generalization often aligns with borrower types, as transaction lending is typically geared towards more transparent and often larger companies, while relationship lending targets opaque and often smaller businesses (Angori et al., 2019). Incorporating specific lending technologies as a factor challenges the implied uniformity of financial technologies and allows for differentiation in credit availability based on the chosen medium (Motta and Sharma, 2020). Transaction lending, which includes financial statement lending, credit scoring, leasing, asset-based lending, factoring, and fixed-asset lending, is the predominant lending technology used by financial institutions. Financial statement lending is primarily employed to assess the most transparent borrowers providing complete and audited financial statements, while other forms of transaction lending cater to more opaque businesses. These technological differences have significant implications for credit availability, as the same borrower may encounter obstacles with one lending technology but face no issues with another. Depending on the chosen lending technology, lenders apply distinct methods related to primary information sources, screening procedures, product terms, and monitoring policies and mechanisms (Motta and Sharma, 2020). Variations between lending technologies can be minor, limited to specific product conditions, or multidimensional, encompassing differences in screening policies, as seen in the case of fixed-asset lending and leasing, where ownership of the underlying asset distinguishes them, or in other instances highlighted by Berger and Udell (2006).

As concluded by Crawford et al. (2018), access to credit is greatly impacted by the availability of standardized and verifiable information collected from the customer, business registries, and credit bureaus. These sources are essential for carrying out credit screening and monitoring processes which enable lenders to assess the creditworthiness of a borrower. The more comprehensive and accurate the information is, the higher is the likelihood of obtaining credit. Therefore, it is essential for individuals and businesses to maintain a good credit history and keep their information up to date so that to increase their chances of accessing credit. Without access to reliable and standardized information, lenders may be hesitant to provide credit, thereby potentially limiting economic opportunities for individuals and businesses. Based on the available quantitative data, lenders are able to quantify the credit risk and make credit decisions. Intuitively, transaction lending, as long as it is based on standardized and verifiable information, might seem homogeneous, but, as demonstrated by Berger and Udell (2006), there exist significant differences between separate transaction lending sub-groups. A number of different transaction lending sub-groups can be identified based on the information source (Eca et al., 2021).

The use of accounting data, to base financing decisions, is one of the most prominent alternatives to relationship lending. Borrowers make lending decisions by assuming that a company would be able to make future loan repayments based on the strength of the borrower's financial statements (Palazuelos et al., 2018). Companies with financial statements that are based on international accounting standards or are reviewed by external auditors are prone to fewer problems arising from information asymmetry. As studies suggest, the majority of businesses do not have their financial statements audited (Lisowsky and Minnis, 2020). To account for intentional or unintentional errors in the retrieved financial statements, banks use alternative sources, such as tax returns. Depending on the country, companies are obliged to provide tax returns annually by providing their sales, expenses, assets and liabilities, which is the data which could be a relatively cheap substitute for financial statements. Although such data is less preferred than financial reports as it does not include the cash flow statement or the granularity in the balance sheet or the income statement (Minnis and Sutherland, 2017). Accounting report data is not the only alternative that can be used to evaluate the financial strength and the credit risk of potential borrowers. By incorporating a set of information on the credit history and financial statements, credit scoring is a widely accepted transaction lending technology. Lenders must decide what credit score source they trust and are willing to use – whether it is calculated internally or bought from an external supplier. Cassar et al. (2015) demonstrated how the use of credit scores can diminish the usefulness of financial statements, though companies using accrual accounting still benefit from a lower interest rate. The popularity of credit scores for SME financing is not necessarily related to their superiority in accuracy over the other lending technologies, but rather due to their wide third-party availability and a significantly cheaper cost of information transfer. In principle, they are a fusion of multiple lending technology factors presented as a single indicator, which is easy to comprehend and transfer (Ciampi et al., 2021). Though transaction lending factors are more commonly employed than the soft information amongst the main SME financing providers, Ferri et al. (2019) demonstrated that the presence of the soft information helps SMEs to access credit, specifically at a time of crisis. A study by Cucculelli and Peruzzi (2017) provides evidence that the amount of the hard data that is required to be provided when trying to access credit can differ depending on the company's ownership structure, indicating that variables are inter-connected. Table 5 summarizes transaction lending factors based on the financial statement data used across different SME access to credit studies.

Table 5. Transaction Lending factors based on financial statement data used for evaluating access to credit. Created by the author.

Factor	Definition	Studies
Cash ratio	Cash to current liabilities ratio.	Degryse et al. (2018); Martí and Quas (2018); Grzelak (2019); Malakauskas and Lakštutienė (2021); Medianovskyi et al. (2023)
Quick ratio	Quick assets to current liabilities ratio.	Xu et al. (2020); Malakauskas and Lakštutienė (2021); Medianovskyi et al. (2023)
Current ratio	Current assets to current liabilities ratio.	Jiménez et al. (2012); Meuleman and De Maeseneire (2012); Florou and Kosi (2015); Adam and Streitz (2016); Cucculelli and Peruzzi (2017); Durguner (2017); Angori et al. (2019); Ferri et al. (2019); Angori et al. (2020); Xu et al. (2020); Malakauskas and Lakštutienė (2021); Zainol Abidin et al. (2021); Medianovskyi et al. (2023)
Debt-to-Equity	Total liabilities to owner's equity ratio.	Elsas (2005); Jiménez et al. (2012); Angori et al. (2019); Malakauskas and Lakštutienė (2021); Zainol Abidin et al. (2021); Banerjee et al. (2021); Medianovskyi et al. (2023)
Tangible assets	Total tangible assets to owner's equity ratio.	Meuleman and De Maeseneire (2012); Florou and Kosi (2015); Adam and Streitz (2016); Degryse et al. (2018); Martí and Quas (2018); Ogura (2018); Angori et al. (2019); Grzelak (2019); Angori et al. (2020); Aoki (2021)
Leverage	Total liabilities to total assets ratio.	Peltoniemi (2007); Bosch and Steffen (2011); Meuleman and De Maeseneire (2012); Florou and Kosi (2015); Adam and Streitz (2016); Kirschenmann (2016); Cucculelli and Peruzzi (2017); Durguner (2017); Martí and Quas (2018); Angori et al. (2020); Xu et al. (2020); Zainol Abidin et al. (2021); Aoki (2021); Banerjee et al. (2021); Berger et al. (2022); Chodorow-Reich et al. (2022)
ROA	Profitability metric measuring net income to total assets.	Cole (1998); Jiménez et al. (2012); Cassar et al. (2015); Florou and Kosi (2015); Cucculelli and Peruzzi (2017); Angori et al. (2019); Ferri et al. (2019); Zainol Abidin et al. (2021); Aoki (2021); Banerjee et al. (2021); Berger et al. (2022)
ROE	Net income to owner's equity ratio.	Chong et al. (2013); Zainol Abidin et al. (2021)
Profitability	Net income to net sales ratio.	Carbo-Valverde et al. (2009); Bosch and Steffen (2011); Florou and Kosi (2015); Adam and Streitz (2016); Cucculelli and Peruzzi (2017); Durguner (2017); Martí and Quas (2018); Ogura (2018); Grzelak (2019); Ferri et al. (2019); Xu et al. (2020); Malakauskas and Lakštutienė (2021); Zainol Abidin et al. (2021); Aoki (2021); Chodorow-Reich et al. (2022); Medianovskyi et al. (2023)
Asset turnover	Net sales to total assets ratio.	Molina and Preve (2012); Meuleman and De Maeseneire (2012); Cassar et al. (2015); Grzelak (2019); Zainol Abidin et al. (2021)
Receivables	Net sales to accounts receivable ratio.	Molina and Preve (2012); Cassar et al. (2015); Durguner (2017); Degryse et al. (2018); Malakauskas and Lakštutienė (2021)
DSCR	Debt-service-coverage-ratio.	Elsas (2005); Carbo-Valverde et al. (2009); Molina and Preve (2012); Adam and Streitz (2016); Degryse et al. (2018); Ogura (2018); Angori et al. (2019); Malakauskas and Lakštutienė (2021); Banerjee et al. (2021); Chodorow-Reich et al. (2022); Medianovskyi et al. (2023)
Coverage	Tangible assets minus current liabilities to total liabilities ratio.	Malakauskas and Lakštutienė (2021); Medianovskyi et al. (2023)
Sales growth	The change in net sales.	Carbo-Valverde et al. (2009); Molina and Preve (2012); Berger et al. (2017); Ogura (2018); Aoki (2021); Malakauskas and Lakštutienė (2021); Aristei and Angori (2022); Chodorow-Reich et al. (2022); Medianovskyi et al. (2023)
Asset growth	The change in current assets.	Martí and Quas (2018); Malakauskas and Lakštutienė (2021); Medianovskyi et al. (2023)
Assets	Total assets amount.	Cole (1998); Peltoniemi (2007); Cassar et al. (2015); Adam and Streitz (2016); Kirschenmann (2016); Berger et al. (2017); Martí and Quas (2018); Angori et al. (2019, 2020); Banerjee et al. (2021)

A wide array of literature has employed transaction lending factors to estimate the access to credit for SME entities. The use of the financial statement data to estimate the business financial health is one of the most common ways of studying factors which impact the company's ability to access credit. Motta and Sharma (2020) evaluated company access to credit and compared transaction lending factors based on the financial statement data with product conditions by considering the information on collateral and the fact of the existence of an audited financial statement. Meanwhile, Berger et al. (2011) analyzed the use of the small business credit scoring (SBCS) transaction lending technology with a focus on smaller community banks. In contrast to the common paradigm that smaller banks use soft information technologies, almost a half of them used some sort of scoring to base their small-bill financing decisions. The majority of the surveyed community banks, which use any sort of scoring, use, exclusively, consumer credit scores (CCS) to evaluate the owner of the company. Only a small part used a combination of SBCS and CCS, while only SBCS was used only 2% of cases. The usage of scoring has a positive effect on credit availability primarily for credits under \$100K, with a diminishing effect over time. Furthermore, the quality of the credit portfolio for banks which started using CCS scoring to issue loans did not deteriorate even if more marginal loans were issued. Jiménez et al. (2012) used a combination of different types of factor groups such as Firm Characteristics and Transaction Lending factors to evaluate access to credit. Transaction lending factors can be grouped based on the information source into two major sub-groups - financial reporting and the credit history. Table 6 summarizes transaction lending factors based on the credit history data that were used across different SME access to credit studies.

Table 6. Transaction Lending factors based on credit history data used for evaluating access to credit. Created by the author.

Variable	Definition	Study
Company delinquencies	Indication whether the company was delinquent on any of its obligations (both internally and externally).	Cole (1998); Cassar et al. (2015); Neuberger and Räthke-Döppner (2015); Kirschenmann (2016); Malakauskas and Lakštutienė (2021); Berger et al. (2022); Medianovskyi et al. (2023)
Owner delinquencies	Indication whether the owner of the company was delinquent on personal obligations.	Cole (1998); Cassar et al. (2015); Malakauskas and Lakštutienė (2021); Medianovskyi et al. (2023)
Defaults	Indication whether the company had any defaults.	Jiménez et al. (2012); Neuberger and Räthke-Döppner (2015); Kirschenmann (2016)
Owner defaults	Indication whether the majority owner had any defaults.	Jiménez et al. (2012)
Credit score	Credit score assigned by the lender.	Sapienza (2004); Elsas (2005); Peltoniemi (2007); Agarwal and Hauswald (2010); Presbitero and Zazzaro (2011); Berger et al. (2011); Florou and Kosi (2015); Adam and Streitz (2016); Ogura (2018); Xu et al. (2020); Aoki (2021); Banerjee et al. (2021); Chodorow-Reich et al. (2022)

As a contrast to transaction lending, relationship lending is the preferred means of lending when the information about the company is limited and the best way to assess creditworthiness is through past relationships or borrower characteristics (Rabetti, 2022). Bank-firm relationships like the intensity of borrower-creditor relationship or the outstanding current account balance with the lender, and thus consistently improve credit terms as well as the availability of credit (Durguner, 2017). The impact of bank-firm relationships is important not only when determining whether a company will ultimately receive the loan but also whether the full requested amount will be received. By focusing on opaque borrowers and small-bill financing products, Kirschenmann (2016) shows that opaque borrowers are indeed more pronounced towards being credit rationed. The resolution of information asymmetry over the course of bank-firm relationships can lead to lowering the rationing effect as the lender is more willing to increase its stakes. Xu et al. (2020) demonstrated that a loss of relationship with a long-term lending provider can have a negative impact on the cost, maturity and the overall availability of loans. Furthermore, firms that are able to hold long-term bank-firm relationships are able to maintain stronger investment and employment growth during times of crisis (Banerjee et al., 2021). The moderating effect of information opacity on the relationship was proxied through the reporting frequency which is one of the key aspects in managing portfolio risks through monitoring. Findings by Durguner (2017) show that the cost of such a relationship loss is considerable for the borrower, mainly through higher loan spreads and the costs increasing together with the borrower's information opacity (see Uzzi (1999)). As banks provide a wide variety of services to a customer, throughout a considerable amount of time, they collect a substantial amount of information not only on the customer's financial needs but also on their behavior. Such knowledge can create benefits for, both, the customer, and the bank (Ongena and Smith, 1998; Belaid et al., 2017; French et al., 2019). On the other hand, more information can act as an increase in the bargaining power for the bank, which in turn makes for the customer the substitution to another bank more costly (Elsas, 2005). Relationship information can be collected through different product provisions, interviews, enquiries to suppliers or clients, but, due to its high variability, most commonly, it is generalized as bank-firm relationship (Ferri et al., 2019). As defined by Ongena and Smith (1998), bank-firm relationship is understood as "the connection between a bank and customer that goes beyond the execution of simple, anonymous, financial transactions." Still, other researches (such as Ma et al. (2019); Aoki (2021)) limit the definition of the bank-firm relationship to be specifically defined by the use of financing products.

Relationship can be quantified through several metrics, the first of which is time. The extent of a relationship depends on the duration of meaningful interactions between the customer and the bank. As the duration of the bank-firm relationship increases, the bank has more opportunities to observe and learn about the customer, has more trust in the customer's business sustainability and continuity, and thus experiences a lower informational asymmetry (Belaid et al., 2017). When evaluating the duration of the bank-firm relationship, it is important to consider possible structural differences which could rise from the historical developments. The average bank-firm relationship differs from one system to another, and its dimensions, from country to country, vary. In Italy, the average bank-firm relationship is 14 years (Gambini and Zazzaro, 2013). Meanwhile, in Germany and Sweden, the average is more than 20 years, and in Japan it is estimated to be over 30 years (Uchida et al., 2008). On the opposite side, there is the United States and Finland where the average bank-firm relationship does not exceed 10 years (Cole, 1998; Peltoniemi, 2007). Some studies argue that long bank-firm relationships could create hold-up problems, which in turn lock-in companies with one or two banks and allows them to hold the monopoly over the customer (Hakimi and Hamdi, 2013; Adam and Streitz, 2016). Though others add that the benefits which are created through bank-firm relationships still outweigh the potential drawbacks (Grzelak, 2019).

A bank-firm relationship could be long but without any meaningful interactions, therefore, it is important to quantify the strength of the existing bank-firm relationship. The intensity of the relationship is defined in terms of breadth of the products and services offered by the bank and used by the company. Such services can span from regular banking products like financing and cash management to other supplementary products like investment banking and insurance (Ongena and Smith, 1998). Berger et al. (2022) further differentiated the scope of the bank-firm relationship into existence and intensity. The intensity can be approximated by calculating the number of institutions providing financing or other services by other institutions (Belaid et al., 2017; Angori et al., 2020). In an attempt to quantify bank-firm relationships, Elsas (2005) determined the factors which indicate that a bank can be considered as the main bank (a.k.a. primary institution or 'Hausbank'): high share of financing; high share of payment transactions; has special, exclusive, or intensive relationship; has long relationship; bank has influence on the company's management. Hausbank customers on average maintain 4-5 banking relationships and have approximately 44% of total debt coming from the Hausbank (Elsas, 2005). Companies that have a limited number of partners indicate the exclusiveness of the bank-firm relationship, which helps to create trust. As underlined by Petersen and Rajan (1995), a financially strong company that has a strong bank-firm relationship has the ability to secure financing from a single lender. Yet, more recent studies show that for cases when a company needs repeated financing, a few banking relationships might be more beneficial than one (Ongena and Smith, 2000; Elyasiani and Goldberg, 2004; Grzelak, 2019). Interestingly, no explicit connection between higher information acquisition and short and long-term financing has been identified. Similar findings hold for underlying collateral, loan syndication or geographical proximity (Elsas, 2005).

Table 7. Relationship lending factors used for evaluating access to credit. Created by the author.

Variable	Definition	Study
Duration	The length of a bank-company relationship in days at the time of application.	Cole (1998); Petersen and Rajan (2002); Elsas (2005); Peltoniemi (2007); Agarwal and Hauswald (2010); Jiménez et al. (2012); Cassar et al. (2015); Neuberger and Räthke-Döppner (2015); Belaid et al. (2017); Cucculelli and Peruzzi (2017); Durguner (2017); Minnis and Sutherland (2017); Angori et al. (2019); Grzelak (2019); Ferri et al. (2019); Angori et al. (2020); Xu et al. (2020); Banerjee et al. (2021); Berger et al. (2022)
Primary institution	Indication whether the bank is the firm's primary financial institution.	Elsas (2005); Cassar et al. (2015); Angori et al. (2019, 2020); Aoki (2021)
Payments	The proportion of incoming and outgoing payments in bank's accounts.	Elsas (2005); Durguner (2017); Angori et al. (2019, 2020)
Rejections	Indication whether the applying company has received any negative decisions during the past 2 years before the application.	Cassar et al. (2015)
Debt share	The proportion of debt held in the bank to total liabilities reported in latest annual financial statements.	Elsas (2005); Peltoniemi (2007); Angori et al. (2019); Ferri et al. (2019); Angori et al. (2020); Berger et al. (2022); Kärnä and Stephan (2022)
Financing contracts	The number of past financing contracts the company had within the bank at the point of application.	Cole (1998); Peltoniemi (2007); Neuberger and Räthke-Döppner (2015); Kirschenmann (2016); Durguner (2017); Minnis and Sutherland (2017)
Other products	Indicates whether the company has any other products in the bank.	Cole (1998); Petersen and Rajan (2002); Peltoniemi (2007); Cassar et al. (2015); Neuberger and Räthke-Döppner (2015); Durguner (2017)
Other relationships	The number of other banks the company has relationship with.	Cole (1998); Peltoniemi (2007); Elsas (2005); Agarwal and Hauswald (2010); Jiménez et al. (2012); Kirschenmann (2016); Cucculelli and Peruzzi (2017); Durguner (2017); Angori et al. (2019); Grzelak (2019); Ferri et al. (2019); Angori et al. (2020); Aoki (2021); Kärnä and Stephan (2022)
Distance	Distance between the bank and the company.	Neuberger and Räthke-Döppner (2015); Durguner (2017); Xu et al. (2020)
Bank ownership	Bank owns or impacts management's decisions in the company.	Neuberger and Räthke-Döppner (2015); Durguner (2017); Xu et al. (2020)

By studying the combined effects of the Transaction Lending and Relationship Lending factor groups on the SME access to credit, it was determined that a complementary effect between the two factor groups exists. Furthermore, it is evident that relative fuzziness between the two factors groups is evident, which suggests that variables belonging to multiple underlying factor groups should be employed when assessing the access to credit.

1.4.2. Lending infrastructure factors

The lending infrastructure encompasses several key components, including the information environment, legal and judicial environment, bankruptcy environment, social environment, tax policies, and regulatory environment. These components collectively shape SME access to credit by influencing the conditions under which lending technologies can operate and function effectively. Additionally, the regulatory environment can exert significant influence on the structures of financial institutions due to various policy constraints (Berger and Udell, 2006).

The information environment is one of the key elements which enables the deployment of quantitative-data-based lending technologies thus directly affecting the credit availability (Brown et al., 2009). Information asymmetry acts as a two-fold problem. On the one hand, due to limited information, lenders have difficulty distinguishing between good and bad creditors, thus leading to credit rationing and constraining company investments. On the other hand, asymmetric information can lead to inverse results by over-lending to non-creditworthy companies. Armstrong et al. (2010) demonstrated the effect of information availability, through financial reporting and accounting standards, on the company governance and, most notably, debt contracting. Limiting the relevant information from credit issuers causes a spurious growth in the reported creditworthiness of an entity, which is then followed by substantial increases in the issued financing. The long-term effect is the opposite as the scores tend to deteriorate further, and the number of delinquencies increases more than the full-information models would have predicted (Musto, 2004). A further study by Dobbie et al. (2020) confirms that the immediate effect of negative information removal has a positive impact on credit scores (PD fell by 3 p.p.), but, in contrast to Musto (2004), demonstrates that, in the long-term, the credit scores do not deteriorate or delinquencies increase. Therefore, adjusting the credit assessment time horizons from the policy perspective can act as a welfare improving factor as the credit accessibility should be improved. Improvement of the accounting standards through standardizing reporting forms and enabling comparability has a positive impact on the company information opaqueness. The implementation of the new International Financial Reporting Standards (IFRS) has a positive impact on the company propensity to raise capital publicly and reduce the cost of debt. Mandatory IFRS reporting has a positive effect on the access to credit, mainly through the bond market due to lower yield spreads. The strength of this effect depends on the underlying country's accounting standards and their similarity to IFRS (Florou and Kosi, 2015).

In addition to the primary information sources such as the customer, credit bureaus, local registries, lenders also use other proxies such as past subsidies or valid rating scores as means of reducing information asymmetry and 'certifying' that a potential borrower is creditworthy (Presbitero and Zazzaro, 2011; Angori et al., 2019, 2020). Factors, like non-resident status, being jobless, which are related to the underlying business, appear to be detrimental for higher credit rationing. Meanwhile, having public financial aid, or a higher invested capital corresponds to factors, which are related to the project, influencing over-lending (Bonnet et al., 2016). Information produced by the certification effect can help improve access to the capital through government issued subsidies, or the availability of external scorings or ratings (Bosch and Steffen, 2011). In financial systems for which balance sheets are marked-to-market, any changes to constituent asset values will be reflected in the balance sheet values, and thus will adjust their net worth. As the net worth changes, financial intermediaries adjust the size of their balance sheets by lowering or increasing the issued capital, thus changing the supply of credit. Adrian and Shin (2010) showed how financial intermediary responses to fluctuations in their net worth affect the market. As the asset price increases, generally, the financial intermediary's balance sheet will grow stronger, and, without adjustments to the held assets, the leverage level will fall. Given the lower level of leverage, capital surplus is created which then can be compensated by employing short-term borrowing and lending it out in order to correct for the initial level of leverage. Evidence from the study indicates that such growth in the credit supply is strongly pro-cyclical. The most notable proof is the financial crisis of 2007 when financial intermediaries created a surplus of credit and lent it out to non-creditworthy borrowers, which later ended-up in borrower inability to repay. It is determined that the financial market liquidity is closely related to the extent of the financial intermediaries' search for borrowers.

The effect of the increased information transparency on the credit accessibility is not only positive. It is necessary to understand not only how precise the information is but also what it means. Johnstone (2016) challenges the wide belief that 'good' financial reporting will lead to a lower uncertainty thus resulting in a lower cost of capital. First, better data can lead decision makers to be more uncertain. Second, when additional data does reduce uncertainty, it could carry unfavorable information, which would have a negative effect on the access to credit through a higher rationing or cost of capital. From the investment-under-uncertainty perspective, financial reports are only a medium for facilitating the accurate future cash-flow estimation, therefore, on average, more information should enable the creation of more accurate future cash flow forecasts, but those forecasts might as well lead to a higher cost of capital. By building on the study by Florou and Kosi (2015), Kalogirou et al. (2021) analyzed the impact of IFRS introduction and the companies which were suffering from high pension deficits. As the lending market prior to changes in the accounting standards mitigated for the non-disclosure of pension fund deficits, the effect of surprise factor was investigated. Analysis models the effect on the company level of leverage, the cost of debt and the debt maturity. The study suggested that financially riskier companies suffered negative consequences with regards to credit availability. By using early adopters, credit issuers adapted to the market opacity, but with the newly available information they fixed the expectation errors and corrected for a higher efficiency in debt contracting.

Understanding the effect of demographics in small business accessibility to credit is important. The majority of developed economies are experiencing aging population, shifts in family structures and population migration from poor regions to richer ones. These changes have a significant impact on the economic age structure and the entrepreneurial activity in the market (Berger and Udell, 2006). Sweeping technological changes lower transportation and information costs, thus minimizing the effect of the geographic distance on the credit supply to a point where geography should become inconsequential. Brei and Von Peter (2018) demonstrated that the effect of globalization exists and is mainly pronounced in cross-border banking. In the domestic setting, the physical distance effect slightly diminished during the past 50 years, but it is still considerable. Banking as a good, inherently, has close to zero transportation costs, and the distance effect can be mainly attributed to information costs. Even if hard data will be costless, information costs will not be eliminated due to the soft information asymmetry and cultural differences. Contrary to Berger and Udell (2006), Neuberger and Räthke-Döppner (2015) found that demographics play a minor role in the cost of borrowing, especially if comparing to relationship lending. The price of financing decreases with the gains of soft information through longer processing times and increases with the hard information about delinquencies or reminders. The entrepreneur's age or marital status have no effect on the cost of borrowing. Low density regions or companies that are far away from the bank are not disadvantaged by higher interest margins.

The regulatory environment can have an adverse effect not only on the lending technology deployment, but also on the existing market set up (Berger and Udell, 2006). Capital requirements work as a central tool for the macro-prudential policy enforcement, by accounting for cyclical variation, it 'cools-off' credit booms, boosts capital and provisions. Aiyar et al. (2014) showed the effect that the macro-prudential regulation has a regulated and non-regulated bank credit supply. The response of the regulated bank loan supply on higher capital requirement adjustments is strongly negative, while for banks which are not subject to local regulation the effect is positive. As the overall effect of the capital requirements on the credit supply depends greatly on the existing financial structures, the overall effect on the overall credit availability can differ by country. Fang et al. (2022) built upon Aiyar et al. (2014) study and showed that increased capital requirements have a significant but short-term negative effect on lending. Six and more months after the increase of the required capital, the lending growth returns to the original levels. The limited effect is attributed to several factors: early announcement of upcoming changes, strong economic growth, and individual bank characteristics. Chen et al. (2017) studied the effect of a monetary policy on the bank perception and tolerance of risk in emerging economies. Evidence suggests that commercial bank riskiness is closely associated with the monetary policy imposed by the central bank. Banks tend to take more risk amid the expansionary monetary policy, during the period of a lower interest. Meanwhile, the extent of the monetary policy on risk taking is less pronounced in concentrated markets and in cases when the policy is clearly communicated.

As evident in the previous parts, the access to credit depends greatly not only on the lending technology but also on the underlying macro conditions which determine the environment in which 'how' and 'by whom' the financing can be provided. The ability to receive credit can deteriorate rapidly if the market is exposed to high levels of uncertainty – an example of which being financial crises (Deyoung et al., 2015). As crises are often followed by tighter monetary conditions and a lower economic growth, they have adverse effects on both credit supply and demand (Gertler and Gilchrist, 1994). Yet, the extent of such effects can vary depending on the type of the crisis and individual borrower characteristics (Berger et al., 2022). One of the most researched cases of economic uncertainty is the credit crunch of the global financial crisis that mainly hit the supply side of credit. Degryse et al. (2018) examined the impact of GFC on the characteristics of the local credit market on the variance of the credit supply and found that during the crisis companies which were closer to banks faced greater credit availability than the ones which were further away. Similar findings were documented earlier by Berger et al. (2022), who also found that small banks seemed to have comparative advantage in providing liquidity insurance for small businesses which were banking with large banks. Peón and Guntín (2021) found evidence that small firms in rural areas suffered a differential negative flow of the bank credit during GFC, especially in the manufacturing and construction sectors. By building on the findings by Deyoung et al. (2015), Beck et al. (2018) suggested that the effect of GFC on credit availability differs depending on the type of lending techniques applied by a bank. Relationship lending has a positive effect on SME access to credit for companies that are smaller and more informationally opaque as well as in the regions which suffered the most. Informational advantage of relationship banking during a time of crisis was also confirmed by Bolton et al. (2016), as such banks were able to offer better continuation-lending terms and suffered lower rates of default. By looking at individual loan applications and connecting them with the strength of the underlying bank balance sheets, Jiménez et al. (2012) showed that, at a time of crisis, banks with weaker balance sheets and a lower capital are more prone to do credit rationing. Furthermore, the initial refusal to receive credit results in companies being unable to close the financing gap by applying to other banks. The effect of external shocks on the credit supply may vary depending on individual loan types. Asset-based lending is quite insensitive to shifts in the monetary policy or financial crises, while most shifts in the aggregate credit supply propagate from cash-flow loans (Ivashina et al., 2022).

Table 8. Lending infrastructure factors used for evaluating access to credit. Created by the author.

Variable	Definition	Study
GDP	GDP per capita.	Brown et al. (2009); Carbo-Valverde et al. (2009); Olivero et al. (2011); Jiménez et al. (2012); Zarutskie (2013); Aiyar et al. (2014); Bertay et al. (2015); Love and Pería (2015); Khan et al. (2016); Chen et al. (2017); Degryse et al. (2018); Fang et al. (2022)
Inflation	The annual change in price index (CPI or HPI).	Brown et al. (2009); Agarwal and Hauswald (2010); Jiménez et al. (2012); Zarutskie (2013); Aiyar et al. (2014); Bertay et al. (2015); Love and Pería (2015); Khan et al. (2016); Chen et al. (2017); Ademosu and Morakinyo (2021); Berger et al. (2022)
Unemployment	Average unemployment rate in markets served by bank.	Berger et al. (2011, 2017); Degryse et al. (2018); Ademosu and Morakinyo (2021); Berger et al. (2022)
Population	Population density.	Carbo-Valverde et al. (2009); Neuberger and Räthke-Döppner (2015)
Regulatory stringency	The extent of regulatory stringency as well as the level of activity.	Khan et al. (2016); Chen et al. (2017)
Rule of law	Measures the extent to which market participants have confidence in and follow the rules of society.	Khan et al. (2016); Chen et al. (2017)

The more recent COVID-19 crisis was unique, in contrast to the Global Financial Crisis (GFC), which was caused by the excessive lending of the banks, it originated from stalled supply lines and seized business operations. To curb the spread of an infectious disease, governments halted major parts of the economy, which acted as an exogenous shock on business liquidity and ultimately on default risk. In the first phase of the crisis, daily credit line draw downs exposed a 'dash for cash', which was induced by a heightened aggregate risk and the extreme precaution of businesses (Acharya and Steffen, 2020) During the GFC, the bank credit line withdrawal rate in the US grew at around 1.2% per week, while during the first three weeks of March, it grew by 6% per week, or about 50 times the average. COVID-19 pandemic was an unprecedented stress test on the bank ability to supply credit, which was controlled through liquidity injections from the Federal Reserve, the strong pre-shock bank capital, and the incidental deposit inflow (Li et al., 2020). Though, 'dash for cash' effect was not uniform across all company sizes, SMEs in contrast to large firms did not 'dash for cash' even if the shock on the demand was similar. Such a phenomenon occurred not because SMEs did not intend to use their issued credit lines but rather due to significant differences in the agreed loan terms. As specified by Chodorow-Reich et al. (2022), SMEs are not able to access liquidity as easily as large firms because: their credit lines were easier to demand by lenders and had shorter maturity terms; they were required to post more collateral; they had higher utilization rates; they paid higher interest rates. As indicated by Greenwald et al. (2021), even though the unused credit line capacity is vast, it is overwhelmingly concentrated amid businesses which are large and not financially constrained. Therefore, any adverse financial conditions tend to crowd-out credit from the most financially constrained businesses, thus further limiting their liquidity and threatening survival. These findings further explain cases when, despite an increase in the aggregate credit supply, the access to credit for SME entities did not improve.

It was determined that the lending infrastructure has a significant impact on SME access to credit, which is influenced by various components including the information environment, legal, judicial and bankruptcy environments, the social environment, tax, and regulatory environments. Information asymmetry may lead to credit rationing and over-lending to non-creditworthy SMEs. However, the improvement of the accounting standards and an increase of the information transparency has a positive impact on the access to credit. The effect of the increased information transparency on SME access to credit is not only positive as it could carry unfavorable information leading to negative effects on the access to credit through higher rationing or the cost of capital.

1.4.3. Financial institution structure factors

A causal chain connects government-enforced policies with the structure of financial institutions, influencing the competition and market presence of these institutions and other lending providers. The structure of financial institutions within a market plays a crucial role in determining the feasibility and profitability of deploying specific lending technologies. However, the existing research literature lacks differentiation regarding the types of lending technologies used, making it challenging to test theories that link financial institution structures to their ability to provide funding to creditworthy transparent and opaque SMEs. For instance, financial institution structures can be assessed by comparing the advantages of large versus small institutions, domestically versus foreign-owned institutions, state-owned versus privately-owned institutions, and the overall level of market competitiveness (Berger and Udell, 2006).

The shrinking number of local bank branch offices is a common trend throughout many countries, led by changes in the customer habits as well as higher efficiency ambitions by the banks. Such a reduction leads to an increase in the physical SME distance to the bank, which can have a negative effect on lending operations, through diminishing soft and unverifiable information flow (Agarwal and Hauswald, 2010). On the other hand, lower soft information usage could be counter-acted by technological advancements and higher reliance on the hard factors (Petersen and Rajan, 2002; Milani, 2014; Kärnä and Stephan, 2022) by investigating the effects of the branch office density in different geographic areas on the SME credit constraints. The study by Kärnä and Stephan (2022) uses company level data from a state-owned Swedish bank and covers the period of 2001-2016. Sweden being one of the leaders in the world in financial technologies and the level of digitalization, poses an interesting context. To assess the credit constraints, the interest rate and the financing amount are regressed on the amount of local bank branches in a geographic region. The findings indicate that regions with a higher branch density and employee numbers have a significant effect on lower interest rates and higher loan amounts for SMEs. The authors concluded that the key driver for such a result is competition, which pushes banks to offer lower interest rates and to increase the available credit supply. Finally, as the credit availability grows, and more SMEs can access credit, the likelihood of default also increases. Domestic and cross-border capital flows play an important role in the credit availability and the overall business cycle. Since the early 1990s, emerging and advanced economies have been experiencing an increase in the magnitude and volatility in capital inflows and outflows. Standard business cycle models suggest that a negative shock in productivity should cause capital inflows to fall and outflows to rise. Davis (2015) analyzed a contradicting phenomenon where capital inflows and outflows are positively correlated and strongly pro-cyclical during a shock on output. The findings suggest that, in order to explain capital flows between two countries, it is necessary to rely on market incompleteness and the lack of diversification. Building on these findings, Wang (2018) studied how bilateral banking flows are impacted by uncertainty, which is measured by stock market volatility. The study uses the *Bank for International Settlements* Locational Banking Statistics data on bilateral banking flow. These findings suggest that, at times of uncertainty, banks tend to reduce their exposures in foreign countries, while an increase of the exposures in local markets leads to a retrenchment. It is evident that such an effect could be attributed to informational asymmetry and lead to the limited company access to credit.

Market Competitiveness is one of the most studied sub-groups of the factors pertaining to the structure of financial institutions. After the recent global financial crisis, the interest shown to this subject has grown as many questioned whether relentless bank competition through the lowering of lending standards was partly to blame (Dell'Ariccia et al., 2012) As governments were dealing with the downfall of financial institutions through bailouts, mergers and guarantee prolongations, it led to concerns related to the future competitiveness of the market and its implications on the SME access to credit (OECD, 2009) The impact of market competition on the access to credit is two-fold: the market power hypothesis states that strong competition in financial markets lowers the cost of financing and improves the availability of credit. On the other hand, competition may have a negative effect on credit due to information asymmetries and agency costs that are created due to the banks having difficulties in building relationships with already opaque SMEs (Petersen and Rajan, 1995; Hauswald and Marquez, 2006; Marquez, 2015). Jayakumar et al. (2018) studied a panel of 32 European countries and concluded that bank competition (as well as stability) are significant long-term drivers of economic growth and lower credit constraints. Banking market concentration is used as a proxy to study the link between bank competition and access to credit (Berger et al., 2017; Angori et al., 2019, 2020). Khan et al. (2016) provides evidence from Southeast Asian countries that a higher bank concentration has a negative effect on the banking competition on multiple levels. Petersen and Rajan (1995), by studying the US market data, determined that SMEs are more likely to receive financing in concentrated credit markets. These results are corroborated by Zarutskie (2006, 2013) who determined that opaque companies in competitive markets had lower levels of debt. A study by Love and Pería (2015) provides consistent international evidence to support the market power hypothesis, which argues that a lower market competition reduces the access to credit and rejects the information hypothesis, which argues that a higher competition diffuses information gains and breaks bank-firm relationships, thus lowering the access to credit. Wang et al. (2020) presented novel evidence that the bank market power has a negative effect on the SME firm ability to access credit by worsening their credit constraints. Such an effect is stronger for businesses which are more information opaque and have higher needs for external funding. On the other hand, Beck et al. (2004) found that companies of all sizes, operating in more concentrated banking markets, face higher financing constraints, while the effect diminishes with the company size. Positive relationship between market concentration and financing constraints was also found by Chong et al. (2013). Arping (2019); Hirata and Ojima (2020) showed that competition in banking can act as a destabilizing mechanism which can make banks act more prudently. With fiercer competition, banks' margins erode thus creating a higher risk of failure; instead of accepting the risk, banks respond by taking less risks. The extent of the destabilizing effect is not distributed equally between credit market industries. By analyzing and comparing the competition in banking and factoring industries in Italy, Degl'Innocenti et al. (2019) noted that, for both industries, an increase in competition has a destabilizing effect, but, for factoring industry, the effect is weaker. Nonetheless, by developing a flexible dynamic model of global banking, Faia et al. (2018) determined that the overall banking risk can decrease due to a higher competition if it originates from foreign expansion to more competitive markets, therefore leading to increased lending volumes.

On the contrary to the previous studies which connect the market concentration to competition, some papers employ the bank pricing behavior as a direct measure to assess the link between competition and the access to credit. Claessens and Tzioumis (2006) used H-statistic (Panzar and Rosse, 1987) which captures the elasticity of bank revenues to input prices to determine that competitive financial sectors are better at providing credit to financially constrained companies. Carbo-Valverde et al. (2009) used the ratio of the bank price mark-up (Lerner index) as a measure of market competitiveness to analyze Spanish SMEs. Their findings suggest that market competition does in fact improve the company access to finance and that concentration is not an appropriate proxy for banking competition. Olivero et al. (2011) used a broader approach for grasping market competition by only focusing on the bank balance and income statement data and connecting it to the loan portfolio growth. Findings suggest that competition has inverse relationship with bank lending, and it lowers the effectiveness of the monetary policy. Banking market competition and its relation to the access to financing is affected by two opposite forces. On the one hand, fierce banking market competition puts pressure on the interest rate, which, in turn, makes lending more affordable for all businesses irrespective of their opacity. On the other hand, it complicates relationship building with SMEs thus increasing the information gap, which leads to higher financing constraints for opaque borrowers (Chong et al., 2013). The perceived higher risk in the emerging and developing markets by cross-border lenders also has an adverse effect on the access to credit through lower credit supply and higher costs of borrowing (Collier and Cust, 2015; Mihalyi et al., 2022)

The participation of state-owned banks in the financial system is a worldwide, pervasive practice, which attributes approx. 20% of the total assets of the whole banking system (Gonzalez-Garcia and Grigoli, 2013). Government's participation in financial markets can be rationalized by two broad views. First, the development-centric approach argues that the existence of state-owned banks can promote the development of certain sectors or regions which normally would not be reached by private-owned banks (Gerschenkron, 1962) The second, a more recent, view, argues that state-owned banks are politically motivated and are used to politicized resource allocation for the sake of fulfilling the goals of personal political agendas (Kornai, 1979, 1980, 1986). States can participate in financial systems in several different ways: through direct subsidies, regulation, persuasion to lend to preferred projects and partial or complete ownership. Bank ownership enables regulators to have complete control over financing choices and leaves the implementation of projects to the private sector. By analyzing individual loan contracts in Italy, Sapienza (2004) compared the cost of borrowing for two sets of companies (with identical characteristics) borrowing from state-owned and private-owned banks. State-owned banks, throughout all risk profiles, charged lower interest rates than private-owned banks, thus improving the company access to credit. However, state-owned banks were found to mostly favor large firms or companies operating in depressed areas. The political agenda of government's participation was also evident as electoral results had a direct effect on the lending behavior of state-owned banks: the stronger is the party in a particular area, the lower is the charged interest. State-owned banks positively impact the availability of credit through the certification effect. Meuleman and De Maeseneire (2012) investigated the effect of R&D subsidies in the company ability to receive financing. The companies which have received R&D subsidies are more likely to successfully raise long-term debt. Martí and Quas (2018)] build upon this study and analyze the impact of participation loans and their effect on SME ability to access external financing. Their findings confirm that the government certification effect exists and that the companies which received participation loans were more likely to access further financing. The certification effect was more potent for those companies which are suffering from higher information asymmetries, such as small, high technology firms. Ogura (2018) considered the importance of state-owned bank lending during the Financial Crisis in Japan. Contrary to private-owned banks which only maintained credit to insolvent companies, state-owned banks issued credits which were used for real investments and were not hoarded in bank accounts. In high income countries, publicly owned banks tend to lend counter-cyclically, however, the credit allocation process remains quite inefficient (Brei and Schclarek, 2013; Bertay et al., 2015). Duprey (2015) confirms that the support of state-owned banks is significantly less cyclical than that of private-owned banks, but also notes that the ability to absorb negative shocks decreases marginally as the extent of the shock increases. The state's participation in bank ownership is not limited to single state as two or more countries can charter Multilateral Development Banks (MDBs) to encourage economic development in poorer countries. Gurara et al. (2020) studied the effect of MDBs on the credit supply, by focusing on the cost-of-borrowing. The participation of MDBs is connected to a higher cost of borrowing, but, at the same time, it enables 'lighter' product conditions, such as longer maturity terms. Such a business model is feasible for MDBs as they have informational advantage over private investors, particularly in assessing the country risk and their capacity to monitor. MDBs tend to ease access to credit for riskier borrowers, mainly, in low-risk countries. Within high-risk markets, there is limited evidence that the availability of funds for riskier borrowers has improved. The availability of low-yield funds for infrastructure investments could be reached by the transformation of MDBs to originate-and-distribute banks, through the participation of long-term institutional investors (Arezki et al., 2017).

On the other hand, state-owned banking lowers the overall banking sector outreach (Beck et al., 2007) and generally does not serve the highly credit constrained companies such as SMEs (see Ongena and Sendeniz-Yunc (2011) for evidence from Turkey). Furthermore, Gonzalez-Garcia and Grigoli (2013) determined that the higher participation of state-owned banks improves access to financing to public sector companies, but it comes at a price of larger fiscal deficits and crowding out of credit from the private sector. Carvalho (2014) provides evidence from emerging markets that the government control over banks has strong influence on the companies' investment decisions as they are closely related to more favorable financing conditions in politically attractive regions. Governments tend to shift lending to politically attractive regions from the unattractive ones. Cao et al. (2023) add that state-owned bank lending to other state-owned companies creates conditions for corporate wealth exploitation by company management. If the loan's creditor and debtor is the same owner, the debt contract can be rewritten at any time, which creates soft budget constraints, corporate governance vacuum, and, consequently, the accumulation of bad loans. Under normal circumstances, the effect from debt financing would act as a means of disciplining managers, especially at times of financial distress. Berger and Roman (2020) find that efficiency and performance of individual state-owned banks, in virtually all cases, is very poor, while also arguing that their continued lending during recessions ensured higher credit availability for businesses and a stronger overall financial system. Companies concentrating on banking relationships that are based on transaction lending are more likely to be denied a credit, while relationship lending helps improve access to financing for all-sized companies (Berger et al., 2022).

Table 9. Financial institution structure factors used for evaluating access to credit. Created by the author.

Variable	Definition	Study
Market concentration	Herfindahl-Hirschman Index in markets served by the bank.	Petersen and Rajan (2002); Elsas (2005); Carbo-Valverde et al. (2009); Berger et al. (2011); Presbitero and Zazzaro (2011); Chong et al. (2013); Zarutskie (2013); Love and Pería (2015); Milani (2014); Khan et al. (2016); Durguner (2017); Berger et al. (2017); Chen et al. (2017); Degryse et al. (2018); Ogura (2018); Angori et al. (2019, 2020); Aristei and Angori (2022)
Deposits	The proportion of deposits controlled by the bank.	Berger et al. (2011); Bertay et al. (2015); Khan et al. (2016); Chen et al. (2017); Degryse et al. (2018); Ogura (2018); Fang et al. (2022)
Income growth	Average income growth in markets served by the bank.	Berger et al. (2011); Bertay et al. (2015); Fang et al. (2022)
Capital	Bank's capital to its risk-weighted assets.	Berger et al. (2011); Olivero et al. (2011); Jiménez et al. (2012); Zarutskie (2013); Aiyar et al. (2014); Bertay et al. (2015); Khan et al. (2016); Belaid et al. (2017); Berger et al. (2017); Chen et al. (2017); Degryse et al. (2018); Berger et al. (2022); Fang et al. (2022)
Liquidity	Bank's ability to meet its short-term obligations and manage unexpected cash outflow.	Berger et al. (2011); Olivero et al. (2011); Jiménez et al. (2012); Aiyar et al. (2014); Bertay et al. (2015); Khan et al. (2016); Berger et al. (2017); Chen et al. (2017); Degryse et al. (2018); Ogura (2018); Berger et al. (2022); Fang et al. (2022)
Total assets	The sum of all assets including loans, investments, cash, and other assets.	Sapienza (2004); Carbo-Valverde et al. (2009); Berger et al. (2011); Olivero et al. (2011); Jiménez et al. (2012); Zarutskie (2013); Aiyar et al. (2014); Bertay et al. (2015); Khan et al. (2016); Chen et al. (2017); Ogura (2018); Berger et al. (2022); Fang et al. (2022)
ROA	Profitability metric measuring net income to total assets.	Carbo-Valverde et al. (2009); Jiménez et al. (2012); Ogura (2018); Fang et al. (2022)
Doubtful loans ratio	The proportion of loans that are at risk of default or are already delinquent.	Sapienza (2004); Jiménez et al. (2012); Bertay et al. (2015); Berger et al. (2017, 2022); Aristei and Angori (2022)
Interest rate	Interest rate at which banks lend or borrow funds from each other in the interbank market.	Agarwal and Hauswald (2010); Jiménez et al. (2012); Khan et al. (2016); Gurara et al. (2020); Ademosu and Morakinyo (2021)
Market share	The proportion of the market share which the bank controls.	Berger et al. (2011); Chong et al. (2013); Bertay et al. (2015); Berger et al. (2022)
Branch concentration	The total count of bank branches located within a specific area or region.	Petersen and Rajan (2002); Carbo-Valverde et al. (2009); Presbitero and Zazzaro (2011); Jiménez et al. (2012); Chong et al. (2013); Milani (2014); Berger et al. (2017); Angori et al. (2019, 2020); Aristei and Angori (2022)
Bank employees	The number of bank employees in a given market.	Petersen and Rajan (2002)
Bank age	The age of the bank.	Berger et al. (2011); Zarutskie (2013)
Ownership	The ownership of the bank is held by foreign investors or the state.	Bertay et al. (2015); Khan et al. (2016); Chen et al. (2017); Ogura (2018)

In conclusion, financial institution structure factors can have an adverse effect on the financial market competitiveness and the availability of funds by directing them towards long-term projects through bank ownership. However, this can also lead to financing inefficient and politicized projects. State-owned banks can benefit SMEs by providing access to credit in areas and sectors where capital would not be available under normal conditions or during recessions when privately-owned banks stop lending. Nevertheless, misappropriation of funds can hinder the effectiveness of state-ownership in supporting SME access to credit (Ongena and Sendeniz-Yunc, 2011; Gonzalez-Garcia and Grigoli, 2013; Cao et al., 2023). The reduction of local bank branches negatively affects SME lending due to a decreased information flow, but technology can help alleviate this issue. Market competitiveness plays a role in credit access, by improving availability overall but potentially hindering relationships with opaque SMEs (Petersen and Rajan, 2002; Agarwal and Hauswald, 2010; Milani, 2014; Kärnä and Stephan, 2022). State-owned banks have a significant role in financing, particularly due to benefiting large firms and depressed areas, but their impact on the overall banking sector is limited. Pricing behavior and government certification also influence credit access.

1.4.4. Firm and product characteristics

Specific firm characteristics can have adverse effects on SME ability to access financing in terms of both 1st degree rationing and 2nd degree rationing in a form of shorter maturity terms, smaller limits, higher collateral requirements, and higher spreads (Chodorow-Reich et al., 2022). As demonstrated by Angori et al. (2019), reducing the information asymmetry through long-term relationships works as one of the best ways for smaller companies to reduce credit constraints. For a smaller business to at least partially alleviate adverse access to credit conditions, it could be beneficial to hold multiple banking relationships, which would not be beneficial to larger companies, for which, having many lenders could have a negative effect on the access to credit. Small and medium-sized companies with a relatively higher share of debt with a single financing provider are less likely to be turned down, as the proven record of successful relationship works to overcome the underlying opacity problems. Mac An Bhaird et al. (2016) demonstrated the importance of age in access to credit as younger companies are more likely to be discouraged from applying.

Mina et al. (2013) demonstrated that knowledge-intensive companies tend to have smaller external capital needs than low knowledge intensity businesses. Furthermore, R&D intensity has a negative effect on company ability to receive financing. Amongst other company specific characteristics, young, service companies tend to have lower external financing need than older, manufacturers. Similarly, companies having growth targets and long pay-off periods tend to need more external financing. Companies with a low profit have lower external financing needs, which is in line with the pecking order theory – companies tend to use the cheapest capital, in this case, internal funds. Lee et al. (2015) add that innovative firms are both more likely to apply for financing and to have difficulties in accessing credit. Such types of companies tend to face absolute credit rationing when the financing application is completely rejected. The inability to access financing for innovative companies acts as a structural problem in a financial system where companies which are directed at growth and new product development are not able to utilize their potential due to credit rationing. As demonstrated by Cheng et al. (2014), firms investing in corporate social responsibility performance are facing significantly lower financial constraints through higher stakeholder engagement and a lower information asymmetry. The companies that advocate for corporate social responsibility tend to enter the positive feedback loop where higher transparency is driven by reporting standards, which tend to change internal controls and improve compliance with regulations.

Aterido and Hallward-Driemeier (2011); de Andrés et al. (2021) argue that the gender is a crucial factor in access to credit as women were determined to face more challenges and constraints in obtaining credit for their SMEs than men. They are less likely to not only apply for credit but also to be approved. A study conducted by Galli et al. (2020) on European SMEs found that female-led firms are more likely to avoid applying for loans. However, when they do apply, they are not subjected to gender discrimination from lenders, except during the upside phase of the economy. On the other hand, a study conducted by Irwin and Scott (2010) found that female SME owners/managers actually find it easier to access credit compared to their male counterparts. The ownership structure is also important in terms of other dimensions such as who is the majority shareholder, whether the manager is the owner and whether other legal entities own the company. Beck et al. (2007) demonstrated that the legal status and the corporate structure of a business, whether it is public, private, or foreign-owned, can impact its ability to borrow money from financial institutions. Sikochi (2020) analyzes how a firm's corporate legal structure and governance affects its borrowing costs. The results show that a more complex legal structure leads to higher loan spreads, which can be partially attributed to recovery risk.

The certification effect is a phenomenon that describes how government subsidies and loans can act as a factor facilitating companies' access to credit. Meuleman and De Maeseneire (2012) investigated the effect of R&D subsidies in a company's ability to receive financing. They found that the companies which have received R&D subsidies are more likely to successfully access external debt. Building upon this study, Martí and Quas (2018) analyzed the impact of participation loans on SMEs' ability to access external financing. Their findings confirm the existence of a government certification effect and show that those companies which received participation loans were more likely to access further financing. The certification effect was more potent for companies suffering from higher information asymmetries, such as small, high technology firms. As demonstrated by Brei and Von Peter (2018), for companies suffering from higher information opaqueness, the impact of the location on credit accessibility can be particularly salient, as these firms may face greater challenges in communicating their creditworthiness to lenders. SMEs located in closer proximity to lending institutions have a higher likelihood of obtaining financial access. This is because financial institutions can easily access qualitative information to assess the credibility of SME applicants. Fatoki and Asah (2011) corroborate this notion and argue that SMEs situated in urban areas have greater success in obtaining credit compared to those in rural regions. Consequently, these studies suggest that a business's location is a significant predictor of credit accessibility. Table 10 summarizes key Firm Characteristics that have been used as independent or control variables.

Table 10. Firm Characteristic factors used to evaluate access to credit. Created by the author.

Factor	Definition	Study
Age	Company age at the time of application.	Cole (1998); Petersen and Rajan (2002); Peltoniemi (2007); Agarwal and Hauswald (2010); Presbitero and Zazzaro (2011); Jiménez et al. (2012); Chong et al. (2013); Cassar et al. (2015); Love and Pería (2015); Neuberger and Räthke-Döppner (2015); Kirschenmann (2016); Cucculelli and Peruzzi (2017); Durguner (2017); Martí and Quas (2018); Angori et al. (2019); Grzelak (2019); Ferri et al. (2019); Angori et al. (2020); Motta and Sharma (2020); Aoki (2021); Malakauskas and Lakštutienė (2021); Zainol Abidin et al. (2021); Kärnä and Stephan (2022); Medianovskyi et al. (2023)
Diversity	Indicates the gender diversity of company owners and managers.	Angori et al. (2019, 2020); Motta and Sharma (2020); Zainol Abidin et al. (2021)
Managing owner	Owner is the manager of the company.	Petersen and Rajan (2002); Cassar et al. (2015); Cucculelli and Peruzzi (2017); Angori et al. (2019, 2020)
Ownership	The set-up of company ownership in terms of the number of owners, majority shareholders.	Petersen and Rajan (2002); Cassar et al. (2015); Love and Pería (2015); Cucculelli and Peruzzi (2017); Angori et al. (2019, 2020); Zainol Abidin et al. (2021); Aristei and Angori (2022)
Governance	The set-up of company governance in terms of the existence of the board and its composition.	Zainol Abidin et al. (2021); Aristei and Angori (2022)
Size	Company's size class as determined by a selected measure.	Cole (1998); Sapienza (2004); Elsas (2005); Brown et al. (2009); Carbo-Valverde et al. (2009); Bosch and Steffen (2011); Presbitero and Zazzaro (2011); Chong et al. (2013); Meuleman and De Maeseneire (2012); Florou and Kosi (2015); Love and Pería (2015); Neuberger and Räthke-Döppner (2015); Kirschenmann (2016); Belaid et al. (2017); Berger et al. (2017); Cucculelli and Peruzzi (2017); Angori et al. (2019); Grzelak (2019); Ferri et al. (2019); Angori et al. (2020); Motta and Sharma (2020); Zainol Abidin et al. (2021); Aoki (2021); Berger et al. (2022); Aristei and Angori (2022); Chodorow-Reich et al. (2022)
Legal form	The legal entity type under which the company is incorporated.	Cole (1998); Petersen and Rajan (2002); Elsas (2005); Peltoniemi (2007); Brown et al. (2009); Bosch and Steffen (2011); Chong et al. (2013); Cassar et al. (2015); Kirschenmann (2016); Berger et al. (2017); Durguner (2017); Grzelak (2019); Gurara et al. (2020); Motta and Sharma (2020); Aristei and Angori (2022)
Location	The classification of the region the company is operating in.	Petersen and Rajan (2002); Jiménez et al. (2012); Milani (2014); Berger et al. (2017); Durguner (2017); Grzelak (2019); Motta and Sharma (2020); Kärnä and Stephan (2022)
Sector	Sector the company is operating in.	Cole (1998); Peltoniemi (2007); Bosch and Steffen (2011); Jiménez et al. (2012); Love and Pería (2015); Neuberger and Räthke-Döppner (2015); Bonnet et al. (2016); Belaid et al. (2017); Durguner (2017); Martí and Quas (2018); Motta and Sharma (2020); Aristei and Angori (2022); Chodorow-Reich et al. (2022); Kärnä and Stephan (2022)
Certification	The company has gone through any form of quality certification.	Presbitero and Zazzaro (2011); Angori et al. (2019, 2020)
Incentives	The company has received any form of public incentives.	Meuleman and De Maeseneire (2012); Bonnet et al. (2016); Martí and Quas (2018); Angori et al. (2019, 2020)
Audit	Indicates whether submitted annual financial statements were audited or not.	Brown et al. (2009); Palazuelos et al. (2018); Motta and Sharma (2020)

Berger and Udell (2006) suggest that a relationship between different loan products and Firm Characteristics exists. Some products may be more appropriate for larger and more established firms, while smaller and younger firms may benefit more from other financing sources, such as trade credit or leasing. Thus, the effectiveness of credit products in meeting the financing needs of firms is contingent on a range of company-specific factors, such as size, age, and creditworthiness which interact with the specific features of the product being considered. For instance, the type of product issued (e.g., loan, line of credit, or trade credit) has been found to be an important factor impacting company access to credit in studies by Agarwal and Hauswald (2010); Zarutskie (2013); Cassar et al. (2015); Kirschenmann (2016). Furthermore, the underlying product conditions such as the size of the pledged collateral (Peltoniemi, 2007; Bosch and Steffen, 2011; Neuberger and Räthke-Döppner, 2015) and the agreement interest rate (Durguner, 2017; Xu et al., 2020; Kärnä and Stephan, 2022) are other factors that have been shown to affect the access to credit. The type of collateral also matters as tangible asset pledging has been found to be positively related to the access to credit in studies by Florou and Kosi (2015); Adam and Streitz (2016); Angori et al. (2019). Finally, the financing amount size as well as the maturity for which the product is issued have been found to be underlying conditions defining SME access to credit (Peltoniemi, 2007; Bosch and Steffen, 2011; Cassar et al., 2015; Adam and Streitz, 2016). Table 11 summarizes variables used for access to credit evaluation across different studies.

Table 11. Product Characteristic factors used for evaluating access to credit. Created by the author.

Variable	Definition	Study
Product	The type of the issued product.	Agarwal and Hauswald (2010); Zarutskie (2013); Cassar et al. (2015); Adam and Streitz (2016); Kirschenmann (2016); Minnis and Sutherland (2017); Gurara et al. (2020); Berger et al. (2022)
Collateral	The size of the pledged collateral.	Peltoniemi (2007); Agarwal and Hauswald (2010); Bosch and Steffen (2011); Zarutskie (2013); Neuberger and Räthke-Döppner (2015); Kirschenmann (2016); Ferri et al. (2019); Gurara et al. (2020); Motta and Sharma (2020); Berger et al. (2022); Chodorow-Reich et al. (2022)
Asset	Indication whether the business pledged any tangible assets like real estate, machinery.	Bosch and Steffen (2011); Florou and Kosi (2015); Adam and Streitz (2016); Kirschenmann (2016); Minnis and Sutherland (2017); Angori et al. (2019); Motta and Sharma (2020); Chodorow-Reich et al. (2022)
Interest rate	The agreement interest rate.	Peltoniemi (2007); Agarwal and Hauswald (2010); Zarutskie (2013); Cassar et al. (2015); Florou and Kosi (2015); Adam and Streitz (2016); Durguner (2017); Minnis and Sutherland (2017); Gurara et al. (2020); Xu et al. (2020); Berger et al. (2022); Chodorow-Reich et al. (2022); Kärnä and Stephan (2022)
Amount	The financing amount size.	Peltoniemi (2007); Cassar et al. (2015); Florou and Kosi (2015); Neuberger and Räthke-Döppner (2015); Adam and Streitz (2016); Minnis and Sutherland (2017); Gurara et al. (2020); Xu et al. (2020); Aoki (2021); Kärnä and Stephan (2022)
Maturity	The length for which the product is issued.	Peltoniemi (2007); Agarwal and Hauswald (2010); Bosch and Steffen (2011); Florou and Kosi (2015); Adam and Streitz (2016); Minnis and Sutherland (2017); Gurara et al. (2020); Xu et al. (2020); Aoki (2021); Berger et al. (2022); Chodorow-Reich et al. (2022)

It has been determined that the SME access to credit can be affected by various company and product characteristics and interactions between them. Amongst more common ones, such as the age and size, these include higher levels of knowledge intensity, the legal structure of the company, location, and the gender of the owner/manager. Women-owned SMEs face more challenges in obtaining credit than their male counterparts, although the situation has improved in recent years. The certification effect, which describes how government subsidies and loans can facilitate companies' access to credit, is also a factor to consider. Finally, different loan product types and product conditions may have an adverse impact on the access to credit and be more appropriate for different company sizes.

1.5. Conceptual Model for Evaluating SME Access to Credit

In order to appropriately evaluate the underlying SME access to credit, it is important to first define the proxy which would be used to quantify it. As indicated in the previous sections, the access to credit proxies can be grouped into two major groups – credit supply (Miao and Wang, 2012; Deyoung et al., 2015; Peón and Guntín, 2021; Chodorow-Reich et al., 2021, 2022; Ćehajić and Košak, 2022; Fang et al., 2022) and credit demand (Chong et al., 2013; Kirschenmann, 2016; Angori et al., 2019, 2020; Aristei and Angori, 2022; Brown et al., 2022). The selection of the proxy mainly depends on the research problem and data availability. Credit demand related studies mostly focus on the factors impacting decisions to request financing and explaining reasons for borrower discouragement. Meanwhile, the studies that use credit supply as a proxy tend to focus on either the bank loan portfolio and macro-specific conditions or the application outcomes and factors impacting the company ability to receive approval or be rationed. Whichever proxy is used, when estimating SME access to credit, it is important to account for both macro- and individual application factors (see Figure 4) (Berger and Udell, 2006).

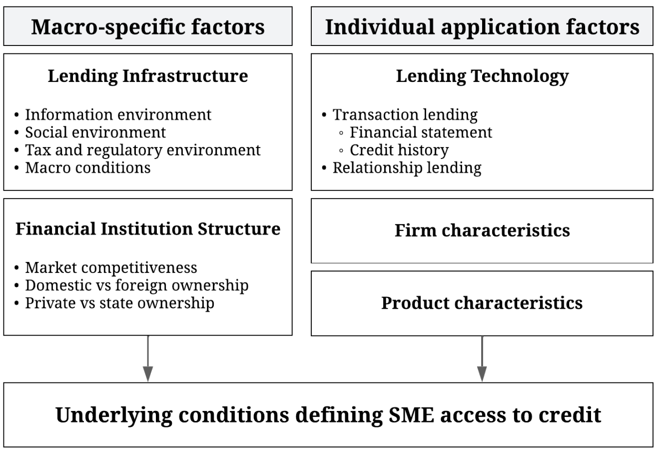


Figure 4. Underlying factor groups for evaluating SME access to credit. Created by the author.

Macro-specific factors, encompassing the Lending Infrastructure and the Financial Institution Structure, establish the fundamental market conditions, often referred to as the ‘rules of the game,’ which are beyond the influence of individual entities. A market characterized by informational transparency, marked-to-market balance sheets, and the presence of active rating agencies tends to exhibit lower financial constraints and, consequently, greater access to credit (Beck et al., 2007; Armstrong et al., 2010; Florou and Kosi, 2015; Dobbie et al., 2020; Deno et al., 2020; Adrian and Shin, 2010; Bosch and Steffen, 2011). However, enhanced market clarity can be a double-edged sword, as it may limit access to credit due to higher perceived risks by lenders (Musto, 2004; Bonnet et al., 2016; Johnstone, 2016; Gross et al., 2020; Kalogirou et al., 2021). Socioeconomic factors, such as marital status, average entrepreneur age, or population density, play a relatively minor role in determining credit accessibility (Neuberger and Räthke-Döppner, 2015; Brei and Von Peter, 2018). Conversely, the tax and regulatory environment closely influences bank lending behavior through capital requirements and active monetary policy (Aiyar et al., 2014; Chen et al., 2017; Fang et al., 2022). Understanding the market participant structure, objectives, and overall competitiveness is crucial, as privately owned firms may have vastly different business objectives compared to state-owned enterprises. State-owned financial institutions, including multilateral development banks, enhance credit availability for riskier companies due to their advantages in assessing country risk and monitoring borrowers (Arezki et al., 2017; Gurara et al., 2020). Additionally, SMEs that secure subsidies from state-owned agencies are more likely to access financing in the future (Lerner, 1999; Meuleman and De Maeseneire, 2012; Martí and Quas, 2018). While domestic- and foreign-owned financial institutions tend to issue financing similarly during regular economic cycles, they may reduce their exposures in foreign countries during uncertain times, affecting SMEs operating in markets dominated by foreign-owned banks (Davis, 2015; Wang, 2018). Lastly, competitive markets, as highlighted by Kärnä and Stephan (2022), have a significantly positive impact on credit accessibility, resulting in lower interest rates and higher loan amounts for SMEs.

Individual application factors, encompassing Lending technology, Firm Characteristics, and Product Characteristics, define the inherent qualities possessed by a potential borrower and those they can influence. The ability of SMEs to access credit is directly affected by lending technology, fully under the control of financial institutions. However, the choice of financing provider and product depends on the company and its decision-makers. Credit accessibility is not evenly distributed among lending providers and products. Transaction lending is typically favored by larger, more transparent companies, increasing their chances of securing financing (Berger and Udell, 2006; Maier, 2016; Kgoroeadira et al., 2019). Firms with strong performance in corporate social responsibility face significantly reduced financial constraints due to higher stakeholder engagement and lower information asymmetry (Cheng et al., 2014). Conversely, for smaller, opaque businesses, the bank-firm relationship, a component of relationship lending, plays a crucial role in mitigating information opacity and reducing the impact of credit rationing (Kirschenmann, 2016; Durguner, 2017; Angori et al., 2019; Xu et al., 2020). Individual firm characteristics, as demonstrated by Angori et al. (2019), can either aid or hinder access to credit; smaller companies benefit from multiple banking relationships, whereas larger ones face a negative effect. Ultimately, firms unable to access financing from traditional providers (banks) can turn to alternative lenders, creating a new loan market with different requirements, evaluation criteria, and decision-making processes (Maier, 2016; Bento et al., 2019; Kgoroeadira et al., 2019; Butticè and Rovelli, 2020; Nisar et al., 2020; Liu et al., 2021; Wasiuzzaman et al., 2021).

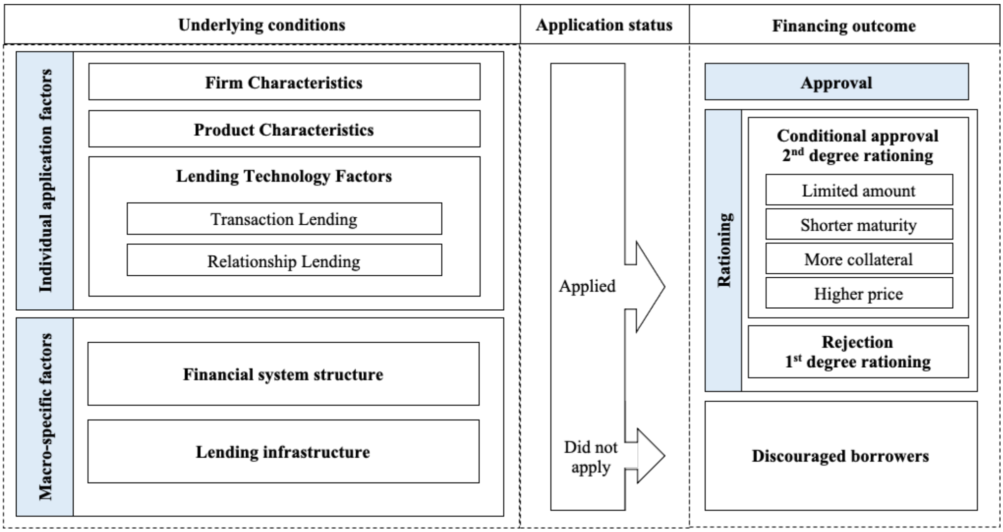


Figure 5. Conceptual SME access to credit model. Created by the author.

Based on previous studies, a conceptual SME access to credit model (see Figure 5) is constructed. The model is composed of three major components:

- Underlying conditions – a group of macro- and individual application factors which define SME's ability to access credit.
- Application status – the decision by an SME entity whether to apply for credit or not.
- Financing outcome – the final outcome which indicates whether the credit was accessed or not.

The conceptual SME access to credit model (see Figure 5) provides key SME access to credit elements. The underlying conditions composed of macro-specific and individual application factors define the key conditions for accessing credit. It is important to consider factors from both macro-specific and individual application factors groups, which might not always be possible due to data limitations. Such studies as Cassar et al. (2015); Adam and Streitz (2016); Kirschenmann (2016); Angori et al.

(2019); Calabrese et al. (2022); Kärnä and Stephan (2022) solve such limitations by constructing individual country models.

1.6. Access to Credit Modelling Techniques and Explainability Methods

Since the access to credit analysis is similar to pattern-recognition problems, algorithms can be used to classify the ability to access credit for individual companies (Kruppa et al., 2013; Pal et al., 2016). The majority of studies that explore different access to credit modelling techniques focus on the perceived creditworthiness of the individual entity. Molina and Preve (2012) find that modelling the access to credit is similar to estimating the financial distress of a company as banks' credit decision making processes are based on evaluating the borrower's ability to make future loan installments. Dastile et al. (2020) elaborate that estimating the access to credit is challenging due to a number of reasons such as: multicollinearity of independent variables – as the access to credit is influenced by a wide range of factors, they are often interdependent and can be difficult to disentangle from one another; data availability – as the availability of data varies between markets and individual entities; what would work in one market might not work in another; human biases – as estimating access to credit involves explaining decisions made by humans, the findings can inherently be 'fuzzy'. Therefore, these factors must be carefully considered and addressed when estimating the access to credit to ensure a fairly and accurately estimated model.

Seminal studies have evaluated access to credit by using such traditional modelling techniques as Discriminant Analysis (DA) and Logistic Regression (LR) (Cox, 1958). Though, often used as benchmarks to compare with state-of-the-art machine-learning techniques, DA and LR are simple by nature and are not well-suited for managing larger datasets and variable inter-dependencies (Barboza et al., 2017). Correa Bahnsen et al. (2016) demonstrated that, for estimated cases when independent variables displayed complex non-linear relationships, the traditional logistic regression did not perform as effectively as state-of-the-art machine learning methods. Despite the fact that the LR model may not match other machine learning models in terms of prediction accuracy, it holds a significant advantage in terms of variable interpretability and stability. In contrast to the traditional estimation techniques, state-of-the-art machine learning techniques such as decision trees (DT), random forest (RF), artificial neural networks (ANN), support vector machine (SVM), k-nearest neighbor (kNN) have become increasingly popular in recent years, as they offer the potential to improve accuracy and reduce bias in credit scoring models (Barboza et al., 2017). SVM can generate functions similar to discriminant analysis. However, unlike discriminant analysis, SVMs are not limited by a series of assumptions and are therefore less restrictive (Cortes and Vapnik, 1995). Silva et al. (2020) predict credit rationing by using machine learning techniques. The estimation is split into static and dynamic parts which depend on data availability. The study concluded that dynamic models are able to achieve a higher accuracy than the static ones but are more susceptible to missing data. Furthermore, feature indiscriminate addition does not necessarily yield a higher accuracy model. Danenas and Garsva (2015) demonstrated that SVM is able to achieve

Table 12. Modelling techniques used for evaluating access to credit. Created by the author.

Modelling technique	Study
Traditional techniques	
Discriminant Analysis	Mahmoudi and Duman (2015); Barboza et al. (2017)
Logistic Regression	Wang et al. (2011); Kruppa et al. (2013); Danenas and Garsva (2015); Datta et al. (2016); Barboza et al. (2017); Ariza-Garzon et al. (2020); Wang et al. (2020); Malakauskas and Lakštutienė (2021); Moscato et al. (2021); Hussin Adam Khatir and Bee (2022); Medianovskyi et al. (2023)
Machine learning techniques	
Decision Tree	Wang et al. (2011); Datta et al. (2016); Trivedi (2020); Wang et al. (2020); Hussin Adam Khatir and Bee (2022)
Random Forest	Wang et al. (2011); Kruppa et al. (2013); Danenas and Garsva (2015); Datta et al. (2016); Barboza et al. (2017); Ariza-Garzon et al. (2020); Silva et al. (2020); Trivedi (2020); Malakauskas and Lakštutienė (2021); Moscato et al. (2021); Hussin Adam Khatir and Bee (2022); Medianovskyi et al. (2023)
K-Nearest Neighbor	Kruppa et al. (2013); Wang et al. (2020); Hussin Adam Khatir and Bee (2022)
Bagged-Nearest Neighbor	Kruppa et al. (2013); Barboza et al. (2017)
Support Vector Machine	Wang et al. (2011); Danenas and Garsva (2015); Datta et al. (2016); Pal et al. (2016); Barboza et al. (2017); Silva et al. (2020); Trivedi (2020)
Naïve Bayes	Trivedi (2020); Wang et al. (2020); Hussin Adam Khatir and Bee (2022)
Artificial Neural Network	Wang et al. (2011); Zhao et al. (2015); Barboza et al. (2017); Dastile and Celik (2021); Hadji Misheva et al. (2021); Malakauskas and Lakštutienė (2021); Hussin Adam Khatir and Bee (2022); Medianovskyi et al. (2023)
Gradient Boosting	Barboza et al. (2017); Bussmann et al. (2020); Qi et al. (2021); Bucker et al. (2022); Medianovskyi et al. (2023)

a classification accuracy comparable to other classifiers such as LR and RF. Trivedi (2020) estimated that DT and RF classification techniques were able to discriminate most accurately in terms of financial distress. Wang et al. (2020), on the other hand, estimated the credit risk by using LR, kNN, decision trees, RF and Naïve Bayes and concluded that RF showed the highest promise in terms of the modelling accuracy. A high RF technique classification accuracy for the access to credit related problem was also determined by Silva et al. (2020); Malakauskas and Lakštutienė (2021); Hussin Adam Khatir and Bee (2022). Medianovskyi et al. (2023), on the other hand, found that Gradient Boosting classification techniques can be efficient and more accurate by also enabling higher model explainability. Finally, Zhao et al. (2015) utilized a multi-layer perceptron (MLP) architecture to estimate an ANN model which was able to outperform other modelling techniques by a significant margin. Though some problems were identified by the authors as MLP tends to underperform with estimations with unbalanced data and is hard to interpret due to the hidden layer(-s). Table 12 summarizes access to credit related studies and their utilized modelling techniques.

The use of better performing machine learning techniques remains limited in highly regulated and high-stakes environments, due to the inherent black-box nature of the algorithms. This has led to a growing demand for transparency from stakeholders in AI. There is a risk of utilizing decisions that lack justification, legitimacy, or the ability to provide comprehensive explanations of their functionality (Preece et al., 2018). Studies that are held to overcome a gap between the machine learning model performance and opaqueness are generally collected under the eXplainable Artificial Intelligence (XAI) topic (Arya et al., 2019; Arrieta et al., 2020). Arrieta et al. (2020) provides a model and discusses the taxonomy on the machine learning model explainability. As a rule of thumb, high model explainability is related to a lower accuracy. As better modelling performance is sought, increased complexity which impedes interpretability is frequently encountered. Until recently, it appeared inevitable that interpretability would suffer as a consequence. However, with the emergence of more advanced methods for explainability, this downward trend could potentially be reversed or even eliminated (Gunning and Aha, 2019).

Table 13. Modelling techniques, demonstrated performance and perceived interpretability. Based on Gunning and Aha (2019); Arrieta et al. (2020).

Modelling technique	Performance	Interpretability
Traditional techniques		
Discriminant Analysis	Low	High
Logistic Regression	Low	High
Machine learning techniques		
Decision Tree	Low	High
Random Forest	High	Moderate
K-Nearest Neighbor	Moderate	Moderate
Bagged-Nearest Neighbor	Moderate	Moderate
Support Vector Machine	High	Low
Naïve Bayes	Moderate	Moderate
Artificial Neural Network	High	Low
Gradient Boosting	High	Moderate

Guidotti et al. (2018) present a classification of the primary issues discussed in the literature concerning the concept of explanation and the type of black box systems. Additionally, the proposed classification of methods for opening black box models provides a valuable perspective on the numerous research questions that remain unanswered. The study defines the following common classes for explainability methods:

- Saliency methods – used to highlight or identify the most important regions of

an input for a model's prediction.
- Neural Network Visualization Methods – used to visualize the internal workings of a neural network to gain insights into how it is making predictions.
- Feature Relevance Methods – used to identify the input features that are most relevant to a model's output.
- Exemplar Methods – used to identify the most representative examples of a particular class or concept.
- Knowledge Distillation Methods – used to transfer knowledge from a larger, more complex model to a smaller, more efficient model.
- High-Level Feature Learning Methods – used to learn abstract features from the input data, typically by using deep neural networks.
- Methods that Provide Rationales – used to generate explanations or justifications for a model's predictions.
- Restricted Neural Network Architectures – architectures that use constraints to limit the complexity and capacity of a model, typically to improve generalization performance.

As noted by Arya et al. (2019), the explainability method should be selected based on the modelling technique that was used and on the problem that is being solved. To evaluate the factor importance and the impact on the model, saliency methods should be used (Medianovskyi et al., 2023). One of the most promising saliency methods for interpreting machine learning models is SHapley Additive exPlanations (SHAP). It is a state-of-the-art method that provides local feature importance attributions. SHAP has been used in various modelling techniques. Kernel SHAP, which was first introduced in the original paper by Lundberg and Lee (2017), is an extension of LIME (Ribeiro et al., 2016) which builds a ridge regression on the sampled points around the explained instance. However, LIME has the drawback of instability where it returns different attributions for the same instance on each call, especially in cases of non-smoothness of the prediction function. Finally, the SHAP feature explainability methodology can be applied to virtually any modelling technique (Shrikumar et al., 2017; Sundararajan et al., 2017; Lundberg et al., 2020).

In conclusion, in order to estimate an appropriate SME access to credit model, it is recommended to use a variety of different modelling techniques and compare them to the benchmark, which provides a standardized and objective basis for evaluating and comparing the performance of different models. Traditional modelling techniques, such as DA and LR, have been used as benchmarks for state-of-the-art machine learning techniques. RF, ANN and GB techniques have consistently indicated high estimation performance in modelling the company access to credit. Machine learning models have the potential to improve accuracy and reduce bias in credit scoring models. On the other hand, the inherent black-box nature of AI models limits the interpretability and thus requires the use of explainability methods. Saliency methods, such as SHAP, are best suited to highlight an individual feature or their interaction importance.

Summary and findings

By conducting the analysis of SME access to credit, it was determined that the definition of SMEs differs across countries and academia. The dissertation adopts the European Commission's definition of an SME, which refers to an entity with less than 250 employees, an annual turnover of less than EUR 50 million, or a balance sheet total of EUR 43 million. Access to credit is defined as the degree to which businesses can obtain bank credit. It has been demonstrated that SMEs play a vital role in the global economy, as they create jobs, foster innovation, and contribute to economic growth. However, they often face challenges in obtaining credit and financing, which can impede their ability to grow and compete. These challenges are exacerbated by such factors as reduced sales, liquidity constraints, and supply chain shocks, which can force SMEs to lay off employees or shut down. Approximately one-third of SMEs cite a lack of access to affordable funding as a major obstacle to growth, resilience, and survival. Furthermore, compared to large companies, SMEs receive worse financing conditions in a form of shorter credit facility maturities, higher collateral requirements, higher utilization rates, and higher interest rates.

It has been determined that access to credit can be assessed by using credit supply or demand proxies, determined by underlying factors. Variable selection for proxies depends on the specific research problem, data availability, and the focus on macro or micro effects. When studying access to credit, it is essential to account for both macro- and individual application factors when using either type of the proxy. These factors can be grouped into lending technology factors, lending infrastructure factors, financial institution structure factors, firm and product characteristics. By connecting the underlying conditions with the possible financing outcomes, a conceptual credit accessibility model has been constructed.

Finally, in order to construct an appropriate model for SME access to credit estimation, it is recommended to utilize a variety of modelling techniques and compare them to the benchmark. Such traditional techniques as DA and LR serve as benchmarks for state-of-the-art machine learning techniques such as Random Forest, Artificial Neural Network, and Gradient Boosting, which have consistently shown high performance in estimating company access to credit. Machine learning models can enhance accuracy and decrease bias in credit scoring models. However, the black-box nature of artificial intelligence models limits their interpretability, thus necessitating the use of explainability methods such as Shapley Additive Explanations to highlight the importance of individual features or their interactions.

2. METHODOLOGY FOR EVALUATING SME ACCESS TO CREDIT

In the second section, the dissertation solves objective 4. A three-stage methodology for evaluating SME access to credit is created. This section defines dependent and independent variables, which are to be used in SME access to credit model development. The section provides the description of techniques and methods used in model development.

2.1. SME Access to Credit Variables and Comparative Analysis

The previous part has established that SMEs are crucial for the economy but often face difficulties in obtaining credit and financing, which hinders their growth and competitiveness. A conceptual access to credit model has been developed by connecting the underlying conditions with the financing outcomes. It has been determined that access to credit can be assessed using the credit supply or demand proxies, which are impacted by underlying factors grouped into macro-specific and individual application factor groups. To model SME access to credit, a variety of modelling techniques can be utilized to enhance accuracy and decrease bias, but explainability methods are necessary to highlight the importance of individual features. This part establishes a research methodology for evaluating the SME access to credit and the factors impacting it.

Dependent variable SME access to credit can be defined either through credit supply or demand proxies. It has been established that there is no universally accepted way of assessing access to credit, as the suitability of both supply- and demand-specific proxies depends on the specific research question, data availability, and research focus on either macro- or micro-level effects. As demonstrated in Figure 4, credit supply-specific proxies include the bank financing portfolio and the received application outcomes (such as approvals, 1^{st} and 2^{nd} degree rationing), while credit demand-specific proxies are comprised of communicated financing needs and complete borrower discouragement. To evaluate the SME access to credit, this dissertation utilizes seminal works by Jiménez et al. (2012), Kirschenmann (2016), and Ćehajić and Košak (2022) and establishes SME access to credit to be in line with the credit supply proxy – application outcomes represented as approvals and rejections (see Table 14).

Table 14. Dependent variable for evaluating SME access to credit. Created by the author.

Dependent variable	Description	Name
Application outcome	Indication whether a financing application was approved (0 = no) or was 1^{st} degree credit rationed (1 = yes)	*Outcome*

By examining access to credit through credit application outcomes in contrast to other proxies, this study will gain insights into the factors that contribute to the SME ability to access credit, including factor interactions (Jiménez et al., 2012). Specifically, this dissertation studies the SME application outcome (*Outcome*) in terms of approvals and 1st degree rationing. The rationing outcome is limited to 1st degree rationing as 2nd degree rationing is relatively difficult to identify and interpret (Jiménez et al., 2012; Kirschenmann, 2016). Therefore, *Outcome* is defined as the target (dependent) variable which is to be used in SME access to credit model estimation.

Independent variables In order to model an SME company's ability to access credit, the data-frame for underlying conditions (independent variables) must be defined. As noted by Malakauskas and Lakštutienė (2021); de Lange et al. (2022) and Medianovskyi et al. (2023), when estimating access to credit through the application outcome proxy, it is important to understand the time of the collected data and appropriately connect it to the underlying request. Furthermore, for the sake of cross-country modelling result comparability, variable selection for the empirical model estimation considers that data-points for each chosen factor are available in all studied countries. As suggested by Berger and Udell (2006), empirical access to credit model must be able to account for both macro-specific and individual application factors. Similarly to the methodology in studies by Cassar et al. (2015); Adam and Streitz (2016); Kirschenmann (2016); Angori et al. (2019); Calabrese et al. (2022); Kärnä and Stephan (2022), to account for macro-specific conditions, this study shall formulate individual country specific SME access to credit models without including macro-specific factors. Such a set-up will mitigate the possibility of the omitted variable bias, which would be present if a cross-country model were developed. As per findings of Pekarskienė and Susnienė (2011); Molendowski and Petraškevičius (2020); Činčikaitė and Meidutė-Kavaliauskienė (2023), concerning the limited extent of research focused on the Baltic States and the relative similarity between the three country economies, the empirical SME access to credit model shall be developed and analyzed for the three Baltic State countries – Estonia (EE), Latvia (LV), and Lithuania (LT). The case of the Baltic States is relevant as all three countries share the same major commercial banks which issue a substantial part of the total external financing. Finally, it is not clear whether the underlying SME access to credit in these countries is uniform and whether the importance of underlying conditions is homogeneous.

Table 15. Variable belonging to the Product Characteristic factor group. Created by the author.

Factor	Description	Measures	Name
Product type	The loan type for which an application was filled in (as per ECB working paper Ivashina et al. (2022))	Asset-based loan, Cash-flow loans, Trade finance, Leasing, Credit Cards	*Product*

As for individual application factors, variables belonging to all defined factor groups must be included; this includes Firm Characteristics, Product Characteristics, and Lending Technology factors (Berger and Udell, 2006). For Product Characteristics, the factor group independent variable (see Table 15) is the applied product type (*Product*), as per Ivashina et al. (2022), which determined that the aggregate credit supply can be driven by individual loan types.

The type of a financial product is an essential factor for the access to credit, as it determines the terms and conditions of the loan or credit facility. Different financial products offer varying levels of risk and return, and lenders often tailor their offerings to the needs and preferences of specific borrowers. Secured loans (such as Asset-based loan and Leasing) may offer lower interest rates but require collateral, while unsecured loans (such as Cash-flow loans and Credit Cards) may be more flexible but come with higher borrowing costs. Similarly, Credit Cards may provide easy access to credit but have higher interest rates and fees. Therefore, the choice of a financial product can significantly impact the cost and availability of a credit for borrowers.

Table 16. Variables belonging to the Firm Characteristics factor group. Created by the author.

Factor	Description	Measures	Name
Age	Company age at the time of application in years	Years	*Age*
Diversity	The share of company's ownership held by female owners (as per Motta and Sharma (2020))	Percentage	*Diversity*
Ownership	The share of company's ownership held by natural persons	Percentage	*Private*
Size	Company's size category as defined in EC (2003)	Micro, Small, Medium	*Segment*
Legal form	The legal entity type under which the company is incorporated	Unlimited liability company, Partnership, Private limited liability company	*Type*
Location	The classification of the level of the region urbanization the company is operating in, based on the EU classification of NUTS level 3 regions (NUTS 2021)	Predominantly rural regions, Intermediate regions, Predominantly urban regions	*Region*
Sector	The economic sector the company is operating in based on NACE	Commerce, Construction, Hotels & Restaurants, Agriculture, Forestry, Fishing, Real Estate, Transportation, Information & Communication, Professional Services, Manufacturing	*Sector*
Audited	Indicates whether the annual financial statements were audited	Yes, No	*Audited*

The Firm Characteristics factor group defines company-specific information that can have a significant impact on a company's ability to access credit (see Table 16).

The independent variables that are used to develop the empirical SME access to credit model include: company's age (*Age*), company's gender diversity (*Diversity*), owner private or legal entity type (*Ownership*), size of the enterprise (*Segment*), company's legal entity form (*Type*), the urbanization level of the region the company is located (*Region*), the economic sector the company is operating in (*Sector*), and the indicator whether the company's financial statements are audited (*Audited*).

The selection of independent variables belonging to the Lending Technology factor group for the SME access to credit model development is based on the studies in Tables 5, 6, and 7. The Lending Technology factors consist of the Relationship Lending and Transaction Lending factor groups. The former type consists of factors defining the information concerning bank-firm relationship. To estimate the empirical SME access to credit model, the study utilizes bank-firm relationship defining variables (see Table 17). The selected variables cover relationship duration (*Relationship*) and intensity (*Payments*), debt share in the bank (*Debt*), the number of financing contracts (*FinContracts*), as well as other held products (*Cards*, *POS*, *Ecommerce*).

Table 17. Variables belonging to Relationship Lending factor group. Created by the author.

Factor	**Description**	**Measures**	**Name**
Duration	The length of a bank-company relationship in days at the time of application	Days	*Relationship*
Payments	The proportion of 12-month incoming payment transactions to the latest net sales	Percentage	*Payments*
Rejections	Indicates whether the applying company has received any negative decisions during the past 2 years before the application	Yes, No	*Rejections*
Debt share	The proportion of debt held in the bank to the total liabilities reported in the latest annual financial statements	Percentage	*Debt*
Contracts	The number of past financing contracts the company had with the bank at the point of application	Count	*FinContracts*
Debit cards	Indicates whether the company had any debit cards before applying for financing	Yes, No	*Cards*
Point of sales	Indicates whether the company had any point of sales products before applying for financing	Yes, No	*POS*
E-commerce	Indicates whether the company had any e-commerce products before applying for financing	Yes, No	*Ecommerce*

The Transaction Lending factor group consists of variables covering the company's financial statements, as well as the company's and its owner's credit history. To estimate the empirical SME access to credit model, the study utilizes Transaction lending variables based on the financial statement data and grouped into Liquidity, Solvency, Profitability and Activity indicator groups described in Tables 18, 19, 20 and 21.

Table 18. Liquidity variables belonging to the Transaction Lending factor group. Based on Walsh (2010); Fridson and Alvarez (2022).

Factor	**Description**	**Name**	
		t	*t-1*
Cash ratio	Cash ratio measures a company's ability to pay off its short-term liabilities with the available cash and cash equivalents. A cash ratio above 1 indicates a company's ability to fully settle the current liabilities with cash, while levels below 0.5 are generally considered as high risk. $$Cash\ Ratio = \frac{Cash\ and\ Cash\ Equivalents}{Current\ Liabilities}$$	*CR*	*pCR*
Quick ratio	The quick ratio, also known as the acid-test ratio, is a financial ratio that measures a company's ability to meet its short-term obligations with its most liquid assets, excluding inventory. A quick ratio of 1.0 or higher is generally considered to be favorable. Conversely, a ratio below 1.0 is generally considered to be weak. $$Quick\ Ratio = \frac{Current\ Assets - Inventory}{Current\ Liabilities}$$	*QR*	*pQR*
Current ratio	The current ratio is a financial ratio that measures a company's ability to pay off its short-term liabilities with its current assets. A current ratio of 1.5 or higher is generally regarded as healthy. Conversely, a current ratio below 1.0 is typically considered unhealthy. $$Current\ Ratio = \frac{Current\ Assets}{Current\ Liabilities}$$	*CuR*	*pCuR*

The liquidity factors as described in Table 18 are crucial to SME access to credit as they reflect the ability of borrowers to fulfill their obligations and repay their debts. The availability of liquid assets can serve as collateral and reduce the perceived risk of lending. Moreover, liquidity provides a buffer against unexpected events or fluctuations in income, which can help borrowers meet their financial commitments without defaulting on their loans. A lack of liquidity can constrain access to credit for SMEs and exacerbate their financial fragility in times of economic stress.

Table 19. Solvency variables belonging to the Transaction Lending factor group. Based on Walsh (2010); Fridson and Alvarez (2022).

Factor	Description	Name	
		t	*t-1*
Debt-to-equity ratio	The debt-to-equity ratio is a financial ratio that measures a company's leverage, or the amount of debt financing relative to equity financing. A favorable debt-to-equity ratio is typically below 1.0, while a ratio of 2.0 or higher is often regarded as indicating increased risk. $$\textit{Debt-to-Equity Ratio} = \frac{\textit{Total Debt}}{\textit{Total Equity}}$$	*DE*	*pDE*
Tangible asset ratio	The tangible asset ratio measures the proportion of a company's tangible assets to its total assets. There is no universally ideal value for the tangible asset ratio as it can vary based on company-specific circumstances. $$\textit{Tangible Asset Ratio} = \frac{\textit{Tangible Assets}}{\textit{Total Assets}}$$	*TA*	*pTA*
Debt ratio	The debt ratio is a financial ratio that measures the proportion of a company's debt to its total assets. A smaller debt ratio indicates low indebtedness, while a relatively good debt ratio typically falls within the range of 1 to 1.5. $$\textit{Debt Ratio} = \frac{\textit{Total Debt}}{\textit{Total Assets}}$$	*DR*	*pDR*
Debt-service-coverage-ratio	The debt service coverage ratio (DSCR) is a financial ratio that measures a company's ability to pay its debt obligations. DSCR of at least 2 is commonly regarded as robust. Traditional lenders often establish minimum DSCR requirements of not less than 1.2. $$\textit{DSCR} = \frac{\textit{Net Operating Income}}{\textit{Total Debt Service}}$$	*DSCR*	*pDSCR*
Asset coverage ratio	The asset coverage ratio is a financial ratio that measures a company's ability to cover its debt obligations with its assets. An asset coverage ratio above 1 is typically considered healthy. $$\textit{Asset Coverage Ratio} = \frac{\text{Total Assets} - \text{Intangible Assets}}{\text{Total Debt}}$$	*ACR*	*pACR*

The solvency factors described in Table 19 are fundamental for the financial stability and sustainability of an organization. They refer to the ability of the entity to meet its long-term financial obligations and maintain its ongoing operations. Solvency is crucial for the access to credit as lenders and investors need to have confidence in the ability of the borrower to repay its debts over time. A solvent SME can provide reassurance to creditors and investors that their investment is safe and will yield a reasonable return. Therefore, a lack of solvency can lead to higher borrowing costs, reduced access to credit, and even bankruptcy in extreme cases, which can have severe consequences for the overall economy. In order to be solvent, a company must generate sufficient earnings, which will be used in covering additional liabilities.

Table 20. Profitability variables belonging to the Transaction Lending factor group. Based on Walsh (2010); Fridson and Alvarez (2022).

Factor	Description	Name	
		t	*t-1*
Return on assets	Return on assets (ROA) is a financial ratio that measures a company's ability to generate profits from its assets. A good ROA is typically considered to be over 5%, but the actual levels to be considered 'good' or 'bad' greatly depend on individual sectors. $Return\ on\ Assets = \frac{Net\ Income}{Total\ Assets}$	*ROA*	*pROA*
Return on equity	Return on equity (ROE) is a financial ratio that measures the amount of net income returned as a percentage of shareholders' equity. A ROE of 15% is generally considered good, but the actual levels to be considered 'good' or 'bad' greatly depend on individual sectors. $Return\ on\ Equity = \frac{Net\ Income}{Shareholders'\ Equity}$	*ROE*	*pROE*
Gross margin ratio	Gross margin ratio is a financial ratio that represents the percentage of sales revenue that exceeds the cost of goods sold. The assessment of a good gross profit margin percentage depends on the industry or nature of sales, but typically an above 10% gross profit margin is considered good. $Gross\ Margin\ Ratio = \frac{Gross\ Profit}{Net\ Sales}$	*GMR*	*pGMR*
Profit margin ratio	Profit margin ratio is a financial ratio that measures the proportion of a company's net income to its net sales revenue. A healthy profit margin typically falls within the range of 7% to 10%. $Profit\ Margin\ Ratio = \frac{Net\ Income}{Net\ Sales}$	*PMR*	*pPMR*

The profitability factors described in Table 20 are a critical metric for any SME, as they reflect the ability to generate returns and sustain long-term growth. Profitability is important for the access to credit, as lenders and investors need to assess the financial health and performance of the borrower or the investment opportunity. Such profitability metrics as the return on investment, net income, or gross margin can provide insight into the revenue and expense structure of the borrower or investment, as well as the potential risks and rewards associated with it. A profitable borrower or investment can demonstrate the ability to generate sufficient cash flows to meet their financial obligations and repay their debts. Therefore, a lack of profitability can make it challenging to obtain credit or attract investment, particularly in competitive or uncertain markets. Furthermore, sustained profitability can improve the creditworthiness and reputation of the borrower or investment and lead to more favorable terms and conditions for future borrowing or investment opportunities.

Table 21. Activity variables belonging to the Transaction Lending factor group. Based on Walsh (2010); Fridson and Alvarez (2022).

Factor	Description	Name	
		t	*t-1*
Asset turnover ratio	Asset turnover ratio is a financial ratio that measures the efficiency of a company's use of its assets to generate revenue. $Asset\ Turnover\ Ratio = \frac{Net\ Sales}{Total\ Assets}$	*ATR*	*pATR*
Receivables turnover ratio	Receivables turnover ratio is a financial ratio that measures the efficiency of a company's management of its accounts receivable. $Receivables\ Turnover\ Ratio = \frac{Net\ Sales}{Accounts\ Receivable}$	*RTR*	*pRTR*
Change in sales	Change in sales is a financial metric that measures the percentage increase or decrease in a company's sales over a period of time. $\Delta Sales = \frac{Sales_t - Sales_{t-1}}{Sales_{t-1}}$	cS	
Change in current assets	Change in current assets is a financial metric that measures the percentage increase or decrease in a company's current assets over a period of time. $\Delta Current\ Assets = \frac{Current\ Assets_t - Current\ Assets_{t-1}}{Current\ Assets_{t-1}}$	cA	

The activity indicators as described in Table 21 are crucial for SME to access credit. These indicators reflect the company's operational efficiency, financial performance, and growth potential, which are the critical factors that lenders consider when assessing creditworthiness. Specifically, the accounts receivable and the asset turnover ratio can indicate the company's ability to manage its cash flow and generate revenue, while changes in sales can provide insight into the market demand and growth prospects. As per Walsh (2010), healthy levels of company activity factors depend greatly on the individual company, sector and other characteristics. In general, higher values indicate a stronger financial standing than the ratios that are closer to zero.

Table 22. Credit history variables belonging to the Transaction Lending factor group. Created by the author.

Factor	**Description**	**Name**
Number of Internal overdues	The count of internal overdues the company owed in the past 2 years.	*IOverC*
Internal overdue sum	The size of internal overdues the company owed in the past 2 years.	*IOverS*
Internal overdue length	The length of internal overdues the company owed in the past 2 years.	*IOverL*
Number of external overdues	The count of external overdues the company owed in the past 2 years.	*EOverC*
External overdue sum	The size of external overdues the company owed in the past 2 years.	*EOverS*
External overdue length	The length of external overdues the company owed in the past 2 years.	*EOverL*
Number of owner's internal overdues	The count of internal overdues the majority owner owed in the past 2 years.	*OIOverC*
Owner's internal overdue sum	The amount of internal overdues the majority owner owed in the past 2 years.	*OIOverS*
Owner's internal overdue length	The length of internal overdues the majority owner owed in the past 2 years.	*OIOverL*
Number of owner's external overdues	The count of external overdues the majority owner owed in the past 2 years.	*OEOverC*
Owner's external overdue sum	The size of external overdues the majority owner owed in the past 2 years.	*OEOverS*
Owner's external overdue length	The length of external overdues the majority owner owed in the past 2 years.	*OEOverL*
Defaults	Indication whether the company had any defaults.	*Defaults*
Owner's defaults	Indication whether the majority owner had any defaults.	*ODefaults*

Independent variables *CR*, *QR*, *CuR*, *DE*, *TA*, *DR*, *ROA*, *ROE*, *GMR*, *PMR*, *ATR*, *RTR*, *DSCR*, *ACR* which are used in the SME access to credit model development are continuous ratios which include two periods: current (t) – which indicates the latest available accounting period, and past (t-1) – which indicates the previous accounting period. The utilization of multi-period financial data enables to account for the period-at-risk, which improves the classifier performance (Shumway, 2001; Malakauskas and Lakštutienė, 2021). Transaction lending independent variables based on the credit history are selected from Table 6 and described in Table 22.

The credit history specific variable sub-set accounts for both the company and the owner credit history, as well as multiple overdue dimensions such as the count of overdues (*IOverC*, *EOverC*, *OIOverC*, *OEOverC*), the overdue size (*IOverS*, *EOverS*, *OIOverS*, *OEOverS*) and the overdue length (*IOverL*, *EOverL*, *OIOverL*, *OEOverL*). Furthermore, the overdue data for both the bank (internal) and other (external) companies is considered as per findings of Cassar et al. (2015); Medianovskyi et al. (2023). Finally, the indication of complete inability to meet financial obligations is considered (*Defaults*, *ODefaults*) as per findings of Neuberger and Räthke-Döppner (2015); Kirschenmann (2016).

Comparative analysis As the access to credit proxy (*Outcome*) and the underlying conditions (see Tables 15, 16, 18, 19, 20, 21 and 22) are defined, the methodology for analyzing the underlying access to credit is established in line with the reporting set-up utilized by ECB reports (ECB, 2022b,a) and a study by the Bank of Lithuania (BoL, 2021). To evaluate the SME access to credit in a given market, it is necessary to consider the changes in the total number of applications as well as the rejection rate throughout a given period. The cross-country comparison of the access to credit is carried out in-line with comparative analysis as conducted by Rupeika-Apoga (2014). The comparative analysis is carried out together with the descriptive statistics of both dependent and independent variables in order to uncover country-specific differences in both the access to credit and the independent variable distributions (see Figure 6).

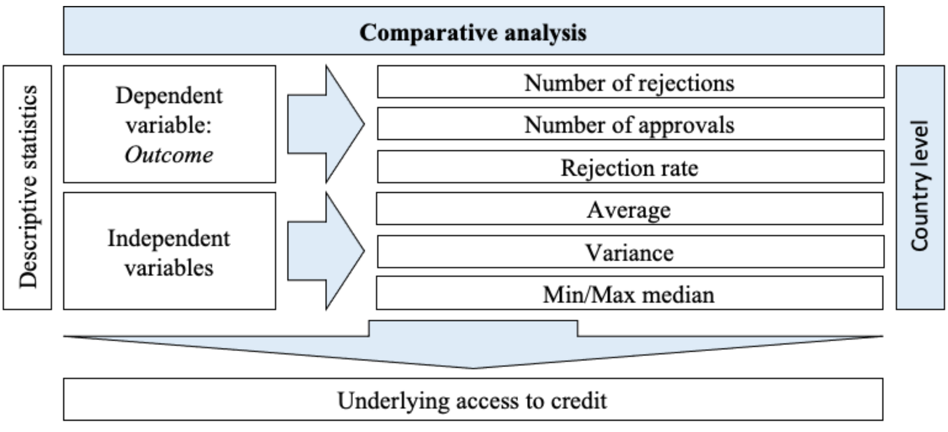


Figure 6. Methodology for evaluating the underlying SME access to credit in individual countries. Based on Rupeika-Apoga (2014); BoL (2021); ECB (2022b,a).

The comparative analysis will provide insights in explaining country-specific differences and unique characteristics. By comparing the actual SME access to credit and the underlying factors, it is possible to identify similarities and differences between countries as well as the potential importance of individual variables. Finally, the determined individual independent variable characteristics will provide insights towards variable dimensionality reduction before model estimation.

2.2. Dimensionality Reduction Procedure

Before creating an SME access to credit model by using state-of-the-art machine learning techniques, it is necessary to perform a dimensionality reduction procedure for each studied country. This procedure aims to decrease the number of features in a dataset while maintaining the maximum amount of relevant information. As noted by Lu et al. (2022), by utilizing a high-dimensionality data feature selection method, the focus on important factors can work towards improving SME entity creditworthiness and an easier access to credit. The purpose of dimensionality reduction is to streamline the complexity of the data, promote model robustness and the interpretability of features, and prevent model over-fitting. The reduction of the dimensionality of the data is a fundamental step in data processing before the model is estimated, which usually is a distinct step in the overall data mining process and can be achieved through feature selection (Shi et al., 2022). As described by Ha and Nguyen (2016), feature selection methods can be broadly classified into two categories: the filter approach, and the wrapper approach. The filter approach considers feature selection as a preliminary step before applying learning algorithms. A disadvantage of this approach is the lack of a direct relationship between the feature selection process and the learning algorithm's performance. The wrapper approach evaluates the feature selection by measuring the learning accuracy. Methods using the wrapper model require dividing all samples into two sets: a training set and a testing set. The algorithm operates on the training set, and the learning outcome is subsequently applied to the testing set to determine the prediction accuracy. However, a disadvantage of this approach is its high computational cost. Hsu and Hsieh (2010) describe feature selection via the correlation coefficient clustering approach for removing similar or redundant features, which involves utilizing correlation coefficient clustering. The features are first collected and then grouped into clusters based on their correlation coefficient values. The most class-specific feature in each cluster is kept, while the others are removed. This way, the features that are most related to the class and least related to each other can be identified.

The correlation coefficient is a statistical measure which reveals the strength and direction of a connection between two variables. Given the adverse nature of variable types (categorical, continuous) throughout the entire variable space as defined in Tables 15, 16, 17, 18, 19, 20, 21 and 22, the selection of correlation measures is important. The Spearman correlation is used to examine non-linear, monotonic relationships (i.e., relationships that consistently increase or decrease), and it does not necessitate normally distributed data (Schober et al., 2018). The Spearman correlation ranges from -1 to +1, with a value of 0 indicating no correlation, and values of -1 or +1 indicating a perfect negative or positive correlation, respectively. It is calculated by ranking the values of each variable, calculating the difference between the ranks for each pair of observations, and then calculating the correlation coefficient based on these rank differences:

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \tag{1}$$

where:
r_s is the Spearman correlation coefficient,
d_i is the difference between the ranks of the ith pair of observations for the two variables being correlated,
n is the number of pairs of observations.

While universal by nature, the Spearman rank correlation cannot be calculated for categorical dummy variable pairs; therefore, for such pairs, the phi coefficient is used (Akoglu, 2018).

$$\phi = \frac{ad - bc}{\sqrt{(a + b)(c + d)(a + c)(b + d)}} \tag{2}$$

where a, b, c, and d are the frequencies of the four possible combinations of the two binary variables.

For evaluating the relationships between variable pairs, this study shall use the phi coefficient for dummy variable pairs, whereas, for all others, the Spearman rank correlation shall be used and presented as a correlation-heat-map. This study performs feature selection by combining the correlation coefficient with clustering analysis as described in Hsu and Hsieh (2010). The Euclidean distance clustering algorithm is described in Figure 7.

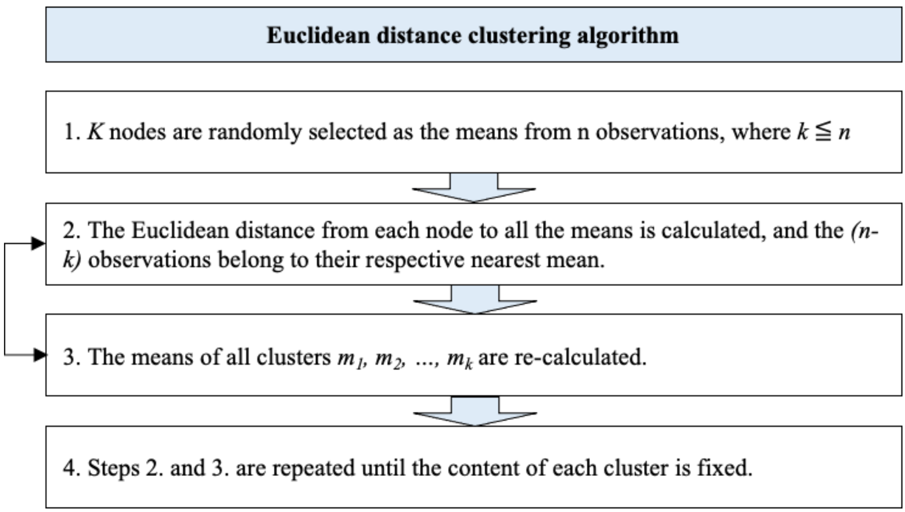


Figure 7. Euclidean distance clustering algorithm. Based on Hsu and Hsieh (2010).

To carry out the feature selection, the most relevant and non-redundant features from the original feature set are selected across different vector distance thresholds. As features in the same cluster are very close to each other, using more than two features of the same kind is not necessary to perform the classification task; on the other hand, as the removal of individual classes is not possible, categorical features are not removed. In order to account for the problem of picking the representative features for each feature cluster, it was proposed by Hsu and Hsieh (2010) to pick the most class-dependent feature by using the correlation coefficients. Ultimately, the problem of selecting the relevant clustering variables is re-framed as a model selection problem; therefore, the representative feature vector is selected based on the marginal modelling performance across different cluster thresholds (as suggested by Fop and Murphy (2018)).

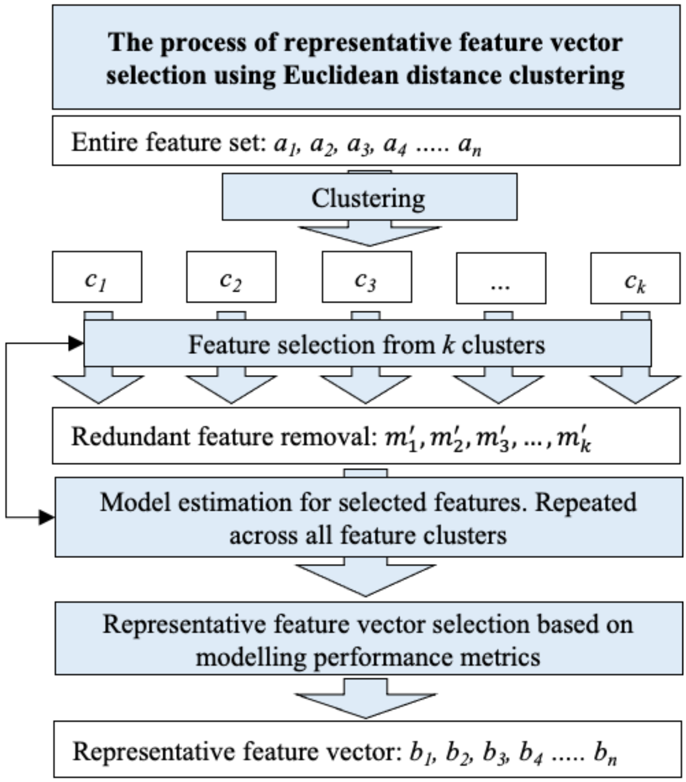


Figure 8. Algorithm for defining the representative feature vector. Based on Hsu and Hsieh (2010) and Fop and Murphy (2018).

In order to account for highly dimensional data, the dimensionality reduction process must be carried out to define the representative feature vector which will be used in SME access to credit model estimation. The process consists of calculating variable pair correlations (see Equations 1 and 2) for identifying highly correlated features and then carrying out Euclidean distance clustering (see Figure 7) which is used for selecting the representative feature vector (see Figure 8).

2.3. SME Access to Credit Modelling Techniques

To create a robust SME access to credit model it is important to select the appropriate modelling techniques and the benchmark. One way of modelling the access to credit is by utilizing a traditional modelling technique as suggested by Cox (1958), specifically, Logistic Regression (LR). While this method is frequently employed as a standard to contrast with advanced machine learning techniques, it is simplistic in design and inadequate for handling larger datasets and complex variable interrelationships (Barboza et al., 2017). Although the LR model may not achieve the same level of prediction accuracy as other machine learning models, it has a significant advantage in terms of the interpretability and stability of its variables. Therefore, for estimating the access to credit, LR will be utilized as the benchmark model to compare the modelling results with state-of-the-art machine learning techniques. Based on the previous studies by Barboza et al. (2017); Dastile et al. (2020); Trivedi (2020); Wang et al. (2020); Malakauskas and Lakštutienė (2021); Hussin Adam Khatir and Bee (2022); Medianovskyi et al. (2023), the empirical SME access to credit shall be estimated by utilizing gradient boosting (GB), random forest (RF), and multi-layer perceptron (MLP) modelling techniques, which have historically demonstrated high discriminatory power in credit accessibility related models.

Logistic regression One of the most commonly used model estimation techniques – Logistic regression (LR), is a statistical method used to analyze and model the relationship between a dependent variable (*Outcome*) and one or more independent variables (predictors) that may be continuous or categorical. It is a type of a generalized linear model that is widely used in classification tasks, such as the SME financing application outcome prediction and financial distress estimation.

As described by Peng et al. (2002) and Kirasich et al. (2018), in logistic regression, the dependent variable (Y) is modeled as a function of the independent variables (X) using a logistic or sigmoid function, which transforms the linear combination of predictors into a probability value between 0 and 1. The logistic function is defined as:

$$P(Y = 1|X) = \frac{1}{1 + e^{-z}} \tag{3}$$

where $P(Y = 1|X)$ is the probability of the dependent variable (Y) taking the value 1 (application was credit rationed) given the values of the independent variables (X), and z is the linear predictor given by:

$$z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_k \tag{4}$$

where $\beta_0, \beta_1, \beta_2, ..., \beta_k$ are the coefficients (parameters) that quantify the impact of each independent variable on the dependent variable, and $X_1, X_2, ..., X_k$ are the values of the corresponding independent variables.

$$\hat{y} = \begin{cases} 1 & \text{if } P(Y=1|X) \geq 0.5 \\ 0 & \text{if } P(Y=1|X) < 0.5 \end{cases} \tag{5}$$

where $\hat{y}$ is the predicted class label for the new observation, $P(Y=1|X)$ is the predicted probability of the dependent variable (Y) taking the value 1 given the values of the independent variables (X), and 0.5 is the threshold value for classification. If the predicted probability is greater than or equal to 0.5, the observation is classified as 1, otherwise it is classified as 0.

Random Forest Random Forest is an ensemble learning method for classification and regression which operates by constructing a multitude of decision trees during the training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees (Breiman, 2001). Tree-based learning algorithms offer several advantages for training models on large datasets, including the ability to handle both quantitative and qualitative input variables. Such models can be robust to redundant or highly correlated variables and can handle outliers or missing values. However, one potential drawback of tree-based models is that they may suffer from poor prediction performance. Decision trees, in particular, are susceptible to over-fitting noise in the training set, which results in models with high variance. Consequently, while these models may be accurate when predicting the same data they were trained on, their performance may not generalize well to datasets with different patterns and variations (Kirasich et al., 2018).

In a Random Forest, each tree is constructed by using a random subset of the training data and a random subset of the features, which ensures that the trees are diverse and not overly correlated. At the prediction time, the class predicted by each tree is obtained, and the final prediction is made by aggregating these individual predictions (see Figure 9).

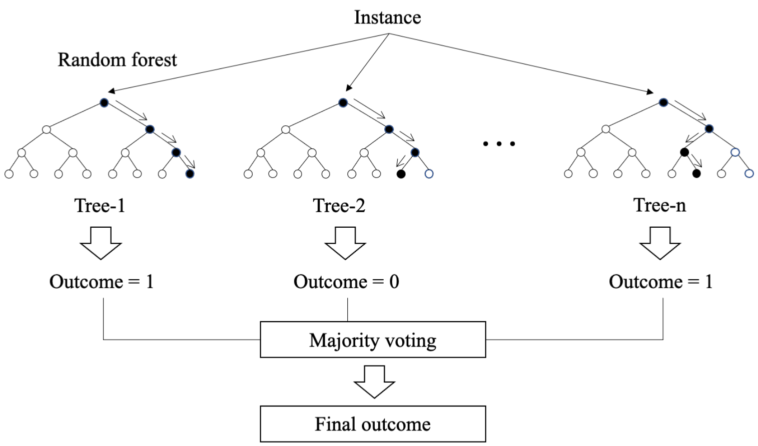


Figure 9. Random Forest algorithm. Based on Kirasich et al. (2018).

One way to determine the feature split at each node is through the computation of entropy (6). Entropy is a metric that measures the level of homogeneity in a subset of data. An entropy value of 1 indicates that the class labels are evenly divided, whereas an entropy of 0 indicates complete homogeneity in the sample.

$$Entropy = -p\log_2(p) - q\log_2(q) \tag{6}$$

where p and q are the probabilities of the two possible outcomes.

For binary classification with only two labels, an entropy of 0 would be obtained if all labels were either 1 or 0, while an entropy of 1 would be observed if a half of the labels were 1 or 0. The entropy is at its maximum when both outcomes are equally probable (i.e., $p = q = 0.5$), in which case the entropy is 1. The entropy is computed for each variable, and then the difference between the entropy before the split (i.e., the parent node) and after the split (i.e., the child node) is calculated for each variable.

Gradient boosting A novel machine learning technique – the gradient boosting technique – is utilized in the access to credit studies by (Barboza et al., 2017; Medianovskyi et al., 2023). As described by Hastie et al. (2009) and González-Recio et al. (2013), it works by improving the performance of a weak learning model $h(x)$ by iteratively adding new models $f_k(x)$ to the ensemble with weights equal to the learning rate. The final model $F(x)$ is a weighted sum of the individual models:

$$F(x) = \sum_{k=1}^{K} \alpha_k f_k(x) \tag{7}$$

where α_k are the weights of the individual models.

The algorithm starts by training a single weak learner $f_1(x)$ on the training set. In the subsequent iterations, new models $f_k(x)$ are added to the ensemble, and their predictions are combined with the previous models by using a weighted sum (see Figure 10). The weights of each model are determined by minimizing the loss function $L(y, F(x))$ which measures the discrepancy between the predicted value $F(x)$ and the actual value y.

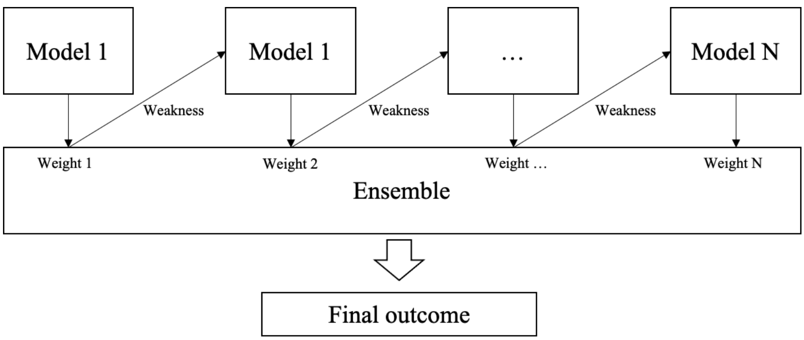


Figure 10. Gradient Boosting algorithm. Based on Kirasich et al. (2018).

The predictions of the base learners are combined by using a weighted sum, with the weights determined by the gradient descent algorithm (Friedman, 2001). One variant of gradient boosting is histogram gradient boosting (HGB), which uses histograms to estimate the features and approximates the gradient boosting algorithm (see Figure 11)

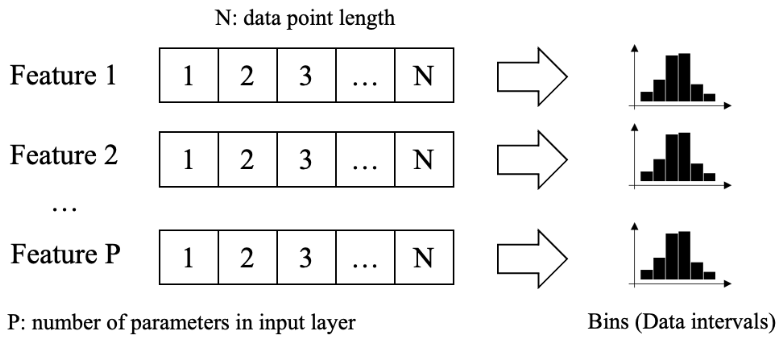


Figure 11. Histogram Gradient Boosting algorithm for feature bundling. Based on Gan et al. (2021).

HGB uses a set of decision trees with a fixed depth and bins to approximate the features (Biau et al., 2020; Gan et al., 2021). The model is trained by minimizing the binary cross-entropy loss function:

$$L(y, F(x)) = -\frac{1}{n}\sum_{i=1}^{n} y_i \log\left(\frac{\exp(F(x_i))}{1+\exp(F(x_i))}\right) + (1-y_i)\log\left(\frac{1}{1+\exp(F(x_i))}\right) \tag{8}$$

where y_i is the binary label of the ith instance, and n is the number of instances in the training set. The model is trained by minimizing this loss function while using the gradient descent with the following update rule:

$$F(x_i) \leftarrow F(x_i) - \eta \frac{\partial L(y_i, F(x_i))}{\partial F(x_i)} \tag{9}$$

where η is the learning rate which controls the step size of the gradient descent. The partial derivative of the loss function with respect to $F(x_i)$ is computed by using the following formula:

$$\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)} = -\frac{y_i - \exp(F(x_i))}{1+\exp(F(x_i))} \tag{10}$$

Multi-layer perceptron One type of the artificial neural network (ANN) which consists of multiple layers of interconnected nodes and artificial neurons is Multi-layer perceptron (MLP). It is a feed-forward neural network, which means that information flows through the network in only one direction, from the input layer to the output layer, without any feedback loops.

The MLP architecture typically consists of an input layer, one or more hidden layers, and an output layer. Each layer contains multiple neurons or processing units, and each neuron is connected to all neurons in the adjacent layers. The connections between neurons are represented by weighted edges, and each neuron applies an activation function to its inputs before passing its output to the next layer (see Figure 12).

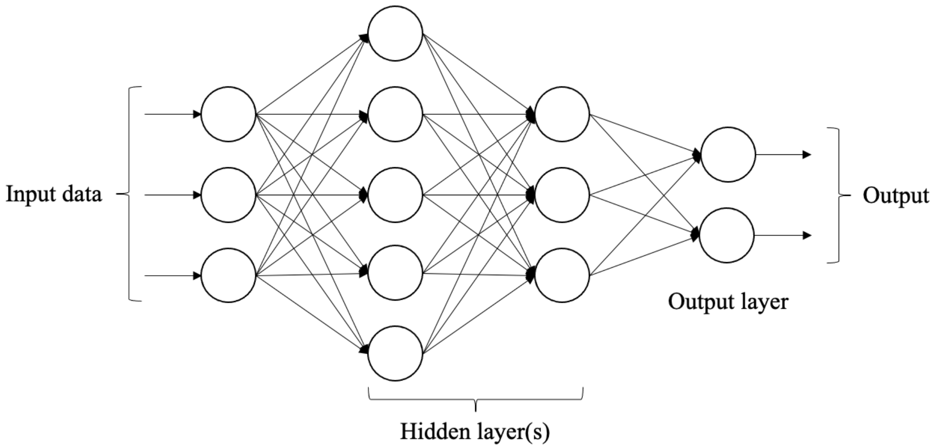


Figure 12. Multi-layer Perceptron structure. Based on Zhao et al. (2015).

During the training phase, the weights of the edges are adjusted by using an optimization algorithm, such as back-propagation, to minimize the difference between the network's output and the desired output. This process of adjusting the weights continues until the network's output becomes sufficiently accurate.

Let $\mathbf{x}$ be the input vector, $\mathbf{W}^{(l)}$ be the weights of the edges between layer $l-1$ and layer l, $\mathbf{b}^{(l)}$ be the bias vector of layer l, f be the activation function, and $\mathbf{y}$ be the output vector. The computations in each layer can be expressed as:

$$\mathbf{z}^{(l)} = \mathbf{W}^{(l)}\mathbf{a}^{(l-1)} + \mathbf{b}^{(l)}, \quad \mathbf{a}^{(l)} = f(\mathbf{z}^{(l)}) \tag{11}$$

where $\mathbf{a}^{(l)}$ is the output of layer l, and $\mathbf{z}^{(l)}$ is the weighted sum of the inputs to layer l, before applying the activation function f.

The output of the network is obtained by applying a sigmoid activation function to the output of the last layer.

2.4. Modelling Performance Evaluation Methods

To ensure that the developed SME access to credit model is robust and accurate, it is important to use the appropriate evaluation methods. The aim of model performance evaluation is to assess how well the model performs on unseen data and to identify any potential issues, such as over-fitting or under-fitting. One of the most common techniques used to evaluate model performance is data partitioning, which involves dividing a dataset into two or more subsets to train and evaluate the model (Morrison et al., 2013). Based on the findings by Gholamy et al. (2018), for the empirical SME access to credit model development, 80% of the dataset is allocated for training, and the remaining 20% is used for testing. The split is often performed randomly to ensure that the data is representative of the entire dataset and to prevent any bias from influencing the model's performance. This allows to assess the model's ability to generalize any new, unseen data. To evaluate the model performance, various metrics, such as confusion matrices, derivative metrics and graphs, are used (Kirasich et al., 2018). In this way, data partitioning and performance metrics are essential tools for evaluating the effectiveness of a machine learning model and ensuring that it meets the desired performance goals.

Confusion matrix A confusion matrix is a table that is commonly used to evaluate the performance of a supervised machine learning model, particularly for binary classification problems. The matrix contains information about the predicted and the actual classification of the data points.

In a binary classification problem, the confusion matrix has four values: true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). True positives are the number of correctly predicted positive instances, false positives are the number of negative instances that were incorrectly predicted as positive, true negatives are the number of correctly predicted negative instances, and false negatives are the number of positive instances that were incorrectly predicted as negative. The confusion matrix is typically arranged as follows:

Table 23. Confusion matrix. Created by the author.

	Predicted Positive	**Predicted Negative**
Actual Positive	True Positives (TP)	False Negatives (FN)
Actual Negative	False Positives (FP)	True Negatives (TN)

From the confusion matrix, various evaluation metrics can be calculated:

- Specificity is a measure of the proportion of actual negative instances that are correctly classified by the model as negative. The formula for specificity is:

$$Specificity = \frac{TN}{TN + FP} \tag{12}$$

- The Negative Predictive Value (NPV) is a measure of the proportion of actual negative instances among those instances that are predicted as negative by the model. The formula for NPV is:

$$NPV = \frac{TN}{TN + FN} \tag{13}$$

- Precision measures the proportion of correctly predicted positive instances out of all predicted positive instances. The formula for Precision is:

$$Precision = \frac{TP}{TP + FP} \tag{14}$$

- Sensitivity (a.k.a. Recall, True Positive Rate, or TPR) measures the proportion of correctly predicted positive instances out of all actual positive instances. The formula for Sensitivity is:

$$Sensitivity = \frac{TP}{TP + FN} \tag{15}$$

- The False Positive Rate (FPR) is a measure of the proportion of actual negative instances that are incorrectly classified as positive by the model. The formula for FPR is:

$$FPR = \frac{FP}{TN + FP} \tag{16}$$

- Accuracy measures the proportion of correctly classified instances. The formula for Accuracy is:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{17}$$

- The F1-score is a combination of precision and Sensitivity that provides an overall measure of the model's performance. The formula for F1-score is:

$$F1 = 2 \cdot \frac{Precision \cdot Sensitivity}{Precision + Sensitivity} \tag{18}$$

ROC and AUC One of the most commonly used measures to evaluate the performance of binary classification models includes the Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC) (Bradley, 1997; Wang et al., 2011, 2020; Hussin Adam Khatir and Bee, 2022; Medianovskyi et al., 2023). The ROC curve is a plot of Sensitivity (15) to FPR (16) across all threshold values. The ROC curve is created by varying the threshold for predicting positive cases and plotting the Sensitivity against FPR. A good classifier will have a ROC curve that is close to the top left corner, where Sensitivity is high and FPR is low (see Figure 13).

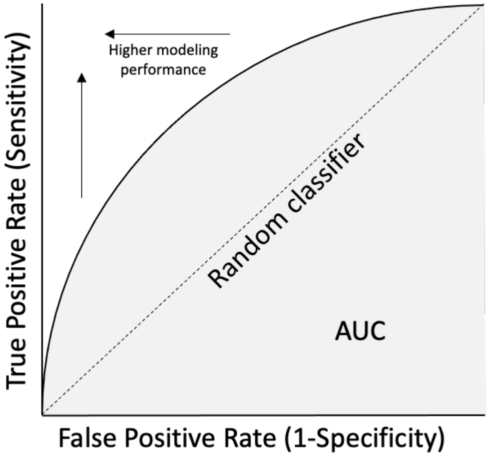


Figure 13. Receiver Operating Characteristic curve. Based on Bradley (1997).

AUC is a single scalar value which measures the area under the ROC curve. AUC ranges from 0 to 1, with a higher value indicating a better performance. An AUC of 0.5 indicates that the model performs no better than random guessing, while an AUC of 1 indicates the perfect classification performance.

The formula for AUC is as follows:

$$AUC = \int_0^1 TPR(FPR^{-1}(t))\mathrm{d}t \tag{19}$$

where $FPR^{-1}(t)$ is the inverse of the FPR function with respect to *t*, which represents the threshold value for separating the positive and negative samples.

While ROC-AUC is most suitable for evaluating the performance of a classifier when the class distribution is balanced or when the cost of false positives and false negatives is roughly equal, it is not the best measure for the cases where the class distribution is imbalance or when the cost of False Positives and False Negatives is significantly different.

Precision-Recall A particularly useful measure in evaluating the performance of classifiers in situations where the classes are imbalanced is Precisions-Recall and Average Precision (AP) (Boyd et al., 2013). AP is a performance metric used in information retrieval and binary classification problems. AP summarizes the precision-recall curve and calculates the average precision for a given set of recall levels (see Figure 14)

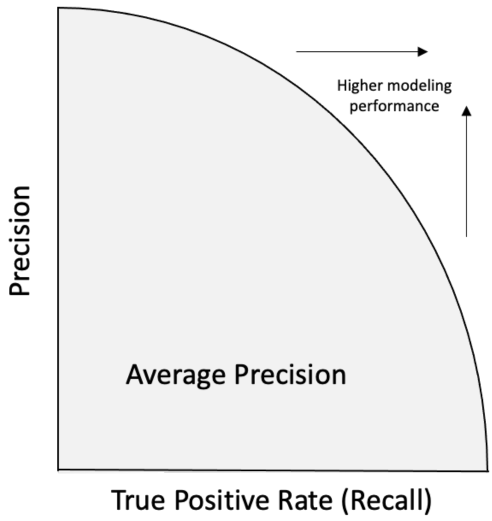


Figure 14. Precision-Recall curve. Based on Boyd et al. (2013).

To calculate AP, first, the precision-recall curve is generated by varying the threshold of the classifier and calculating the precision and recall for each threshold. Then, the precision values are interpolated for each recall level, and the area under the interpolated precision-recall curve is calculated. Finally, the area is normalized by dividing it by the total area under the perfect recall curve (i.e., the area where recall equals 1.0).

The formula for calculating AP is:

$$AP = \sum_{n=1}^{N} (R_n - R_{n-1}) P_n \tag{20}$$

where N is the total number of distinct recall values, R_n is the recall value at the nth threshold, R_{n-1} is the recall value at the $(n-1)$th threshold, and P_n is the maximum precision value obtained for any threshold greater than or equal to R_n.

2.5. Feature Explainability Methods

Feature explainability is an important aspect of machine learning models, as it helps to understand how the model makes its predictions. It provides insights into the importance of different features in the model, and how they contribute to the final prediction (Chen and Bharodia, 2019). In terms of the access to credit, it provides insights towards underlying conditions which determine the actual ability to access credit. There are several methods collected under the XAI topic, available for feature explainability, each with its own strengths and weaknesses (Arya et al., 2019). The choice of the appropriate method depends on the particular use case and the desired level of interpretability. Modern XAI frameworks together with ML models should be used to analyze an importance drift of the factors which affect the access to credit for

SME companies. The approach to using explanations, and, specifically, the attribution of the importance of characteristics, to analyze changes in the data is relatively new, but it has already been presented in the literature by Duckworth et al. (2021); Seiffer and Gerling (2021); Saarela and Jauhiainen (2021). In the context of finance-related models, Chen and Bharodia (2019) explores interpretations of the credit risk model, whereas Castelnovo et al. (2021) address the problem of the data drift for the Loan Assessment model. The empirical model shall utilise the following feature explainability techniques:

SHAP SHapley Additive exPlanations (SHAP) is a method for explaining the predictions of machine learning models. It provides a way to assign an importance score to each feature in a prediction, thereby indicating how much that feature contributed to the final prediction (Lundberg and Lee, 2017). The SHAP approach is model-agnostic, which means that it can be applied to any machine learning model, and that it considers all possible combinations of features (Sundararajan et al., 2017; Lundberg et al., 2020). It is based on the concept of Shapley values which come from the cooperative game theory and measure the contribution of each player in a game to a particular outcome.

To calculate the mean absolute SHAP values for a particular prediction, the SHAP method first creates a reference dataset of similar instances. For each feature, it computes the difference between the contribution of that feature in the prediction and its contribution in the reference dataset. This difference is multiplied by a weight that reflects the number of possible feature combinations which include that feature. The SHAP values for all features are then summed up to get the final importance score for each feature. The SHAP method can be represented by the following equations:

Let $f(x)$ be the model's prediction for input x, and let S be a subset of features. Then, the SHAP value for feature i is defined as:

$$\phi_i(x) = \sum_{S \subseteq 1,2,\ldots,p \setminus i} \frac{|S|!(p - |S| - 1)!}{p!} [f(x_S \cup i) - f(x_S)] \tag{21}$$

where p is the number of features, and x_S is the instance with all features in S set to their reference values. This formula calculates the difference in the model's prediction when feature i is included in the input, compared to when it is excluded, for all possible subsets of features. The weight of each term reflects the number of possible feature combinations which include feature i.

The final SHAP value for feature i is obtained by averaging the $\phi_i(x)$ values across many instances:

$$SHAP_i = \frac{1}{N} \sum_{j=1}^{N} \phi_i(x_j) \tag{22}$$

where N is the number of instances.

The SHAP values can be visualized in a SHAP plot, which shows the contribution of each feature to the final prediction for a particular instance. Such a plot is necessary to understand how the model arrived at its decision, and which features were the most important in making that particular decision.

SHAP values are also used to plot SHAP dependence plots showing the relationship between a feature and the predicted output of a machine learning model, while considering the impact of other features (Lundberg et al., 2020). Such visualization helps to identify non-linear relationships between features and the model prediction, and to detect interactions between features. They are useful for understanding how the model makes predictions and how different features interact with each other. The SHAP dependence plot shows the values of the feature on the x-axis and the corresponding SHAP values on the y-axis. Each point on the plot represents a specific instance in the dataset. The color of the point represents the value of a second feature, which can be selected by the user. The plot shows the relationship between the selected feature and the model prediction, while adjusting for the impact of the second feature.

Partial dependence plots A supplementary measure to SHAP dependence plots for evaluating feature interactions with the model is Partial Dependence Plots (PDPs). PDPs are a tool for visualizing and interpreting the relationship between the dependent and the independent variables in a machine learning model (Goldstein et al., 2015). In essence, a partial dependence plot shows the marginal effect of a predictor variable on the predicted outcome of a model, while holding all other predictors constant. They provide a clear visualization of how the model's predicted outcome is changing as the value of the predictor variable is changing simultaneously. PDPs are used to identify important predictors and non-linear relationships between the predictors and the predicted outcome.

$$\mathrm{PDP}(x) = \frac{1}{n}\sum_{i=1}^{n} \hat{f}(x_1, x_2, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n) \tag{23}$$

where $\hat{f}$ is the fitted model, n is the number of observations in the dataset, and $x_1, x_2, \ldots, x_n$ are the values of the other predictor variables in the dataset.

Intuitively, the PDP for a predictor variable shows how the model's predicted outcome is changing as the value of that predictor variable is also changing, while holding all other predictors constant.

Permutation feature importance Permutation Feature Importance (PFI) is a method for assessing the importance of features in a machine learning model by measuring the decrease in the model performance when a feature's values are randomly shuffled. The algorithm states that important features will have a larger impact on the model's performance, and thus their shuffling will result in a greater decrease in terms of performance. Formally, PFI can be described as:

Let X be the input matrix with shape (n, d), where n is the number of instances and d is the number of features, and let y be the target vector with shape $(n,)$. Let $f(\cdot)$ be the trained model which maps inputs to predictions, and let $S \subset 1, 2, \ldots, d$ be the set of features to be evaluated for importance. The permutation feature importance FI_j for feature $j \in S$ is defined as the decrease in the model performance when the values of feature j are randomly shuffled across all instances:

$$PFI_j = \frac{1}{n_1}\sum_{i=1}^{n}(y_i - f(X_i))^2 - \frac{1}{n_1}\sum_{i=1}^{n}(y_i - f(\pi_j(X_i)))^2, \tag{24}$$

where $\pi_j(\cdot)$ is a random permutation of the values of feature j.

To estimate the permutation feature importance of all features in S, the above formula can be computed for each feature $j \in S$. A larger value of FI_j indicates that the model's performance is more sensitive to changes in feature j, and thus it is more important.

Variable importance grouping Variable grouping is a technique used in statistical modelling and data analysis to reduce the complexity of high-dimensional datasets (Gregorutti et al., 2015). It involves combining similar or related variables into groups based on their shared characteristics. For this study, the selected features shall be grouped into pre-defined variables groups – Firm Characteristics, Product Characteristics, Lending Technology. This process can help to simplify data analysis, reduce the computational burden, and improve the interpretability of statistical models by reducing the number of the features being considered. Variable importance grouping is a technique which involves grouping related variables and computing their collective importance to a model. This can be achieved by summing up the SHAP and PFI values of individual features into pre-defined groups (Au et al., 2022). The formula for grouping mean absolute SHAP and PFI values for a group of features can be expressed as follows:

$$GroupedSHAP_g = \sum j \in gSHAP_j \tag{25}$$

$GroupedSHAP_g$ represents the collective SHAP value for group g of related variables. The sum of $SHAP_j$ overall j in g calculates the total SHAP value for group g.

$$GroupedPFI_g = \sum j \in gPFI_j \tag{26}$$

In this equation, $GroupedPFI_g$ represents the collective PFI value for group g of related variables. The sum of PFI_j overall j in g calculates the total PFI value for group g. By grouping the variable importance into pre-defined groups, we shall define what variable groups are the most important for SMEs when trying to access to credit.

Summary and findings

The methodology for creating an empirical model to evaluate the SME access to credit has been created. It has been determined that the SME access to credit (the dependent variable) is represented as the outcome of the submitted financing application (*Outcome*). The independent variables that will be used to model *Outcome* are grouped into Firm characteristics, Product characteristics, and Lending technology variable groups. To carry out the evaluation of the underlying SME access to credit, first, a comparative analysis for the studied countries must be conducted. Next, a representative feature vector should be defined by carrying out the dimensionality reduction process. Finally, the machine-learning model should be estimated by utilizing state-of-the-art machine learning techniques. The selected techniques are Random Forest, Gradient Boosting, Multi-layer Perceptron, and bench-marked to Logistic Regression. The best performing modelling technique is utilized for the further model dissemination for individual variable importance and interaction evaluation (see Figure 15).

The model for evaluating the SME access to credit consists of three stages. Stage I involves the execution of comparative analysis to evaluate the underlying access to credit in each specific country. In Stage II, a process of dimensionality reduction is implemented to account for variable correlations and determine representative feature vectors for each country. Stage III is conducted to develop an actual SME access to credit model, which would allow for the assessment of the importance and significance of the variables and their interactions.

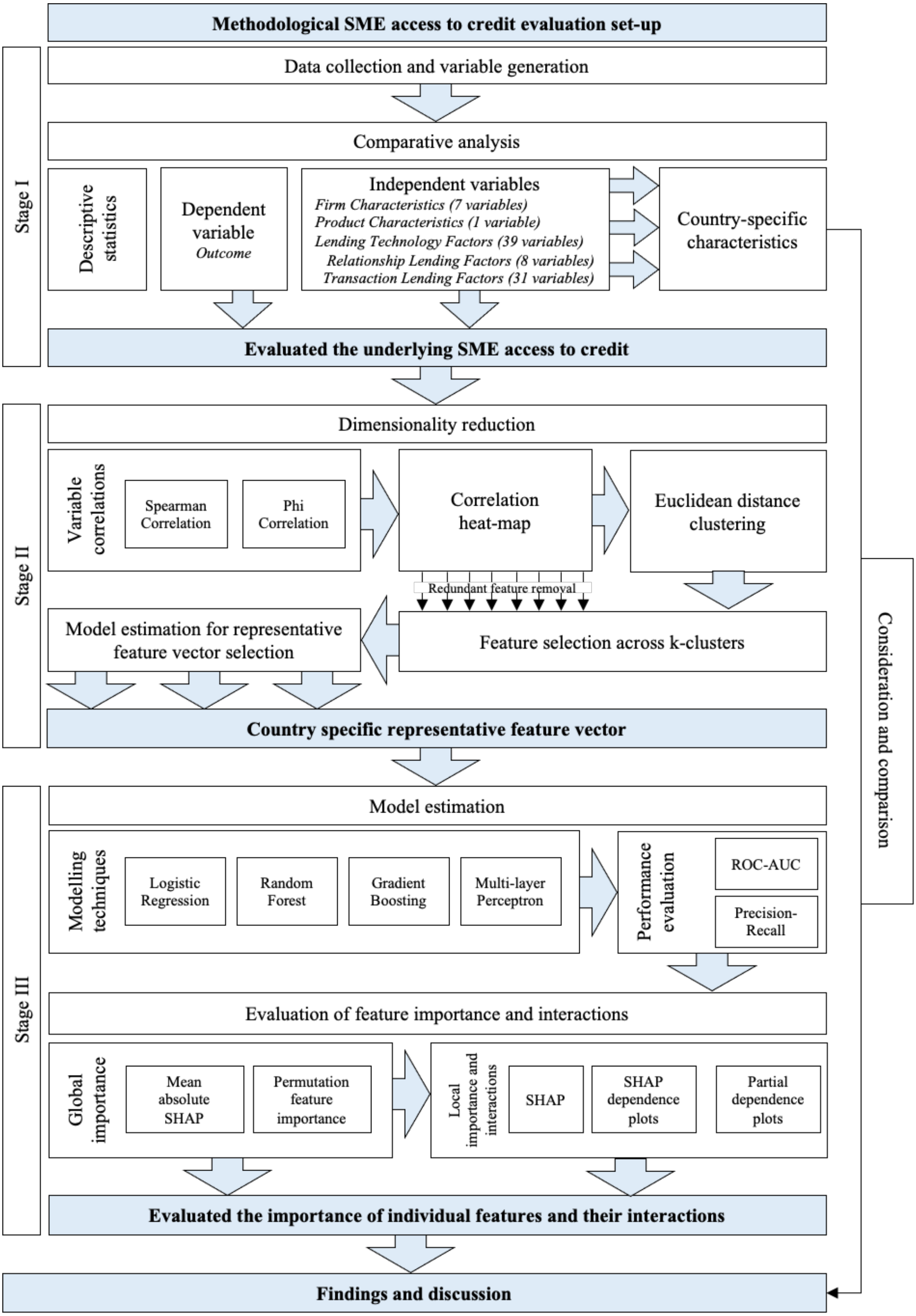


Figure 15. Model for evaluating SME access to credit. Created by the author.

3. EMPIRICAL MODEL FOR EVALUATING SME ACCESS TO CREDIT

In the third section, the dissertation solves objective 5. The SME access to credit is evaluated empirically in a country-specific setting. In Stage I, a comparative analysis is carried out to define the underlying access to credit in each individual country. In Stage II, the dimensionality reduction process is carried out to define country-specific representative feature vectors. In Stage III, an SME access to credit model is created, and the variable and interaction importance is evaluated. The findings and summary are presented at the end of the section.

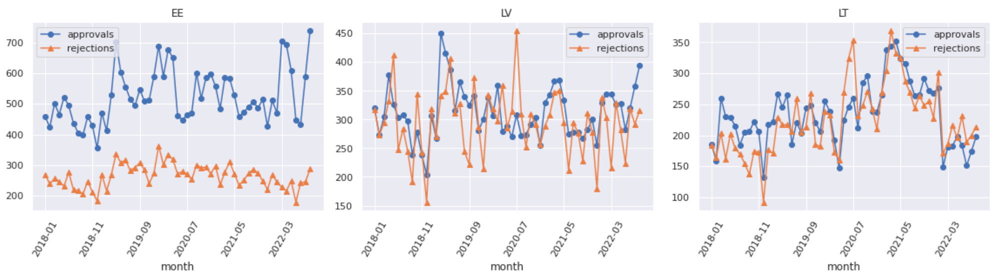
3.1. Comparative Analysis for Evaluating the Underlying SME Access to Credit

The comparative analysis is carried out to uncover and compare between countries the underlying SME access to credit and the underlying factor values. This is the first stage for the empirical SME access to credit evaluation, as described in Section Tables 24, 26, 25, 27, 28, 29, 30, 31, 32, 33 and 34: the total number of observations (Count), the average value (Mean); for continuous and dummy variables: the standard deviation (STD), the minimum value (Min), the median value (Median) and the maximum value (Max). The data set is composed of financing applications received by a financial institution between the beginning of 2018 and the end of 2022. It includes information on the application's outcome (*Outcome*) and the underlying conditions: Firm Characteristics (*Age*, *Diversity*, *Private*, *Segment*, *Type*, *Region*, *Sector*, *Audited*), Product Characteristics (*Product*), Lending Technology factors, consisting of the Relationship Lending factor group (*Relationship*, *Payments*, *Rejections*, *Debt*, *FinContracts*, *Cards*, *POS*, *Ecommerce*), and the Transaction Lending factor group which is further grouped into financial statement-based factors based on Liquidity (*CR*, *QR*, *CuR*, *pCR*, *pQR*, *pCuR*), Solvency (*DE*, *TA*, *DR*, *DSCR*, *ACR*, *pDE*, *pTA*, *pDR*, *pDSCR*, *pACR*), Profitability (*ROA*, *ROE*, *GMR*, *PMR*, *pROA*, *pROE*, *pGMR*, *pPMR*), Activity (*ATR*, *RTR*, *CS*, *CiA*, *pATR*, *pRTR*), and the credit history based factors for the company (*IOverC*, *IOverS*, *IOverL*, *EOverC*, *EOverS*, *EOverL*) and the owner (*OIOverC*, *OIOverS*, *OIOverL*, *OEOverC*, *OEOverS*, *OEOverL*, *Defaults*, *ODefaults*). The total number of the retrieved records is close to 300 000. By selecting only applications received from SMEs, the data is refined to 120 000 records (29 000 in LT, 39 000 in LV and 51 000 in EE) which is used to estimate country-specific models. Each application is classified into two possible outcome values: no rationing – when financing was issued, and rationing – when financing was not issued. *Outcome* is the dependent variable used in the model and is encoded as a binary value of 1 when rationed and 0 when approved (see Table 24).

Table 24. Descriptive statistics of the dependent variable for evaluating access to credit. Created by the author.

Variable	Country	Count	Mean	STD	Min	Median	Max
Outcome							
	Estonia	50 998	0.343	0.475	0	0	1
	Latvia	38 924	0.487	0.500	0	0	1
	Lithuania	28 917	0.488	0.500	0	0	1

Throughout the observed period, most financing applications were received in EE – 50 998, followed by LV – 38 924 and LT – 28 917. The mean *Outcome* values suggest that the observed SME access to credit in EE is significantly different in comparison to LV and LT, with an average rejection rate at 34%. On average, SMEs that submit a financing application in LV or LT are 14 p.p. less likely to be approved than their EE counterparts. It is evident that, for the given dataset, the observed rejection rate (*Outcome*) is heterogeneous across the three studied countries. These findings show that, for an average SME application, the access to credit is higher in EE and lower in LV and LT. To evaluate and compare the incoming application flow and the actual outcomes across countries, Figure 16 is plotted.



(a) EE (b) LV (c) LT

Figure 16. The total number of approved and rationed applications throughout the observed period. Created by the author.

Figure 16 demonstrates the count of retrieved applications per month distributed over the observed period and grouped by the outcome. The total number of incoming applications per month varies between countries, with the biggest flow being in EE between 700 and 1,000, followed by LV between 350 and 700, and LT between 250 and 700 applications per month. In EE and LT, there was an upward facing trend for the total number of incoming applications throughout the period, whilst for LV the number is constant. Due to the relatively similar variance of the total application number across the studied countries, it is suggested that common market trends exist. To determine whether the underlying access to credit is uniform throughout different periods of time, the rejection ratio between the rejected applications and the total applications is plotted (see Figure 17).

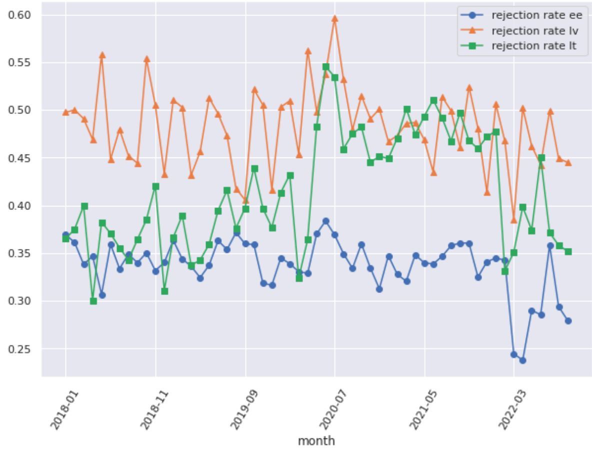


Figure 17. Rejection rate as observed throughout the studied period across countries. Created by the author.

Figure 17 shows that the rejection rate is not uniform across all three Baltic States. The rejection rate throughout different time intervals varies both between countries and specific periods. In LT, out of the three observed countries, the range is the widest – between 30 and 55% (EE – between 25 and 38%, LV – between 40 and 60%). In LV, uniformly throughout the whole period, an average SME requesting financing has a significantly harder time to access credit than an SME in EE. Meanwhile, in both of these countries, the series is relatively stationary. In LT, there were significant shifts in the underlying access to credit between the end of 2019 and the beginning of 2020 when the average rejection rate shifted from 35–40% to 45–50%. In the observed timeline, two periods of interest can be identified (the rejection rate shifted by more than two standard deviations for more than one month). First, the beginning of COVID-19 (between March 2020 and August 2020) when the rejection rate grew significantly in LV (to 60%) and LT (to 55%), and only slightly in EE (to 39%). Second, in February and March of 2022, aligning with the Ukrainian War, when the rejection rate decreased to an all-time low for all three countries (EE – <25%, LV – <40%, LT – <25%). It is not clear whether the shift in the SME access to credit was triggered by a change in the risk appetite on the supply side or a deterioration in the average quality of the submitted applications on the demand side, yet the absolute demand in terms of the number of financing requests did not change (see Figure 16). It can be concluded that, for the evaluated countries, the SME access to credit is not uniform throughout different time periods and can go through substantial positive and negative shifts in the rejection rate for extended periods of time. To determine how the underlying factor composition differs across the studied countries factor groups are individually analyzed starting with the Firm Characteristics factor group (see Tables 25 and 26).

Table 25. Descriptive statistics of continuous and dummy variables belonging to the Firm Characteristics factor group. Created by the author.

Variable	Count	Mean	SD	Min	Median	Max
Estonia						
Age	50 998	12.4	6.5	0	11.6	41.3
Diversity	50 998	0.198	0.399	0	0	1
Private	50 998	0.87	0.31	0	1	1
Audited	50 998	0.172	0.377	0	0	1
Latvia						
Age	38 924	13.9	8.0	0	12.6	31.5
Diversity	38 924	0.217	0.412	0	0	1
Private	38 924	0.90	0.28	0	1	1
Audited	39 924	0.332	0.471	0	0	1
Lithuania						
Age	28 917	13.7	8.2	0	12.2	68.7
Diversity	28 917	0.243	0.429	0	0	1
Private	28 917	0.86	0.33	0	1	1
Audited	28 917	0.177	0.382	0	0	1

Table 25 presents the descriptive statistics for three continuous variables *Age*, *Diversity*, *Private* and one dummy variable *Audited* representing the Firm Characteristics factor group. The mean and standard deviation of *Age* across the three countries suggests that the majority of SMEs applying for financing are relatively young, with a mean age ranging from 12.4 years in EE to 13.9 years in LV (LT – 13.7). However, there is a high standard deviation for *Age* in all the three countries, which indicates that there is a considerable amount of variation in the ages of SMEs applying for financing (with the youngest being a few days old to 68-year-old businesses). The mean and standard deviation of the variable *Diversity* suggests that female ownership in SMEs is relatively low in all three countries, with a mean proportion ranging from 20% in EE to 24% in Lithuania. It is worth noting that the proportion of female-owned companies varies considerably within each country, as indicated by the high standard deviation. The distribution of natural ownership (*Private*) in companies is relatively uniform across the three countries – ranging from 86% in LT, 87% in EE, and 90% in LV. The proportion of applications from companies providing audited financial statements ranges between 17%–18% in EE and LT to 33% in LV, which suggests that the proportion of applications received from larger companies is relatively bigger in LV. These findings are also substantiated by the fact that 1/3 of the received applications in LV were received from SMEs with audited financial statements (*Audited*), which indicates that companies in LV are larger and thus are prone to lower information asymmetry than in EE and LT, where the share of audited financial statements is 17%. Table 26 provides further insights to categorical variables belonging to the Firm Characteristics factor group.

Table 26. Descriptive statistics of categorical variables belonging to the Firm Characteristics factor group. Created by the author.

	Estonia		Latvia		Lithuania	
Variable	**Count**	**Comp.**	**Count**	**Comp.**	**Count**	**Comp.**
Segment						
Micro	37 368	0.733	22 248	0.572	13 342	0.461
Small	11 844	0.232	13 062	0.336	10 668	0.369
Medium	1 786	0.035	3 614	0.093	4 907	0.170
Type						
Unlimited liability	-	0	157	0.004	956	0.033
Partnership	88	0.002	34	0.001	2 025	0.070
Private limited liability	50 910	0.998	38 733	0.995	25 936	0.897
Region						
Predominantly rural	13 477	0.264	11 084	0.285	5 413	0.187
Intermediate	17 025	0.334	9 133	0.235	9 066	0.314
Predominantly urban	20 496	0.402	18 707	0.481	14 438	0.499
Sector						
Commerce	10 369	0.203	7 616	0.196	8 504	0.294
Construction	9 775	0.192	2 988	0.077	3 396	0.117
Hotels & Restaurants	1 599	0.031	476	0.012	551	0.019
Agriculture, Forestry, Fishing	3 471	0.068	5 716	0.147	904	0.031
Real Estate	2 707	0.053	1 402	0.036	973	0.034
Transportation	4 194	0.082	2 957	0.076	4 716	0.163
IT & Communication	1 411	0.028	1 452	0.037	1 114	0.039
Professional Services	6 987	0.137	4 820	0.124	4 212	0.146
Not provided	4 232	0.083	5 773	0.148	1 015	0.035
Manufacturing	6 253	0.123	5 724	0.147	3 532	0.122

Table 26 provides descriptive statistics for the *Segment*, *Type*, *Region*, *Sector* categorical variables belonging to the Firm Characteristics factor group. In terms of the company size segmentation *Segment*, micro-enterprises constitute the majority of applications in all the three countries, with EE having the highest proportion of micro-enterprises at 73%, followed by LV at 57% and 46% in LT. Medium-sized firms have the lowest representation in all the three countries. Still, the proportion of larger company applications differs significantly across the three countries as Small and Medium applications account for 54% of all applications in LT, while only accounting for 27% in EE (with LV standing at 43%). This indicates that an average application is received from a larger business in LT and smaller companies in EE. If considering the legal entity type *Type*, the majority of firms in all the three countries are private limited liability companies, with Estonia having the highest proportion at 99.8%. Unlimited liability companies and partnerships have a very low representation in all the three countries. Notably, LT has relatively more diverse applications in terms of company types as 3% are Unlimited liability companies and 7% Partnerships. In terms of application distribution by *Region*, the majority of applications across all the three countries come from

Predominantly urban regions (EE – 40%, LV – 48%, LT – 50%). Across all the three countries, Predominantly rural regions correspond to the lowest proportion of applications in LT (18.7%), while in LV it is second most common case (28.4%). Even though the extent of individual differences exists between the *Region* categories in EE and LT, the two countries are relatively similar. In terms of the economic sector *Sector*, across all the three countries, the largest proportion of applications are from companies within the Ecommerce sector, which accounts for 20%, of applications in EE and LV, while in LT it covers 29%. Construction is the second most common economic sector in EE at 19% yet it is only the fifth in LV and LT (8% and 12% accordingly). The Hotels and Restaurants sector across all the three countries accounted for the lowest proportion of applications. Notably, a significant proportion of applicants (EE at 8%, LV at 15%, LT at 4%) did not provide their sector to the official company registry. It is determined that *Age*, *Diversity*, *Private* variables are all uniformly distributed across the countries, while significant differences exist between the three markets when it comes to the *Segment*, *Type*, *Region*, *Sector* variables. Next, the variable belonging to the Product Characteristics factor group is analyzed to determine the distribution of the incoming applications in terms of the product requested (see Table 27).

Table 27. Descriptive statistics of variable belonging to the Product Characteristic factor group. Created by the author.

	Estonia		Latvia		Lithuania	
Variable	**Count**	**Comp.**	**Count**	**Comp.**	**Count**	**Comp.**
Product						
Credit Card	11 535	0.226	2 853	0.073	4 132	0.143
Asset-based loans	7 402	0.145	2 140	0.055	1 054	0.036
Leasing	4 245	0.083	2 147	0.055	16 384	0.567
Trade Finance	6 922	0.136	3 243	0.083	1 825	0.063
Cash-flow loans	20 894	0.410	28 541	0.733	5 522	0.191

Table 27 provides descriptive statistics for the *Product* variable, which provides insights on the proportion of the requested products by country. It is evident that the composition of product applications is significantly different across the countries. The most commonly applied *Product* in EE and LV is Cash-flow loans constituting 41% and 73% of all applications. In LT, Cash-flow loan applications make up only 19% of the total flow, while the most commonly applied for is Leasing at 57% (which is the least common product type in EE and LV, at 8% and 6%, respectively). Credit Card is the second most requested product type in EE (23%) and third in LV (7%) and LT (14%). Comparative analysis of the descriptive statistics for the *Product* variable indicates heterogeneity between countries, which indicates a significant difference in the demanded products between the observed countries. To determine the underlying set-up of the Lending Technology factors, the Relationship Lending and Transaction Lending factor group variables are analyzed (see Tables 28, 29, 30, 31, 32, 33 and 34.

Table 28. Descriptive statistics of variables belonging to the Relationship Lending factor group. Created by the author.

Name	Count	Mean	SD	Min	Median	Max
Estonia						
Relationship	50 998	11.4	6.0	0	11	23
Payments	50 998	1.544	16	0	1.148	2467
Rejections	50 998	0.093	0.291	0	0	1
Debt	50 998	0.559	4.370	0	0.184	383
FinContracts	50 998	2.5	4.4	0	1	108
Cards	50 998	0.811	0.391	0	1	1
POS	50 998	0.171	0.377	0	0	1
Ecommerce	50 998	0.044	0.204	0	0	1
Latvia						
Relationship	38 924	10.5	5.4	0	10	24
Payments	38 924	1.445	12	0	1.096	1338
Rejections	38 924	0.140	0.347	0	0	1
Debt	38 924	0.251	0.788	0	0.084	65
FinContracts	38 924	2.8	6.5	0	1	226
Cards	38 924	0.817	0.387	0	1	1
POS	38 924	0.162	0.368	0	0	1
Ecommerce	38 924	0.031	0.173	0	0	1
Lithuania						
Relationship	28 917	10.0	5.5	0	10	21
Payments	28 917	1.222	7	0	1.064	872
Rejections	28 917	0.058	0.233	0	0	1
Debt	28 917	0.216	1.268	0	0.025	125
FinContracts	28 917	3.0	6.4	0	1	100
Cards	28 917	0.702	0.457	0	1	1
POS	28 917	0.134	0.340	0	0	1
Ecommerce	28 917	0.021	0.142	0	0	1

The Lending Technology factors are the most variable-rich factor group that has an impact on the SME access to credit, whereas, a part of it – the Relationship Lending factor group consisting of *Relationship*, *Payments*, *Rejections*, *Debt*, *FinContracts*, *Cards*, *POS* and *Ecommerce* variables – cover the pre-existing relationship between the bank and the SME. As evidently shown in Table 28, across all the three countries, the absolute majority of financing applications were received from SMEs with a pre-existing relationship with the bank. The length of the company-bank relationship (*Relationship*), at the time of application, was the highest in EE at 11.4 years, followed by LV at 10.5 years and the lowest in LT at 10 years. The average relationship lengths are not only long but intense as, for an average applicant, the average number of financial contracts (*FinContracts*) is ca. 2.5 in EE and 3 in LT (and 2.8 in LV). Only a minority of the applying SMEs had previous applications rejected, most being in LV at 14%, and the least in LT at 6% (with EE at 9%). The mean *Payments* represents the ratio between the incoming payments and sales, which indicates whether the applying company is actively using bank payment services since the higher is the ratio, the bigger proportion of sales is going through bank accounts. The biggest proportion of payments to sales is in EE (1.5), followed by LV at 1.4 and LT at 1.2. Even though the difference between the countries is substantial, the overall intensity of the relationship is high as the ratio is above 1. The findings for high relationship intensity across the countries is also substantiated by other product usage. More than 81% of the applying companies in LV and EE (with LT at 70%) had a debit card. Furthermore, 17% of the applying SMEs in EE also had a payment collection product (*POS*), followed by 16% in LV and 13% in LT. 4% of EE applicants used e-commerce solution (*Ecommerce*), while in LV and LT, the proportion was lower at 3% and 2%, accordingly. Finally, the average *Debt* in LV and LT was at 22–25%, while in EE it is much higher – at 60%, which shows a significantly higher proportion of financing provided by the bank to the average applicant in EE. It is evident that the length and intensity of the bank-firm relationship is not equal across countries and is strongest in EE and slightly weaker in LV and LT. Companies that are applying for financing in EE, on average, have longer relationships, are holding more products, making more payments, and have received less rejections in the past. Both in LV and LT, bank-firm relationships for the applying customers are slightly weaker, but still very strong. The most significant difference between the countries has been determined to be *Debt*. Next, the composition of the Transaction Lending factor group variables is analyzed throughout the different financial ratio (see Tables 29, 30, 31 and 32) and credit history (see Tables 33 and 34) variable groups.

Table 29. Descriptive statistics of Liquidity variables belonging to the Transaction Lending factor group. Created by the author.

Name	Count	Mean	SD	Min	Median	Max
Estonia						
CR	48 505	2.168	10.767	0	0.362	354
QR	50 998	4.262	16.454	0	1.301	567
CuR	50 998	5.187	18.355	0	1.828	721
pCR	48 387	2.283	11.127	0	0.366	353
pQR	50 998	4.447	17.552	0	1.299	587
pCuR	50 998	5.353	19.729	0	1.796	642
Latvia						
CR	38 099	1.243	7.400	0	0.193	365
QR	38 924	2.418	10.548	0	0.904	463
CuR	38 924	3.496	13.495	0	1.461	539
pCR	38 147	1.202	7.313	0	0.181	258
pQR	38 924	2.386	10.954	0	0.887	492
pCuR	38 924	3.527	15.550	0	1.407	633
Lithuania						
CR	26 562	1.469	6.423	0	0.301	275
QR	28 917	3.065	9.988	0	1.250	499
CuR	28 917	4.101	12.329	0	1.843	697
pCR	26 846	1.510	7.180	0	0.275	313
pQR	28 914	3.110	10.196	0	1.207	330
pCuR	28 914	4.159	12.869	0	1.799	521

Table 29 provides the descriptive statistics for the financial statement Liquidity variables belonging to the Transaction Lending factor group for the current (*CR*, *QR*, *CuR*) and previous (*pCR*, *pQR*, *pCuR*) financial reporting periods. The liquidity of an average SME entity's application is the highest in EE (*CuR* at 5.19, *QR* at 4.26, *CR* at 2.17) and the lowest in LV (*CuR* at 3.5, *QR* at 2.4, *CR* at 1.2). A similar takeaway is also evident for the median values, which helps to diffuse the effect of outliers on the mean values. The median value for *CuR* is the highest in LT. By comparing *CuR*, *QR*, *CR* with *pCuR*, *pQR*, *pCR*, it is determined that, across all the three countries, there have not been any significant changes in the Liquidity variables between the past and the current periods. It is important to note that, based on the theoretically acceptable Liquidity variable thresholds in Table 18 (CR – >0.5; QR – >1.0; CuR – >1.5), some applications do not meet the minimal thresholds to receive financing. By considering that financing is issued based on a multitude of factors, the adverse effect of sub-par liquidity levels can potentially be compensated by other factors. To analyze the underlying ability to meet financial obligations across the retrieved financing applications, Solvency variables are evaluated (see Table 30).

Table 30. Descriptive statistics of Solvency variables belonging to the Transaction Lending factor group. Created by the author.

Name	Count	Mean	SD	Min	Median	Max
Estonia						
DE	48 505	2.168	10.767	0.000	0.362	354
TA	43 042	0.335	0.284	0.000	0.262	1
DR	50 998	0.470	0.327	0.001	0.443	12
DSCR	50 998	1.264	6.726	-18.240	0.270	224
ACR	50 998	6.095	21.934	0.016	2.234	937
pDE	50 998	2.474	11.761	-177.421	0.777	443
pTA	43 042	0.335	0.284	0.000	0.262	1
pDR	50 998	0.481	0.368	0.001	0.453	17
pDSCR	50 998	1.475	8.223	-15.319	0.264	232
pACR	50 998	6.206	22.918	0.030	2.190	734
Latvia						
DE	38 924	2.420	14.388	-160.247	1.061	381
TA	37 360	0.390	0.292	0.000	0.346	1
DR	38 924	0.685	0.817	0.001	0.586	20
DSCR	38 924	1.184	6.006	-17.959	0.199	209
ACR	38 924	3.217	11.627	0.034	1.696	782
pDE	38 924	2.772	16.978	-170.616	1.111	452
pTA	36 834	0.395	0.292	0.000	0.356	1
pDR	38 924	0.716	0.826	0.002	0.612	20
pDSCR	38 924	1.082	5.939	-14.881	0.175	231
pACR	38 924	3.056	11.377	0.002	1.622	554
Lithuania						
DE	28 917	2.011	10.246	-162.641	0.878	298
TA	27 392	0.295	0.248	0.000	0.231	1
DR	28 917	0.536	0.512	0.001	0.493	18
DSCR	28 917	1.242	4.418	-17.465	0.278	197
ACR	28 917	4.253	16.671	0.055	2.008	908
pDE	28 914	1.894	11.332	-173.562	0.885	437
pTA	26 740	0.305	0.253	0.000	0.241	1
pDR	28 914	0.564	0.596	0.002	0.510	20
pDSCR	28 914	1.259	5.159	-15.155	0.252	216
pACR	28 914	4.133	12.862	0.005	1.248	539

Table 30 provides the descriptive statistics for the financial statement Solvency variables belonging to the Transaction Lending factor group for the current (*DE*, *TA*, *DR*, *DSCR*, *ACR*) and previous (*pDE*, *pTA*, *pDR*, *pDSCR*, *pACR*) financial reporting periods. Variables *DE*, *TA*, *DR* have the highest values, whereas *DSCR*, *ACR* have the lowest values in LV, which indicates that the average applying SME in LV is more leveraged than in EE and LT. EE and LT share similar mean values for *DE*, *DR*, *TA*, and *DSCR* factors, while *ACR* is on average stronger in EE at 6.206 (whereas LT stands at 3.38). Similarly to the Liquidity variables, the Solvency variables are uniform across the past and current periods. It has been determined that Solvency is not equally distributed across the countries, and the SMEs that are applying for financing in LV are more leveraged in comparison to EE and LT. For some applications, the Solvency ratio levels are outside the theoretical financing levels as defined in Table 19 (DE – >2; DR – >1.5; DSCR – >1.2; ACR – <1). By considering that financing is issued based on a multitude of factors, the adverse effect of sub-par solvency factor levels can either lead to a rejection or be considered irrelevant if compensated by other factors. To analyze the underlying SME profitability across the retrieved financing applications, Profitability variables are evaluated (see Table 31).

Table 31. Descriptive statistics of Profitability variables belonging to the Transaction Lending factor group. Created by the author.

Name	Count	Mean	SD	Min	Median	Max
Estonia						
ROA	50 998	0.155	0.339	-6.167	0.113	8
ROE	50 998	0.352	2.207	-50.343	0.241	82
GMR	50 997	0.787	0.266	-9.419	0.870	1
PMR	50 997	0.103	0.339	-12.076	0.070	1
pROA	50 998	0.161	0.379	-9.172	0.113	7
pROE	50 998	0.396	2.000	-48.819	0.249	76
pGMR	50 994	0.778	0.323	-14.752	0.868	1
pPMR	50 994	0.097	0.434	-16.890	0.069	1
Latvia						
ROA	38 924	0.244	0.701	-6.390	0.111	9
ROE	38 924	0.646	3.365	-48.769	0.302	83
GMR	38 924	0.294	0.327	-9.742	0.225	1
PMR	38 924	0.112	0.334	-9.979	0.064	1
pROA	38 924	0.227	0.696	-9.088	0.101	7
pROE	38 924	0.658	3.554	-48.992	0.300	84
pGMR	38 924	0.284	0.393	-16.145	0.221	1
pPMR	38 924	0.096	0.446	-16.488	0.058	1
Lithuania						
ROA	28 917	0.264	0.491	-5.740	0.129	9
ROE	28 917	0.656	2.734	-45.204	0.293	88
GMR	28 916	0.369	0.307	-10.857	0.318	1
PMR	28 916	0.128	0.313	-12.668	0.070	1
pROA	28 914	0.248	0.527	-9.553	0.119	7
pROE	28 914	0.598	2.511	-48.582	0.288	73
pGMR	28 912	0.356	0.317	-8.445	0.309	1
pPMR	28 912	0.116	0.370	-13.930	0.066	1

Table 31 provides the descriptive statistics for the financial statement Profitability variables belonging to the Transaction Lending factor group for the current (*ROA*, *ROE*, *GMR*, *PMR*) and previous (*pROA*, *pROE*, *pGMR*, *pPMR*) financial reporting periods. Depending on the selected factor, the profitability across the countries differs significantly. The most significant difference is for the mean *GMR*, *pGMR* values; in

EE, the mean value for *GMR* is 0.79, which is significantly higher than in LV at 0.29 and LT at 0.36. The profitability is relatively homogeneous for *PMR* across all the three countries ranging between 10% in EE and 13% in LT (with 11% in LV). In terms of *ROE* and *ROA*, LT has the highest mean values followed by LV and EE. The same findings are manifested across the median variable values. Similarly, the Liquidity and Solvency variables are relatively uniform across the past and current periods. For some applications, the Profitability levels are outside the theoretical financing thresholds as defined in Table 20 (ROA - >0.05; ROE - >0.15; GMR - >0.1; PMR - >0.07). By considering the volatile nature of SME profitability, the sustained losses do not necessarily mean that an application cannot receive financing, as specific reasons are not clear; nor do they indicate what the profitability may be in the future. It has been determined that the net profitability across the countries is relatively equal, while the operating profit is significantly higher in EE. To analyze the efficiency of the company collection policies for the retrieved applications, Activity variables are evaluated (see Table 32).

Table 32. Descriptive statistics of Activity variables belonging to the Transaction Lending factor group. Created by the author.

Name	Count	Mean	SD	Min	Median	Max
Estonia						
ATR	50 998	2.277	2.200	0.000	1.734	47
RTR	48 171	65.193	619.470	0.001	9.603	29341
pATR	50 998	2.336	2.363	0.000	1.757	62
pRTR	47 768	57.959	487.194	0.000	9.586	21477
CS	50 998	1.040	35.486	-1	0.101	4848
CiA	50 998	1.845	119.151	-1	0.118	16366
Latvia						
ATR	38 924	2.645	2.818	0.000	1.969	48
RTR	35 681	118.963	809.403	0.008	15.131	28467
pATR	38 924	2.690	3.103	0.000	1.973	61
pRTR	35 344	117.388	705.627	0.006	14.843	20958
CS	38 924	1.584	72.821	-0.999	0.092	12966
CiA	38 924	1.729	70.431	-1	0.103	7384
Lithuania						
ATR	28 917	2.332	2.087	0.000	1.905	47
RTR	26 582	63.381	694.630	0.000	7.949	29647
pATR	28 914	2.387	2.456	0.000	1.894	61
pRTR	27 038	53.467	484.077	0.000	7.776	21302
CS	28 917	1.591	32.389	-1	0.131	2046
CiA	28 917	1.184	69.509	-0.998	0.152	11799

Table 32 provides the descriptive statistics for the financial statement Activity variables belonging to the Transaction Lending factor group for current (*ATR*, *RTR*, *CS*, *CiA*) and previous (*pATR*, *pRTR*) financial reporting periods. The average *ATR* and *RTR* values are similar in EE and LT, while in LV they are higher; *RTR* is almost

120 days, which indicates that, for companies in LV, accounts receivable takes approximately 2x longer to be collected than in the two other countries. This suggests a larger dependency on informal trade credit provided by business partners. Close to zero Activity levels indicate low to non-existent sales activity, which could have an adverse effect on accessing credit, but it does not necessarily mean that financing cannot be issued without considering other factors. Activity ratios for the current and previous periods are relatively uniform as for the Liquidity, Solvency, Profitability variable groups. Transaction lending factors based on the financial statement data are notably different across the countries, specifically, in LV. The financial ratios for an average application in LV have been determined to be less liquid, have lower Activity ratios and are significantly more leveraged than in EE and LT (note that *PMR* is similar across the three markets). A relatively weaker average financial standing of an incoming SME application in LV is potentially alleviated through the lower information opaqueness due to a higher fraction of larger company applications and a larger number of audited financial statements. Next, the underlying company and owner's credit history across the received financing applications are analyzed (see Tables 33 and 34).

Table 33. Descriptive statistics of company's credit history variables belonging to the Transaction Lending factor group. Created by the author.

Name	Count	Mean	SD	Min	Median	Max
Estonia						
IOverC	50 998	0.405	1.670	0	0	60
IOverS	50 998	0.382	0.962	0	0	6
IOverL	50 998	2	6	0	0	347
EOverC	50 998	1.008	1.874	0	0	85
EOverS	50 998	1.079	1.650	0	0	6
EOverL	50 998	51	104	0	0	730
Defaults	50 998	0.029	0.167	0	0	1
Latvia						
IOverC	38 924	0.236	1.484	0	0	115
IOverS	38 924	0.225	0.780	0	0	5
IOverL	38 924	1	5	0	0	168
EOverC	38 924	0.074	0.797	0	0	37
EOverS	38 924	0.038	0.340	0	0	5
EOverL	38 924	6	48	0	0	720
Defaults	38 924	0.052	0.222	0	0	1
Lithuania						
IOverC	28 917	0.265	1.576	0	0	110
IOverS	28 917	0.258	0.823	0	0	6
IOverL	28 917	1	7	0	0	568
EOverC	28 917	0.558	3.681	0	0	551
EOverS	28 917	0.654	1.186	0	0	6
EOverL	28 917	23	71	0	0	726
Defaults	28 917	0.039	0.194	0	0	1

Table 33 provides the descriptive statistics for company credit history variables (*IOverC*, *IOverS*, *IOverL*, *EOverC*, *EOverS*, *EOverL*, *Defaults*) belonging to the Transaction Lending factor group. For both internal and external overdues, their count, amount and the length of the overall overdue levels for the applying companies is highest in EE and lowest in LV. Even though the general overdue levels are the highest in EE, whereas the mean *Defaults* values indicate the lowest default numbers (0.029) in comparison to LV (0.052) and LT (0.039). It was determined that each country throughout different overdue evaluation dimensions has varying overdue levels. Furthermore, the connection between the overdue levels and the default rates is not clear as the country with the highest overdue metrics (EE) has the lowest number of defaults, while LV with the lowest overdue levels has the highest overdue levels. Next the owner's credit history factors are analyzed (see Table 34).

Table 34. Descriptive statistics of the company owner's credit history variables belonging to the Transaction Lending factor group. Created by the author.

Name	**Count**	**Mean**	**SD**	**Min**	**Median**	**Max**
Estonia						
OIOverC	50 998	0.314	1.674	0	0	47
OIOverS	50 998	0.178	0.601	0	0	5
OIOverL	50 998	1	8	0	0	713
OEOverC	50 998	0.082	0.560	0	0	27
OEOverS	50 998	0.033	0.343	0	0	6
OEOverL	50 998	5	39	0	0	727
ODefaults	50 998	0.001	0.036	0	0	1
Latvia						
OIOverC	38 924	0.150	1.037	0	0	39
OIOverS	38 924	0.097	0.450	0	0	5
OIOverL	38 924	1	8	0	0	591
OEOverC	38 924	0.060	0.526	0	0	25
OEOverS	38 924	0.028	0.271	0	0	4
OEOverL	38 924	8	56	0	0	730
ODefaults	38 924	0.002	0.046	0	0	1
Lithuania						
OIOverC	28 917	0.258	1.698	0	0	55
OIOverS	28 917	0.124	0.509	0	0	6
OIOverL	28 917	1	13	0	0	725
OEOverC	28 917	0.096	0.570	0	0	28
OEOverS	28 917	0.124	0.548	0	0	8
OEOverL	28 917	7	43	0	0	729
ODefaults	28 917	0.003	0.059	0	0	1

Table 34 provides the descriptive statistics for the company credit history variables (*OIOverC*, *OIOverS*, *OIOverL*, *OEOverC*, *OEOverS*, *OEOverL*, *ODefaults*) belonging to the Transaction Lending factor group. In line with the findings concerning the applying company's credit history, the owner's overdue levels are also the highest in EE and the lowest in LV, while the default levels are significantly lower than those of the applying companies.

Comparative analysis of the dependent variable *Outcome* and the independent variable factor groups was carried out, and it was determined that the underlying SME access to credit is not uniformly distributed across the studied countries – it is higher in EE and lower in LV and LT. These findings add to the already existing literature on the Baltic States and indicate that there exist some country specifics concerning the underlying access to credit for SME entities. The underlying access to credit is not a constant, and it varies throughout different periods of time. Furthermore, it has been concluded that the underlying Firm Characteristics, Product Characteristics, Transaction Lending and Relationship Lending factor groups are not homogeneous across the incoming applications flow for the three studied countries. The analysis has demonstrated that applications, which are exceeding the theoretically acceptable Liquidity, Solvency, Profitability and Activity levels, exist, which suggests that, in order to evaluate the SME access to credit, the multitude of factors and their relationships should be considered. In EE, an average SME application is from a younger, smaller, and less diversely owned company that has a stronger bank-firm relationship, stronger financial health and relatively more overdues which do not materialize as defaults. In LV, an average application is received from an older and larger company that is more likely to provide audited financial statements, has relatively strong bank-firm relationships but weaker financial health and a higher probability to have historic defaults. Finally, in LT, an average SME application is from a larger company which requested a leasing product and has a relatively shorter and less intense bank-firm relationship.

To determine the impact of individual factors and their interactions on the SME access to credit, Stage II, the dimensionality reduction procedure, must be carried out to define a representative feature vector which shall be used for model estimation in Stage III.

3.2. Definition of the Representative Feature Vector for SME Access to Credit Estimation

In Stage II, a dimensionality reduction procedure is carried out for each individual country. The goal of this procedure is to reduce the number of features used in the model while retaining as much information as possible. The motivation behind dimensionality reduction is to simplify the complexity of the data while ensuring model robustness, feature interpretability and avoiding over-fitting. The procedure follows the methodology as described in Section 2.2.

To identify closely related variables, correlations are calculated for all variable pairs across all the three countries. The Spearman correlation is used as a statistical measure of the strength and direction of the monotonic relationship between two variables. For the sake of estimating correlations between two categorical (nominal) variables, one-hot (dummy) encoding is carried out. As the correlation between encoded categorical variables will only repeat synthetically created variable relations, they will be ignored in any further dimensionality reduction. Correlation values for categorical dummy variable pairs are calculated by using phi-statistic. Figure 18 demonstrates the feature correlation heat-map for LT (see Annex 1 for EE, Annex 2 for LV).

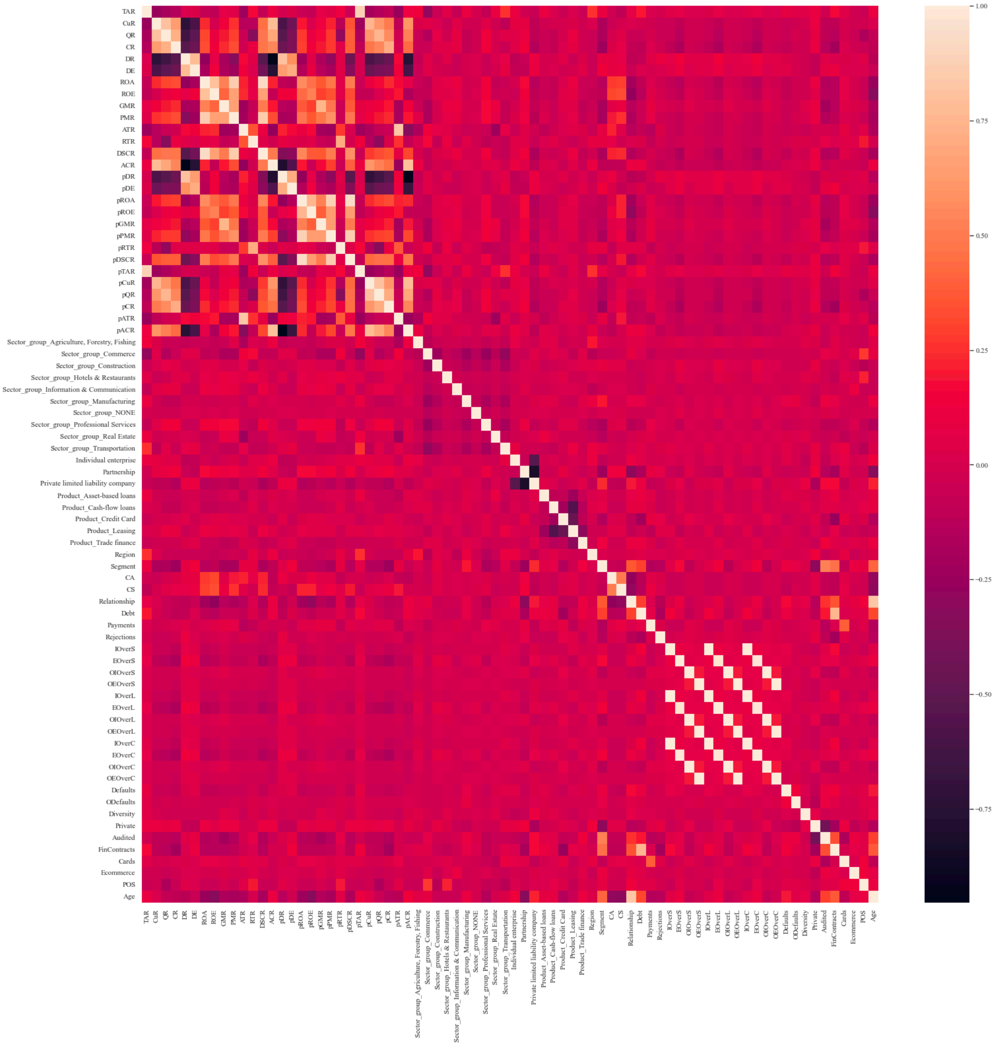


Figure 18. Feature correlation heat-map for the Lithuanian dataset. Created by the author.

It has been determined that, across all the three countries, some feature groups have strong feature pair correlations, which, if included into model estimation, would hinder feature interpretability and would not provide any substantial improvement in the model accuracy. Particularly strong correlations are observed between variable pairs which are derived from or are measuring similar indicators, such as the Transaction Lending factor group that is mostly based on the financial statement data (*CR*, *QR*, *CuR*, *DE*, *TA*, *DR*, *ROA*, *ROE*, *GMR*, *PMR*, *ATR*, *RTR*, *DSCR*, *ACR*, *CS*, *CiA*, *Audited*, *pCR*, *pQR*, *pCuR*, *pDE*, *pTA*, *pDR*, *pROA*, *pROE*, *pGMR*, *pPMR*, *pATR*, *pRTR*, *pDSCR*, *pACR*, *pAudited*) and the credit history factors for the company (*IOverC*, *IOverS*, *IOverL*, *EOverC*, *EOverS*, *EOverL*) or the owner (*OIOverC*, *OIOverS*, *OIOverL*, *OEOverC*, *OEOverS*, *OEOverL*, *Defaults*, *ODefaults*). It has also been determined that a very strong correlation exists between *Age* and *Relationship* as both features have a direct linear dependency on one another by measuring the time since a particular event. A correlation heat-map is limited in its ability to understand the non-linear relationships between variables. Therefore, hierarchical clustering is used to complement the information obtained from a correlation heat-map. Hierarchical clustering on Spearman rank-order correlations is performed and visualized as a dendrogram for LT in Figure 19 (see Annex 3 for EE, Annex 4 for LV).

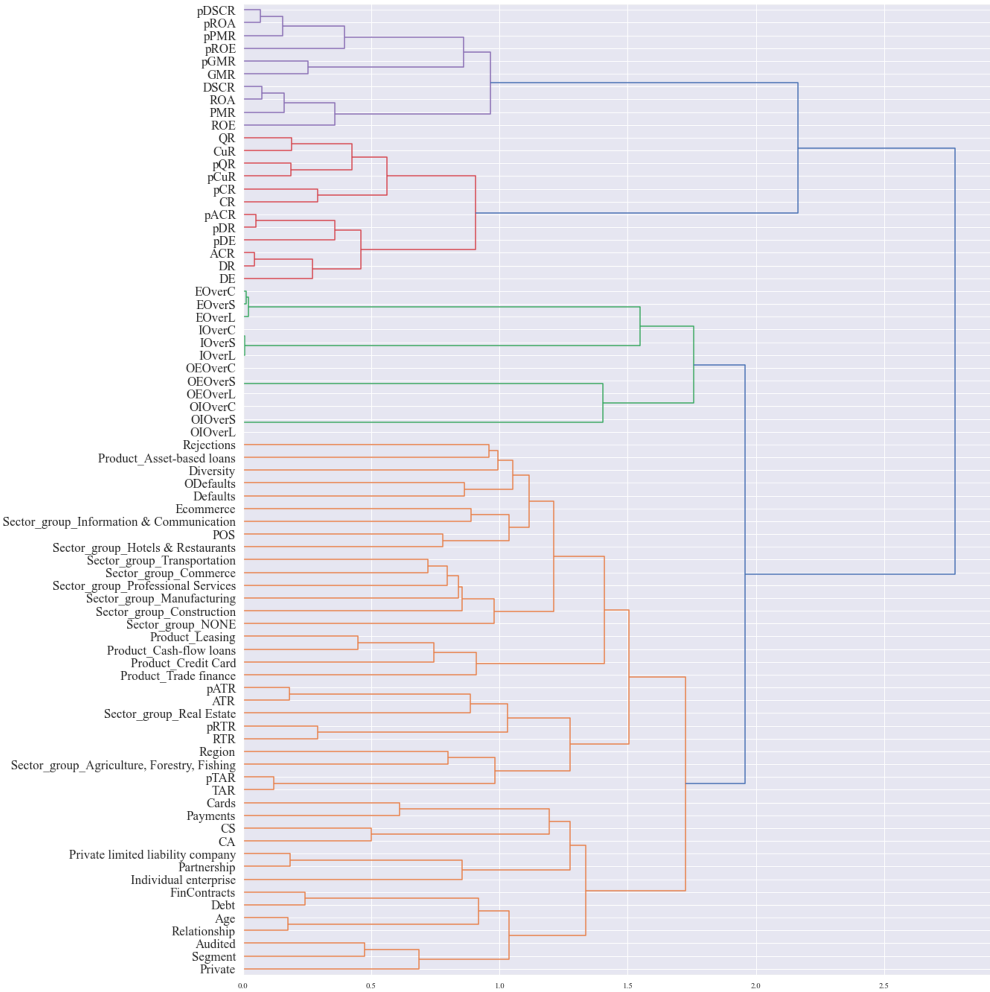


Figure 19. Hierarchical clustering dendrogram for the Lithuanian dataset. Created by the author.

The purpose of these dendrograms is, across all the three countries, to group variables into specific groups for estimated distances between the features. In line with the findings from feature correlations in Figure 18, Figure 19 demonstrates that the variables belonging to specific factor groups are part of single clusters, specifically, the Transaction Lending factor group factors based on the financial statement data, such as Liquidity (*CR*, *QR*, *CuR*, *pCR*, *pQR*, *pCuR*), Solvency (*DE*, *TA*, *DR*, *DSCR*, *ACR*, *pDE*, *pTA*, *pDR*, *pDSCR*, *pACR*), Profitability (*ROA*, *ROE*, *GMR*, *PMR*, *pROA*, *pROE*, *pGMR*, *pPMR*) and credit history factors for the company (*IOverC*, *IOverS*, *IOverL*, *EOverC*, *EOverS*, *EOverL*) and the owner (*OIOverC*, *OIOverS*, *OIOverL*, *OEOverC*, *OEOverS*, *OEOverL*, *Defaults*, *ODefaults*). Furthermore, it has been determined that, across all the three countries, the estimated distance between variables is small between relatively all the current and previous period variables, *Segment* and *Audited* variables, *Debt* and *FinContracts* variables, *CS* and *CiA* variables indicating that the features are closely related, and thus feature selection must be carried out. In LV, the *DE* feature shares close distance only to *pDE*, which suggests that it potentially has additional information and should be kept throughout the different Euclidean distance thresholds.

By combining the findings from the correlations heat-map in Figure 18 and the hierarchical cluster dendrogram in Figure 19, feature selection is carried out to define the representative feature vector. In cases when two variables of a similar nature are compared, the leading feature is selected based on the continuous variable variance and relationship with the rejection rate (the application outcome), the value frequency is plotted together with the rejection rate. By combining all the variables, feature vectors for each country are defined across the pre-defined correlation distance thresholds (0.00, 0.25, 0.50, 0.75, 1.00, and 1.25). Country-specific feature vectors are presented across the studied factor groups: Firm Characteristics (see Table 35), Product Characteristics (see Table 36), Lending Technology factors (for Relationship Lending factors, see Table 37, for Transaction Lending factors, see Tables 38 and 39).

Table 35. Firm Characteristic factor group's feature vectors throughout different correlation distance thresholds. Created by the author.

	Clustering distance threshold					
Variable	**0**	**0.25**	**0.5**	**0.75**	**1.00**	**1.25**
Continuous						
Age	●					
Diversity	●	●	●	●	LV	
Private	●	●	●	EE, LV		
Audited	●	●				
Categorical						
Segment	●	●	●	●	●	●
Type	●	●	●	●	●	●
Region	●	●	●	●	●	●
Sector	●	●	●	●	●	●

Country feature vectors for the Firm Characteristic factor group across all the clustering distance thresholds include all categorical and continuous variables, such as *Diversity* and *Private* for the clustering distance threshold up to 0.75. As *Age* shares most of the information with *Relationship*, it is excluded from the feature vectors.

Table 36. Product characteristic factor group's feature vectors throughout different correlation distance thresholds. Created by the author.

	Clustering distance threshold					
Variable	**0**	**0.25**	**0.5**	**0.75**	**1.00**	**1.25**
Product	●	●	●	●	●	●

As the study focuses on the importance of *Product* on the SME access to credit and the fact that the feature is categorical, it is automatically part of the representative feature vector independent of clustering distance threshold selections.

Table 37. Relationship lending factor group's feature vectors throughout different correlation distance thresholds. Created by the author.

	Clustering distance threshold					
Variable	**0**	**0.25**	**0.5**	**0.75**	**1.00**	**1.25**
Relationship	●	●	●	●	EE, LV	LV
Payments	●	●	●	●	EE, LT	LT
Rejections	●	●	●	●		
Debt	●	●	LT	LT		
FinContracts	●	●	●	●	●	●
Cards	●	●	●	EE, LV		
POS	●	●	●	LT		
Ecommerce	●	●	●	●		

Country feature vectors for the Relationship Lending factor group across all the clustering distance thresholds includes *FinContracts*. In terms of other variables, *Relationship* in LV and *Payments* in LT are part of the feature vectors across all clustering thresholds.

Table 38. Transaction lending factor group's based on financial statement data feature vectors throughout different correlation distance thresholds. Created by the author.

	Clustering distance threshold					
Variable	**0**	**0.25**	**0.5**	**0.75**	**1**	**1.25**
Liquidity indicators						
CR	●	●	LV			
QR	●					
CuR	●	●	●	●		
pCR	●	●				
pQR	●					
pCuR	●	●				
Solvency indicators						
DE	●	LV, LT	LV	LV	LV	
TA	●	●	●	●		
DR	●	●	●	●	●	●
DSCR	●	●				
ACR	●					
pDE	●	LV, LT				
pTA	●					
pDR	●	●				
pDSCR	●					
pACR	●					
Profitability indicators						
ROA	●	●	●	●	EE, LT	EE, LT
ROE	●	●	LV			
GMR	●	●	●	●	EE, LV	EE
PMR	●	EE				
pROA	●	●	●	●	EE	EE
pROE	●	●	LV			
pGMR	●	LT				
pPMR	●	EE				
Activity indicators						
ATR	●	●	●	●	LT	LT
RTR	●	●	●	●	●	LV
pATR	●					
pRTR	●	●				
CS	●	●	●	●	LV, LT	
CiA	●	●	EE, LV			

One of the most correlated feature groups is the financial statement variables belonging to the Lending factor group (see Table 38). In line with the findings from correlation heat-maps, the previous period variables are closely correlated to the current period variables, thus providing limited information for the SME access to credit model development. Some exceptions exist, such as for *ROA* and *pROA*.

Table 39. Transaction lending factor group's based on credit history data feature vectors throughout different correlation distance thresholds. Created by the author.

	Clustering distance threshold					
Variable	**0**	**0.25**	**0.5**	**0.75**	**1**	**1.25**
Company's credit history						
IOverC	●					
IOverS	●	●	●	●	●	●
IOverL	●					
EOverC	●					
EOverS	●	●	●	●	●	●
EOverL	●	LV				
Defaults	●	●	●	●	EE, LT	
Owner's credit history						
OIOverC	●					
OIOverS	●	●	●	●	●	●
OIOverL	●	EE				
OEOverC	●					
OEOverS	●	●	●	●	●	●
OEOverL	●	EE, LV	EE			
ODefaults	●	●	●	●		

Similarly to the financial statement variables, credit history variables are highly correlated; therefore, feature vectors across different clustering thresholds include a variable indicating the size of the overdues and defaults for both the company (*IOverS*, *EOverS*, *Defaults*) and its owners (*OIOverS*, *OEOverS*, *ODefaults*).

A representative feature vector for the correlation distance threshold at 0.00 represents the full set of available variables (61 variables in total). As the correlation distance threshold increases (at increments of 0.25), the number of variables in the representative feature vector gets smaller (48 at 0.25, 36 at 0.50, 31 at 0.75, 22 at 1.00). At the correlation distance threshold of 1.25, across all the three countries, *Segment*, *Type*, *Region*, *Sector*, *Product*, *FinContracts*, *DR*, *IOverS*, *EOverS*, *OIOverS*, *OEOverS* features are preserved. Meanwhile, some features, such as *GMR* and *pROA* for EE, *Relationship* and *RTR* for LV, *Payments* and *ATR* for LT, *ROA* for both EE and LT - are kept for specific countries. As the number of features is decreasing, the information that is available for the model to estimate the target variable *Outcome* becomes scarce. Therefore, to select the appropriate representative feature vector, the SME access to credit models is estimated by using modelling techniques as defined in Section 2.3 by iterating representative feature vectors throughout pre-defined hierarchical clustering thresholds (based on Section 2.2). Table 40 demonstrates the accuracy

Table 40. Modelling accuracy fall-off throughout clustering distance thresholds across different modelling techniques. Created by the author.

	LGB	**RF**	**MLP**	**LR**
ROC-AUC				
0.00				
0.25	-0.003	0.001	0.009	0.003
0.50	-0.010	0.006	0.003	0.010
0.75	**0.000**	**0.001**	**0.000**	**0.000**
1.00	-0.022	-0.010	-0.017	0.013
1.25	-0.004	0.001	0.004	0.026
Precision				
0.00				
0.25	-0.002	0.003	0.018	-0.001
0.50	-0.009	0.006	0.000	0.011
0.75	**0.000**	**0.001**	**0.005**	**0.000**
1.00	-0.034	-0.029	-0.027	0.019
1.25	-0.006	0.004	0.001	0.035

fall-off throughout different clustering distance thresholds for the selected modelling techniques in terms of ROC-AUC and Precision.

Table 40, throughout the selected clustering distance thresholds, demonstrates the average model accuracy fall-off using the changes of the ROC-AUC and Precision measures. The removal of correlated features does not have a uniform effect on the modelling accuracy, as the effect varies throughout different thresholds and employed modelling techniques. It has been determined that, for the thresholds of 0.25 to 0.75, the accuracy in terms of ROC-AUC and Precision is lower for the HGB (AUC – -0.013, Precision – -0.01) modelling technique, whilst the accuracy increased for RF (ROC – 0.008, Precision – 0.01), MLP (ROC – 0.012, Precision – 0.023) and LR (ROC – 0.013, Precision – 0.01) modelling techniques. Across the LGB, RF and MLP modelling techniques, the model accuracy fell significantly from the threshold of 0.75 to 1.00. LR, in contrast to other modelling techniques, had its accuracy improved with a lower number of variables. It is concluded that, by considering the fall-off in the ROC-AUC and Precision model accuracy metrics at the level of 1.00, the representative feature vector at the clustering threshold of 0.75 is selected for the model estimation and variable exploration in Stage III (see Table 41).

Table 41. Representative feature vector for the country specific SME access to credit model. Created by the author.

Variable	EE	LV	LT
Firm characteristics			
Diversity	●	●	●
Private	●	●	
Segment	●	●	●
Type	●	●	●
Region	●	●	●
Sector	●	●	●
Product Characteristics			
Product	●	●	●
Lending Technology factors			
Relationship lending factors			
Relationship	●	●	●
Payments	●	●	●
Rejections	●	●	●
Debt			●
FinContracts	●	●	●
Cards	●	●	
POS			●
Ecommerce	●	●	●
Transaction lending factors			
Based on financial statement data			
CuR	●	●	●
DE	●	●	●
TA	●	●	●
DR	●	●	●
ROA	●	●	●
GMR	●	●	●
ATR	●	●	●
RTR	●	●	●
CS	●	●	●
pROA	●	●	●
Based on credit history data			
IOverS	●	●	●
EOverS	●	●	●
Defaults	●	●	●
OIOverS	●	●	●
OEOverS	●	●	●
ODefaults	●	●	●

Stage II is carried out and finalized by determining the representative feature vector (see Table 41) that is composed of variables which shall be employed in Stage III for developing the country specific SME access to credit model. It is evident that the composition of the representative feature vectors across the studied countries is similar, but not the same, which confirms the findings of other studies that country-specific model estimation is necessary. It has been found that certain features within the Firm Characteristics and Transaction Lending factor group are highly correlated. As a result, the past period financial ratios and the credit history variables, which did not provide any additional information, were excluded from the representative feature vectors.

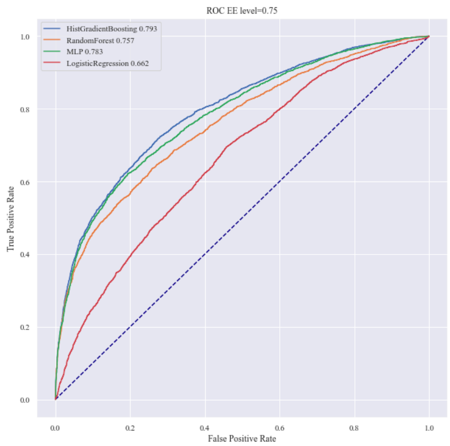
3.3. Country-specific SME Access to Credit Models

In Stage III, to evaluate the SME access to credit and to evaluate the feature importance and interactions in EE, LV and LT, country-specific models are created. *Output* is estimated by using independent variables from the selected representative feature vector. The representative feature vector consists of at least one variable belonging to the underlying factor group: Firm Characteristics (*Diversity*, *Private* in EE and LV, *Segment*, *Type*, *Region*, *Sector*), Product Characteristics (*Product*), Relationship lending factors (*Relationship*, *Payments*, *Rejections*, *Debt* in LT, *FinContracts*, *Cards* in EE and LV, *POS* in LT, *Ecommerce*), and Transaction lending factors (*CuR, DE* in LV, *TA*, *DR*, *ROA*, *GMR*, *ATR*, *RTR*, *CS*, *pROA*, *IOverS*, *EOverS*, *OIOverS*, *OEOverS*, *Defaults*, *ODefaults*). Country-specific SME access to credit models is estimated by using HGB, RF, MLP, and LR modelling techniques. Accuracy evaluation is carried out as described in Section 2.4 based on the confusion matrix derivative ratios and on the ROC-AUC and Precision-Recall (Prec-Rec.) metrics. Each individual model's accuracy is summarized in Table 42.

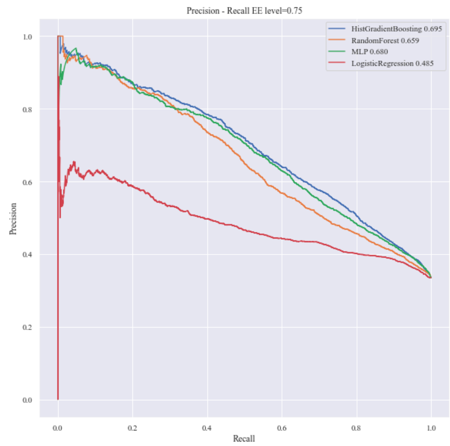
Table 42. The accuracy of country specific SME access to credit modelling techniques. Created by the author.

	EE		**LV**		**LT**	
	ROC-AUC	**Avr. Prec.**	**ROC-AUC**	**Avr. Prec.**	**ROC-AUC**	**Avr. Prec.**
HGB	**0.793**	**0.695**	**0.803**	**0.817**	**0.816**	**0.810**
RF	0.756	0.658	0.771	0.791	0.750	0.757
MLP	0.783	0.680	0.794	0.802	0.789	0.777
LR	0.662	0.485	0.657	0.635	0.700	0.654

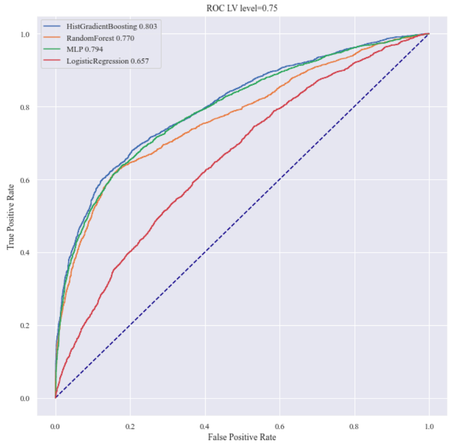
Uniformly across all the three country SME access to credit models, the best performing modelling technique is HGB. On the other hand, the worst performing modelling technique is the benchmark – LR, which demonstrated the lowest ROC-AUC and Avr. Prec. values across all the three country models. Overall, the highest modelling accuracy was reached in LT (ROC-AUC – 0.816, Avr. Prec. – 0.810) and LV (ROC – 0.804, Avr. Prec. – 0.818) depending on the selected metric. Meanwhile in EE, the SME access to credit modelling accuracy was lowest with ROC-AUC at 0.796 and Avr. Prec. at 0.697. To evaluate the modelling performance across all the discriminatory threshold values, ROC and Precision-Recall curves are used (see Figure 20).



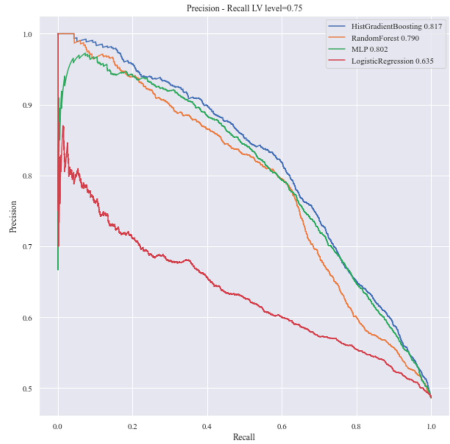
(a) Estonia – ROC



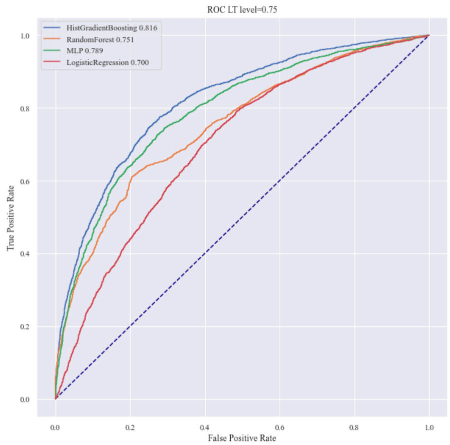
(b) Estonia – Precision-Recall



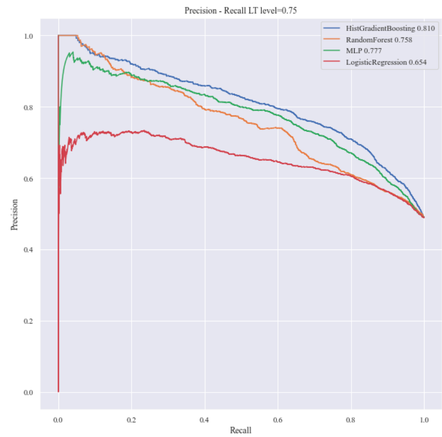
(c) Latvia – ROC



(d) Latvia – Precision-Recall



(e) Lithuania – ROC



(f) Lithuania – Precision-Recall

Figure 20. ROC and Precision-Recall curves for the estimated SME access to credit models. Created by the author.

Throughout most of the discriminatory threshold values, the HGB modelling technique outperformed RF, MLP, and LR. Meanwhile, for all the three country models, the RF modelling technique had the highest accuracy at extreme Avr. Prec. values (>0.9). In comparison to the other techniques, LR is consistently unperformed across all discriminatory threshold values. It has been determined that the SME access to credit modelling performance in terms of the overall modelling accuracy is uniform across all the three countries. Overall, the best modelling technique for estimating the SME access to credit is HGB as it demonstrated the highest accuracy both in terms of the overall ROC-AUC and Avr. Prec. values and throughout most discriminatory threshold values. To evaluate the HGB model performance for a single discriminatory cut-off threshold (0.5), model-specific confusion matrices and derivative metrics are utilized (see Table 43).

Table 43. Confusion matrices and derived accuracy metrics for country-specific SME access to credit models. Created by the author.

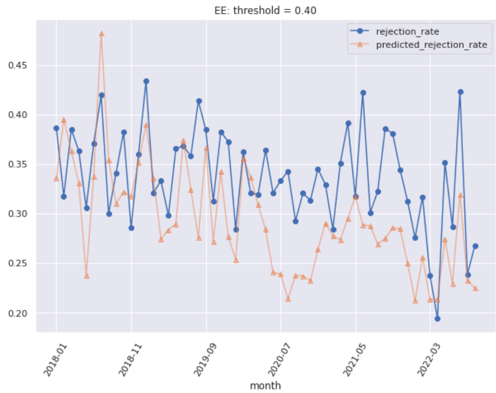
(a) Confusion matrices

	Predicted	***Approval***	***Rejection***
	Actual		
Estonia	*Approval*	4616	1007
	Rejection	1339	1995
Latvia	*Approval*	2163	701
	Rejection	1414	2684
Lithuania	*Approval*	1776	477
	Rejection	874	2062

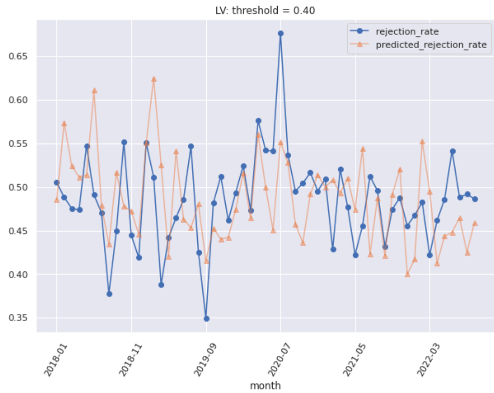
(b) Accuracy ratios

Accuracy metric	**Estonia**	**Latvia**	**Lithuania**
Specificity	0.598	0.655	0.702
NPV	0.665	0.793	0.812
Precision	0.775	0.605	0.670
Sensitivity	0.821	0.755	0.788
FPR	0.402	0.345	0.298
Accuracy	0.738	0.696	0.740
F1	0.797	0.672	0.724

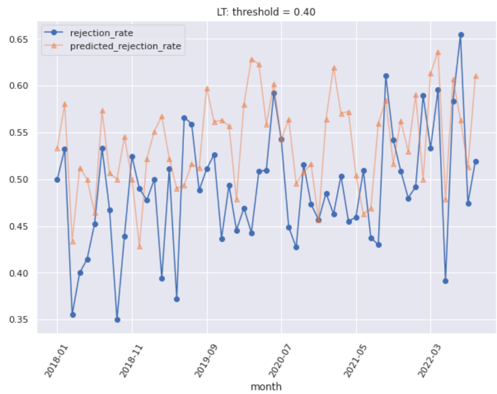
The confusion matrices and the derived accuracy ratios (see 43 provide a more detailed breakdown of the performance for each country model. It has been determined that the EE model, in comparison to the other country models, has the highest Sensitivity and Precision values, at 0.82 and 0.78, accordingly, whilst having the lowest Specificity (0.6) and NPV (0.67) values. It has been determined that the estimated SME access to credit model for EE tends to classify cases as approvals at the cost of lower Specificity and NPV predictions. The LT model, on the other hand, has the highest Specificity (0.7) and NPV (0.81), which indicates that the model is more likely to classify a case as credit rationing, while maintaining a high overall accuracy of 0.74. The LV model has the lowest Sensitivity (0.76) and Precision (0.61) and the overall Accuracy (0.7) values, which indicates that the model is the least likely to classify a case as an Approval while at the same time having the worst accuracy in differentiating the predicted approvals from the actual approvals. Finally, the predicted and actual rejection rates are compared throughout different periods (see Figure 21).



(a) Estonia



(b) Latvia



(c) Lithuania

Figure 21. SME access to credit model predictions and the actual rejection rate for the test sample. Created by the author.

Figure 21 demonstrates the dynamics of the predicted and actual rejection rate across the test dataset. Depending on the specific country and period, the predicted rejection rate tends to both under-estimate and over-estimate the SME access to credit. In EE, for the entirety of 2020, the model estimated a significantly lower rejection rate than the actual one (see Figure 21a), while in LT the model throughout the whole period tended to over-estimate the rejection rate (see Figure 21c). Uniformly across the countries, the SME access to credit models are able to accurately follow the actual access to credit values and trends.

It has been determined that the best performing modelling technique is HGB which outperformed both the benchmark LR and the other state-of-the-art machine learning techniques, namely, RF and MLP. It has also been found that there exist some differences between countries as the individual model accuracy is not uniform across countries since the SME access to credit model for EE is able to better identify credit rationing cases, while, for LT, it is able to better identify the cases when financing is issued. As the best performing modelling technique has been determined, the importance of the variables and their interactions shall be evaluated next.

3.4. The Importance of Features and Interactions in Country-specific SME Access to Credit Models

The final step of Stage III in evaluating the SME access to credit is to evaluate the relationship between the features and the SME access to credit model's prediction. This step will demonstrate the importance of the individual features and their interactions, as well as the impact they have on the SME ability to access credit. The individual feature impact on the predicted SME access to credit is estimated by employing the mean absolute SHAP and the PFI feature importance measures to identify which features are the most important for the model's predictions. The first measure is the mean absolute SHAP (SHapley Additive exPlanations), which is a method that estimates the importance of each feature by calculating the average difference in the model's predictions when a feature is present or absent. The second measure is Permutation Feature Importance (PFI), which is a method that estimates the importance of each feature by measuring the change in the model's performance when the values of a feature are randomly permuted. It is worth noting that the PFI and the mean absolute SHAP values are not directly comparable, but they both give an idea of the feature's importance. PFI is based on how much the model's performance decreases when a feature is removed, while the mean absolute SHAP is based on how much each feature contributes to the model's output. The individual importance of features is presented in Table 44.

For each country-specific SME access to credit model, Table 44 lists the features in order of their importance sorted by the mean absolute SHAP values. It is necessary to evaluate both the absolute importance of individual features on the company's ability to access credit as well as their relative, model-specific order. According to the highest mean absolute SHAP (EE – 0.55, LV – 0.53, LT – 0.96) and the PFI value (EE – 0.11, LV – 0.14, LT – 0.32), the number of financial contracts *FinContracts* is one of the most important features when estimating the SME access to credit across all the three countries. Proportionally the largest difference in terms of the feature importance is in LT, where the importance of *FinContracts* in terms of the mean absolute SHAP accounts for almost a half of the mean absolute SHAP across all features summed together. *Product* is the most important feature in LV and the second most important in EE, while in LT it is only number 6. Although *Product* in LV has the highest mean absolute SHAP, it does not have the highest PFI, which suggests that the feature contributes the most

Table 44. The importance of features in terms of the mean absolute SHAP and PFI for country-specific SME access to credit models. Created by the author.

Estonia			Latvia			Lithuania		
Variable	**SHAP**	**PFI**	**Variable**	**SHAP**	**PFI**	**Variable**	**SHAP**	**PFI**
FinContracts	0.549	0.107	*Product*	0.907	0.063	*FinContracts*	0.957	0.316
Product	0.399	0.054	*FinContracts*	0.527	0.144	*Type*	0.296	0.003
Rejections	0.266	0.043	*Rejections*	0.283	0.033	*Debt*	0.230	0.044
Relationship	0.135	0.006	*ATR*	0.167	0.022	*DR*	0.179	0.015
Payments	0.129	0.006	*Payments*	0.156	0.009	*Payments*	0.147	0.006
IOverS	0.123	0.005	*DR*	0.087	0.004	*Rejections*	0.132	0.012
Segment	0.096	0.001	*Sector*	0.081	0.004	*Product*	0.114	0.072
ATR	0.087	0.003	*ROA*	0.080	0.002	*ROA*	0.091	0.003
TA	0.072	0.003	*IOverS*	0.067	0.002	*CS*	0.069	0.002
pROA	0.063	0.001	*Relationship*	0.065	0.002	*Sector*	0.064	0.002
ROA	0.063	0.002	*CuR*	0.060	0.001	*pROA*	0.058	0.002
GMR	0.057	0.001	*Type*	0.055	0.000	*RTR*	0.058	0.003
RTR	0.057	0.002	*CS*	0.049	0.002	*GMR*	0.057	0.002
DR	0.051	0.002	*RTR*	0.038	0.000	*Relationship*	0.056	0.003
Sector	0.036	0.003	*TA*	0.033	0.001	*Segment*	0.042	0.002
OIOverS	0.035	0.001	*GMR*	0.032	0.000	*EOverS*	0.040	0.001
CS	0.033	0.000	*pROA*	0.031	0.000	*ATR*	0.039	0.001
CuR	0.032	0.001	*DE*	0.024	0.000	*TA*	0.038	0.001
EOverS	0.030	0.001	*Cards*	0.023	0.000	*IOverS*	0.036	0.000
Region	0.026	0.000	*Diversity*	0.013	0.000	*POS*	0.032	0.000
Private	0.017	0.000	*Private*	0.009	0.001	*CuR*	0.023	0.001
Cards	0.017	0.000	*OEOverS*	0.005	0.001	*OEOverS*	0.016	0.000
Ecommerce	0.007	0.000	*OIOverS*	0.005	0.000	*OIOverS*	0.012	0.001
OEOverS	0.003	0.000	*Defaults*	0.005	0.000	*Region*	0.011	0.000
Defaults	0.003	0.000	*Segment*	0.005	0.000	*Defaults*	0.005	0.000
Type	0.002	0.000	*EOverS*	0.004	0.000	*Diversity*	0.001	0.000
ODefaults	0.000	0.000	*Ecommerce*	0.001	0.000	*ODefaults*	0.000	0.000
Diversity	0.000	0.000	*ODefaults*	0.000	0.000	*Ecommerce*	0.000	0.000

towards the outcome, while *FinContracts* contributes the most towards the modelling performance. *Rejections* is the third most important feature in EE and LV, while in LT it is only the fifth. Uniformly, across all the three countries, owner defaults (*ODefaults*) had zero values for both the mean absolute SHAP and PFI, which indicates that this feature is not important in the SME access to credit model predictions. EE and LV have a relatively similar distribution of factors in terms of importance, whereas, in LT, factors like *Type* (the mean absolute SHAP – 0.3, PFI – 0.002), *Debt* (the mean absolute SHAP – 0.23, PFI – 0.04) and *Payments* (the mean absolute SHAP – 0.15, PFI – 0.01) are amongst the top 5 most important features. This suggests the presence of a notable differences between the countries in factor importance for estimating the access to credit for SMEs. It should be noted that, for all the three country models, most Lending technology factors (*CS*, *CuR*, *DR*, *RTR*, *GMR*, *ROA*, *pROA*, *TA*, *ATR*) are not individually important in terms of PFI (<0.005). This can be explained by the fact that the importance of an individual variable is watered down by including a bigger proportion of variables belonging to the same factor group.

It has been determined that the importance and impact of individual factors on the SME access to credit differs amongst the countries and depends on the individual model specifics and potentially existing variable interactions. To understand how the factor groups impact the estimated model's predictions, factor grouping by mean absolute SHAP and PFI values is carried out. Table 45 provides a summarized view on how each group of variables contributes towards the SME access to credit and modelling performance.

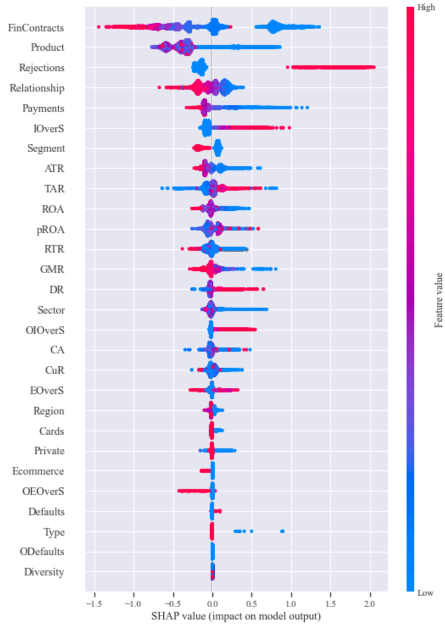
Table 45. The importance of factor groups in terms of the mean absolute SHAP and PFI for country-specific SME access to credit models. Created by the author.

	Estonia		**Latvia**		**Lithuania**	
	SHAP	**PFI**	**SHAP**	**PFI**	**SHAP**	**PFI**
Firm Characteristics	0.176	0.005	0.163	0.005	0.414	0.006
Product Characteristics	0.399	0.054	0.907	0.063	0.114	0.072
Lending technology factors	1.812	0.186	1.743	0.226	2.128	0.407
Relationship lending factors	1.103	0.162	1.055	0.189	1.407	0.375
Transaction lending factors	0.709	0.024	0.687	0.037	0.722	0.032
Financial statement based	0.515	0.016	0.600	0.033	0.611	0.029
Liquidity variables	0.032	0.001	0.060	0.001	0.023	0.001
Solvency variables	0.123	0.005	0.143	0.005	0.217	0.015
Profitability variables	0.183	0.005	0.142	0.003	0.206	0.007
Activity variables	0.177	0.006	0.254	0.023	0.166	0.006
Credit history variables	0.195	0.008	0.087	0.004	0.110	0.003
Company variables	0.156	0.007	0.077	0.003	0.082	0.002
Owner variables	0.039	0.001	0.010	0.001	0.028	0.001

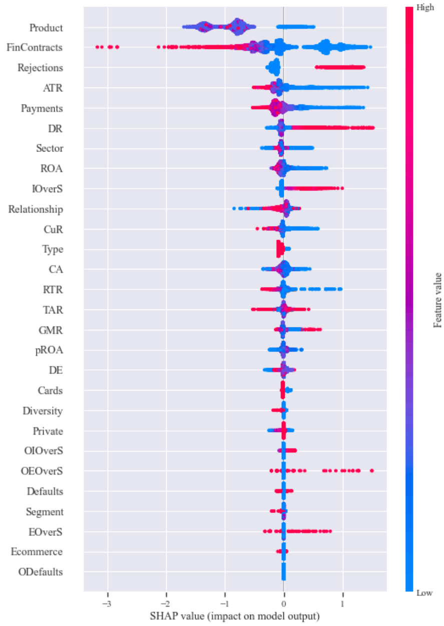
For all the three countries, according to the mean absolute SHAP (EE – 1.8, LV – 1.7 and LT – 2.1) and PFI (EE – 0.19, LV, – 0.23, LT – 0.4), the most important factor group is Lending technology factors. Relationship lending factors, as part of the Lending Technology factor group, accounts for more than 60% of the total mean absolute SHAP importance and more than 87% of PFI, which indicates a significant contribution both towards the outcome when applying for credit and the modelling performance. The contribution of the Credit history variables, as part of the Lending Technology factor group, is similar across the countries in terms of the relative proportion. In EE and LV, the least important variable group is the Firm characteristics, which accounts for approximately 8% and 6% respectively, of the total mean absolute SHAP (in terms of PFI, the impact on modelling performance is smaller, at around 2% in both countries). In LT, the least important factor group is the Product characteristics, which has a mean absolute SHAP of 0.4 and a PFI of 0.006 (corresponding to 16% of the feature impact in terms of the mean absolute SHAP and 1% in terms of PFI). Even though these groups give a sense of the relative impact and importance of groups of variables in the context of the entire set of variables used in the model, it assumes that the importance of a variable is only determined by its group, and does not consider the possible interactions and dependencies between the variables. It is evident that the structure and extent of the variable importance in each SME access to credit model differs between countries. It has been determined that Lending technology uniformly across all the three countries is the most important factor group, while the importance and impact of Firm characteristics and Product characteristics is not uniform and varies across different markets.

To evaluate how the impact of individual variables changes across different values, SHAP plots are used (see Figure 22), where the effects of the features on the resulting predictions are visualized. Each dot on the feature row represents a single instance in the dataset, distributed on the x-axis according to the SHAP value for that feature value. The red and blue dots can be observed to have a distinct difference, as high (red) and low (blue) feature values, on the SME access to credit model's outcome. As demonstrated in Figure 22, the direction of the impact can be seen by the distribution of dots (observations) to the zero (neutral) line: to the right (high SHAP values) – a stronger negative effect (to be rationed), to the left (low SHAP values) – a stronger positive effect (to be approved). The missing values are colored in grey. The features are ranked by importance, with the most important features for the prediction at the top.

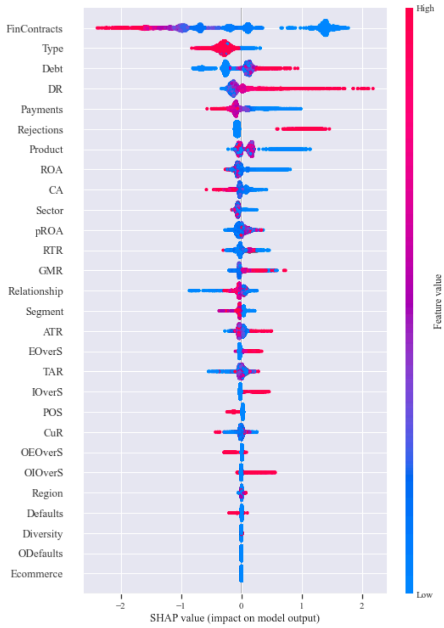
The extent of each variable's impact on the model's prediction is not uniform across the countries. In EE (see Figure 22a), the past rejections (*Rejections*) is a very strong and distinct indication that an application will be rationed, whilst in LV (see Figure 24c) and LT (see Figure 22c), the impact, though negative, is similar to the share of payment transactions (*Payments*). In LV and LT, high *DR* values have a strong negative effect on the SME access to credit, while in EE the impact is moderate. On the other end, across all the three countries, a higher number of past financial contracts (*FinContracts*) and long bank-firm relationships (*Relationship*) are both very strong indications for the better SME access to credit. *Relationship* demonstrates a stable shift from the negative to positive impact on the SME access to credit in EE. While a similar trend is observed in LV and LT, it is evident that notable breaks in the relationship duration exist as the companies that have newly formed bank-firm relationships tend to access credit easier than the ones with pre-existing relationships, which suggests that a different customer acquisition strategy exists. Figure 22 also indicates that, in the case of Transaction lending factors, which are related to the financial health of the SME, better values do not necessarily mean a higher ability to access credit. Factors like *RTR* (in EE and LT), *GMR* (in LV and LT), *ATR* (in LV and LT), *TA* (in all the three countries), *CuR* (in EE and LT) have 'breaks' in how individual feature values are contributing to the company's ability to access credit. Such inconsistencies are specifically evident at extremes for both high and low values which tend to reduce the ability to access credit or, in some cases, have a neutral impact, which suggests that the feature's importance is not uniform across all values. *Sector* is a factor of moderate impact across all three countries, with slightly higher negative impact towards the SME access to credit for some specific sectors in EE. An indication of a company's internal overdues (*IOverS*) is



(a) Estonia



(b) Latvia



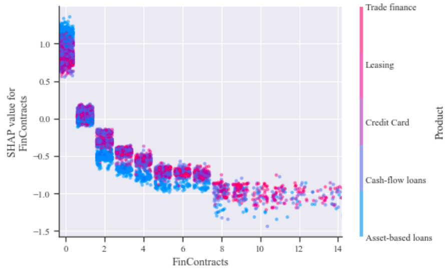
(c) Lithuania

Figure 22. SHAP plots for variables for country-specific SME access to credit models. Created by the author.

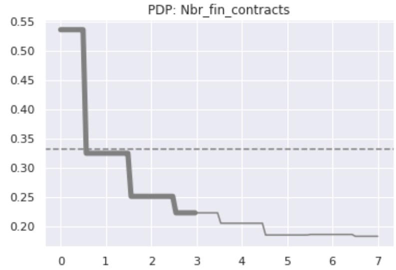
a clear-cut indication that an application is going to be rationed, with the impact being significantly more negative throughout higher overdue amounts. Whilst non-existing external over-dues (*EOverS*) have neither a positive nor a negative effect on the SME access to credit, higher overdue amounts depending on individual cases may have a positive or negative impact on the model's predictions. Similar findings stand for all the three countries in terms of owner overdues (*OIOverS* and *OEOverS*), for which, the impact on prediction is limited and does not provide substantial indication towards any direction suggesting for potential existence of feature interactions. *Defaults*, *ODefaults* and *Diversity* had a close-to-zero impact on the prediction outcome across all the three countries.

It has been determined that the extent of the variable impact on the SME access to credit outcome is not uniform across different feature values and studied countries. Depending on the country, different *Product* selection has a different effect on the SME access to credit outcome. These findings enhance the existing body of research (see Table 11 related to the significance of products in the SME access to credit. Furthermore, it has been demonstrated that a new bank-firm relationship is relatively more likely to be approved than some older ones. Some common trends also exist, e.g., the more intense is the bank-firm relationship (a higher number of financial contracts *FinContracts*, more products, e.g., *Cards*, *POS*), the higher is the likelihood that a company's credit will be approved. This aligns with and supports the conclusions drawn in other studies (see Table 7 on the importance of the bank-firm relationships and also add that completely new relationships could support the SME access to credit. Though, Figure 22 provides insights towards the impact of the individual variables values on the SME access to credit outcome, it does not provide information on the relationships between variables.

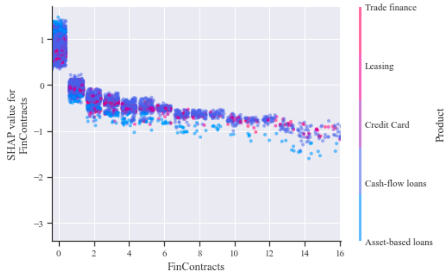
A SHAP dependence plot is used to demonstrate the relationship between a single feature and the model's predictions, while accounting for the values and effects of the other feature. The plot shows the expected impact on the model's predictions across all feature values, whilst individual points show the actual predictions for different instances. A partial dependence plot (PDP), on the other hand, shows the relationship between a single feature and the model's predictions, while averaging out the effects of the other features. The plot shows the expected value of the model's predictions for different values of the feature, but it does not show the individual predictions for different instances. Figures 23, 24 and 25 demonstrate SHAP and PDP dependence plots for high importance *FinContracts*, *Product* and *DR* variables for the three studied countries (see Annexes 5, 6 and 7 for all other feature SHAP and PDP plots).



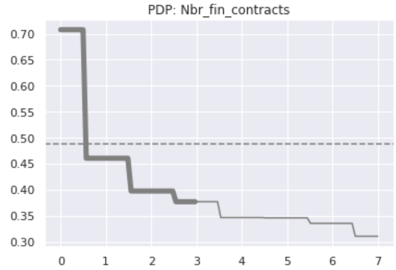
(a) Estonia – SHAP



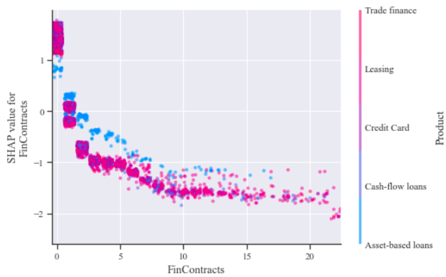
(b) Estonia – PDP



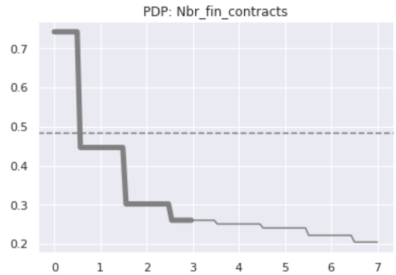
(c) Latvia – SHAP



(d) Latvia – PDP



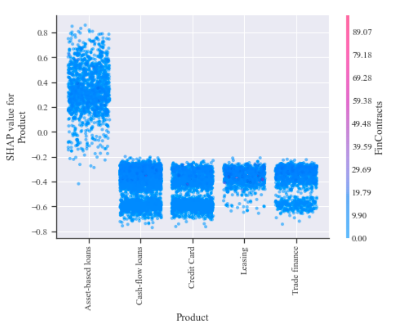
(e) Lithuania – SHAP



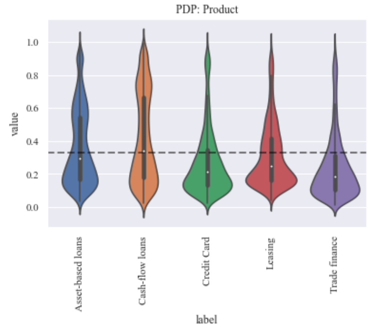
(f) Lithuania – PDP

Figure 23. SHAP dependence and PDP plots for the *FinContracts* variable. Created by the author.

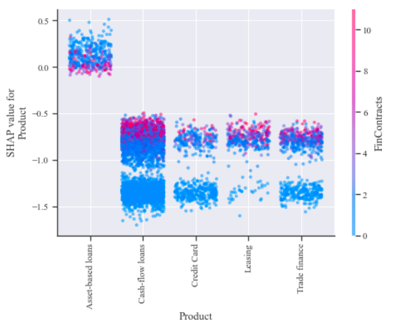
The impact of a number of financial contracts on the SME access to credit is not uniform across different feature values and observed countries. A company that did not have any financial contracts in the past is more than 20 p.p. (EE – 20 p.p., LV – 25 p.p., and LT – 30 p.p.) less likely to be approved than a company that had one (see Figure 23b). In all the three countries, the impact of financial contracts on the SME access to credit is neutral with one contract and is positive with more than one. The positive effect diminishes as the number of financial contracts grows to different levels across countries: for EE, any application with more than 5 historic financing agreements would have only a marginally positive effect, while in LV and LT it is less with 4 and 3, accordingly.



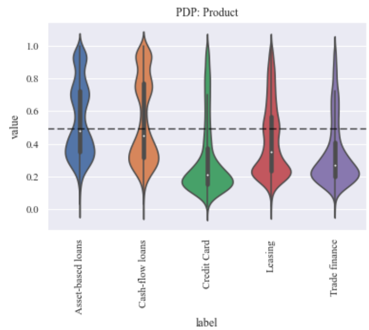
(a) Estonia – SHAP



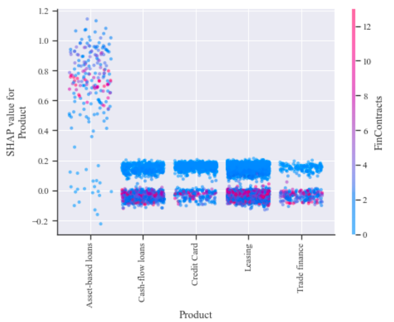
(b) Estonia – PDP



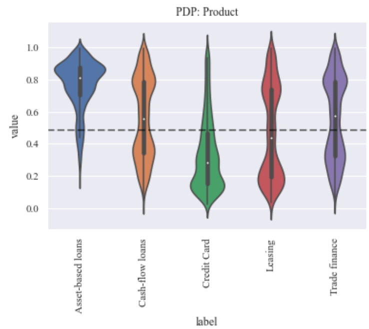
(c) Latvia – SHAP



(d) Latvia – PDP



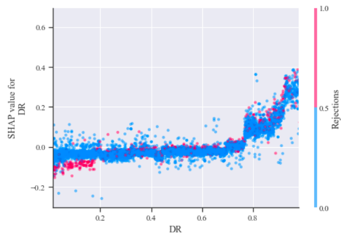
(e) Lithuania – SHAP



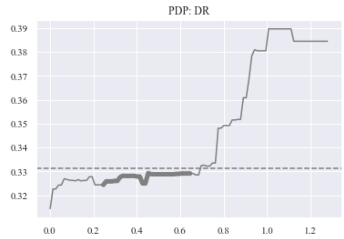
(f) Lithuania – PDP

Figure 24. SHAP dependence and PDP plots for the *Product* variable. Created by the author.

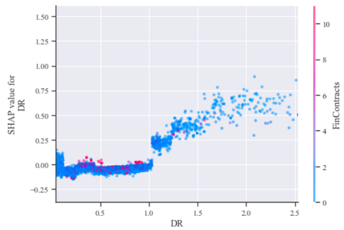
The requested product is one of the most important factors for all the three countries' SME access to credit models, while the extent of the impact varies depending on the individual product that was requested. Across all the three countries, on average, applications for Asset-based loans tend to be more credit-constrained than other products. Figure 24b indicates that the variance of the effect of individual products on credit rationing is high for Cash-flow loans. The impact of other products on the SME access to credit is uniformly positive across the three countries, whilst PDPs for Product values show a significant presence of outliers. Such feature behavior suggests that feature interactions could exist, notably, between the number of financial contracts and products (see Figure 24a). Even though *Relationship* in LV and LT indicates an opposite effect on the SME access to credit for completely new relationships than in EE, the general trend for the majority of cases is shared across all the three countries – the older a bank-firm relationship is, the easier it is to access credit. The threshold when the impact on the SME access to credit becomes positive differs amongst the countries; in EE, it is around 11 years, in LV, it is the highest at around 15 years, while in LT it is the lowest, at around 7 years.



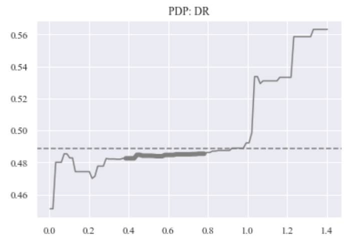
(a) Estonia – SHAP



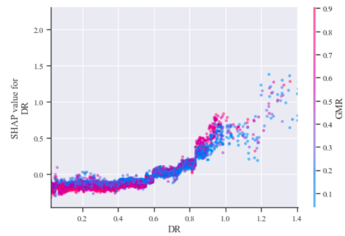
(b) Estonia – PDP



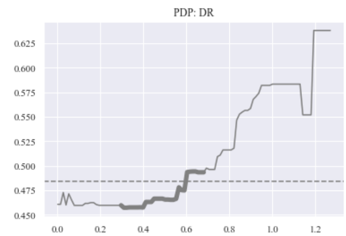
(c) Latvia – SHAP



(d) Latvia – PDP



(e) Lithuania – SHAP

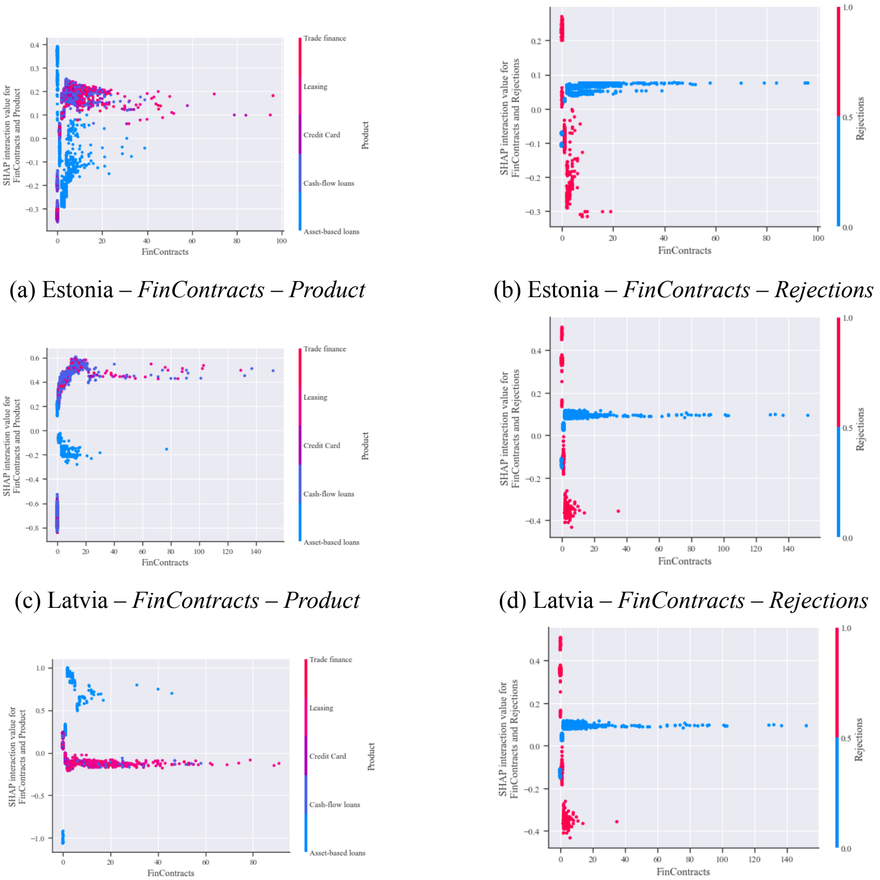


(f) Lithuania – PDP

Figure 25. SHAP dependence and PDP plots for the *DR* variable. Created by the author.

Even though *DR* is one of the most important Transaction Lending factors in LT, the feature is less important in LV and EE, but still it shares similarities amongst all the three countries. The feature provides a neutral indication towards the SME access to credit up until a certain threshold (EE at 0.7, LV at 1.0, LT at 0.6), after which, the ability to access credit starts to diminish. The key difference amongst the countries is the extent to which the feature impacts the underlying probability to access credit – in LT, an application that has sub-threshold debt ratio values will on average be 10 p.p.

more likely to be approved, whilst in LV and EE, the difference is approximately 7 p.p. Similarly to *DR*, *Sector*, whilst having the biggest impact in LV, shows a similar behavior across all the three countries. SMEs operating in the Agriculture, Fishing and Forestry sector tend to suffer from a higher probability to be credit-rationed, while other sectors are relatively uniform. Even though SHAP dependence plot is able to demonstrate the relationship between two features, it does not show, whether variable interactions exists or have any affect on the predicted *Outcome*. To determine how features interact with each other and impact the model's predictions, SHAP interactions are estimated and plotted across all variable pairs, with the strongest being between *FinContracts*, *Product*, *Rejections* (see Figure 26). For other features interaction pairs, see Annexes 8, 9 and 10.



(a) Estonia – *FinContracts – Product*

(b) Estonia – *FinContracts – Rejections*

(c) Latvia – *FinContracts – Product*

(d) Latvia – *FinContracts – Rejections*

(e) Lithuania – *FinContracts – Product*

(f) Lithuania – *FinContracts – Rejections*

Figure 26. SHAP interaction plots for *FinContracts-Product* and *FinContracts-Rejections* variable pairs. Created by the author.

As suggested in the earlier findings on variable interrelations, country-specific SME access to credit models share a common feature interaction between the number of financial contracts *FinContracts* and *Product* and between *FinContracts* and *Rejections*. Although the interaction exists, it has been determined that its impact on the model's predictions is not uniform across all products. Figure 26 demonstrates how the number of financial contracts and the product interaction behaves in terms of model prediction. The dependence is specifically notable for Asset-based loans (see Figures 24a, 24c and 24e). In EE and LV, it has been determined that an interaction for an application from an SME which has more than one financing contract and is requesting an Asset-based loan has a positive contribution towards accessing credit in comparison to other products. Meanwhile, in LT, the effect is opposite as SMEs with existing contracts are less likely to be approved for an Asset-based loan. Those SMEs which do not have any historic financing contracts and are applying for a Credit Card or Cash-flow loan tend to be less rationed than the ones that are applying for other products (see Figures 26a, 26c and 26e). Figures 26b, 26d and 26f indicate that an interaction between *FinContracts* and *Rejections* exists in all the three countries. This interaction has a high negative impact on the SME access to credit when there are no past financial contracts. The negative effect diminishes with at least one financial contract in the past until it starts to become positive as the number of financial contracts grows. Such an interaction suggests that the relationship with the bank in terms of having past financial contracts negates any previous negative information and can work towards improving the ability for SMEs to access credit.

Summary and findings

The conceptual model to evaluate the SME access to credit has been implemented empirically and was carried out for Estonia, Latvia and Lithuania. The empirical study was carried out to evaluate the SME access to credit and the impact of the underlying factors and their groups. The empirical study was conducted in three stages, each concluding with key findings. In Stage I, by utilizing comparative analysis, the underlying SME access to credit has been analyzed. It has been determined that the SME access to credit is not uniform across the studied countries and different periods of time. It is evident that, on average, the SME access to credit is similar in Latvia and Lithuania, while being significantly higher in Estonia. The study has shown that the demand for credit is not uniform across the countries as significant differences exist throughout all underlying factor groups. An average financing application in Estonia is filled-in for a Cash-flow loan product by a smaller and a younger company which has stronger firm-bank relationships and a stronger financial health but has a bigger proportion of liabilities on the balance sheet. In Lithuania, an average application is filled-in for a Leasing product by a company that is average in size, operates in a more urban location, has weaker firm-bank relationships and a relatively strong financial health and a lower number of liabilities. Finally, in Latvia, an average application is filled-in for a Cash-flow loan product by an older company that is bigger in size, operates in a more rural location, has a relatively strong firm-bank relationship but has the weakest financial health. In Stage II, by conducting the dimensionality reduction procedure, country-specific representative feature vectors have been defined. It has been determined that some features within factor groups, specifically, the Transaction Lending factor group, are highly correlated. Therefore, some past period financial ratios and credit history variables which were not contributing additional information were not included in the representative feature vectors. In Stage III, by utilizing state-of-the-art machine learning techniques, the SME access to credit model has been developed. It has been determined that the best modelling technique for estimating the SME access to credit is Gradient Boosting, with the highest modelling accuracy reached in Lithuania and Latvia. Furthermore, the model has demonstrated that the importance and impact of individual factors is not uniform across the countries. The extent of the factor importance depends on the individual factor values and interactions, which suggests a non-linear relationships between the underlying factors and the SME access to credit. It has been determined that, across all the three countries, the most important factor group for estimating the SME access to credit is the Lending technology factors, namely the Relationship Lending factor group. In terms of the impact of individual factors on the SME access to credit, across all the countries, the number of financial contracts is one of the most important factors, which, depending on the number of contracts that a company could have, may result in both positive (when more than one contact in the past exists) or negative effect on the SME ability to access credit. It has been determined that the overall importance of individual features also depends on the specific values, specifically, when financial statement-related factors are at extremes. It has been suggested that, in order to properly evaluate the SME access to credit, it is important to evaluate not only the individual features, but also the entirety of factors and their interactions. Furthermore, not only individual feature values are important, but also their interactions with other features. Strong feature interactions have been identified between the number of financial contracts and the product and between the number of financial contracts and rejections pairs. It has been indicated that those SMEs which have had financing contracts in the past are not suffering from a significantly lower ability to access credit due to past rejections as the negative effect of individual features is canceled out by the positive effect of the feature interaction. These findings add to the existing body of research on the SME access to credit evaluation by demonstrating the existence of non-linear relationships between the underlying conditions and the SME ability to access credit.

CONCLUSIONS

1. SMEs are an integral part of the world economy; therefore, it is important to ensure their financial health specifically by ensuring sufficient ability to access credit. In contrast to larger enterprises, SMEs have difficulties in accessing credit, which negatively impacts their growth opportunities. The constrained ability to access credit has adverse effects on the SME growth in a form of limited international activities, constrained innovation, and lower productivity. Access to credit is vital for the growth and development of SMEs; therefore, it is important to understand the impact of the underlying factors, thus lowering information asymmetries for the companies and regulators.

2. The analysis of SME access to credit and underlying factors has led to 2 major conclusions:

   2.1. It has been determined that, for evaluating the SME access to credit, it is crucial to first define the proxy which would be used to measure it. Previous research has identified two major groups of access to credit proxies: credit supply and credit demand. The selection of the proxy depends on the research problem and data availability. Credit demand studies mainly focus on the factors that influence financing decisions and the reasons for borrower discouragement. On the other hand, credit supply studies tend to focus on either the bank loan portfolio and macro-specific conditions, or on the application outcomes and factors affecting a company's ability to receive approval or be rationed. The latter is divided into several groups which depend on what decision was made: partial approval (or 2nd degree rationing) occurs when the decision to grant credit is positive, but it is rationed through heightened requirements for the financing amount, price, term, or collateral. Rejection (or 1st degree rationing) occurs when the crediting decision is negative, and the company does not receive any loan.

   2.2. The ability for an SME to access credit is defined by the underlying conditions which are split into macro-specific and individual application factors. The macro-specific factors consisting of the Lending Infrastructure and Financial Institution Structure, are beyond the control of individual entities and set the underlying market conditions. A market with strong accounting standards, marked-to-market balance sheets, and active rating agencies can have lower financial constraints and a higher access to credit. Competitive markets tend to have a significant positive effect on access to credit through lower interest rates and higher loan amounts for SMEs. Individual application factors define the underlying characteristics of a potential borrower that are defined by the Lending Technology, Firm Characteristics, and Product Characteristics factor groups. The ability of an SME to access credit is contingent on a range of factors belonging to these factor groups.

A limited number of studies are investigating the SME access to credit in terms of the applied product; therefore, it is not clear whether the choice to apply for a specific financing product can affect the final financing decision. Finally, it has been concluded that it is necessary to consider both macro-specific and individual application factors when estimating SME access to credit.

3. It has been found that the use of state-of-the-art machine learning models remains limited in assessing a company's access to credit. To develop an effective model for assessing the SME access to credit, a range of modelling techniques were used and compared to the benchmark. Such traditional techniques as Discriminant Analysis and Logistic Regression have historically served as benchmarks; however, recent studies have shown that machine learning techniques, such as Random Forest, Multi-layer Perceptron, and Gradient Boosting consistently outperform them. The use of machine learning models has the potential to improve the accuracy and reduce the modelling bias, but the inherent black-box nature of these models means that interpretability may be limited, and that it would require the use of explainability methods. Saliency methods, such as Shapley Additive Explanations, are particularly effective in highlighting the importance of individual features or their interactions.

4. A three-stage SME access to credit evaluation methodology has been developed. It has been defined that the proxy for measuring the SME access to credit is the financing application outcome, which serves as the dependent variable. In order to account for Macro-specific factors, the model is developed for individual countries. The independent variables are grouped into Firm Characteristics, Product Characteristics, and Lending Technology factor groups. In Stage I, the underlying SME access to credit is evaluated based on the comparative analysis of the studied countries. In Stage II, the variable dimensionality reduction procedure is carried out to define a representative feature vector. Finally, in Stage III, state-of-the-art machine learning techniques are employed to estimate the SME access to credit model and evaluate the impact and importance of individual factors, their groups and interactions on ensemble predictions.

5. By applying the SME access to credit evaluation methodology in an empirical setting, three major findings have been drawn:

   5.1. Based on the comparative analysis of the dependent variable and independent variable factor groups, it has been determined that the underlying access to credit for SMEs is not uniformly distributed across the studied countries. The SME access to credit has been found to be higher in Estonia and lower in Latvia and Lithuania. Furthermore, the analysis has revealed that the underlying Firm Characteristics, Product Characteristics, Lending

Technology factor groups are not homogeneous across the incoming applications flow for the three studied countries. The relative significance of individual factors is influenced by the actual values, specifically, for the financial statement-related factors at their extremes. This demonstrates that, in order to properly evaluate the SME access to credit, it is important to evaluate the entirety of factors and their interactions. In Estonia, the average SME application is received from a younger, smaller, and less diversely owned company which has a stronger bank-firm relationship, stronger financial health, and relatively more overdues which do not materialize as defaults. In Latvia, the average application is received from an older and larger company that is more likely to provide audited financial statements, has relatively strong bank-firm relationships, but weaker financial health and a higher probability of historic defaults. Finally, in Lithuania, an average SME application is received from a larger company which requested a leasing product and has a relatively shorter and less intense bank-firm relationship.

5.2. The country-specific SME access to credit models have been estimated by utilizing Random Forest, Multi-layer Perceptron, Gradient Boosting state-of-the-art machine learning techniques and the traditional Logistic Regression which was employed as a benchmark. It has been determined that the Gradient Boosting modelling technique is the most accurate for estimating the SME access to credit, with the highest accuracy being observed in Lithuania and Latvia. Logistic regression, which is one of the most commonly employed modelling techniques, demonstrated the worst classification accuracy.

5.3. It has been determined that the most critical factor group for estimating the SME access to credit is the Lending technology factors, specifically, the Relationship Lending factor group. Among the individual factors, the number of financial contracts is one of the most important variables across all countries, with the extent of the impact on the prediction depending on the actual financial contract number. Additionally, the study has shown that not only individual feature values but also their interactions with other features are critical. Strong feature interactions have been identified between the number of financial contracts and the product and between the number of financial contracts and rejections pairs, which indicates that SMEs that have had financing contracts in the past do not suffer from lower access to credit due to past rejections.

6. Based on the conducted analysis and the empirical model's findings, it is evident that the evaluation of the SME access to credit requires comprehensive modelling, utilization of a wide array of factors and state-of-the-art modelling techniques. The empirical application of the SME access to credit model and the

use of state-of-the-art modelling techniques has demonstrated that the importance of the underlying factors is not uniform across the actual feature values. The non-linear relationship between some feature values and the SME access to credit is evident. For example, the duration of the bank-firm relationship holds a general notion that longer relationships yield higher ability to access credit, whilst it has been demonstrated that, in fact, new bank-firm relationships have a significantly higher likeliness to access credit. These findings provide valuable insights to not only companies which are trying to access credit (in terms of establishing new bank-firm relationships) but also to regulators who are searching for measures to improve the SME ability to access credit. Further studies could follow multiple directions, such as the improvement of the SME access to credit modelling by utilizing a wider array of the underlying factors ranging from the company owner's personal characteristics to the metrics related to company sustainability. Another potential direction could be the integration of discouraged borrowers to formulate a general SME access to credit model, which would be able to evaluate the access to credit from both the credit demand and supply angles.

SANTRAUKA

ĮVADAS

Temos aktualumas. Mažos ir vidutinės įmonės (MVĮ) vaidina svarbų vaidmenį pasaulio ekonomikoje. Vienas iš pagrindinių iššūkių, su kuriuo jos susiduria, yra ribotos galimybės gauti finansavimą iš išorės kredito teikėjų (Muller ir kt., 2022). Ribotos galimybės gauti išorinį kreditą mažina įmonių pardavimus, likvidumą, sukuria tiekimo grandinės trikdžius, o tai neigiamai veikia augimo galimybes ir gali sąlygoti darbuotojų atleidimus bei bankrotus (Khan, 2022). Palyginti su didelėmis įmonėmis, MVĮ gali gauti kreditus mažiau palankesnėmis finansavimo sąlygomis, tokiomis kaip trumpesni kredito sutarties terminai, didesni užstato reikalavimai ir didesnės palūkanų normos (Chodorow-Reich ir kt., 2022). Tokios nepalankios sąlygos gali būti siejamos su informacijos asimetrija, dėl kurios finansavimo teikėjai negali tinkamai įvertinti įmonės kreditingumo. Ribotos MVĮ galimybės gauti kreditą yra siejamos su įvairiais individualiais veiksniais, įskaitant ribotą užstato prieinamumą, silpnesnę finansinę būklę ir didesnį jautrumą pokyčiams ne tik konkrečioje pramonės šakoje, bet ir visoje rinkoje (Angori ir kt., 2019). Moksliniuose tyrimuose taip pat analizuojami ir makrospecifiniai veiksniai[1], tokie kaip skolinimo infrastruktūra ir finansų institucijų struktūra, lemiantys kredito rinkos sąlygas, kurios taip pat gali turėti įtakos MVĮ galimybėms gauti kreditą. Konkurencingos rinkos paprastai sudaro palankesnes sąlygas kredito prieinamumui, nes MVĮ gali gauti tiek didesnes paskolų sumas, tiek paskolos gali būti su mažesnėmis palūkanų normomis (Kärnä and Stephan, 2022). Siekiant skatinti MVĮ augimą, labai svarbu gerinti MVĮ kredito prieinamumą, todėl būtina didinti informacijos skaidrumą, identifikuoti išorinius ir vidinius veiksnius ir įvertinti pagrindines sąlygas, darančias įtaką kredito prieinamumui mažoms ir vidutinėms įmonėms.

Šioje disertacijoje siūlomas kredito prieinamumo mažoms ir vidutinėms įmonės vertinimo modelis, skirtas identifikuoti kredito prieinamumą lemiančius veiksnius, kurie yra svarbūs mažoms ir vidutinėms įmonėms, taip pat ir valstybinėms teisinio reguliavimo institucijoms. Kredito prieinamumo ir jį lemiančių veiksnių mažoms ir vidutinėms įmonėms kompleksinis vertinimas leistų padidinti MVĮ informacijos skaidrumą ir pagerinti galimybes gauti išorinį kreditą.

Mokslinė problema ir jos ištyrimo lygis. Atlikta literatūros analizė parodė, jog kredito prieinamumas MVĮ gali būti nagrinėjamas keletu būdų, atsižvelgiant į jį veikiančius veiksnius. Šioje disertacijoje kredito prieinamumas yra suprantamas kaip MVĮ įmonių galimybės gauti kreditą iš tradicinių komercinių bankų. Tyrimai, nagrinėjantys kredito prieinamumą MVĮ, gali būti suskirstyti į tiriančius kredito paklausą ir tiriančius kredito pasiūlą (Maier, 2016; Angori ir kt., 2019; Altavilla ir kt., 2021). Kredito paklausos tyrimuose nagrinėjami veiksniai, turintys įtakos skolininko sprendimams

[1] Angliškas terminas 'macro-specific factors'. Kadangi lietuvių kalboje nėra visuotinai priimto termino, šioje disertacijoje terminas 'macro-specific factors' verčiamas kaip 'makrospecfiniai veiksniai'.

kreiptis į kredito institucijas siekiant gauti paskolą (Mac An Bhaird ir kt., 2016; Nguyen ir kt., 2021; Altavilla ir kt., 2021). Kredito pasiūlos tyrimuose daugiausia dėmesio skiriama banko paskolų portfelio formavimui esant tam tikroms makrospecifinėms sąlygoms (Bolton ir kt., 2016; Peón ir Guntín, 2021; Altavilla ir kt., 2021) arba paraiškų teikimo rezultatų analizei ir veiksniams, identifikuojantiems įmonės galimybes gauti paskolos patvirtinimą arba atmetimą (Kirschenmann, 2016; Chodorow-Reich ir kt., 2021). Kredito pasiūlos ar kredito paklausos pasirinkimas kredito prieinamumui MVĮ vertinti priklauso nuo tyrimo problemos ir duomenų prieinamumo (Lee ir kt., 2015). Nors MVĮ kredito prieinamumo vertinimas per banko paskolų portfelį yra naudingas vertinant makroekonomikos poveikį kreditų pasiūlai, tačiau gali būti neatsižvelgta į konkrečias kiekvienai įmonei būdingas ypatybes, kurios gali būti labai svarbios norint suprasti veiksnius, turinčius įtakos gauti kreditą. Kredito prieinamumo MVĮ vertinimas per kreditavimo paraiškų rezultatus gali būti detalesnis, tačiau priklauso nuo turimų duomenų prieinamumo ir patikimumo (Kirschenmann, 2016). Mokslininkai, tyrinėdami MVĮ galimybes gauti kreditą, analizuodami paraiškos ir sprendimo rezultatus, gali suprasti konkrečias priežastis, dėl kurių sumažėja ar padidėja galimybė gauti kreditą, taip pat išskirti atvejus, kai išduodami vadinamieji sąlyginiai patvirtinimai. Kai finansavimo paraiškos atmetamos, tai vadinama pirmo laipsnio normavimu, o kai finansavimo paraiškos patvirtinamos, tačiau pakoreguojamos produkto sąlygos, tokios kaip finansavimo suma, kaina, terminas ir (arba) prašomas užstatas, yra vadinama antrojo laipsnio normavimu (Jiménez ir kt., 2012; Molina ir Preve, 2012; Berger ir kt., 2022).

MVĮ kredito prieinamumui daro įtaką įvairūs veiksniai, kurie skirstomi į makrospecifinius ir individualius paraiškos veiksnius (Berger ir Udell, 2006). Makrospecifiniai veiksniai yra svarbūs nustatant visos kredito rinkos skolinimo ir skolinimosi galimybes, taigi ir MVĮ prieigą prie kreditų (Dobbie ir kt., 2020; Angori ir kt., 2020). Plačiai mokslininkų tiriami ir individualūs paraiškos veiksniai, kurie nurodo pagrindines potencialaus skolininko savybes ir gali būti suskirstyti į skolinimo technologijų, įmonės charakteristikų ir produkto charakteristikų veiksnių grupes (Berger ir Udell, 2006). Šie veiksniai yra būdingi atskiroms įmonėms ir yra svarbūs nustatant MVĮ galimybes gauti kreditą. Mokslinėje literatūroje skolinimo technologijų veiksniai apibrėžiami kaip technologinės priemonės, kurias skolintojai naudoja vertindami MVĮ kreditingumą, ir yra skirstomi į dvi grupes: sandorio skolinimo[2] ir santykių skolinimo[3] technologijas. Sandorio skolinimo metu atsižvelgiama į įmonių finansinių ataskaitų duomenis ir kredito istoriją (Motta ir Sharma, 2020), todėl šis skolinimo technologijų tipas paprastai naudojamas didesniems ir skaidresniems skolininkams, kurie teikia audituotas ir išsamias finansines ataskaitas (Palazuelos ir kt., 2018; Ferri ir kt., 2019). Santykių skolinimas yra tinkamiausias būdas, kai informacija apie įmonę yra ribota, o kreditingumas gali būti vertinamas remiantis ankstesniais banko ir įmonės santy-

[2] Angliškas terminas 'transaction lending'. Kadangi lietuvių kalboje nėra visuotinai priimto termino, šioje disertacijoje terminas 'transaction lending' verčiamas kaip 'sandorio skolinimas'.

[3] Angliškas terminas 'relationship lending'. Kadangi lietuvių kalboje nėra visuotinai priimto termino, šioje disertacijoje terminas 'relationship lending' verčiamas kaip 'santykių skolinimas'.

kiais (Durguner, 2017; Rabetti, 2022). Šie technologiniai skirtumai turi didelę įtaką MVĮ galimybei gauti kreditą, nes individualus skolininkas gali užsitikrinti finansavimą pagal veiksnius, priklausančius tiek vienai, tiek kitai skolinimo technologijai (Angori ir kt., 2019; Chodorow-Reich ir kt., 2022). Kita individualių paraiškos veiksnių grupė – įmonės charakteristikos – yra unikali kiekvienai įmonei ir gali apimti tokius veiksnius, kaip įmonės dydis, amžius, sektorius (Mina ir kt., 2013) bei įmonės valdymo struktūra (Aterido ir Hallward-Driemeier, 2011; Sikochi, 2020; de Andrés ir kt., 2021). Paskutinė individualių paraiškos veiksnių grupė – produkto charakteristikos – yra paremta finansavimo produkto savybėmis, tokiomis kaip reikalingo užstato suma (Gurara ir kt., 2020; Berger ir kt., 2022), palūkanų norma (Xu ir kt., 2020; Kärnä ir Stephan, 2022), ir sutarties terminu (Minnis ir Sutherland, 2017; Aoki, 2021). Ferri ir kt. (2019) pastebi, kad atliekant kredito prieinamumo MVĮ vertinimą reikėtų naudoti kintamuosius, priklausančius abiem veiksnių grupėms. Iš analizuotų empirinių tyrimų nustatyta, jog kredito prieinamumo mažoms ir vidutinėms įmonėms vertinimo būdai ir metodai yra riboti, apimantys tik pavienes veiksnių grupes, pasigendama kompleksinio bendro veiksnių grupių vertinimo, juo labiau kad, vystantis finansinėms technologijoms, tobulėjant dirbtiniam intelekto pritaikomumui ir atsirandant vis naujiems mašininio mokymosi metodams, kredito prieinamumo vertinimas nėra baigtinis ir sisteminis procesas.

Mokslinėje literatūroje nagrinėjamos skirtingos galimybės naudotis kredito prieinamumo modeliavimo metodais, daugiausia dėmesio skiriant paraišką pateikusių įmonių kreditingumui vertinti (Molina ir Preve, 2012; Kruppa ir kt., 2013; Pal ir kt., 2016). Pažymima, jog tinkamai įvertinti kredito prieinamumą yra sudėtinga dėl nepriklausomų kintamųjų daugialypiškumo, duomenų prieinamumo ir žmogiškojo faktoriaus vertinant įmonės skolinimosi galimybes (Dastile ir kt., 2020). Todėl, siekiant užtikrinti tikslaus kredito prieinamumo modelio sukūrimą, svarbu atsižvelgti tiek į skirtingas veiksnių grupes, tiek į tinkamą mašininio mokymosi metodo pasirinkimą. Analizuoti moksliniai tyrimai atskleidė, kad, modeliuojant kredito prieinamumą MVĮ, makrospecifiniai veiksniai gali būti ir neįtraukiami, jeigu yra kuriami individualūs šalių modeliai (Angori ir kt., 2020; Calabrese ir kt., 2022; Kärnä ir Stephan, 2022). Siekiant ištirti kredito prieinamumą MVĮ, empiriniuose tyrimuose yra naudojami įvairūs mašininio mokymosi metodai, pradedant tradiciniais metodais, tokiais kaip diskriminacinė analizė (Barboza ir kt., 2017) ir logistinė regresija (Wang ir kt., 2020; Malakauskas ir Lakštutienė, 2021; Medianovskyi ir kt., 2023), iki pažangiausio mašininio mokymosi technikų, tokių kaip sprendimų medžiai (Trivedi, 2020), atsitiktiniai miškai (Medianovskyi ir kt., 2023), dirbtiniai neuroniai tinklai (Hadji Misheva ir kt., 2021), atraminių vektorių klasifikatorius (Silva ir kt., 2020) ir k-artimiausio kaimyno klasifikatorius (Barboza ir kt., 2017; Malakauskas ir Lakštutienė, 2021; Medianovskyi ir kt., 2023). Nors tradiciniai metodai, tokie kaip logistinė regresija, turi pranašumą dėl paprastesnės interpretacijos ir modelių stabilumo, jie nėra tinkami didesniems duomenų masyvams bei esamiems ryšiams tarp kintamųjų modeliuoti (Correa Bahnsen ir kt., 2016). Kita vertus, mašininio mokymosi metodai tampa vis populiaresni, nes jie suteikia galimybę pagerinti modeliavimo tikslumą ir sumažinti kredito vertinimo modelių šališkumą. Mokslininkai, naudodami mašininio mokymosi metodus, tyrė kredito prieinamumą MVĮ ir nustatė metodus, galinčius pasiekti didelį tikslumą, tokius kaip atsitiktiniai miškai, gradialinis nusileidimas, daugiasluoksnis perceptronas (Barboza ir kt., 2017; Malakauskas ir Lakštutienė, 2021; Medianovskyi ir kt., 2023). Mokslinė literatūros analizė atskleidė, jog kiekvienas modeliavimo metodas turi savo privalumų ir trūkumų, o tinkamiausios technikos pasirinkimas priklauso nuo konkrečios mokslinės problemos (Preece ir kt., 2018). Mašininio mokymosi metodams būdinga riboto interpretavimo savybė lemia, kad, norint įvertinti kintamųjų svarbą, reikia naudoti paaiškinamumo metodus, tokius kaip Šaplio papildomi paaiškinimai[4] (SHAP) (Arya ir kt., 2019; Arrieta ir kt., 2020). Mokslinai tyrimai atskleidė, kad mašininio mokymosi metodų naudojimas gali pagerinti kredito prieinamumo MVĮ vertinimo tikslumą ir nustatyti netiesines priklausomybes ir kintamųjų sąveiką.

Tyrimai atskleidė, jog, nors egzistuoja didelė apimtis mokslinių tyrimų, analizuojančių atskirų veiksnių grupių daromą įtaką vertinant kredito prieinamumą MVĮ, tačiau nėra aiški kiekvienos veiksnių grupės, atskirų veiksnių bei jų sąveikos svarba. Be to, empirinių tyrimų rezultatai neidentifikuoja mašininio mokymosi metodų, atitinkančių tikslumui keliamus reikalavimus ir tinkančius kredito prieinamumui MVĮ vertinti, todėl atsiranda poreikis iš kelių mašininio mokymosi metodų nustatyti tiksliausią ir tinkantį kredito prieinamumui MVĮ vertinti. Akivaizdu, kad MVĮ kredito prieinamumo vertinimas yra daugiamatė problema, kuriai reikia sisteminio požiūrio nustatant MVĮ prieigą prie kreditų, identifikuojant pagrindinius veiksnius ir pasirenkant tinkamus mašininio mokymosi metodus. Kadangi mokslinėje literatūroje pasigendama kompleksinio kredito prieinamumo MVĮ vertinimo tyrimų taikant mašininio mokymosi metodus, šioje disertacijoje kredito prieinamumas mažoms ir vidutinėms įmonėms kompleksiškai vertinamas pirmą kartą, taikant Baltijos šalių pavyzdį. Disertacijos temos aktualumas ir darbo problema grindžiama poreikiu sukurti modelį, kuris padėtų kompleksiškai įvertinti kredito prieinamumą mažoms ir vidutinėms įmonėms ir nustatyti jį lemiančių veiksnių bei jų sąveikos svarbą. Disertacijoje nagrinėjama tyrimo problema – kaip įvertinti mažų ir vidutinių įmonių kredito prieinamumą?

Tyrimo objektas – veiksniai, lemiantys mažų ir vidutinių įmonių kredito prieinamumą.

Tyrimo tikslas – sukurti kompleksinį kredito prieinamumo mažoms ir vidutinėms įmonėms vertinimo modelį ir jį empiriškai patikrinti.

[4] Angliškas terminas ‘Shapley Additive Explanations’. Kadangi lietuvių kalboje nėra bendrai priimto termino, šioje disertacijoje ‘Shapley Additive Explanations’ verčiamas kaip ‘Šaplio papildomi paaiškinimai’.

Tyrimo uždaviniai:

1. Išanalizuoti kredito prieinamumo svarbą mažoms ir vidutinėms įmonėms.
2. Nustatyti rodiklius, naudojamus matuojant MVĮ prieigą prie kredito, ir pagrindinius veiksnius, lemiančius MVĮ galimybę gauti kreditą.
3. Išanalizuoti mašininio mokymosi metodus kredito prieinamumui mažoms ir vidutinėms įmonėms vertinti.
4. Sukurti kompleksinį kredito prieinamumo mažoms ir vidutinėms įmonėms vertinimo modelį.
5. Patikrinti kredito prieinamumo mažoms ir vidutinėms įmonėms vertinimo modelio taikymo galimybes empiriškai pagal Baltijos šalių pavyzdį.

Tyrimo metodai. Siekiant nustatyti MVĮ kredito prieinamumo svarbą, tam įtakos turinčius veiksnius ir modeliavimo metodus, atliekamas analitinis tyrimas, apimantis mokslinės literatūros sisteminimą, palyginimą, apibendrinimą, analizę ir sintezę. Norint empiriškai įvertinti MVĮ kredito prieinamumą, naudojami šie metodai: aprašomoji statistinių duomenų analizė, skirta lyginamajai analizei nustatant esamą MVĮ kredito prieinamumą, kintamųjų skaičiaus mažinimo procesas, kuriame naudojami koreliaciniai šilumos žemėlapiai, ir euklidinis atstumo klasterizavimas reprezentatyviam veiksnių rinkiniui sukurti. Taip pat naudojami mašininio mokymosi metodai (logistinė regresija, atsitiktiniai miškai, daugiasluoksnis perceptronas ir gradialinis nusileidimas), skirti įvertinti MVĮ kredito prieinamumui, ROC ir Vidutinio tikslumo kreivės, kad būtų galima įvertinti apskaičiuotų modelių našumą, vidutinis absoliutus SHAP ir PFI svarbos metrika, skirti globaliai kintamųjų svarbai įvertinti, SHAP ir dalinės priklausomybės grafikai, skirti vietiniai kintamųjų svarbai ir jų sąveikai įvertinti.

Tyrimo informaciniai šaltiniai ir duomenų bazė. Teorinei darbo daliai parengti ir modeliui sudaryti buvo naudojami moksliniai tyrimai, publikuoti moksliniuose leidiniuose ir įtraukti į šias duomenų bazes: Elsevier, CA Web of Science, Scopus, EBSCO, Emerald Management, Springer, Google Scholar. Taip pat darbe naudoti vienos Baltijos šalyse veikiančios kredito įstaigos paraiškų duomenys. Duomenys apima MVĮ kredito paraiškas, gautas 2018–2022 metų periodu.

Tyrimo apribojimai

1. Sukurtas kredito prieinamumo MVĮ modelis neatsigręžia į formalius ir neformalius ryšius tarp paraišką teikiančių MVĮ ir kitų didesnių grupės įmonių, traktuodamas jas kaip vieną kategoriją. Šis homogenizavimas gali turėti įtakos mokslinių tyrimų rezultatams bei išvadoms, ypač MVĮ, kurios turi stiprų grupės įmonių palaikymą, nes joms gali būti taikomi kitokie kredito standartai, palyginti su nepriklausomomis MVĮ.

2. Sukurtas kredito prieinamumo MVĮ modelis naudoja tik finansavimo paraiškų patvirtinimo ir 1-ojo laipsnio normavimo rezultatus, kas apriboja kredito prieinamumo įvertinimą ir neatsižvelgia į sąlyginius patvirtinimus (2-ojo laipsnio normavimas) ir atvejus, kai potencialūs skolininkai buvo visiškai atgrasyti nuo kredito paraiškos užpildymo. Atsižvelgiant į juos, kredito prieinamumo MVĮ vertinimas galėtų būti išsamesnis ir suteikti platesnės informacijos apie įmonių galimybes gauti kreditą.

Mokslinis naujumas ir tyrimo išvadų reikšmė. Disertacijos rezultatai papildo ir praplečia mokslinę literatūrą, susijusią su kredito prieinamumo mažoms ir vidutinėms įmonėms vertinimu, ir suteikia naujų žinių apie mašininio mokymosi metodų panaudojimo galimybes. Šioje disertacijoje kredito prieinamumo vertinimo tarpkryptiškumas, sujungiant ekonomikos ir matematikos mokslo kryptis, parodo nagrinėjamos mokslinės problemos sudėtingumą, aktualumą bei svarbą teoriniu bei praktiniu požiūriais. Siūlomas originalus kredito prieinamumo mažoms ir vidutinėms įmonės vertinimo modelis yra skirtas kredito prieinamumą lemiantiems veiksniams identifikuoti, kurie yra svarbūs mažoms ir vidutinėms įmonėms, taip pat ir valstybinėms teisinio reguliavimo institucijoms. Disertacijos mokslinis naujumas atsiskleidžia kompleksinio kredito prieinamumo MVĮ vertinimo metodikoje, kurioje atsižvelgiama ne tik į pagrindines veiksnių grupes, bet ir į atskirus veiksnius bei jų sąveiką. Kuriant konceptualų modelį ir atrenkant tiksliausius mašininio mokymosi metodus, disertacijoje siūlomas trijų etapų MVĮ kredito prieinamumo vertinimo modelis, kuris empiriškai pritaikytas Baltijos šalių pavyzdžiu. Sukurtas modelis sujungia naujausius MVĮ kredito prieinamumo tyrimus, pasiūlo veiksnių atrankos procedūras ir naudoja naujausius modeliavimo bei paaiškinamumo metodus. Šis modelis yra originalus ir pirmą kartą taikomas empiriniuose tyrimuose, kas leidžia visapusiškai įvertinti kredito prieinamumą MVĮ ir identifikuoti tiek individualių veiksnių, tiek veiksnių sąveikos svarbą. Sukurtas kredito prieinamumo MVĮ vertinimo modelis – tai įrankis, leidžiantis įvertinti MVĮ galimybes gauti kreditą, identifikuojantis pagrindinius kredito prieinamumą lemiančius veiksnius bei padedantis priimti argumentuotus sprendimus galimiems kredito pareiškėjams, vertinant savo galimybes gauti kreditą, skolintojams, vertinant gautas paraiškas, taip pat ir valstybės reguliacinėms institucijoms, skatinančioms smulkaus ir vidutinio verslo plėtrą.

Empirinio tyrimo rezultatai atskleidžia, kad kredito prieinamumas MVĮ skirtingose šalyse nėra vienodas, o pagrindinių veiksnių grupių ir atskirų veiksnių svarba ir poveikis taip pat skiriasi. Disertacijos išvados rodo, kad veiksnių bei jų sąveikos svarba MVĮ kredito prieinamumui tarp skirtingų šalių yra nevienoda, todėl politikos formuotojai ir finansų institucijos, kurdami politiką ir produktus, padedančius MVĮ gauti kreditą, turi atsižvelgti į konkrečius šalies veiksnius. Be to, tyrimo rezultatai suteikia įžvalgų apie veiksnius, kurie yra labai svarbūs, kai MVĮ siekia gauti kreditą. Tolesni tyrimai gautų disertacijos išvadų pagrindu galėtų pratęsti mokslinius tyrimus, susijusius su kredito prieinamumu, ir padėti kurti veiksmingesnę politiką ir finansinius produktus, siekiant pagerinti MVĮ galimybes gauti kreditą. Šis modelis yra nepriklausomas nuo šalies specifikos, todėl gali būti taikomas įvairiose šalyse.

Loginė disertacijos struktūra. Disertaciją sudaro įvadas, 3 dalys, išvados ir literatūros sąrašas. Disertacija parengta pasitelkus literatūros šaltinius. Įvade pristatomas disertacijos temos aktualumas, mokslinė problema ir jos ištyrimo lygis, tyrimo objektas, tikslas ir uždaviniai, taip pat naudojami tyrimo metodai, disertacijos mokslinis naujumas ir galimas praktinis rezultatų pritaikymas. Pateikta disertacijos loginė struktūra. Pirmoje disertacijos dalyje analizuojama kredito prieinamumo svarba mažoms ir vidutinėms įmonėms, nustatomi MVĮ kredito prieinamumo matavimo indikatoriai ir pagrindiniai veiksniai, lemiantys MVĮ galimybę gauti kreditą, analizuojami kredito prieinamumo MVĮ mašininio mokymosi modeliavimo ir paaiškinamumo metodai. Pirmos dalies pabaigoje suformuojamas kredito prieinamumo vertinimo konceptualus modelis. Antroje disertacijos dalyje aprašoma modelio kūrimo metodika ir galutinai suformuojamas originalus kredito prieinamumo MVĮ vertinimo modelis. Trečioje dalyje kredito prieinamumo MVĮ vertinimo modelis yra empiriškai patikrintas Baltijos šalių pavyzdžiu. Gauti reikšminiai rezultatai sisteminami ir apibendrinami. Disertacijos pabaigoje pateikiamos bendros išvados. Disertacijoje pateikiami disertacijos santrauka, literatūros sąrašas, disertacijos autoriaus mokslinių publikacijų disertacijos tema sąrašas, mokslinių konferencijų, kuriose buvo pristatyti disertacijos tyrimo rezultatai, sąrašas, autoriaus gyvenimo aprašymas ir priedai.

Disertacijos apimtis: 184 puslapiai be priedų. Darbą sudaro 45 lentelės, 26 paveikslai, 10 priedų. Literatūros sąraše – 273 nuorodos.

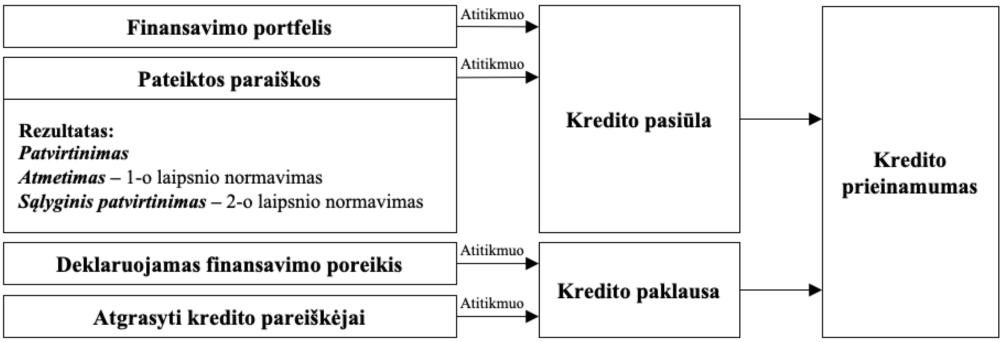
1. KREDITO PRIEINAMUMO MAŽOMS IR VIDUTINĖMS ĮMONĖMS LEMIANČIŲ VEIKSNIŲ VERTINIMO TEORINIS PAGRINDIMAS

Pirmoje disertacijos dalyje yra sprendžiami 1-as, 2-as ir 3-ias uždaviniai. Šioje dalyje analizuojama kredito prieinamumo svarba mažoms ir vidutinėms įmonėms, nustatomi kredito prieinamumo MVĮ matavimo indikatoriai ir pagrindiniai veiksniai, lemiantys MVĮ galimybę gauti kreditą, analizuojami kredito prieinamumo MVĮ vertinimo mašininio mokymosi metodai.

Disertacijoje mažos ir vidutinės įmonės yra apibrėžiamos pagal Europos Komisijos apibrėžties reikalavimus, kur MVĮ įvardijama kaip įmonė, kurioje dirba mažiau nei 250 darbuotojų, o metinė apyvarta nesiekia 50 mln. eurų arba balanso suma nesiekia 43 mln. eurų. Kredito prieinamumas yra apibrėžiamas kaip įmonių galimybė gauti kreditą iš komercinio banko. Kredito prieinamumo mažoms ir vidutinėms įmonėms vertinimas yra viena iš pagrindinių temų tiek akademikams, tiek reguliacinės politikos formuotojams visame pasaulyje. Mokslinius tyrimus, nagrinėjančius kredito prieinamumo vertinimą, galima suskirstyti į dvi kategorijas: tiriančius kredito pasiūlą ir tiriančius kredito paklausą. Kredito pasiūlos tyrimus galima išskaidyti į dvi pagrindines kryptis, analizuojančias bankų paskolų portfelį (Miao ir Wang, 2012; Molina ir Preve, 2012; Deyoung ir kt., 2015; Bolton ir kt., 2016; Peón ir Guntín, 2021; Altavilla ir kt., 2021) arba analizuojančias kreditavimo paraiškų rezultatus (Jiménez ir kt., 2012; Kirschenmann, 2016; Berger ir kt., 2022; Chodorow-Reich ir kt., 2021). Kredito prieinamumo vertinimas per bankų paskolų portfelį paprastai naudojamas makroekonominiam poveikiui įvertinti, nes yra lengviau kiekybiškai įvertinamas ir stebimas, tačiau jis yra nepakankamas, kai reikia įvertinti atskirus kredito prieinamumą lemiančius veiksnius. Šis būdas neatsižvelgia į įmonei būdingus individualius veiksnius, kurie gali būti labai svarbūs norint suprasti ir įvertinti kredito prieinamumą. Tačiau kredito prieinamumo vertinimas, tiriant atskirų finansavimo paraiškų rezultatus, gali būti ir itin detalus, bet taip pat priklauso nuo duomenų šaltinyje pateiktos detalios informacijos ir duomenų rinkinio patikimumo (Lee ir kt., 2015). Kredito prieinamumo vertinimas per kredito paraiškos rezultatus leidžia suprasti sudėtingesnio arba paprastesnio kredito prieinamumo priežastingumą.

Galimybė gauti kreditą gali būti vertinama ne tik tais atvejais, kai potencialus skolininkas bandė gauti finansavimą, bet ir tais atvejais, kai įmonė buvo atgrasyta nuo paraiškos užpildymo (Mac An Bhaird ir kt., 2016; Nguyen ir kt., 2021). Todėl kredito prieinamumo vertinimas, tiriant kredito paklausą, gali padėti nustatyti konkrečius veiksnius, nulemiančius įmonės finansavimo poreikį arba atgrasančius nuo kredito paraiškos pateikimo (Angori ir kt., 2019). Šis kredito prieinamumo vertinimo būdas leidžia atsižvelgti ne tik į kreditavimo poreikius, pateiktus kredito paraiškoje, bet ir į veiksnius, kurie atgraso įmones nuo kredito paraiškos užpildymo (Mac An Bhaird ir kt., 2016; Nguyen ir kt., 2021; Altavilla ir kt., 2021).

Neigiami finansavimo paraiškos rezultatai paprastai skirstomi į du tipus: 1-ojo laipsnio normavimas, kai finansavimo paraiška visiškai atmetama (Jiménez ir kt., 2012), ir 2-ojo laipsnio normavimas (Berger ir kt., 2021), kai finansavimo paraiška patvirtinama, bet koreguojamos kredito gavimo sąlygos, tokios kaip finansuojama suma, paskolos palūkanos, terminas ir užstatas. Tam tikrais atvejais, pakoregavus produkto sąlygas į sudėtingesnes, įmonė gali negauti kredito (Kirschenmann, 2016; Durguner, 2017).



27 pav. Kredito prieinamumo vertinimo rodikliai. Sudaryta autoriaus

MVĮ kredito prieinamumui įtaką daro įvairūs veiksniai, kurie skirstomi į makrospecifinius ir individualius paraiškos veiksnius (Berger ir Udell, 2006). Makrospecifiniai veiksniai, tokie kaip skolinimo infrastruktūra ir finansų institucijų struktūra, sąlygoja pagrindines skolinimo rinkos sąlygas, taip pat prieigą prie kreditų MVĮ (Dobbie ir kt., 2020; Angori ir kt., 2020). Tyrimai rodo, kad rinka, kurioje įmonėms taikomi griežti apskaitos standartai (Florou ir Kosi, 2015; Deno ir kt., 2020), sudaromos finansinės ataskaitos pagal rinkos vertę (Adrian ir Shin, 2010) ir veikia reitingų agentūros, turi mažesnius finansinius suvaržymus (Bosch ir Steffen, 2011), o tai didina MVĮ galimybes gauti kreditą. Be to, nustatyta, jog kredito rinkos, kuriose vyrauja didelė konkurencija, pasižymi didesniu kredito prieinamumu, todėl MVĮ gali gauti didesnius kreditus su mažesnėmis palūkanų normomis (Love ir Pería, 2015; Wang ir kt., 2020). Plačiai mokslininkų tiriami individualūs paraiškos veiksniai nurodo pagrindines potencialaus skolininko savybes ir gali būti suskirstyti į skolinimo technologijų, įmonės charakteristikų ir produkto charakteristikų veiksnių grupes (Berger ir Udell, 2006). Šie veiksniai būdingi atskiroms įmonėms ir atlieka lemiamą vaidmenį nustatant MVĮ galimybes gauti kreditą. Mokslinėje literatūroje skolinimo technologijų veiksniai apibrėžiami kaip technologinės priemonės ir sąrankos, kurias skolintojai naudoja vertindami MVĮ kreditingumą, ir yra skirstomi į dvi grupes: sandorio skolinimą ir santykių skolinimą. Sandorio skolinimo metu atsižvelgiama į įmonių finansinių ataskaitų duomenis ir kredito istoriją (Motta ir Sharma, 2020), todėl šis skolinimo technologijų tipas paprastai naudojamas įvertinti didesniems ir skaidresniems skolininkams, kurie teikia audituotas ir išsamias finansines ataskaitas (Palazuelos ir kt., 2018; Ferri ir kt., 2019). Santykių skolinimas yra geriausias būdas, kai yra ribota informacija apie įmonę, o kreditingumas gali būti vertinamas remiantis ankstesniais banko ir įmonės santykiais (Durguner, 2017; Rabetti, 2022). Šie technologiniai skirtumai turi didelę įtaką MVĮ galimybei gauti kreditą, nes individualus skolininkas gali užtikrinti finansavimą pagal veiksnius, priklausančius tiek vienai, tiek kitai skolinimo technologijai (Angori ir kt., 2019; Chodorow-Reich ir kt., 2022). Kita individualių paraiškos veiksnių grupė yra įmonės charakteristikos, kurios yra unikalios kiekvienai įmonei ir gali apimti veiksnius, tokius kaip dydis, amžius, sektorius (Mina ir kt., 2013) ir valdymo struktūra (Aterido ir Hallward-Driemeier, 2011; Sikochi, 2020; de Andrés ir kt., 2021). Mokslininkų tyrimai rodo, kad jaunesnėms MVĮ gali būti sunkiau gauti kreditą nei geriau įsitvirtinusioms įmonėms (Mac An Bhaird ir kt., 2016). Paskutinė individualių paraiškos veiksnių grupė yra produkto charakteristikos, paremtos finansavimo produkto savybėmis, tokiomis kaip reikalingo užstato suma (Gurara ir kt., 2020; Berger ir kt., 2022), palūkanų norma (Xu ir kt., 2020; Kärnä ir Stephan, 2022) ir sutarties terminas (Minnis ir Sutherland, 2017; Aoki, 2021). Mokslininkai nustatė, kad tinkamo finansavimo produkto pasirinkimas taip pat gali turėti didėlę įtaką MVĮ kredito prieinamumui (Adam ir Streitz, 2016; Gurara ir kt., 2020). Ferri ir kt. (2019) nustatė, kad reikėtų naudoti kintamuosius, priklausančius abiem veiksnių grupėms, vertinant MVĮ kredito prieinamumą. Pasigendama mokslinių tyrimų, apimančių kredito prieinamumo MVĮ vertinimą, todėl būtina sujungti įvertintas veiksnių grupes ir identifikuoti tuos rodiklius, kurie yra svarbūs mažoms ir vidutinėms įmonėms, gaunant finansavimą iš kredito įstaigų. Kredito prieinamumą lemiantys veiksniai ir juos nagrinėjantys tyrimai yra suskirstyti į veiksnių grupes ir pateikti 46, 47, 48, 49, 50, 51, 52 lentelėse.

46 lentelė. Santykių skolinimo technologijų veiksniai. Sudaryta autoriaus

Veiksnys	Apibrėžimas	Tyrimai
Trukmė	Banko ir įmonės santykių trukmė.	Cole (1998); Petersen ir Rajan (2002); Elsas (2005); Peltoniemi (2007); Agarwal ir Hauswald (2010); Jiménez ir kt. (2012); Cassar ir kt. (2015); Neuberger ir Räthke-Döppner (2015); Belaid ir kt. (2017); Cucculelli ir Peruzzi (2017); Durguner (2017); Minnis ir Sutherland (2017); Angori ir kt. (2019); Grzelak (2019); Ferri ir kt. (2019); Angori ir kt. (2020); Xu ir kt. (2020); Banerjee ir kt. (2021); Berger ir kt. (2022)
Pagrindinis partneris	Informacija, ar bankas yra pagrindinė įmonės finansų institucija.	Elsas (2005); Cassar ir kt. (2015); Angori ir kt. (2019, 2020); Aoki (2021)
Mokėjimai	Įeinančių ir išeinančių mokėjimų santykis banko sąskaitose.	Elsas (2005); Durguner (2017); Angori ir kt. (2019, 2020)
Atmetimai	Informacija, ar paraišką pateikusi įmonė iki paraiškos pateikimo gavo neigiamų sprendimų.	Cassar ir kt. (2015)
Skolos dalis	Banke turimų skolų ir visų įsipareigojimų, nurodytų finansinėse ataskaitose, dalis.	Elsas (2005); Peltoniemi (2007); Angori ir kt. (2019); Ferri ir kt. (2019); Angori ir kt. (2020); Berger ir kt. (2022); Kärnä ir Stephan (2022)
Finansavimo sutartys	Ankstesnių finansavimo sutarčių, kurias įmonė turėjo banke, skaičius.	Cole (1998); Peltoniemi (2007); Neuberger ir Räthke- Döppner (2015); Kirschenmann (2016); Durguner (2017); Minnis ir Sutherland (2017)
Kiti produktai	Informacija, ar įmonė turėjo ne finansavimo produktų.	Cole (1998); Petersen ir Rajan (2002); Peltoniemi (2007); Cassar ir kt. (2015); Neuberger ir Räthke-Döppner (2015); Durguner (2017)
Kiti santykiai	Kitų bankų, su kuriais įmonė palaiko ryšius, skaičius.	Cole (1998); Peltoniemi (2007); Elsas (2005); Agarwal ir Hauswald (2010); Jiménez ir kt. (2012); Kirschenmann (2016); Cucculelli ir Peruzzi (2017); Durguner (2017); Angori ir kt. (2019); Grzelak (2019); Ferri ir kt. (2019); Angori ir kt. (2020); Aoki (2021); Kärnä ir Stephan (2022)
Atstumas	Atstumas tarp artimiausio banko padalinio ir įmonės.	Neuberger ir Räthke-Döppner (2015); Durguner (2017); Xu ir kt. (2020)
Banko valdymas	Bankas valdo arba daro įtaką vadovybės sprendimams įmonėje.	Neuberger ir Räthke-Döppner (2015); Durguner (2017); Xu ir kt. (2020)

46 lentelėje pateikiami santykių skolinimo technologijų veiksniai, jų apibrėžimai bei susiję tyrimai. Pastebėta, kad banko ir įmonės santykių trukmė yra vienas iš dažniausiai tiriamų veiksnių, lemiančių kredito prieinamumą MVĮ.

47 lentelė. Sandorio skolinimo technologijų veiksniai, paremti finansinių ataskaitų duomenimis. Sudaryta autoriaus

Veiksnys	Apibrėžimas	Tyrimai
Grynųjų pinigų santykis	Grynųjų pinigų ir trumpalaikių įsipareigojimų santykis.	Degryse ir kt. (2018); Martí ir Quas (2018); Grzelak (2019); Malakauskas ir Lakštutienė (2021); Medianovskyi ir kt. (2023)
Skubaus padengimo santykis	Greitas turto ir trumpalaikių įsipareigojimų santykis.	Xu ir kt. (2020); Malakauskas ir Lakštutienė (2021); Medianovskyi ir kt. (2023)
Bendrasis likvidumo santykis	Trumpalaikio turto ir trumpalaikių įsipareigojimų santykis.	Jiménez ir kt. (2012); Meuleman ir De Maeseneire (2012); Florou ir Kosi (2015); Adam ir Streitz (2016); Cucculelli ir Peruzzi (2017); Durguner (2017); Angori ir kt. (2019); Ferri ir kt. (2019); Angori ir kt. (2020); Xu ir kt. (2020); Malakauskas ir Lakštutienė (2021); Zainol Abidin ir kt. (2021); Medianovskyi ir kt. (2023)
Skolos-nuosavybės santykis	Bendras įsipareigojimų ir savininko nuosavybės santykis.	Elsas (2005); Jiménez ir kt. (2012); Angori ir kt. (2019); Malakauskas ir Lakštutienė (2021); Zainol Abidin ir kt. (2021); Banerjee ir kt. (2021); Medianovskyi ir kt. (2023)
Materialusis turtas	Bendras materialiojo turto ir savininko nuosavybės santykis.	Meuleman ir De Maeseneire (2012); Florou ir Kosi (2015); Adam ir Streitz (2016); Degryse ir kt. (2018); Martí ir Quas (2018); Ogura (2018); Angori ir kt. (2019); Grzelak (2019); Angori ir kt. (2020); Aoki (2021)
Įsiskolinimas	Bendrų įsipareigojimų ir viso turto santykis.	Peltoniemi (2007); Bosch ir Steffen (2011); Meuleman ir De Maeseneire (2012); Florou ir Kosi (2015); Adam ir Streitz (2016); Kirschenmann (2016); Cucculelli ir Peruzzi (2017); Durguner (2017); Martí ir Quas (2018); Angori ir kt. (2020); Xu ir kt. (2020); Zainol Abidin ir kt. (2021); Aoki (2021); Banerjee ir kt. (2021); Berger ir kt. (2022); Chodorow-Reich ir kt. (2022)
Turto grąža	Grynųjų pajamų ir viso turto santykis.	Cole (1998); Jiménez ir kt. (2012); Cassar ir kt. (2015); Florou ir Kosi (2015); Cucculelli ir Peruzzi (2017); Angori ir kt. (2019); Ferri ir kt. (2019); Zainol Abidin ir kt. (2021); Aoki (2021); Banerjee ir kt. (2021); Berger ir kt. (2022)
Nuosavybės grąža	Grynųjų pajamų ir nuosavo kapitalo santykis.	Chong ir kt. (2013); Zainol Abidin ir kt. (2021)
Pelningumas	Grynųjų pajamų ir grynųjų pardavimų santykis.	Carbo-Valverde ir kt. (2009); Bosch ir Steffen (2011); Florou ir Kosi (2015); Adam ir Streitz (2016); Cucculelli ir Peruzzi (2017); Durguner (2017); Martí ir Quas (2018); Ogura (2018); Grzelak (2019); Ferri ir kt. (2019); Xu ir kt. (2020); Malakauskas ir Lakštutienė (2021); Zainol Abidin ir kt. (2021); Aoki (2021); Chodorow-Reich ir kt. (2022); Medianovskyi ir kt. (2023)
Turto apyvarta	Grynųjų pardavimų ir viso turto santykis.	Molina ir Preve (2012); Meuleman ir De Maeseneire (2012); Cassar ir kt. (2015); Grzelak (2019); Zainol Abidin ir kt. (2021)
Gautinos sumos	Grynųjų pardavimų ir gautinų sumų santykis.	Molina ir Preve (2012); Cassar ir kt. (2015); Durguner (2017); Degryse ir kt. (2018); Malakauskas ir Lakštutienė (2021)
DSCR	Skolos aptarnavimo ir padengimo santykis.	Elsas (2005); Carbo-Valverde ir kt. (2009); Molina ir Preve (2012); Adam ir Streitz (2016); Degryse ir kt. (2018); Ogura (2018); Angori ir kt. (2019); Malakauskas ir Lakštutienė (2021); Banerjee ir kt. (2021); Chodorow-Reich ir kt. (2022); Medianovskyi ir kt. (2023)
Aprėptis	Materialaus turto atėmus trumpalaikius įsipareigojimus ir įsipareigojimų santykis.	Malakauskas ir Lakštutienė (2021); Medianovskyi ir kt. (2023)
Pardavimų augimas	Grynųjų pardavimų pokytis.	Carbo-Valverde ir kt. (2009); Molina ir Preve (2012); Berger ir kt. (2017); Ogura (2018); Aoki (2021); Malakauskas ir Lakštutienė (2021); Aristei ir Angori (2022); Chodorow-Reich ir kt. (2022); Medianovskyi ir kt. (2023)
Turto augimas	Trumpalaikio turto pokytis.	Martí ir Quas (2018); Malakauskas ir Lakštutienė (2021); Medianovskyi ir kt. (2023)
Turtas	Bendra turto suma.	Cole (1998); Peltoniemi (2007); Cassar ir kt. (2015); Adam ir Streitz (2016); Kirschenmann (2016); Berger ir kt. (2017); Martí ir Quas (2018); Angori ir kt. (2019, 2020); Banerjee ir kt. (2021)

47 lentelėje pateikiami sandorio skolinimo technologijų veiksniai, paremti finansinių ataskaitų duomenimis, jų apibrėžimai bei susiję tyrimai. Tai yra viena iš plačiausiai naudojamų veiksnių grupių vertinant verslo kreditingumą, kuri atspindi įmonės finansinę padėtį bei galimybes grąžinti kreditą.

48 lentelė. Sandorio skolinimo technologijų veiksniai, paremti kredito istorijos duomenimis. Sudaryta autoriaus

Veiksnys	Apibrėžimas	Tyrimai
Įmonės nemokos	Informacija, ar įmonė nevykdė kokių nors savo įsipareigojimų (tiek viduje, tiek išorėje).	Cole (1998); Cassar ir kt. (2015); Neuberger ir Räthke- Döppner (2015); Kirschenmann (2016); Malakauskas ir Lakštutienė (2021); Berger ir kt. (2022); Medianovskyi ir kt. (2023)
Savininkų nemokos	Informacija, ar įmonės savininkai nevykdė kokių nors įsipareigojimų (tiek viduje, tiek išorėje).	Cole (1998); Cassar ir kt. (2015); Malakauskas ir Lakštutienė (2021); Medianovskyi ir kt. (2023)
Įmonės nevykdomi įsipareigojimai	Informacija, ar įmonė kada nors nevykdė įsipareigojimų.	Jiménez ir kt. (2012); Neuberger ir Räthke-Döppner (2015); Kirschenmann (2016)
Savininkų nevykdomi įsipareigojimai	Informacija, ar įmonės savininkai kada nors nevykdė įsipareigojimų.	Jiménez ir kt. (2012)
Rizikos balas	Įmonės rizikos įvertinimas.	Sapienza (2004); Elsas (2005); Peltoniemi (2007); Agarwal ir Hauswald (2010); Presbitero ir Zazzaro (2011); Berger ir kt. (2011); Florou ir Kosi (2015); Adam ir Streitz (2016); Ogura (2018); Xu ir kt. (2020); Aoki (2021); Banerjee ir kt. (2021); Chodorow-Reich ir kt. (2022)

48 lentelėje pateikiami sandorio skolinimo technologijų veiksniai, paremti kredito istorijos duomenimis, jų apibrėžimai bei susiję tyrimai. Ši veiksnių grupė atspindi įmonės bei jos savininkų informaciją apie nemokas bei nevykdomus įsipareigojimus.

49 lentelė. Skolinimo infrastruktūros veiksniai. Sudaryta autoriaus

Veiksniai	Apibrėžimas	Tyrimai
BVP	Bendrasis vidaus produktas.	Brown ir kt. (2009); Carbo-Valverde ir kt. (2009); Olivero ir kt. (2011); Jiménez ir kt. (2012); Zarutskie (2013); Aiyar ir kt. (2014); Bertay ir kt. (2015); Love ir Pería (2015); Khan ir kt. (2016); Chen ir kt. (2017); Degryse ir kt. (2018); Fang ir kt. (2022)
Infliacija	Metinis kainų indekso pokytis.	Brown ir kt. (2009); Agarwal ir Hauswald (2010); Jiménez ir kt. (2012); Zarutskie (2013); Aiyar ir kt. (2014); Bertay ir kt. (2015); Love ir Pería (2015); Khan ir kt. (2016); Chen ir kt. (2017); Ademosu ir Morakinyo (2021); Berger ir kt. (2022)
Nedarbas	Vidutinis nedarbo lygis.	Berger ir kt. (2011, 2017); Degryse ir kt. (2018); Ademosu ir Morakinyo (2021); Berger ir kt. (2022)
Gyventojų tankumas	Gyventojų tankumas.	Carbo-Valverde ir kt. (2009); Neuberger ir Räthke-Döppner (2015)
Reguliavimo griežtumas	Reguliavimo griežtumo mastas ir veiklos lygis.	Khan ir kt. (2016); Chen ir kt. (2017)
Įstatymo taisyklė	Matuoja, kiek rinkos dalyviai pasitiki visuomenės taisyklėmis ir jų laikosi.	Khan ir kt. (2016); Chen ir kt. (2017)

49 lentelėje pateikiami skolinimo infrastruktūros veiksniai, jų apibrėžimai bei susiję tyrimai. Didžioji dalis tyrimų nagrinėja sąsają tarp BVP ir infliacijos lygio bei kredito prieinamumo MVĮ.

50 lentelė. Finansų įstaigos struktūros veiksniai. Sudaryta autoriaus.

Veiksnys	Apibrėžimas	Tyrimai
Rinkos koncentracija	Herfindahl-Hirschman indeksas banko aptarnaujamose rinkose.	Petersen ir Rajan (2002); Elsas (2005); Carbo-Valverde ir kt. (2009); Berger ir kt. (2011); Presbitero ir Zazzaro (2011); Chong ir kt. (2013); Zarutskie (2013); Love ir Pería (2015); Milani (2014); Khan ir kt. (2016); Durguner (2017); Berger ir kt. (2017); Chen ir kt. (2017); Degryse ir kt. (2018); Ogura (2018); Angori ir kt. (2019, 2020); Aristei ir Angori (2022)
Indėliai	Banko kontroliuojamų indėlių dalis.	Berger ir kt. (2011); Bertay ir kt. (2015); Khan ir kt. (2016); Chen ir kt. (2017); Degryse ir kt. (2018); Ogura (2018); Fang ir kt. (2022)
Pajamų augimas	Banko vidutinis pajamų augimas banko aptarnaujamose rinkose.	Berger ir kt. (2011); Bertay ir kt. (2015); Fang ir kt. (2022)
Kapitalas	Banko kapitalo ir pagal riziką įvertinto turto santykis.	Berger ir kt. (2011); Olivero ir kt. (2011); Jiménez ir kt. (2012); Zarutskie (2013); Aiyar ir kt. (2014); Bertay ir kt. (2015); Khan ir kt. (2016); Belaid ir kt. (2017); Berger ir kt. (2017); Chen ir kt. (2017); Degryse ir kt. (2018); Berger ir kt. (2022); Fang ir kt. (2022)
Likvidumas	Banko gebėjimas vykdyti savo trumpalaikius įsipareigojimus ir valdyti pinigų srautus.	Berger ir kt. (2011); Olivero ir kt. (2011); Jiménez ir kt. (2012); Aiyar ir kt. (2014); Bertay ir kt. (2015); Khan ir kt. (2016); Berger ir kt. (2017); Chen ir kt. (2017); Degryse ir kt. (2018); Ogura (2018); Berger ir kt. (2022); Fang ir kt. (2022)
Bendras turtas	Viso turto, įskaitant paskolas, investicijas, grynuosius pinigus ir kitą turtą, suma.	Sapienza (2004); Carbo-Valverde ir kt. (2009); Berger ir kt. (2011); Olivero ir kt. (2011); Jiménez ir kt. (2012); Zarutskie (2013); Aiyar ir kt. (2014); Bertay ir kt. (2015); Khan ir kt. (2016); Chen ir kt. (2017); Ogura (2018); Berger ir kt. (2022); Fang ir kt. (2022)
Turto grąža	Banko pelningumas, matuojant grynąsias pajamas ir visą turtą.	Carbo-Valverde ir kt. (2009); Jiménez ir kt. (2012); Ogura (2018); Fang ir kt. (2022)
Abejotinų paskolų santykis	Paskolų, kurioms gresia įsipareigojimų neįvykdymas arba kurios jau yra pradelstos, dalis.	Sapienza (2004); Jiménez ir kt. (2012); Bertay ir kt. (2015); Berger ir kt. (2017, 2022); Aristei ir Angori (2022)
Palūkanų norma	Palūkanų norma, kuria bankai skolina arba skolinasi vienas iš kito lėšas tarpbankinėje rinkoje.	Agarwal ir Hauswald (2010); Jiménez ir kt. (2012); Khan ir kt. (2016); Gurara ir kt. (2020); Ademosu ir Morakinyo (2021)
Rinkos dalis	Banko kontroliuojamos rinkos dalis.	Berger ir kt. (2011); Chong ir kt. (2013); Bertay ir kt. (2015); Berger ir kt. (2022)
Padalinių koncentracija	Bendras banko padalinių, esančių tam tikroje vietovėje ar regione, skaičius.	Petersen ir Rajan (2002); Carbo-Valverde ir kt. (2009); Presbitero ir Zazzaro (2011); Jiménez ir kt. (2012); Chong ir kt. (2013); Milani (2014); Berger ir kt. (2017); Angori ir kt. (2019, 2020); Aristei ir Angori (2022)
Banko darbuotojai	Banko darbuotojų skaičius tam tikroje rinkoje.	Petersen ir Rajan (2002)
Amžius	Banko amžius.	Berger ir kt. (2011); Zarutskie (2013)
Valdymas	Banko nuosavybės teisės priklauso užsienio investuotojams arba valstybei.	Bertay ir kt. (2015); Khan ir kt. (2016); Chen ir kt. (2017); Ogura (2018)

50 lentelėje pateikiami finansų įstaigos struktūros veiksniai, jų apibrėžimai bei susiję tyrimai. Tai yra gausiausia makrospecifinių veiksnių grupė, kuri apibrėžia finansų rinką, jos konkurencingumą, koncentraciją ir kokybę.

51 lentelė. Įmonės charakteristikos veiksniai. Sudaryta autoriaus

Veiksnys	Apibrėžimas	Tyrimai
Amžius	Įmonės amžius paraiškos pateikimo metu.	Cole (1998); Petersen ir Rajan (2002); Peltoniemi (2007); Agarwal ir Hauswald (2010); Presbitero ir Zazzaro (2011); Jiménez ir kt. (2012); Chong ir kt. (2013); Cassar ir kt. (2015); Love ir Pería (2015); Neuberger ir Räthke- Döppner (2015); Kirschenmann (2016); Cucculelli ir Peruzzi (2017); Durguner (2017); Martí ir Quas (2018); Angori ir kt. (2019); Grzelak (2019); Ferri ir kt. (2019); Angori ir kt. (2020); Motta ir Sharma (2020); Aoki (2021); Malakauskas ir Lakštutienė (2021); Zainol Abidin ir kt. (2021); Kärnä ir Stephan (2022); Medianovskyi ir kt. (2023)
Įvairovė	Informacija apie įmonės savininkų ir vadovų lyčių įvairovę.	Angori ir kt. (2019, 2020); Motta ir Sharma (2020); Zainol Abidin ir kt. (2021)
Valdantis savininkas	Savininkas yra įmonės vadovas.	Petersen ir Rajan (2002); Cassar ir kt. (2015); Cucculelli ir Peruzzi (2017); Angori ir kt. (2019, 2020)
Akcininkų struktūra	Įmonės nuosavybės struktūra pagal savininkų skaičių, didžiųjų akcininkų skaičių.	Petersen ir Rajan (2002); Cassar ir kt. (2015); Love ir Pería (2015); Cucculelli ir Peruzzi (2017); Angori ir kt. (2019, 2020); Zainol Abidin ir kt. (2021); Aristei ir Angori (2022)
Valdymas	Bendrovės valdymo struktūra atsižvelgiant į valdybos egzistavimą ir jos sudėtį.	Zainol Abidin ir kt. (2021); Aristei ir Angori (2022)
Dydis	Įmonės dydis.	Cole (1998); Sapienza (2004); Elsas (2005); Brown ir kt. (2009); Carbo-Valverde ir kt. (2009); Bosch ir Steffen (2011); Presbitero ir Zazzaro (2011); Chong ir kt. (2013); Meuleman ir De Maeseneire (2012); Florou ir Kosi (2015); Love ir Pería (2015); Neuberger ir Räthke-Döppner (2015); Kirschenmann (2016); Belaid ir kt. (2017); Berger ir kt. (2017); Cucculelli ir Peruzzi (2017); Angori ir kt. (2019); Grzelak (2019); Ferri ir kt. (2019); Angori ir kt. (2020); Motta ir Sharma (2020); Zainol Abidin ir kt. (2021); Aoki (2021); Berger ir kt. (2022); Aristei ir Angori (2022); Chodorow- Reich ir kt. (2022)
Teisinė forma	Juridinio asmens tipas, pagal kurį įmonė yra registruota.	Cole (1998); Petersen ir Rajan (2002); Elsas (2005); Peltoniemi (2007); Brown ir kt. (2009); Bosch ir Steffen (2011); Chong ir kt. (2013); Cassar ir kt. (2015); Kirschenmann (2016); Berger ir kt. (2017); Durguner (2017); Grzelak (2019); Gurara ir kt. (2020); Motta ir Sharma (2020); Aristei ir Angori (2022)
Vieta	Regiono, kuriame veikia įmonė, klasifikacija.	Petersen ir Rajan (2002); Jiménez ir kt. (2012); Milani (2014); Berger ir kt. (2017); Durguner (2017); Grzelak (2019); Motta ir Sharma (2020); Kärnä ir Stephan (2022)
Sektorius	Įmonės veiklos sritis.	Cole (1998); Peltoniemi (2007); Bosch ir Steffen (2011); Jiménez ir kt. (2012); Love ir Pería (2015); Neuberger ir Räthke-Döppner (2015); Bonnet ir kt. (2016); Belaid ir kt. (2017); Durguner (2017); Martí ir Quas (2018); Motta ir Sharma (2020); Aristei ir Angori (2022); Chodorow-Reich ir kt. (2022); Kärnä ir Stephan (2022)
Sertifikavimas	Įmonė praėjo bet kokią kokybės sertifikavimo formą.	Presbitero ir Zazzaro (2011); Angori ir kt. (2019, 2020)
Subsidijos	Bendrovė gavo bet kokių viešųjų subsidijų.	Meuleman ir De Maeseneire (2012); Bonnet ir kt. (2016); Martí ir Quas (2018); Angori ir kt. (2019, 2020)
Auditas	Nurodo, ar pateiktos metinės finansinės ataskaitos buvo audituotos, ar ne.	Brown ir kt. (2009); Palazuelos ir kt. (2018); Motta ir Sharma (2020)

51 lentelėje pateikiami įmonės charakteristikos veiksniai, jų apibrėžimai bei susiję tyrimai. Ši veiksnių grupė atspindi bendrą informaciją apie įmonę, jos amžių, lokaciją, tipą, sektorių ir informacijos skaidrumą.

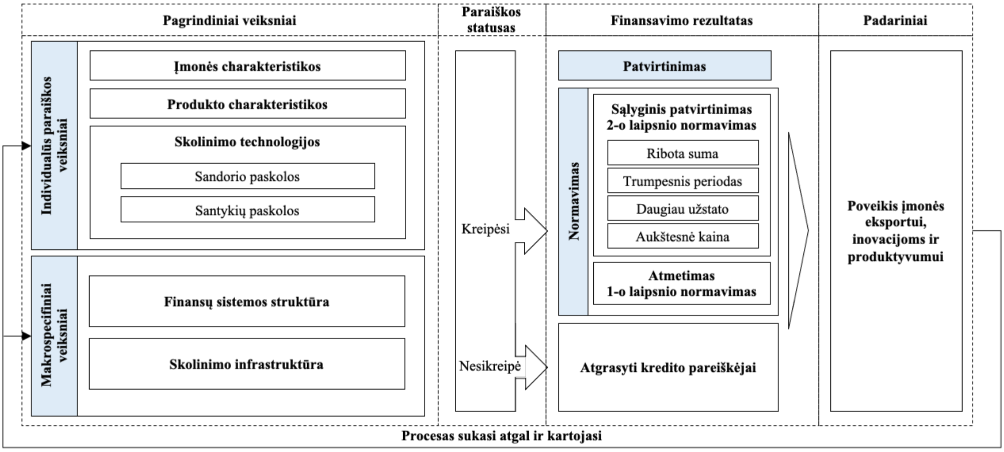
52 lentelė. Produkto charakteristikos veiksniai. Sudaryta autoriaus

Veiksnys	**Apibrėžimas**	**Tyrimai**
Produktas	Išduodamo produkto tipas.	Agarwal ir Hauswald (2010); Zarutskie (2013); Cassar ir kt. (2015); Adam ir Streitz (2016); Kirschenmann (2016); Minnis ir Sutherland (2017); Gurara ir kt. (2020); Berger ir kt. (2022)
Užstatas	Įkeisto turto dydis.	Peltoniemi (2007); Agarwal ir Hauswald (2010); Bosch ir Steffen (2011); Zarutskie (2013); Neuberger ir Räthke- Döppner (2015); Kirschenmann (2016); Ferri ir kt. (2019); Gurara ir kt. (2020); Motta ir Sharma (2020); Berger ir kt. (2022); Chodorow-Reich ir kt. (2022)
Turtas	Informacija, ar įmonė įkeitė kokį nors materialųjį turtą.	Bosch ir Steffen (2011); Florou ir Kosi (2015); Adam ir Streitz (2016); Kirschenmann (2016); Minnis ir Sutherland (2017); Angori ir kt. (2019); Motta ir Sharma (2020); Chodorow-Reich ir kt. (2022)
Palūkanų norma	Sutarties palūkanų norma.	Peltoniemi (2007); Agarwal ir Hauswald (2010); Zarutskie (2013); Cassar ir kt. (2015); Florou ir Kosi (2015); Adam ir Streitz (2016); Durguner (2017); Minnis ir Sutherland (2017); Gurara ir kt. (2020); Xu ir kt. (2020); Berger ir kt. (2022); Chodorow-Reich ir kt. (2022); Kärnä ir Stephan (2022)
Paskolos suma	Finansavimo sumos dydis.	Peltoniemi (2007); Cassar ir kt. (2015); Florou ir Kosi (2015); Neuberger ir Räthke-Döppner (2015); Adam ir Streitz (2016); Minnis ir Sutherland (2017); Gurara ir kt. (2020); Xu ir kt. (2020); Aoki (2021); Kärnä ir Stephan (2022)
Terminas	Finansavimo sutarties termino ilgis.	Peltoniemi (2007); Agarwal ir Hauswald (2010); Bosch ir Steffen (2011); Florou ir Kosi (2015); Adam ir Streitz (2016); Minnis ir Sutherland (2017); Gurara ir kt. (2020); Xu ir kt. (2020); Aoki (2021); Berger ir kt. (2022); Chodorow- Reich ir kt. (2022)

52 lentelėje pateikiami produkto charakteristikos veiksniai, jų apibrėžimai bei susiję tyrimai. Ši veiksnių grupė atspindi informaciją apie produktą, jo kaštus, užstatą bei grąžinimo terminus.

Remiantis atlikta mokslinės literatūros analize, sukonstruotas konceptualus kredito prieinamumo MVĮ vertinimo modelis (žr. 28 pav.). Modelį sudaro trys pagrindiniai komponentai:

- Pagrindinės sąlygos – makrospecifiniai ir individualūs paraiškos veiksniai, kurie apibrėžia esamą kredito prieinamumą.
- Prašymo statusas – MVĮ sprendimas kreiptis dėl kredito ar nesikreipti.
- Finansavimo rezultatas – galutinis rezultatas, nurodantis, ar buvo gautas kreditas, ar nebuvo gautas.



28 pav. Konceptualus kredito prieinamumo MVĮ modelis. Sudaryta autoriaus

Konceptualus kredito prieinamumo MVĮ modelis (žr. 28 pav.) pateikia pagrindinius MVĮ kredito prieinamumo komponentus. Pagrindiniai veiksniai, susidedantys iš makrospecifinių ir individualių paraiškos veiksnių, apibrėžia pagrindines kredito gavimo sąlygas. Todėl, norint empiriškai įvertinti kredito prieinamumą, svarbu atsižvelgti tiek į makrospecifinius, tiek į individualius paraiškų veiksnius. Šių veiksnių įtraukimas gali būti neįmanomas dėl tokių priežasčių, kaip duomenų trūkumas, todėl tokie tyrimai, kaip Cassar ir kt. (2015); Adam ir Streitz (2016); Kirschenmann (2016); Angori ir kt. (2019); Calabrese ir kt. (2022); Kärnä ir Stephan (2022), išsprendžia šiuos apribojimus sukurdami individualius šalių modelius.

Norint sukurti tinkamą kredito prieinamumo MVĮ vertinimo modelį, rekomenduojama naudoti mašininio mokymosi metodus. Tradiciniai metodai, tokie kaip diskriminacinė analizė (DA) ir logistinė regresija (LR), yra dažnai naudojami norint įvertinti prieigą prie kredito, tačiau jie nėra pritaikyti didesniems duomenų rinkiniams ir sudėtingiems nelinijiniams ryšiams. Šiuolaikiniai mašininio mokymosi metodai, tokie kaip sprendimų medžiai, atsitiktiniai miškai, dirbtiniai neuroniniai tinklai (ANN), atraminių vektorių klasifikatorius (SVM), k-artimiausio kaimyno klasifikatorius (kNN), siekiant pagerinti modeliavimo tikslumą yra vis plačiau taikomi moksliniuose tyrimuose (Barboza ir kt., 2017). Mokslinės literatūros analizė atskleidė, jog naujausius mašininio mokymosi metodus gali būti sunku interpretuoti (Preece ir kt., 2018). Kadangi mašininio mokymosi metodams būdingas ribotas paaiškinamumas, siekiant užtikrinti modelių paaiškinamumą, reikia naudoti paaiškinamumo metodus, tokius kaip SHAP, kurie tinkami atskiroms savybėms arba jų sąveikų svarbai nustatyti (Shrikumar ir kt., 2017; Sundararajan ir kt., 2017; Lundberg ir kt., 2020).

53 lentelė. Kredito prieinamumo MVĮ modeliavimo metodai. Sudaryta autoriaus

Metodas	Tyrimai
Tradiciniai metodai	
Diskriminacinė analizė	Mahmoudi ir Duman (2015); Barboza ir kt. (2017)
Logistinė regresija	Wang ir kt. (2011); Kruppa ir kt. (2013); Danenas ir Garsva (2015); Datta ir kt. (2016); Barboza ir kt. (2017); Ariza-Garzon ir kt. (2020); Wang ir kt. (2020); Malakauskas ir Lakštutienė (2021); Moscato ir kt. (2021); Hussin Adam Khatir ir Bee (2022); Medianovskyi ir kt. (2023)
Mašininio mokymosi metodai	
Sprendimų medis	Wang ir kt. (2011); Datta ir kt. (2016); Trivedi (2020); Wang ir kt. (2020); Hussin Adam Khatir ir Bee (2022)
Atsitiktiniai miškai	Wang ir kt. (2011); Kruppa ir kt. (2013); Danenas ir Garsva (2015); Datta ir kt. (2016); Barboza ir kt. (2017); Ariza-Garzon ir kt. (2020); Silva ir kt. (2020); Trivedi (2020); Malakauskas ir Lakštutienė (2021); Moscato ir kt. (2021); Hussin Adam Khatir ir Bee (2022); Medianovskyi ir kt. (2023)
K-artimiausio kaimyno	Kruppa ir kt. (2013); Wang ir kt. (2020); Hussin Adam Khatir ir Bee (2022)
Supakuoto-artimiausio kaimyno	Kruppa ir kt. (2013); Barboza ir kt. (2017)
Atraminių vektorių klasifikatorius	Wang ir kt. (2011); Danenas ir Garsva (2015); Datta ir kt. (2016); Pal ir kt. (2016); Barboza ir kt. (2017); Silva ir kt. (2020); Trivedi (2020)
Naïve Bayes	Trivedi (2020); Wang ir kt. (2020); Hussin Adam Khatir ir Bee (2022)
Dirbtiniai neuroniniai tinklai	Wang ir kt. (2011); Zhao ir kt. (2015); Barboza ir kt. (2017); Dastile ir Celik (2021); Hadji Misheva ir kt. (2021); Malakauskas ir Lakštutienė (2021); Hussin Adam Khatir ir Bee (2022); Medianovskyi ir kt. (2023)
Gradialinis nusileidimas	Barboza ir kt. (2017); Bussmann ir kt. (2020); Qi ir kt. (2021); Bucker ir kt. (2022); Medianovskyi ir kt. (2023)

Norint tinkamai įvertinti kredito prieinamumą MVĮ, rekomenduojama naudoti įvairius modeliavimo metodus ir palyginti juos su apsibrėžtu etalonu, kuris dažniausiai yra vienas iš tradicinių modeliavimo metodų. Susisteminti tyrimai rodo, jog mašininio mokymosi metodai gali pagerinti kredito vertinimo modelių tikslumą ir sumažinti jų šališkumą. Kita vertus, mašininiu mokymusi pagrįstiems metodams būdingas juodosios dėžės pobūdis, kuris riboja modelių paaiškinamumą, todėl būtina naudoti paaiškinamumo metodus, tokius kaip SHAP.

2. KREDITO PRIEINAMUMO MVĮ VERTINIMO METODOLOGIJA

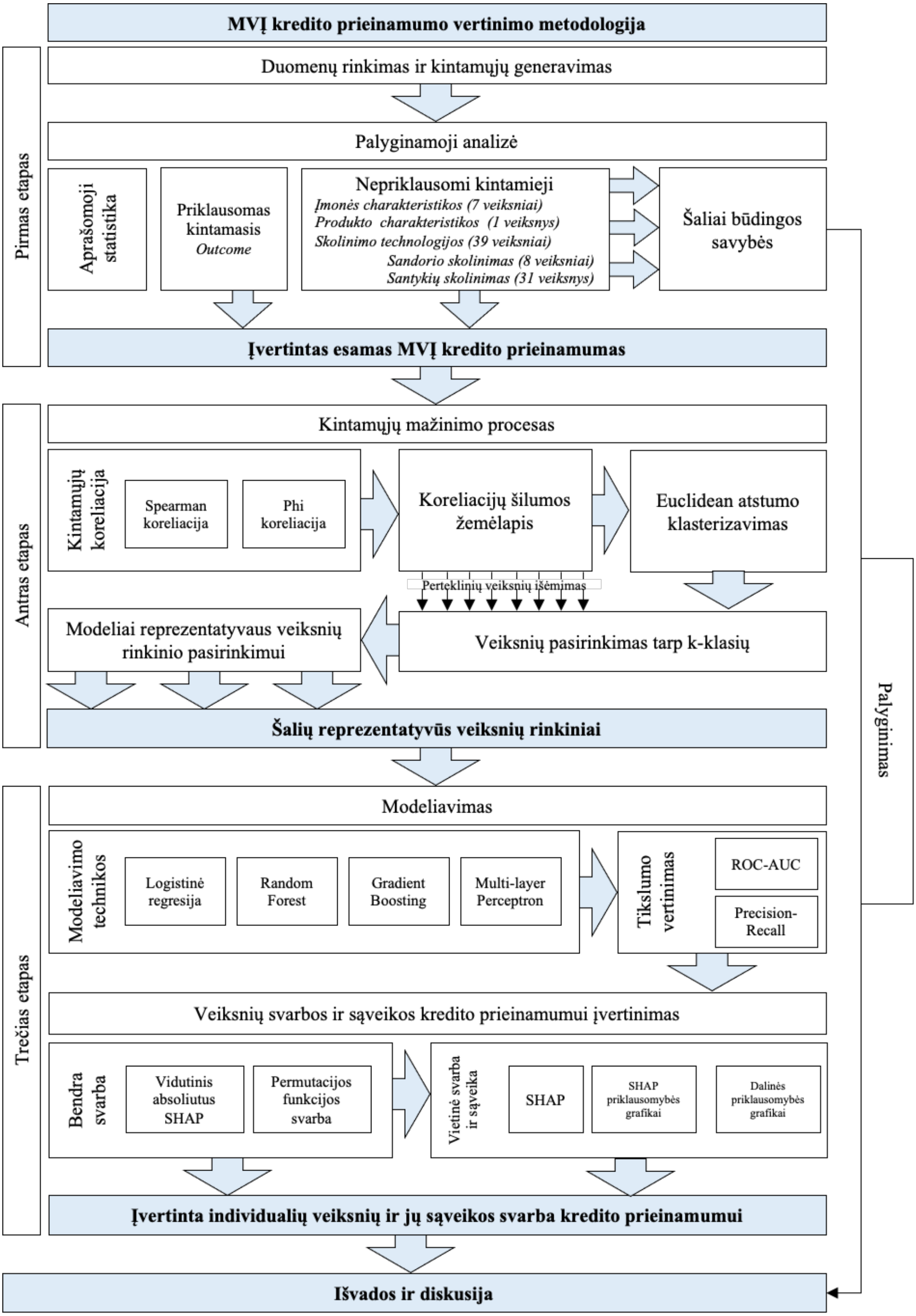
Antroje dalyje sprendžiamas 4-as uždavinys. Sukurtas trijų etapų kredito prieinamumo MVĮ vertinimo modelis. Šioje dalyje apibrėžiami priklausomi ir nepriklausomi kintamieji, kurie naudojami kuriant kredito prieinamumo MVĮ vertinimo modelį, pateikiamas modelio kūrimo technikų ir metodų pagrindimas.

Pirmiausiai, siekiant sukurti kredito prieinamumo MVĮ vertinimo modelį, nustatomas indikatorius, apibrėžiantis MVĮ kredito prieinamumą, bei jį sąlygojantys veiksniai. Disertacijoje kredito prieinamumas yra vertinamas per kredito pasiūlą naudojantis pateiktų finansavimo paraiškų rezultatais (*Outcome*). Šis būdas leidžia įvertinti veiksnius, nulemiančius kredito prieinamumą MVĮ (Jimeneze ir kt., 2012). Kiekvienos finansavimo paraiškos rezultatas gali būti dviejų tipų – patvirtinimas arba atmetimas (1-ojo laipsnio normavimas). Kintamasis *Outcome* yra apibrėžiamas kaip priklausomas kintamasis, kuris naudojamas kredito prieinamumui MVĮ vertinti. Mokslinės literatūros analizės rezultatais yra pagrįsti ir atrinkti 61 nepriklausomi kintamieji (veiksnys), kurie yra naudojami modeliuojant finansavimo paraiškos rezultatą (*Outcome*), yra sugrupuoti į įmonės charakteristikų, produkto charakteristikų ir skolinimo technologijų kintamųjų grupes:

- Įmonės charakteristikų veiksniai (8 kintamieji) – amžius (*Age*), lyčių lygybė (*Diversity*), tiesioginis valdymas (*Private*), dydis (*Segment*), juridinio asmens tipas (*Type*), regionas (*Region*), sektorius (*Sector*), auditas (*Audited*).
- Produkto charakteristikos veiksnys – produkto tipas (*Product*).
- Skolinimo technologijų veiksniai (52 kintamieji):
  - Santykių skolinimo technologijų veiksniai (8 kintamieji) – banko ir įmonės dalykinių santykių trukmė (*Relationship*), mokėjimai sąskaitose (*Payments*), atmetimai (*Rejections*), paskolų dalis banke (*Debt*, finansiniai kontraktai (*FinContracts*, debetinės kortelės (*Cards*), mokėjimų surinkimų produktai (*POS*), e-komercijos produktai (*Ecommerce*).
  - Sandorio skolinimo technologijų veiksniai (44 kintamieji):
    - ∗ Likvidumas (6 kintamieji) – grynųjų pinigų santykis (esamo periodo – *CR*, praėjusio periodo – *pCR*), skubaus padengimo santykis (esamo periodo – *QR*, praėjusio periodo – *pQR*), bendrasis likvidumo santykis (esamo periodo – *CuR*, praėjusio periodo – *pCuR*).
    - ∗ Mokumas (10 kintamųjų) – skolos ir nuosavybės santykis (esamo periodo – *DE*, praėjusio periodo – *pDE*), materialiojo turto santykis (esamo periodo – *TA*, praėjusio periodo – *pTA*), skolos santykis (esamo periodo – *DR*, praėjusio periodo – *pDR*), skolos aptarnavimo padengimo koeficientas (esamo periodo – *DSCR*, praėjusio periodo – *pDSCR*), turto padengimo koeficientas (esamo periodo – *ACR*, praėjusio periodo – *pACR*).

* Pelningumas (8 kintamieji) – turto grąža (esamo periodo – *ROA*, praėjusio periodo – *pROA*), nuosavybės grąža (esamo periodo – *ROE*, praėjusio periodo – *pROE*), bendros maržos koeficientas (esamo periodo – *GMR*, praėjusio periodo – *pGMR*), pelno maržos koeficientas (esamo periodo – *PMR*, praėjusio periodo – *pPMR*).
* Apyvartumas (6 kintamieji) – sąskaitos apyvartos koeficientas (esamo periodo – *ATR*, praėjusio periodo – *pATR*), gautinų sumų apyvartumo koeficientas (esamo periodo – *RTR*, praėjusio periodo – *pRTR*), pardavimų pokytis (*cS*), trumpalaikio turto pokytis (*cSA*).
* Kredito istorija (14 kintamųjų):
  - Įmonės (7 kintamieji) – vidinių nemokų skaičius (*IOverC*), vidinių nemokų suma (*IOverS*), vidinių nemokų trukmė (*IOverL*), išorinių nemokų skaičius (*EOverC*), išorinių nemokų suma (*E-OverS*), išorinių nemokų trukmė (*EOverL*), įsipareigojimų nevykdymas (*ODefaults*).
  - Savininkų (7 kintamieji) – vidinių nemokų skaičius (*OIOverC*), vidinių nemokų suma (*OIOverS*), vidinių nemokų trukmė (*OI-OverL*), išorinių nemokų skaičius (*OEOverC*), išorinių nemokų suma (*OEOverS*), išorinių nemokų trukmė (*OEOverL*), įsipareigojimų nevykdymas (*ODefaults*).

Kredito prieinamumo MVĮ vertinimo metodologija yra suskirstyta į tris etapus. I etape atliekama palyginamoji tiriamų Baltijos šalių analizė pasitelkiant aprašomąją kintamųjų statistiką, išskiriant šalims būdingas savybes. II etape atliekama kintamųjų mažinimo procedūra, kuria apibrėžiami reprezentatyvūs veiksnių rinkiniai. Procedūra atliekama apskaičiuojant veiksnių koreliacijos koeficientus ir pasitelkiant ‘Euclidean’ atstumo klasterizavimą bei koreliacijos šilumos žemėlapius. III etape sukuriamas kredito prieinamumo MVĮ vertinimo modelis naudojantis atrinktais mašininio mokymosi metodais – atsitiktiniai miškai, gradialinis nusileidimas, daugiasluoksnis perceptronas ir palyginamuoju etalonu – logistine regresija. Tiksliausias mašininio mokymosi metodas naudojamas vertinant kredito prieinamumą lemiančius veiksnius, įvertinant individualią veiksnių svarbą ir tarpusavio sąveiką. Kredito prieinamumo mažoms ir vidutinėms įmonėms vertinimo modelis yra pateiktas 3-iame paveiksle.



29 pav. Kredito prieinamumo MVĮ vertinimo modelis. Sudaryta autoriaus

3. EMPIRINIS KREDITO PRIEINAMUMO MAŽOMS IR VIDUTINĖMS ĮMONĖMS VERTINIMAS PAGAL BALTIJOS ŠALIŲ PAVYZDĮ

Trečioje disertacijos dalyje sprendžiamas 5-as uždavinys. Kredito prieinamumas mažoms ir vidutinėms įmonėms vertinimo empirinis tyrimas atliekamas Lietuvoje, Latvijoje ir Estijoje. I etape atliekama palyginamoji analizė, siekiant nustatyti esamą kredito prieinamumą mažoms ir vidutinėms įmonėms kiekvienoje Baltijos šalyje. II etape kintamųjų mažinimo procedūra atliekama siekiant apibrėžti konkrečiai šaliai būdingus reprezentatyvius veiksnių rinkinius. III etape sukuriamas kredito prieinamumo mažoms ir vidutinėms įmonėms modelis ir įvertinami individualūs veiksniai bei jų tarpusavio sąveikos svarba.

I etape, atlikus palyginamąją priklausomo kintamojo (*Outcome*) ir nepriklausomų kintamųjų analizę, nustatyta, kad esamas kredito prieinamumas MVĮ nėra tolygiai pasiskirstęs tirtose šalyse – jis didesnis Estijoje ir mažesnis Latvijoje ir Lietuvoje. Be to, esamas kredito prieinamumas nėra pastovus skirtingų laikotarpių atžvilgiu. Nustatyta, kad įmonės charakteristikų, produkto charakteristikų, sandorio skolinimo technologijų ir santykių skolinimo technologijų veiksnių grupės nėra vienodai pasiskirsčiusios Baltijos šalyse. Nustatyta, jog egzistuoja paraiškos kreditui, kurios viršija visuotinai priimtinus likvidumo, mokumo, pelningumo ir apyvartumo lygius, o tai rodo, kad tokiais atvejais, norint gauti kreditą, reikia atsižvelgti į veiksnių bei jų sąveikos visumą. Estijoje vidutinė MVĮ paraiška yra gaunama iš jaunesnės, mažesnės įmonės, kuri turi tvirtesnius banko ir įmonės santykius, stipresnę finansinę būklę ir santykinai daugiau pradelstų mokėjimų, kurie nesibaigia skolinių įsipareigojimų neįvykdymu. Latvijoje vidutinė paraiška gaunama iš ilgiau veikiančios ir didesnės įmonės, kuri teikia audituotas finansines ataskaitas, turi santykinai tvirtus ryšius su banku, tačiau yra silpnesnės finansinės būklės ir su didesne tikimybe, kad neįvykdys skolinių įsipareigojimų. Lietuvoje vidutinė MVĮ paraiška yra iš didesnės įmonės, kuri pateikė paraišką lizingo produktui gauti, turi santykinai trumpesnius ir ne tokius intensyvius banko ir įmonės santykius.

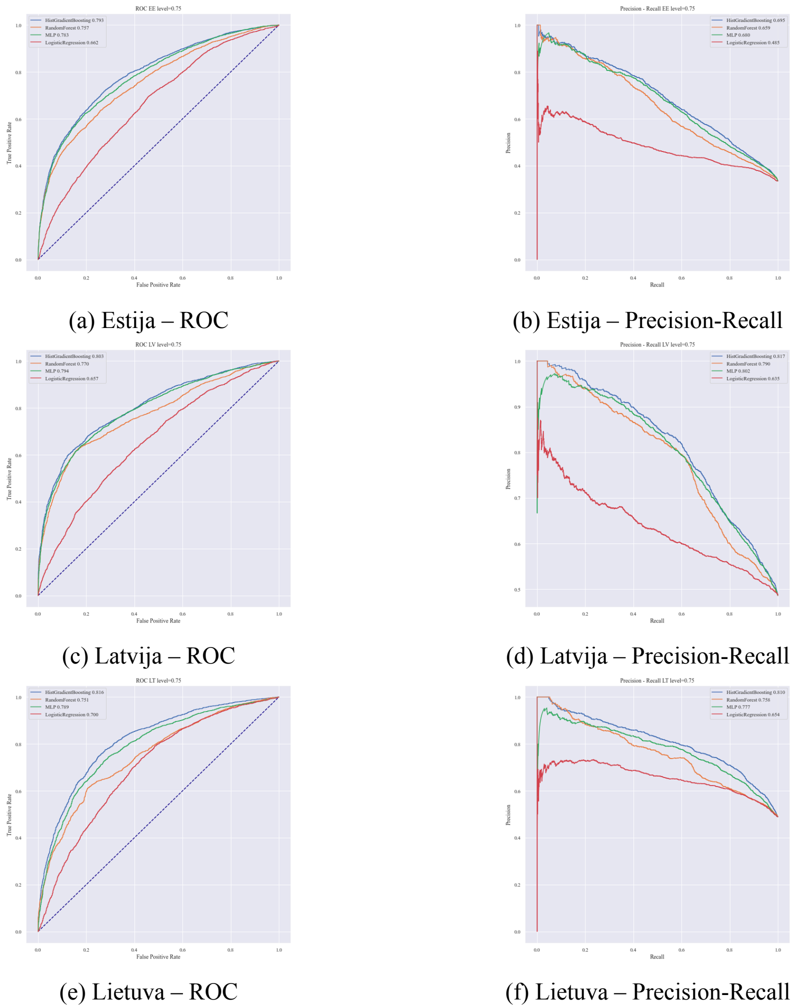
II etape kintamųjų mažinimo procedūra yra atliekama kiekvienai šaliai, siekiant nustatyti reprezentatyvius veiksnių rinkinius. Šios procedūros tikslas – supaprastinti kuriamą kredito prieinamumo vertinimo modelį, kartu užtikrinant modelio tikslumą ir veiksnių interpretavimo galimybes.

54 lentelė. Reprezentatyvūs veiksnių rinkiniai šalių kredito prieinamumo MVĮ vertinimo modeliams. Sudaryta autoriaus

Veiksnys	Estija	Latvija	Lietuva
Įmonės charakteristikų veiksniai			
Diversity	●	●	●
Private	●	●	
Segment	●	●	●
Type	●	●	●
Region	●	●	●
Sector	●	●	●
Produkto charakteristikos veiksnys			
Product	●	●	●
Skolinimo technologijų veiksniai			
Santykių skolinimo technologijų veiksniai			
Relationship	●	●	●
Payments	●	●	●
Rejections	●	●	●
Debt			●
FinContracts	●	●	●
Cards	●	●	
POS			●
Ecommerce	●	●	●
Sandorio skolinimo technologijų veiksniai			
Pagal finansines ataskaitas			
CuR	●	●	●
DE	●	●	●
TA	●	●	●
DR	●	●	●
ROA	●	●	●
GMR	●	●	●
ATR	●	●	●
RTR	●	●	●
CS	●	●	●
pROA	●	●	●
Pagal kredito istoriją			
IOverS	●	●	●
EOverS	●	●	●
Defaults	●	●	●
OIOverS	●	●	●
OEOverS	●	●	●
ODefaults	●	●	●

Nustatyta, kad reprezentatyvių veiksnių rinkinių sudėtis tiriamose šalyse yra panaši, tačiau išskirtina tai, kad tam tikri įmonės charakteristikų ir sandorio skolinimo technologijų veiksniai yra susiję. Dėl to praėjusio laikotarpio finansiniai veiksniai ir kredito istorijos veiksniai nesuteikė papildomos informacijos modeliuojant kredito prieinamumą MVĮ ir nebuvo įtraukti į reprezentatyvų veiksnių rinkinį.

III etape, siekiant įvertinti kredito prieinamumą MVĮ, jį lemiančių veiksnių svarbą ir sąveiką, sukurti Estijos, Latvijos ir Lietuvos šalių kredito prieinamumo MVĮ modeliai. Visų trijų šalių atveju tiksliausia kredito prieinamumo MVĮ modeliavimo metodika yra gradialinis nusileidimas. O mažiausiai tikslus naudotas mašininio mokymosi metodas buvo etalonas – logistinė regresija, kuri pademonstravo žemiausias ROC-AUC ir Avr. Prec. vertes visuose trijuose šalių modeliuose. Bendras didžiausias modeliavimo tikslumas buvo pasiektas Lietuvoje (ROC-AUC – 0,816, Avr. Prec. – 0,810) ir Latvijoje (ROC-AUC – 0,804, Avr. Prec. – 0,818). Estijoje kredito prieinamumo MVĮ modelio tikslumas buvo žemiausias, kur ROC-AUC – 0,796 ir Avr. Prec. – 0,697 (žr. 30 pav.).



(a) Estija – ROC

(b) Estija – Precision-Recall

(c) Latvija – ROC

(d) Latvija – Precision-Recall

(e) Lietuva – ROC

(f) Lietuva – Precision-Recall

30 pav. ROC ir Precision-Recall kreivės Baltijos šalių kredito prieinamumo MVĮ mode-liams. Sudaryta autoriaus

Paskutinis III etapo žingsnis kredito prieinamumo MVĮ modelyje yra nustatyti kredito prieinamumą lemiančių veiksnių įtaką. Šis žingsnis parodo atskirų veiksnių ir jų sąveikos svarbą bei poveikį kredito prieinamumui MVĮ. Atskirų veiksnių įtaka kredito prieinamumui MVĮ įvertinama naudojant vidutinius absoliučiuosius SHAP ir PFI, siekiant nustatyti, kurie veiksniai yra svarbiausi. Vidutinis absoliutus SHAP yra metodas, įvertinantis kiekvieno veiksnio svarbą apskaičiuojant vidutinę modelio prognozę, kai veiksnys yra arba jo nėra. Antrasis metodas – PFI įvertina kiekvieno veiksnio svarbą matuojant modelio našumo pokytį, kai veiksnių reikšmės yra atsitiktinai keičiamos. Svarbu paminėti, kad PFI ir vidutinės absoliučios SHAP reikšmės nėra tiesiogiai palyginamos, tačiau jos abi leidžia įvertinti veiksnių svarbą. Individuali veiksnių svarba pateikta 55 lentelėje.

55 lentelė. Baltijos šalių kredito prieinamumo MVĮ modelio individuali veiksnių svarba. Sudaryta autoriaus

Estija			Latvija			Lietuva		
Veiksnys	**SHAP**	**PFI**	**Veiksnys**	**SHAP**	**PFI**	**Veiksnys**	**SHAP**	**PFI**
FinContracts	0,549	0,107	*Product*	0,907	0,063	*FinContracts*	0,957	0,316
Product	0,399	0,054	*FinContracts*	0,527	0,144	*Type*	0,296	0,003
Rejections	0,266	0,043	*Rejections*	0,283	0,033	*Debt*	0,230	0,044
Relationship	0,135	0,006	*ATR*	0,167	0,022	*DR*	0,179	0,015
Payments	0,129	0,006	*Payments*	0,156	0,009	*Payments*	0,147	0,006
IOverS	0,123	0,005	*DR*	0,087	0,004	*Rejections*	0,132	0,012
Segment	0,096	0,001	*Sector*	0,081	0,004	*Product*	0,114	0,072
ATR	0,087	0,003	*ROA*	0,080	0,002	*ROA*	0,091	0,003
TA	0,072	0,003	*IOverS*	0,067	0,002	*CS*	0,069	0,002
pROA	0,063	0,001	*Relationship*	0,065	0,002	*Sector*	0,064	0,002
ROA	0,063	0,002	*CuR*	0,060	0,001	*pROA*	0,058	0,002
GMR	0,057	0,001	*Type*	0,055	0,000	*RTR*	0,058	0,003
RTR	0,057	0,002	*CS*	0,049	0,002	*GMR*	0,057	0,002
DR	0,051	0,002	*RTR*	0,038	0,000	*Relationship*	0,056	0,003
Sector	0,036	0,003	*TA*	0,033	0,001	*Segment*	0,042	0,002
OIOverS	0,035	0,001	*GMR*	0,032	0,000	*EOverS*	0,040	0,001
CS	0,033	0,000	*pROA*	0,031	0,000	*ATR*	0,039	0,001
CuR	0,032	0,001	*DE*	0,024	0,000	*TA*	0,038	0,001
EOverS	0,030	0,001	*Cards*	0,023	0,000	*IOverS*	0,036	0,000
Region	0,026	0,000	*Diversity*	0,013	0,000	*POS*	0,032	0,000
Private	0,017	0,000	*Private*	0,009	0,001	*CuR*	0,023	0,001
Cards	0,017	0,000	*OEOverS*	0,005	0,001	*OEOverS*	0,016	0,000
Ecommerce	0,007	0,000	*OIOverS*	0,005	0,000	*OIOverS*	0,012	0,001
OEOverS	0,003	0,000	*Defaults*	0,005	0,000	*Region*	0,011	0,000
Defaults	0,003	0,000	*Segment*	0,005	0,000	*Defaults*	0,005	0,000
Type	0,002	0,000	*EOverS*	0,004	0,000	*Diversity*	0,001	0,000
ODefaults	0,000	0,000	*Ecommerce*	0,001	0,000	*ODefaults*	0,000	0,000
Diversity	0,000	0,000	*ODefaults*	0,000	0,000	*Ecommerce*	0,000	0,000

55 lentelėje pateikiami veiksniai pagal jų svarbą, suklasifikuojant veiksnius pagal vidutines absoliučias SHAP vertes. Pagal didžiausią vidutinį absoliutų SHAP (Estija – 0,55, Latvija – 0,53, Lietuva – 0,96) ir PFI vertę (Estija – 0,11, Latvija – 0,14, Lietuva –

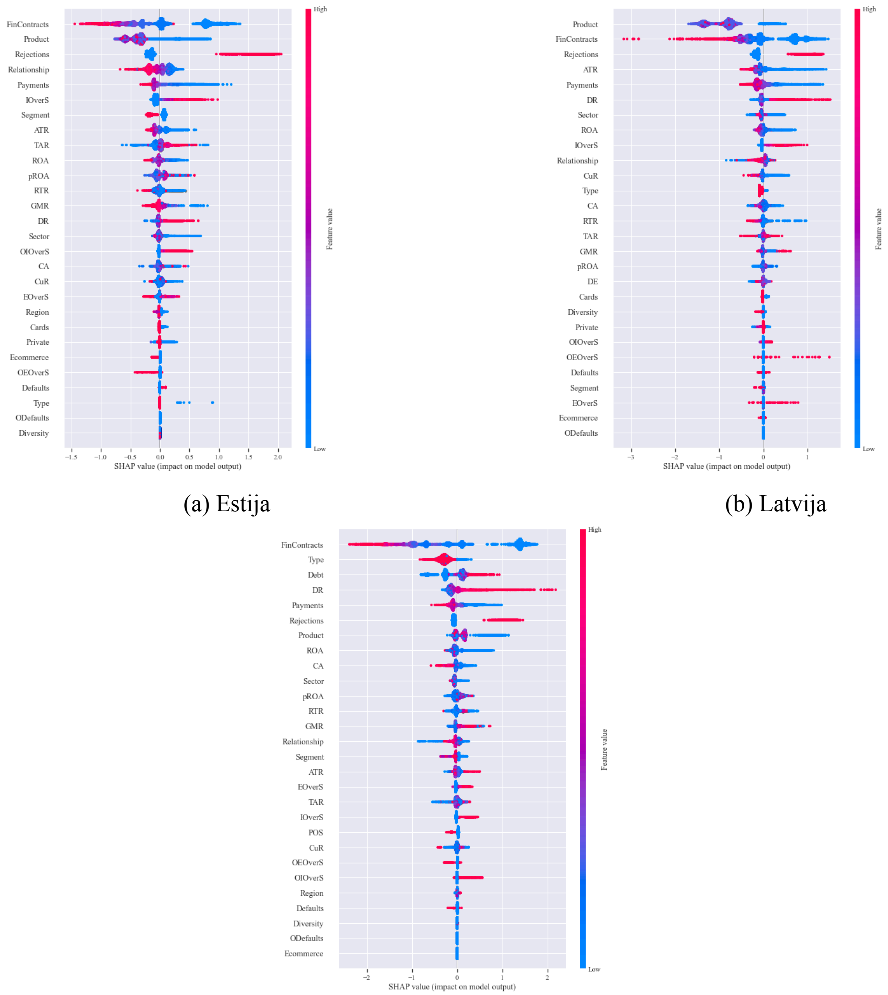
0,32) finansinių sutarčių skaičius *FinContracts* yra vienas iš svarbiausių veiksnių vertinant kredito prieinamumą mažoms ir vidutinėms įmonėms visose trijose šalyse. Didžiausias veiksnių svarbos skirtumas yra Lietuvoje, kur *FinContracts* svarba vidutiniu absoliučiu SHAP atžvilgiu sudaro beveik pusę vidutinio absoliutaus SHAP sudėjus visus veiksnius kartu. *Product* yra svarbiausias veiksnys Latvijoje ir antras pagal svarbą Estijoje (Lietuvoj ji yra tik 6-as). Nors *Product* Latvijoje turi didžiausią vidutinį absoliutų SHAP, jis neturi didžiausio PFI, o tai rodo, kad šis veiksnys labiausiai prisideda prie kredito prieinamumo rezultato, o *FinContracts* labiausiai prisideda prie modeliavimo tikslumo. *Rejections* yra 3-ias svarbiausias veiksnys Estijoje ir Latvijoj, o Lietuvoje tik 5-as. Visose trijose šalyse savininko įsipareigojimų neįvykdymo (*ODefaults*) vertės buvo lygios nuliui, tai rodo, kad šis veiksnys nėra svarbus kredito prieinamumui MVĮ. Estija ir Latvija turi santykinai panašų veiksnių pasiskirstymą pagal svarbą, o Lietuvoje tokie veiksniai, kaip *Type* (vidutinis absoliutus SHAP – 0,3, PFI – 0,002), *Debt* (vidutinis absoliutus SHAP – 0,23, PFI – 0,04) ir *Payments* (vidutinis absoliutus SHAP – 0,15, PFI – 0,01), yra tarp 5-ių svarbiausių veiksnių. Reikėtų pažymėti, kad visų trijų šalių modeliuose dauguma skolinimo technologijų veiksnių (*CS, CuR, DR, RTR, GMR, ROA, pROA, TA, ATR*) nėra individualiai svarbūs PFI (<0,005) atžvilgiu. Tai galima paaiškinti tuo, kad atskirų veiksnių svarba susilpnėja įtraukiant didesnę dalį kintamųjų, priklausančių tai pačiai veiksnių grupei.

Nustatyta, kad atskirų veiksnių svarba ir įtaka MVĮ galimybėms gauti kreditą įvairiose šalyse skiriasi ir priklauso nuo individualaus modelio specifikos bei galimai esamų sąveikų. Norint suprasti, kaip veiksnių grupės lemia modeliuojamą kredito prieinamumą MVĮ, atliekamas veiksnių grupavimas pagal vidutines absoliučias SHAP ir PFI reikšmes. 56 lentelėje pateikiamas apibendrintas vaizdas, kaip kiekviena veiksnių grupė prisideda prie modeliuojamo kredito prieinamumo MVĮ.

56 lentelė. Veiksnių grupių svarba kredito prieinamumo MVĮ modeliui. Sudaryta autoriaus

	Estija		**Latvija**		**Lietuva**	
	SHAP	**PFI**	**SHAP**	**PFI**	**SHAP**	**PFI**
Įmonės charakteristikų veiksniai	0,176	0,005	0,163	0,005	0,414	0,006
Produkto charakteristikos veiksnys	0,399	0,054	0,907	0,063	0,114	0,072
Skolinimo technologijų veiksniai	1,812	0,186	1,743	0,226	2,128	0,407
Santykių skolinimo technologijų veiksniai	1,103	0,162	1,055	0,189	1,407	0,375
Sandorio skolinimo technologijų veiksniai	0,709	0,024	0,687	0,037	0,722	0,032
Pagal finansines ataskaitas	0,515	0,016	0,600	0,033	0,611	0,029
Likvidumas	0,032	0,001	0,060	0,001	0,023	0,001
Mokumas	0,123	0,005	0,143	0,005	0,217	0,015
Pelningumas	0,183	0,005	0,142	0,003	0,206	0,007
Apyvartumas	0,177	0,006	0,254	0,023	0,166	0,006
Pagal kredito istoriją	0,195	0,008	0,087	0,004	0,110	0,003
Įmonės	0,156	0,007	0,077	0,003	0,082	0,002
Savininkų	0,039	0,001	0,010	0,001	0,028	0,001

Nustatyta, jog skolinimo technologijų veiksniai yra svarbiausia faktorių grupė visose trijose Baltijos šalyse, iš kurių santykių skolinimo technologijų veiksniai sudaro daugiau nei 60% vidutinio absoliutaus SHAP ir daugiau nei 87% PFI. Kredito istorijos veiksniai turi panašų indėlį įvairiose šalyse. Įmonės charakteristika yra mažiausiai svarbi veiksnių grupė Estijoje ir Latvijoje, o produkto charakteristikos veiksnys yra mažiausiai svarbus Lietuvoje. Veiksnių grupių svarba ir poveikis įvairiose rinkose skiriasi, tačiau skolinimo technologijos svarba nustatyta visose rinkose. Vertinant individualių veiksnių svarbą neatsižvelgiama į galimą kintamųjų sąveiką ir priklausomybę, todėl norint įvertinti, kaip kinta atskirų veiksnių svarba priklausomai nuo jų verčių, naudojami SHAP grafikai (žr. 31 pav.).



(a) Estija

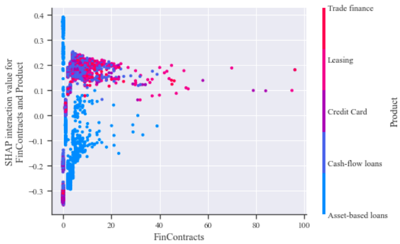
(b) Latvija

(c) Lietuva

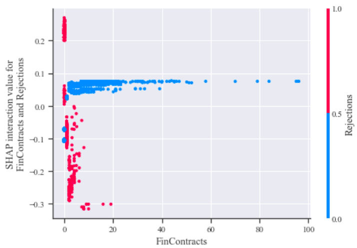
31 pav. Veiksnių svarba kredito prieinamumo MVĮ modeliui, priklausomai nuo jų verčių. Sudaryta autoriaus

Nustatyta, kad veiksnių poveikio kredito prieinamumui MVĮ mastas nėra vienodas pagal visas veiksnių reikšmes tiriamose Baltijos šalyse. Priklausomai nuo šalies, skirtingas *Product* pasirinkimas turi skirtingą poveikį kredito prieinamumui. Nustatyta, kad nauji banko ir įmonės santykiai galimai turės didesnį teigiamą efektą kredito prieinamumui MVĮ nei jau esami santykiai. Taip pat nustatyta, jog kuo intensyvesni yra banko ir įmonės santykiai (daugiau finansinių sutarčių (*FinContracts*), daugiau produktų (*Cards*, *POS*)), tuo didesnė tikimybė, kad įmonės paraiška bus patvirtina.

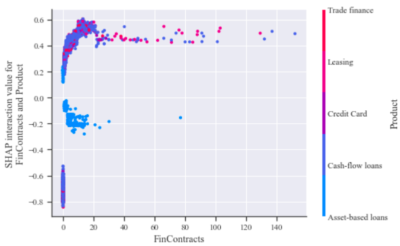
Paskutinis šio etapo žingsnis yra nustatyti ir įvertinti veiksnių sąveiką (žr. 32 pav.).



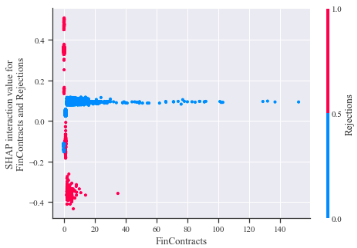
(a) Estija – *FinContracts – Product*



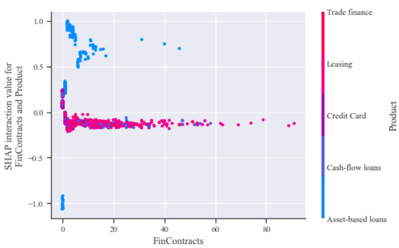
(b) Estija – *FinContracts – Rejections*



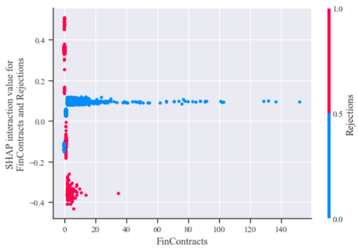
(c) Latvija – *FinContracts – Product*



(d) Latvija – *FinContracts – Rejections*



(e) Lietuva – *FinContracts – Product*



(f) Lietuva – *FinContracts – Rejections*

32 pav. SHAP sąveikos grafikai *FinContracts–Product* ir *FinContracts–Rejections* veiksnių poroms. Parengta autoriaus

Nustatyta, kad visų šalių kredito prieinamumo MVĮ modeliai rodo finansinių sutarčių skaičiaus, produkto ir istorinių atmetimų sąveikos svarbą. Šios sąveikos poveikis kredito prieinamumo modeliui nėra vienodas pagal visas produktų reikšmes. 32 paveiksle parodyta, kaip finansinių sutarčių ir produkto sąveika veikia kredito prieinamumo modelį, o tai turi didelį poveikį turtu pagrįstoms paskoloms. Estijoje ir Latvijoje daugiau nei vienos finansavimo sutarties turėjimas ir turtu pagrįstos paskolos prašymas teigiamai prisideda prie kredito gavimo, o Lietuvoje šis ryšys yra priešingas. MVĮ, kurios neturi ankstesnių finansinių sutarčių ir kurios kreipiasi dėl kredito kortelės arba pinigų srautų paskolos, paprastai yra mažiau normuotos nei tos, kurios kreipiasi dėl kitų produktų. Sąveika tarp finansinių sutarčių ir atmetimų egzistuoja visose trijose šalyse ir turi didelį neigiamą poveikį kredito prieinamumui MVĮ, kai nėra ankstesnių finansinių sutarčių. Neigiamas efektas mažėja sudarius bent vieną finansinę sutartį, kol augant finansinių sutarčių skaičius tampa teigiamas. Tai rodo, kad ankstesnės finansinės sutartys su banku gali pagerinti MVĮ galimybes gauti kreditą, nepaisant ankstesnės neigiamos informacijos.

IŠVADOS

1. Mažos ir vidutinės įmonės yra neatsiejama pasaulio ekonomikos dalis, todėl svarbu užtikrinti jų finansinę būklę kartu užtikrinant galimybę gauti kreditą. Priešingai nei didesnėms įmonėms, mažoms ir vidutinėms įmonėms sunku gauti kreditą, o tai neigiamai veikia jų augimo galimybes. Ribotos galimybės gauti kreditą daro neigiamą poveikį mažų ir vidutinių įmonių augimui tarptautinėse rinkose, riboja inovacijas ir sąlygoja žemesnį produktyvumą. Galimybė gauti kreditą yra svarbi sąlyga MVĮ augimui ir plėtrai, todėl svarbu suprasti pagrindinius veiksnius, turinčius įtakos kredito prieinamumui, taip sumažinant informacijos asimetriją įmonėms ir reguliavimo institucijoms.

2. Kredito prieinamumo MVĮ ir jį lemiančių veiksnių analizė leido padaryti 2 pagrindines išvadas:

   2.1. Nustatyta, jog kredito prieinamumo vertinimas gali būti atliekamas tiriant kredito pasiūlą arba paklausą. Konkretaus būdo pasirinkimas priklauso nuo tyrimo problemos ir duomenų prieinamumo. Kredito paklausos tyrimuose daugiausia dėmesio yra skiriama veiksniams, turintiems įtakos įmonių finansavimo poreikiams ir kredito pareiškėjų atgrasymo priežastims. Kita vertus, kreditų pasiūlos tyrimuose daugiausia dėmesio skiriama tiriant dvi kryptis: banko paskolų portfelį ir su juo susijusias makroekonomines sąlygas arba finansavimo paraiškų rezultatus ir juos lemiančius veiksnius, kurie daro įtaką įmonės paraiškos patvirtinimui, atmetimui (1-o laipsnio normavimas) arba sąlyginiam patvirtinimui (2-o laipsnio normavimas), kai sprendimas suteikti kreditą yra teigiamas, tačiau normuojamas per padidintus reikalavimus finansavimo sumai, kainai, terminui ar užstatui.

   2.2. Kredito prieinamumą mažoms ir vidutinėms įmonėms lemiantys veiksniai yra skirstomi į makrospecifinius ir individualios paraiškos veiksnius. Makrospecifiniai veiksniai, susidedantys iš skolinimo infrastruktūros ir finansų institucijų struktūros, yra nepriklausomi nuo individualių įmonių ir apibrėžia pagrindines rinkos sąlygas. Nustatyta, jog įmonės, veikiančios rinkose, kuriose taikomi griežti apskaitos standartai, ruošiamos finansinės ataskaitos pagal rinkos vertę ir veikia reitingų agentūros, turi mažesnius finansinius suvaržymus ir didesnę prieigą prie kredito. Konkurencingos rinkos paprastai daro didelį teigiamą poveikį galimybei gauti kreditą dėl mažesnės palūkanų normos ir didesnės vidutinės paskolų sumos. Individualūs paraiškos veiksniai nurodo pagrindines konkrečios paskolos charakteristikas, kurias apibrėžia skolinimo technologijų, įmonės charakteristikų ir produkto charakteristikų veiksnių grupės. Mažoms ir vidutinėms įmonėms galimybę gauti kreditą riboja įvairūs veiksniai, priklausantys makrospecifinėms ir individualios paraiškos veiksnių grupėms, tačiau tyrimų, kompleksiškai vertinančių kredito prieinamumą MVĮ, nėra. Nustatyta, jog nedaug tyrimų tiria mažo ir vidutinio verslo galimybes gauti kreditą atsižvelgiant į pageidaujamą gauti banko produktą, todėl vis dar lieka neaišku, ar įmonės pasirinkimas kreiptis dėl konkretaus finansavimo produkto gali turėti įtakos galutiniam finansavimo sprendimui. Todėl daroma išvada, kad, vertinant kredito prieinamumą mažoms ir vidutinėms įmonėms būtina atsižvelgti į makrospecifinius ir į individualius paraiškos veiksnius, vertinant ir pasirinkto finansavimo produkto svarbą.

3. Nustatyta, kad šiuolaikinių mašininio mokymosi metodų naudojimas išlieka ribotas vertinant įmonių kredito prieinamumą. Siekiant sukurti veiksmingą kredito prieinamumo mažoms ir vidutinėms įmonėms vertinimo modelį, turi būti naudojami įvairūs modeliavimo metodai ir lyginami su etalonu. Nors tradiciniai metodai, tokie kaip diskriminacinė analizė ir logistinė regresija istoriškai buvo naudojami kaip etalonas, naujausi tyrimai parodė, kad mašininio mokymosi metodai, tokie kaip atsitiktiniai miškai, daugiasluoksnis perceptronas, gradialinis nusileidimas, yra tikslesni. Nustatyta, jog modernių mašininio mokymosi metodų naudojimas gali pagerinti tikslumą ir sumažinti modeliavimo šališkumą. Tačiau šiems metodams būdinga ribota rezultatų interpretacija, todėl, norint įvertinti kredito prieinamumą lemiančius veiksnius, reikia naudoti paaiškinamumo metodus, tokius kaip SHAP, kurie yra ypač veiksmingi pabrėžiant atskirų veiksnių ar jų sąveikos svarbą.

4. Sukurtas originalus kredito prieinamumo mažoms ir vidutinėms įmonėms vertinimo kompleksinis modelis, kurio įgyvendinimas apima tris tyrimo etapus. Apibrėžta, kad kredito prieinamumas mažoms ir vidutinėms įmonėms yra vertinamas atsižvelgiant į finansavimo paraiškos rezultatą, kuris yra priklausomas kintamasis. Nepriklausomi kintamieji yra sugrupuoti į įmonės charakteristikų, produkto charakteristikų ir skolinimo technologijų veiksnių grupes. I etape esamas kredito prieinamumas mažoms ir vidutinėms įmonėms yra įvertinamas remiantis palyginamąja analize. II etape kintamųjų mažinimo procedūra atliekama reprezentatyviam veiksnių rinkiniui apibrėžti. III etape naudojami pažangiausi

mašininio mokymosi metodai, siekiant įvertinti mažų ir vidutinių įmonių prieigos prie kredito modelį ir įvertinti atskirų veiksnių, jų grupių ir sąveikos poveikį bei įtaką kredito prieinamumui.

5. Empiriškai patikrintus kompleksinį kredito prieinamumo mažoms ir vidutinėms įmonėms vertinimo modelį Baltijos šalyse, padarytos trys pagrindinės išvados:

    5.1. Remiantis palyginamąja priklausomo kintamojo ir nepriklausomų kintamųjų veiksnių grupių analize, nustatyta, kad esamas mažų ir vidutinių įmonių kredito prieinamumas nėra vienodas tirtose Baltijos šalyse. Nustatyta, kad mažų ir vidutinių įmonių kredito prieinamumas yra didesnis Estijoje, o mažesnis – Latvijoje ir Lietuvoje. Be to, analizė atskleidė, kad pagrindinės įmonės charakteristikų, produktų charakteristikų, skolinimo technologijų veiksnių grupės nėra vienodai pasiskirsčiusios tiriamose šalyse. Santykinę su finansinėmis ataskaitomis susijusių veiksnių svarbą lemia ir veiksnių faktinės vertės, ypač kai jos siekia kraštutinius dydžius. Todėl, siekiant tinkamai įvertinti kredito prieinamumą MVĮ, svarbu atsižvelgti į veiksnių bei jų sąveikos visumą. Estijoje vidutinė mažos ir vidutinės įmonės paraiška yra gaunama iš jaunesnės ir mažesnės įmonės, kuri turi stipresnius banko ir įmonės ryšius, stipresnę finansinę būklę ir santykinai daugiau pradelstų mokėjimų, kurie nepriveda prie rimtesnio įsipareigojimų nevykdymo. Latvijoje vidutinė paraiška gaunama iš senesnės ir didesnės įmonės, kuri dažniau pateikia audituotas finansines ataskaitas, turi santykinai tvirtus ryšius su banku ir įmonėmis, tačiau turi silpnesnę finansinę būklę ir didesnę istorinių įsipareigojimų nevykdymo tikimybę. Lietuvoje vidutinė mažos ir vidutinės įmonės paraiška gaunama iš didesnės įmonės, kuri teikia paraišką gauti lizingo produktą ir turi santykinai trumpesnius ir ne tokius intensyvius santykius su banku.

    5.2. Baltijos šalių kredito prieinamumas mažoms ir vidutinėms įmonėms buvo vertinamas naudojant atsitiktinių miškų, daugiasluoksnio perceptrono, gradialinio nusileidimo pažangiausius mašininio mokymosi metodus ir tradicinę logistinę regresiją, kuri buvo naudojama kaip etalonas. Nustatyta, kad, vertinant kredito prieinamumą mažoms ir vidutinėms įmonėms, tiksliausias modeliavimo metodas yra gradialinis nusileidimas, o didžiausias modelio tikslumas nustatytas Lietuvoje ir Latvijoje. Logistinė regresija, kuri yra viena iš dažniausiai naudojamų mašininio mokymosi metodų ir buvo pasirinkta kaip etalonas, pasiekė prasčiausią modeliavimo tikslumą.

    5.3. Nustatyta, kad svarbiausia veiksnių grupė, vertinant kredito prieinamumą mažoms ir vidutinėms įmonėms, yra skolinimo technologijų veiksniai, konkrečiai – santykių skolinimo technologijų veiksnių grupė. Tarp individualių veiksnių nustatyta, kad finansinių sutarčių skaičius yra vienas iš svarbiausių kintamųjų visose šalyse, jo įtaka modelio rezultatui priklauso nuo faktinio finansinių sutarčių skaičiaus. Be to, tyrimas parodė, kad svarbios ne tik atskirų veiksnių reikšmės, bet ir jų sąveika su kitais veiksniais. Buvo nustatyta stipri sąveika tarp finansinių sutarčių skaičiaus ir paraiškos produkto bei finansinių sutarčių skaičiaus ir istorinių paraiškų atmetimų porų. Tai rodo, kad mažų ir vidutinių įmonių, kurios praeityje turėjo finansavimo sutartis, kredito prieinamumas nesuprastėja dėl istorinių paraiškų atmetimų.

6. Remiantis empirinio tyrimo rezultatais ir gautomis išvadomis, akivaizdu, kad kredito prieinamumo mažoms ir vidutinėms įmonėms analizavimas reikalauja kompleksinio vertinimo, plataus spektro veiksnių panaudojimo ir pažangiausių mašininio mokymosi metodų taikymo. Empirinis kredito prieinamumo MVĮ modelio taikymas ir naujausių modeliavimo metodų naudojimas parodė, kad pagrindinių veiksnių svarba faktinėse veiksnių reikšmėse nėra vienoda. Nustatyta, kad egzistuoja netiesinis ryšys tarp kai kurių veiksnių ir MVĮ galimybės gauti kreditą. Pavyzdžiui, ankstesniuose tyrimuose pastebėta, jog banko ir įmonės santykių trukmė lemia, kad ilgesni santykiai suteikia daugiau galimybių gauti kreditą, o šioje disertacijoje buvo nustatyta, kad visiškai nauji banko ir įmonės santykiai lemia daug didesnę tikimybę gauti kreditą nei vidutinės trukmės. Šios išvados suteikia vertingų įžvalgų tiek įmonėms, kurios bando gauti kreditą, tiek reguliuotojams, ieškantiems priemonių pagerinti MVĮ galimybes gauti kreditą. Tolesni tyrimai galėtų būti atliekami keliomis kryptimis, praplečiant tyrimą ir įtraukiant dar platesnį pagrindinių veiksnių spektrą, pradedant nuo įmonės savininko asmeninių savybių ir baigiant su įmonės tvarumu susijusiais duomenimis. Kita galima kryptis būtų atgrasytų kredito pareiškėjų integravimas, siekiant suformuoti bendrinį kredito prieinamumo mažoms ir vidutinėms įmonėms vertinimo modelį, pagal kurį būtų galima įvertinti kredito prieinamumą tiek kredito paklausos, tiek kredito pasiūlos atžvilgiu.

REFERENCES

Acharya, V. V. and Steffen, S. (2020). The risk of being a fallen angel and the corporate dash for cash in the midst of covid. *SSRN Electronic Journal*.

Adam, T. R. and Streitz, D. (2016). Hold-up and the use of performance-sensitive debt. *Journal of Financial Intermediation*, 26:47–67.

Ademosu, A. and Morakinyo, A. (2021). Financial system and smes access to finance: A market-oriented approach. *Studia Universitatis Vasile Goldis Arad Seria Stiinte Economice*, 31(3):21–36. Ademosu, Akinwande Morakinyo, Akinola 2285-3065.

Adrian, T. and Shin, H. S. (2010). Liquidity and leverage. *Journal of Financial Intermediation*, 19(3):418–437.

Agarwal, S. and Hauswald, R. (2010). Distance and Private Information in Lending. *The Review of Financial Studies*, 23(7):2757–2788.

Aghion, P., Askenazy, P., Berman, N., Cette, G., and Eymard, L. (2012). Credit constraints and the cyclicality of r&d investment: Evidence from france. *Journal of the European Economic Association*, 10(5):1001–1024.

Aiyar, S., Calomiris, C. W., and Wieladek, T. (2014). Does macro-prudential regulation leak? evidence from a uk policy experiment. *Journal of Money, Credit and Banking*, 46(s1):181–214.

Akande, J. O., Nzimande, N., Chisasa, J., and Sunde, T. (2021). Bank competition or concentration: Which is more important for access to finance in africa? *African Review of Economics and Finance-Aref*, 13(1):321–347. Akande, Joseph Olorunfemi Nzimande, Ntokozo Chisasa, Joseph Sunde, Tafirenyika Sunde, Tafirenyika/ABB-8803-2021 Sunde, Tafirenyika/0000-0002-9124-9383 2410-4906.

Akoglu, H. (2018). User's guide to correlation coefficients. *Turk J Emerg Med*, 18(3):91–93.

Altavilla, C., Boucinha, M., Holton, S., and Ongena, S. (2021). Credit supply and demand in unconventional times. *Journal of Money, Credit and Banking*.

Altomonte, C. and Békés, G. (2016). *Measuring Competitiveness in Europe: Resource Allocation, Granularity and Trade*. Number 12051 in Blueprints. Bruegel.

Amadasun, D. O. E. and Mutezo, A. T. (2022). Influence of access to finance on the competitive growth of smes in lesotho. *Journal of Innovation and Entrepreneurship*, 11(1):56.

Angori, G., Aristei, D., and Gallo, M. (2019). Lending technologies, banking relationships, and firms' access to credit in italy: the role of firm size. *Applied Economics*, 51(58):6139–6170.

Angori, G., Aristei, D., and Gallo, M. (2020). Banking relationships, firm-size heterogeneity and access to credit: Evidence from european firms. *Finance Research Letters*, 33:101231.

Aoki, Y. (2021). The effect of bank relationships on bond spreads: Additional evidence from japan. *Journal of Corporate Finance*, 68:101937.

Arezki, R., Bolton, P., Peters, S., Samama, F., and Stiglitz, J. (2017). From global savings glut to financing infrastructure. *Economic Policy*, 32(90):221–261.

Aristei, D. and Angori, G. (2022). Heterogeneity and state dependence in firms' access to bank credit. *Small Business Economics*, 59:47–78.

Ariza-Garzon, M. J., Arroyo, J., Caparrini, A., and Segovia-Vargas, M.-J. (2020). Explainability of a machine learning granting scoring model in peer-to-peer lending. *IEEE Access*, 8:64873–64890.

Armstrong, C. S., Guay, W. R., and Weber, J. P. (2010). The role of information and financial reporting in corporate governance and debt contracting. *Journal of Accounting and Economics*, 50(2-3):179–234.

Arping, S. (2019). Competition and risk taking in banking: The charter value hypothesis revisited. *Journal of Banking & Finance*, 107:105609.

Arrieta, A. B., Díaz-Rodríguez, N., Ser, J. D., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., Chatila, R., and Herrera, F. (2020). Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai. *Information Fusion*. General overview
Levels of transparency: interpretability, decomposability, algorithmic transparency.

Arrow, K. J. (1962). The economic implications of learning by doing. *The Review of Economic Studies*, 29(3):155–173.

Arya, V., Bellamy, R. K. E., Chen, P.-Y., Dhurandhar, A., Hind, M., Hoffman, S. C., Houde, S., Liao, Q. V., Luss, R., Mojsilović, A., Mourad, S., Pedemonte, P., Raghavendra, R., Richards, J., Sattigeri, P., Shanmugam, K., Singh, M., Varshney, K. R., Wei, D., and Zhang, Y. (2019). One explanation does not fit all: A toolkit and taxonomy of ai explainability techniques. *arXiv*.

Aterido, R. and Hallward-Driemeier, M. (2011). Whose business is it anyway? *Small Business Economics*, 37(4):443–464.

Au, Q., Herbinger, J., Stachl, C., et al. (2022). Grouped feature importance and combined features effect plot. *Data Mining and Knowledge Discovery*, 36:1401–1450.

Ayalew, M. M. and Xianzhi, Z. (2019). Bank competition and access to finance: Evidence from african countries. *Journal of Industry, Competition and Trade*, 19(1):155–184.

Banerjee, R. N., Gambacorta, L., and Sette, E. (2021). The real effects of relationship lending. *Journal of Financial Intermediation*.

Barboza, F., Kimura, H., and Altman, E. (2017). Machine learning models and bankruptcy prediction. *EXPERT SYSTEMS WITH APPLICATIONS*, 83:405–417.

Basiglio, S., Vincentiis, P., Isaia, E., and Rossi, M. (2022). Women-led firms and credit access. a gendered story? *Italian Economic Journal*.

Beck, T. (2002). Financial development and international trade. *Journal of International Economics*, 57(1):107–131.

Beck, T., Degryse, H., De Haas, R., and Van Horen, N. (2018). When arm's length is too far: Relationship banking over the credit cycle. *Journal of Financial Economics*, 127(1):174–196.

Beck, T., Demirguc-Kunt, A., and Maksimovic, V. (2004). Bank competition and access to finance: International evidence. *Journal of Money Credit and Banking*, 36(3):627–648. Beck, T Demirguc-Kunt, A Maksimovic, V Conference on Bank Concentration and Competition May 21-23, 2003 Cleveland, OH Fed Reserve Bank Cleveland 1538-4616 2.

Beck, T., Demirguc-Kunt, A., and Martinez Peria, M. S. (2007). Reaching out: Access to and use of banking services across countries. *Journal of Financial Economics*, 85(1):234–266.

Beck, T., Demirgüç-Kunt, A., and Honohan, P. (2009). Access to financial services: Measurement, impact, and policies. *The World Bank Research Observer*, 24(1):119–145.

Belaid, F., Boussaada, R., and Belguith, H. (2017). Bank-firm relationship and credit risk: An analysis on tunisian firms. *Research in International Business and Finance*, 42:532–543.

Bellone, F., Musso, P., Nesta, L., and Schiavo, S. (2010). Financial constraints and firm export behaviour. *World Economy*, 33(3):347–373.

Bento, N., Gianfrate, G., and Groppo, S. V. (2019). Do crowdfunding returns reward risk? evidences from clean-tech projects. *Technological Forecasting and Social Change*, 141:107–116.

Berger, A., Bouwman, C., Norden, L., Roman, R., Udell, G., and Wang, T. (2022). Is a friend in need a friend indeed? how relationship borrowers fare during the covid-19 crisis. *SSRN Electronic Journal*.

Berger, A. N., Bouwman, C. H. S., and Kim, D. (2017). Small bank comparative advantages in alleviating financial constraints and providing liquidity insurance over time. *The Review of Financial Studies*, 30(10):3416–3454.

Berger, A. N., Cowan, A. M., and Frame, W. S. (2011). The surprising use of credit scoring in small business lending by community banks and the attendant effects on credit availability, risk, and profitability. *Journal of Financial Services Research*, 39(1-2):1–17.

Berger, A. N. and Roman, R. A. (2020). *TARP and Other Bank Bailouts and Bail-ins around the World: Connecting Wall Street, Main Street, and the Financial System*. Academic Press.

Berger, A. N. and Udell, G. F. (2006). A more complete conceptual framework for sme finance. *JOURNAL OF BANKING & FINANCE*, 30(11):2945–2966.

Bertay, A. C., Demirgüç-Kunt, A., and Huizinga, H. (2015). Bank ownership and credit over the business cycle: Is lending by state banks less procyclical? *Journal of Banking & Finance*, 50:326–339.

Biau, G., Cadre, B., Dudoit, S., Janoueix-Lerosey, I., Kermanshahi, S. H., and Robins, J. M. (2020). Histgradientboosting: Efficient and accurate tree-based gradient boosting. *arXiv preprint arXiv:2006.08138*.

Block, J. H., Colombo, M. G., Cumming, D. J., and Vismara, S. (2018). New players in entrepreneurial finance and why they are there. *Small Business Economics*, 50(2):239–250.

BoL (2021). Smulkiojo ir vidutinio verslo finansavimo galimybių tyrimas. *Analize ir Tyrimai*, Nr. 11 / 2021.

Bolton, P., Freixas, X., Gambacorta, L., and Mistrulli, P. E. (2016). Relationship and transaction lending in a crisis. *The Review of Financial Studies*, 29(10):2643–2676.

Bond, Harhoff, and Reenen, V. (2005). Investment, r&d and financial constraints in britain and germany. *Annales d'Économie et de Statistique*, 79/80:433.

Bonnet, J., Cieply, S., and Dejardin, M. (2016). Credit rationing or overlending? an exploration into financing imperfection. *Applied Economics*, 48(57):5563–5580.

Bosch, O. and Steffen, S. (2011). On syndicate composition, corporate structure and the certification effect of credit ratings. *Journal of Banking & Finance*, 35(2):290–299.

Bougheas, S. (2004). Internal vs external financing of r&d. *Small Business Economics*, 22(1):11–17.

Boyd, K., Eng, K. H., and Page, C. D. (2013). Area under the precision-recall curve: Point estimates and confidence intervals. In Blockeel, H., Kersting, K., Nijssen, S., and Železný, F., editors, *Machine Learning and Knowledge Discovery in Databases*, volume 8190 of *Lecture Notes in Computer Science*, pages 451–466. Springer Berlin Heidelberg.

Bradley, A. P. (1997). The use of the area under the roc curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30(7):1145–1159.

Brei, M. and Schclarek, A. (2013). Public bank lending in times of crisis. *Journal of Financial Stability*, 9(4):820–830.

Brei, M. and Von Peter, G. (2018). The distance effect in banking and trade. *Journal of International Money and Finance*, 81:116–137.

Breiman, L. (2001). Random forests. *Machine learning*, 45(1):5–32.

Brooks, W. and Dovis, A. (2020). Credit market frictions and trade liberalizations. *Journal of Monetary Economics*, 111:32–47.

Brown, J. R., Martinsson, G., and Petersen, B. C. (2013). Law, stock markets, and innovation. *The Journal of Finance*, 68(4):1517–1549.

Brown, M., Jappelli, T., and Pagano, M. (2009). Information sharing and credit: Firm-level evidence from transition countries. *Journal of Financial Intermediation*, 18(2):151–172.

Brown, M., Ongena, S., Popov, A., and Yeşin, P. (2011). Who needs credit and who gets credit in eastern europe? *Economic Policy*, 26(65):93–130.

Brown, R., Linares-Zegarra, J. M., and Wilson, J. O. S. (2022). Innovation and borrower discouragement in smes. *Small Business Economics*.

Bucker, M., Szepannek, G., Gosiewska, A., and Biecek, P. (2022). Transparency, auditability, and explainability of machine learning models in credit scoring. *Journal of the Operational Research Society*, 73(1).

Buckley, P. J. and Prescott, K. (1989). The structure of british industry's sales in foreign markets. *Managerial and Decision Economics*, 10(3):189–208.

Bussmann, N., Giudici, P., Marinelli, D., and Papenbrock, J. (2020). Explainable machine learning in credit risk management. *SSRN Electronic Journal*.

Bustos, P. (2011). Trade liberalization, exports, and technology upgrading: Evidence on the impact of mercosur on argentinian firms. *The American Economic Review*, 101(1):304–340.

Butticè, V. and Rovelli, P. (2020). "fund me, i am fabulous!" do narcissistic entrepreneurs succeed or fail in crowdfunding? *Personality and Individual Differences*, 162:110037.

Byiers, B. and et al. (2010). Credit demand in mozambican manufacturing. *Journal of International Development*, 22(1):37–55.

Caggese, A. (2019). Financing constraints, radical versus incremental innovation, and aggregate productivity. *American Economic Journal: Macroeconomics*, 11(2):275–309.

Calabrese, R., Degl'Innocenti, M., and Zhou, S. (2022). Expectations of access to debt finance for smes in times of uncertainty. *Journal of Small Business Management*, 60(6):1351–1378.

Cao, Y., Fisman, R., Lin, H., and Wang, Y. (2023). Soes and soft incentive constraints in state bank lending. *American Economic Journal: Economic Policy*, 15(1):174–95.

Carbo-Valverde, S., Rodriguez-Fernandez, F., and Udell, G. F. (2009). Bank market power and sme financing constraints. *Review of Finance*, 13(2):309–340.

Carvalho, D. (2014). The real effects of government-owned banks: Evidence from an emerging market. *The Journal of Finance*, 69(2):577–609.

Cassar, G., Ittner, C. D., and Cavalluzzo, K. S. (2015). Alternative information sources and information asymmetry reduction: Evidence from small business debt. *Journal of Accounting and Economics*, 59(2-3):242–263.

Cassiman, B. and Golovko, E. (2011). Innovation and internationalization through exports. *Journal of International Business Studies*, 42(1):56–75.

Castellani, D. and Zanfei, A. (2007). Internationalisation, innovation and productivity: How do firms differ in italy? *The World Economy*, 30(1):156–176.

Castelnovo, A., Malandri, L., Mercorio, F., Mezzanzanica, M., and Cosentini, A. (2021). Towards fairness through time. In , e. a., editor, *Machine Learning and Principles and Practice of Knowledge Discovery in Databases*, volume 1524 of *Communications in Computer and Information Science*, Cham. Springer.

Chaney, T. (2016). Liquidity constrained exporters. *Journal of Economic Dynamics and Control*, 72(C):141–154.

Chen, M., Wu, J., Jeon, B. N., and Wang, R. (2017). Monetary policy and bank risk-taking: Evidence from emerging economies. *Emerging Markets Review*, 31:116–140.

Chen, W. and Bharodia, N. (2019). What can we learn from what a machine has learned? interpreting credit risk machine learning models. *Journal of Risk Model Validation*, 15(2).

Cheng, B., Ioannou, I., and Serafeim, G. (2014). Corporate social responsibility and access to finance. *Strategic Management Journal*, 35(1):1–23.

Chodorow-Reich, G., Darmouni, O., Luck, S., and Plosser, M. (2021). Bank liquidity provision across the firm size distribution. *Journal of Financial Economics*.

Chodorow-Reich, G., Darmouni, O., Luck, S., and Plosser, M. (2022). Bank liquidity provision across the firm size distribution. *Journal of Financial Economics*, 144(3):908–932.

Chong, T. T.-L., Lu, L., and Ongena, S. (2013). Does banking competition alleviate or worsen credit constraints faced by small- and medium-sized enterprises? evidence from china. *Journal of Banking & Finance*, 37(9):3412–3424.

Ciampi, F., Giannozzi, A., Marzi, G., and Altman, E. I. (2021). Rethinking sme default prediction: a systematic literature review and future perspectives. *Scientometrics*, 126(3):2141–2188.

Claessens, S. and Tzioumis, K. (2006). Measuring firms' access to finance. In *Access to Finance: Building Inclusive Financial Systems*.

Coakley, J., Lazos, A., and Linares-Zegarra, J. M. (2018). Follow-on equity crowdfunding. *SSRN Electronic Journal*.

Cole, R. A. (1998). The importance of relationships to the availability of credit. *Journal of Banking & Finance*, 22(6-8):959–977.

Collier, P. and Cust, J. (2015). Investing in africa's infrastructure: Financing and policy options. *Annual Review of Resource Economics*, 7(1):473–493.

Correa Bahnsen, A., Aouada, D., Stojanovic, A., and Ottersten, B. (2016). Feature engineering strategies for credit card fraud detection. *Expert Systems with Applications*, 51:134–142.

Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3):273–297.

Cox, D. R. (1958). The regression analysis of binary sequences. *Journal of the Royal Statistical Society. Series B (Methodological)*, 20(2):215–242.

Crawford, G. S., Pavanini, N., and Schivardi, F. (2018). Asymmetric information and imperfect competition in lending markets. *The American Economic Review*, 108(7):1659–1701.

Cucculelli, M. and Peruzzi, V. (2017). Bank screening technologies and the founder effect: Evidence from european lending relationships. *Finance Research Letters*, 20:229–237.

Czarnitzki, D. and Hottenrott, H. (2011). R&d investment and financing constraints of small and medium-sized firms. *Small Business Economics*, 36(1):65–83.

Danenas, P. and Garsva, G. (2015). Selection of support vector machines based classifiers for credit risk domain. *Expert Systems with Applications*, 42(6):3194–3204.

Dastile, X. and Celik, T. (2021). Making deep learning-based predictions for credit scoring explainable. *IEEE Access*, 9:50426–50440.

Dastile, X., Celik, T., and Potsane, M. (2020). Statistical and machine learning models in credit scoring: A systematic literature survey. *Applied Soft Computing Journal*, 91.

Datta, A., Sen, S., and Zick, Y. (2016). Algorithmic transparency via quantitative input influence: Theory and experiments with learning systems. In *2016 IEEE SYMPOSIUM ON SECURITY AND PRIVACY (SP)*, IEEE Symposium on Security and Privacy, pages 598–617. IEEE; IEEE Comp Soc. IEEE Symposium on Security and Privacy (SP), San Jose, CA, MAY 23-25, 2016.

Davis, J. S. (2015). The cyclicality of (bilateral) capital inflows and outflows. *Federal Reserve Bank of Dallas, Globalization and Monetary Policy Institute Working Papers*, 2015(247).

de Andrés, P., Gimeno, R., and Mateos de Cabo, R. (2021). The gender gap in bank credit access. *Journal of Corporate Finance*, 71:101782.

de Lange, P., Melsom, B., Vennerød, C., and Westgaard, S. (2022). Explainable ai for credit assessment in banks. *Journal of Risk and Financial Management*, 15(12):556.

Degl'Innocenti, M., Fiordelisi, F., Girardone, C., and Radić, N. (2019). Competition and risk-taking in investment banking. *Financial Markets, Institutions & Instruments*, 28(2):241–260.

Degryse, H., Matthews, K., and Zhao, T. (2018). Smes and access to bank credit: Evidence on the regional propagation of the financial crisis in the uk. *Journal of Financial Stability*, 38:53–70.

Dell'Ariccia, G., Igan, D., and Laeven, L. (2012). Credit booms and lending standards: Evidence from the subprime mortgage market. *Journal of Money, Credit and Banking*, 44(2/3):367–384.

Deno, S., Loy, T., and Homburg, C. (2020). What happens if private accounting information becomes public? small firms' access to bank debt. *Entrepreneurship Theory and Practice*, 44(6):1091–1111.

Deyoung, R., Gron, A., Torna, G., and Winton, A. (2015). Risk overhang and loan portfolio decisions: Small business loan supply before and during the financial crisis. *The Journal of Finance*, 70(6):2451–2488.

Dobbie, W., Goldsmith□Pinkham, P., Mahoney, N., and Song, J. (2020). Bad credit, no problem? credit and labor market consequences of bad credit reports. *The Journal of Finance*, 75(5):2377–2419.

Duckworth, C., Chmiel, F. P., Burns, D. K., Zlatev, Z. D., White, N. M., Daniels, T. W., Kiuber, M., and Boniface, M. J. (2021). Using explainable machine learning to characterise data drift and detect emergent health risks for emergency department admissions during covid-19. *Scientific Reports*, 11(1):1–10.

Duprey, T. (2015). Do Publicly Owned Banks Lend Against the Wind? *International Journal of Central Banking*, 11(2):65–112.

Durguner, S. (2017). Do borrower-lender relationships still matter for small business loans? *Journal of International Financial Markets, Institutions and Money*, 50:98–118.

EC (2003). Commission recommendation of 6 may 2003 concerning the definition of micro, small and medium-sized enterprises.

Eca, A., Ferreira, M. A., Porras Prado, M., and Rizzo, A. E. (2021). The real effects of fintech lending on smes: Evidence from loan applications. *SSRN Electronic Journal*.

ECB (2022a). Euro area bank lending survey: Third quarter of 2022. *European Central Bank*.

ECB (2022b). Survey on the access to finance of enterprises in the euro area. *European Central Bank*.

Elsas, R. (2005). Empirical determinants of relationship lending. *Journal of Financial Intermediation*, 14(1):32–57.

Elyasiani, E. and Goldberg, L. G. (2004). Relationship lending: a survey of the literature. *Journal of Economics and Business*, 56(4):315–330.

Faia, E., Ottaviano, G. I., and Laffitte, S. (2018). Foreign expansion, competition and bank risk. *CEPR Discussion Paper*, DP13150.

Fang, X., Jutrsa, D., Peria, S. M., Presbitero, A. F., and Ratnovski, L. (2022). Bank capital requirements and lending in emerging markets: The role of bank characteristics and economic conditions. *Journal of Banking & Finance*, 135:105806.

Fatoki, O. and Asah, F. (2011). The impact of firm and entrepreneurial characteristics on access to debt finance by smes in king williams' town, south africa. *International Journal of Business and Management*, 6(8):170–179.

Fazzari, S. M., Hubbard, R. G., Petersen, B. C., Blinder, A. S., and Poterba, J. M. (1988). Financing constraints and corporate investment. *Brookings Papers on Economic Activity*, 1988(1):141.

Ferrando, A. and Ruggieri, A. (2018). Financial constraints and productivity: Evidence from euro area companies. *International Journal of Finance & Economics*, 23:257–282.

Ferri, G., Murro, P., Peruzzi, V., and Rotondi, Z. (2019). Bank lending technologies and credit availability in europe: What can we learn from the crisis? *Journal of International Money and Finance*, 95:128–148.

Florou, A. and Kosi, U. (2015). Does mandatory ifrs adoption facilitate debt financing? *Review of Accounting Studies*, 20(4):1407–1456.

Fop, M. and Murphy, T. B. (2018). Variable selection methods for model-based clustering. *Statistics Surveys*, 12:18–65.

Frank, M. Z., Goyal, V., and Shen, T. (2020). The pecking order theory of capital structure.

French, J. J., Yan, J., and Yasuda, Y. (2019). Relationships matter: the impact of bank-firm relationships on mergers and acquisitions in japan. *Journal of Financial Services Research*, 56(3):259–305.

Fridson, M. and Alvarez, F. (2022). *Financial Statement Analysis: A Practitioner's Guide*. John Wiley & Sons, Inc., fifth edition.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5):1189–1232.

Galli, E., Mascia, D. V., and Rossi, S. P. S. (2020). Bank credit constraints for women-led smes: Self-restraint or lender bias? *European Financial Management*, 26(4):1147–1188.

Gambini, A. and Zazzaro, A. (2013). Long-lasting bank relationships and growth of firms. *Small Business Economics*, 40(4):977–1007.

Gan, M., Pan, S., Chen, Y., Cheng, C., Pan, H., and Zhu, X. (2021). Application of the machine learning lightgbm model to the prediction of the water levels of the lower columbia river. *J. Mar. Sci. Eng.*, 9(5):496.

Gatti, R. and Love, I. (2008). Does access to credit improve productivity? evidence from bulgaria. *The Economics of Transition*, 16(3):445–465.

Gerschenkron, A. (1962). *Economic Backwardness in Historical Perspective*. Harvard University Press, Cambridge, United States.

Gertler, M. and Gilchrist, S. (1994). Monetary policy, business cycles, and the behavior of small manufacturing firms. *The Quarterly Journal of Economics*, 109(2):309–340.

Gholamy, A., Kreinovich, V., and Kosheleva, O. (2018). Why 70/30 or 80/20 relation between training and testing sets: A pedagogical explanation. Technical report, Technical Report: UTEP-CS-18-09.

Giovannini, A. and Moran, J. (2013). Finance for growth. Technical report, Report of the High Level Expert Group on SME and Infrastructure Financing.

Goldstein, A., Kapelner, A., Bleich, J., and Pitkin, E. (2015). Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation. *Journal of Computational and Graphical Statistics*, 24(1):44–65.

Gonzalez-Garcia, J. and Grigoli, F. (2013). State-owned banks and fiscal discipline. *IMF Working Paper*, 13/206.

González, J. L., Munro, L., Gourdon, J., Mazzini, E., and Andrenelli, A. (2019). Participation and benefits of smes in gvcs in southeast asia. *OECD Trade Policy Papers*, 231.

González-Recio, O., Jiménez-Montero, J., and Alenda, R. (2013). The gradient boosting algorithm and random boosting for genome-assisted evaluation in large data sets. *Journal of Dairy Science*, 96(1):614–624.

Gorodnichenko, Y. and Schnitzer, M. (2013). Financial constraints and innovation: Why poor countries don't catch up. *Journal of the European Economic Association*, 11(5):1115–1152.

Greenaway, D., Guariglia, A., and Kneller, R. (2007). Financial factors and exporting decisions. *Journal of International Economics*, 73(2):377–395.

Greenwald, D. L., Krainer, J., and Paul, P. (2021). The credit line channel. *Federal Reserve Bank of San Francisco Working Paper*, 2020(26).

Gregorutti, B., Michel, B., and Saint-Pierre, P. (2015). Grouped variable importance with random forests and application to multiple functional data analysis. *Computational Statistics & Data Analysis*, 90:15–35.

Gross, T., Notowidigdo, M. J., and Wang, J. L. (2020). The marginal propensity to consume over the business cycle. *American Economic Journal-Macroeconomics*, 12(2):351–384.

Grzelak, M. (2019). The hold-up problem and banking relationships: Evidence from the polish sme sector. *Prague Economic Papers*, 28(6):670–687.

Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., and Pedreschi, D. (2018). A survey of methods for explaining black box models. *ACM Comput. Surv.*, 51(5).

Gunning, D. and Aha, D. (2019). Darpa's explainable artificial intelligence (xai) program. *AI Magazine*, 40(2):44–58.

Gurara, D., Presbitero, A., and Sarmiento, M. (2020). Borrowing costs and the role of multilateral development banks: Evidence from cross-border syndicated bank lending. *Journal of International Money and Finance*, 100:102090.

Ha, V.-S. and Nguyen, H.-N. (2016). Credit scoring with a feature selection approach based deep learning. *MATEC Web of Conferences*, 54:05004.

Hadji Misheva, B., Hirsa, A., Osterrieder, J., Kulkarni, O., and Fung Lin, S. (2021). Explainable ai in credit risk management. *SSRN Electronic Journal*.

Hakimi, A. and Hamdi, H. (2013). The duration of relationship banking and the performance of tunisian firms: An empirical test. *Journal of Applied Business Research (JABR)*, 30(1):59.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer Series in Statistics. Springer, 2 edition.

Hauswald, R. and Marquez, R. (2006). Competition and Strategic Information Acquisition in Credit Markets. *The Review of Financial Studies*, 19(3):967–1000.

Himmelberg, C. P. and Petersen, B. (1994). R&d and internal finance: A panel study of small firms in high-tech industries. *The Review of Economics and Statistics*, 76(1):38–51.

Hirata, W. and Ojima, M. (2020). Competition and bank systemic risk: New evidence from Japan's regional banking. *Pacific-Basin Finance Journal*, 60(C).

Hsu, H.-H. and Hsieh, C.-W. (2010). Feature selection via correlation coefficient clustering. *Journal of Software*, 5(12):1371–1377.

Hussin Adam Khatir, A. A. and Bee, M. (2022). Machine learning models and data-balancing techniques for credit scoring: What is the best combination? *Risks*, 10(9).

Irwin, D. and Scott, J. M. (2010). Barriers faced by smes in raising bank finance. *International Journal of Entrepreneurial Behavior & Research*, 16(3):245–259.

Ivashina, V., Laeven, L., and Moral-Benito, E. (2022). Loan types and the bank lending channel. *Journal of Monetary Economics*, 126:171–187.

Jayakumar, M., Pradhan, R. P., Dash, S., Maradana, R. P., and Gaurav, K. (2018). Banking competition, banking stability, and economic growth: Are feedback effects at work? *Journal of Economics and Business*, 96:15–41.

Jiménez, G., Ongena, S., Peydró, J.-L., and Saurina, J. (2012). Credit supply and monetary policy: Identifying the bank balance-sheet channel with loan applications. *The American Economic Review*, 102(No. 5 (AUGUST 2012)):2301–2326.

Johnstone, D. (2016). The effect of information on uncertainty and the cost of capital. *Contemporary Accounting Research*, 33(2):752–774.

Kalogirou, F., Kiosse, P. V., and Pope, P. F. (2021). Pension deficits and corporate financial policy: Does accounting transparency matter? *European Accounting Review*, pages 1–25.

Kaplan, S. N. and Zingales, L. (1997). Do investment-cash flow sensitivities provide useful measures of financing constraints? *The Quarterly Journal of Economics*, 112(1):169–215.

Keller, W. (2004). International technology diffusion. *Journal of Economic Literature*, 42(3):752–782.

Kgoroeadira, R., Burke, A., and Van Stel, A. (2019). Small business online loan crowdfunding: who gets funded and what determines the rate of interest? *Small Business Economics*, 52(1):67–87.

Khan, H. H., Ahmad, R. B., and Gee, C. S. (2016). Bank competition and monetary policy transmission through the bank lending channel: Evidence from asean. *International Review of Economics & Finance*, 44:19–39.

Khan, M. A. (2022). Barriers constraining the growth of and potential solutions for emerging entrepreneurial smes. *Asia Pacific Journal of Innovation and Entrepreneurship*.

Kirasich, K., Smith, T., and Sadler, B. (2018). Random forest vs logistic regression: Binary classification for heterogeneous datasets. *SMU Data Science Review*, 1(3):Article 9.

Kirschenmann, K. (2016). Credit rationing in small firm-bank relationships. *Journal of Financial Intermediation*, 26:68–99.

Kletzer, K. and Bardhan, P. (1987). Credit markets and patterns of international trade. *Journal of Development Economics*, 27(1):57–70.

Kornai, J. (1979). Resource-constrained versus demand-constrained systems. *Econometrica*, 47(4):801–819.

Kornai, J. (1980). *Economics of Shortage*. North-Holland, Amsterdam.

Kornai, J. (1986). The soft budget constraint. *Kyklos*, 39(1):3–30.

Kruppa, J., Schwarz, A., Arminger, G., and Ziegler, A. (2013). Consumer credit risk: Individual probability estimates using machine learning. *Expert Systems with Applications*, 40(13):5125–5131.

Kärnä, A. and Stephan, A. (2022). Do firms in rural regions lack access to credit? local variation in small business loans and firm growth. *Regional Studies*, 56(11):1919–1933.

Lee, N., Sameen, H., and Cowling, M. (2015). Access to finance for innovative smes since the financial crisis. *Research Policy*, 44(2):370–380.

Lerner, J. (1999). The government as venture capitalist: The long□run impact of the sbir program. *The Journal of Business*, 72(3):285–318.

Li, L., Strahan, P. E., and Zhang, S. (2020). Banks as lenders of first resort: Evidence from the covid-19 crisis. *SSRN Electronic Journal*.

Lisowsky, P. and Minnis, M. (2020). The silent majority: Private u.s. firms and financial reporting choices. *Journal of Accounting Research*, 58(3):547–588.

Liu, Y., Chen, Y., and Fan, Z.-P. (2021). Do social network crowds help fundraising campaigns? effects of social influence on crowdfunding performance. *Journal of Business Research*, 122:97–108.

Love, I. and Pería, M. S. M. (2015). How bank competition affects firms' access to finance. *The World Bank Economic Review*, 29(3):413–448.

Lu, Y., Yang, L., Shi, B., and et al. (2022). A novel framework of credit risk feature selection for smes during industry 4.0. *Annals of Operations Research*.

Lundberg, S. and Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*.

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., and Lee, S.-I. (2020). From local explanations to global understanding with explainable ai for trees. *Nature Machine Intelligence*, 2(1):56–67.

Ma, Z., Stice, D., and Williams, C. (2019). The effect of bank monitoring on public bond terms. *Journal of Financial Economics*, 133(2):379–396.

Mac An Bhaird, C., Vidal, J. S., and Lucey, B. (2016). Discouraged borrowers: Evidence for eurozone smes. *Journal of International Financial Markets, Institutions and Money*, 44:46–55.

Mahmoudi, N. and Duman, E. (2015). Detecting credit card fraud by modified fisher discriminant analysis. *Expert Systems with Applications*, 42(5):2510–2516.

Maier, E. (2016). Supply and demand on crowdlending platforms: connecting small and medium-sized enterprise borrowers and consumer investors. *Journal of Retailing and Consumer Services*, 33:143–153.

Malakauskas, A. and Lakštutienė, A. (2021). Financial distress prediction for small and medium enterprises using machine learning techniques. *Engineering Economics*, 32(1):4–14.

Mancusi, M. L. and Vezzulli, A. (2014). R&d and credit rationing in smes. *Economic Inquiry*, 52(3):1153–1172.

Manole, V. and Spatareanu, M. (2010). Trade openness and income – a re-examination. *Economics Letters*, 106(1):1–3.

Manova, K. (2013). Credit constraints, heterogeneous firms, and international trade. *The Review of Economic Studies*, 80(2):711–744.

Manova, K., Wei, S.-J., and Zhang, Z. (2015). Firm exports and multinational activity under credit constraints. *The Review of Economics and Statistics*, 97(3):574–588.

Manzoor, F., Wei, L., and Siraj, M. (2021). Small and medium-sized enterprises and economic growth in pakistan: An ardl bounds cointegration approach. *Heliyon*, 7(2):e06340.

Marquez, R. (2015). Competition, Adverse Selection, and Information Dispersion in the Banking Industry. *The Review of Financial Studies*, 15(3):901–926.

Martí, J. and Quas, A. (2018). A beacon in the night: government certification of smes towards banks. *Small Business Economics*, 50(2):397–413.

Mazanai, M. and Fatoki, O. (2012). Access to finance in the sme sector: A south african perspective. *Asian Journal of Business and management*, 4:58–67.

Medianovskyi, K., Malakauskas, A., Lakstutiene, A., and Yahia, S. B. (2023). Interpretable machine learning for sme financial distress prediction. In Laouar, M. R., Balas, V. E., Lejdel, B., Eom, S., and Boudia, M. A., editors, *12th International Conference on Information Systems and Advanced Technologies "ICISAT 2022"*, pages 454–464, Cham. Springer International Publishing.

Meuleman, M. and De Maeseneire, W. (2012). Do r&d subsidies affect smes' access to external financing? *Research Policy*, 41(3):580–591.

Miao, J. and Wang, P. (2012). Sectoral bubbles and endogenous growth. *SSRN Electronic Journal*.

Mihalyi, D., Hwang, J., Rivetti, D., and Cust, J. (2022). Resource-backed loans in sub-saharan africa. *Policy Research Working Paper*, No. 9923.

Milani, C. (2014). Borrower–lender distance and loan default rates: Macro evidence from the italian local markets. *Journal of Economics and Business*, 71(C):1–21.

Mina, A., Lahr, H., and Hughes, A. (2013). The demand and supply of external finance for innovative firms. *Industrial and Corporate Change*, 22(4):869–901.

Minetti, R. and Zhu, S. C. (2011). Credit constraints and firm export: Microeconomic evidence from italy. *Journal of International Economics*, 83(2):109–125.

Minnis, M. and Sutherland, A. (2017). Financial statements as monitoring mechanisms: Evidence from small commercial loans. *Journal of Accounting Research*, 55(1):197–233.

Molendowski, E. and Petraškevičius, V. (2020). International competitive positions of the baltic states – changes and determinants in the post-accession period. *Journal of Business Economics and Management*, 21(3):706–724.

Molina, C. A. and Preve, L. A. (2012). An empirical analysis on the effect of financial distress on trade credit. *Financial Management*, 41(1 (SPRING 2012)):187–205.

Morrison, R. E., Bryant, C. M., Terejanu, G., Prudhomme, S., and Miki, K. (2013). Data partition methodology for validation of predictive models. *Computers & Mathematics with Applications*, 66(10):2114–2125. ICNC-FSKD 2012.

Moscato, V., Picariello, A., and Sperlí, G. (2021). A benchmark of machine learning approaches for credit score prediction. *Expert Syst. Appl.*, 165:113986.

Motta, V. and Sharma, A. (2020). Lending technologies and access to finance for smes in the hospitality industry. *International Journal of Hospitality Management*, 86:102371.

Mulkay, B., Hall, B. H., and Mairesse, J. (2001). Firm level investment and r&d in france and the united states: A comparison. In Bundesbank, D., editor, *Investing Today for the World of Tomorrow*, pages 229–273, Berlin, Heidelberg. Springer Berlin Heidelberg.

Muller, P., Ladher, R., Booth, J., Mohamed, S., Gorgels, S., Priem, M., Blagoeva, T., Martinelle, A., and Milanesi, G. (2022). Annual report on european smes 2021/2022: Smes and environmental sustainability. *EUROPEAN COMMISSION*, JUNE.

Musto, D. K. (2004). What happens when information leaves a market? evidence from post-bankruptcy consumers. *The Journal of Business*, 77(4):725–748.

Myers, S. C. and Majluf, N. S. (1984). Corporate financing and investment decisions when firms have information that investors do not have. *Journal of Financial Economics*, 13(2):187–221.

Neuberger, D. and Räthke-Döppner, S. (2015). The role of demographics in small business loan pricing. *Small Business Economics*, 44(2):411–424.

Nguyen, H. T., Nguyen, H. M., Troege, M., and Nguyen, A. T. H. (2021). Debt aversion, education, and credit self-rationing in smes. *Small Business Economics*.

Nisar, T. M., Prabhakar, G., and Torchia, M. (2020). Crowdfunding innovations in emerging economies: Risk and credit control in peer□to□peer lending network platforms. *Strategic Change*, 29(3):355–361.

OECD (2009). Competition and the financial crisis. *OECD Discussion Papers*, 1.

OECD (2022). *Financing SMEs and Entrepreneurs*. OECD Publishing.

Ogura, Y. (2018). The objective function of government-controlled banks in a financial crisis. *Journal of Banking & Finance*, 89:78–93.

Olivero, M. P., Li, Y., and Jeon, B. N. (2011). Competition in banking and the lending channel: Evidence from bank-level data in asia and latin america. *Journal of Banking & Finance*, 35(3):560–571.

Ongena, S. and Sendeniz-Yunc, . (2011). Which firms engage small, foreign, or state banks? and who goes islamic? evidence from turkey. *Journal of Banking & Finance*, 35(12):3213–3224.

Ongena, S. and Smith, D. C. (2000). What determines the number of bank relationships? cross-country evidence. *Journal of Financial Intermediation*, 9(1):26–56.

Ongena, S. R. and Smith, D. C. (1998). Bank relationships: A review. *SSRN Electronic Journal*.

Pal, R., Kupka, K., Aneja, A., and Militky, J. (2016). Business health characterization: A hybrid regression and support vector machine analysis. *Expert Systems with Applications*, 49:48–59.

Palazuelos, E., Forcadell, F. J., and Gutiérrez, C. (2018). Accounting information quality and trust as determinants of credit granting to smes: The role of external audit. *Small Business Economics*, 51(4):861–877.

Panzar, J. C. and Rosse, J. N. (1987). Testing for ”monopoly” equilibrium. *The Journal of Industrial Economics*, 35(4):443–456.

Pekarskienė, I. and Susnienė, R. (2011). Assessment of the level of globalization in the baltic states. *Inžinerinė ekonomika*, 22(1):58–68.

Peltoniemi, J. (2007). The benefits of relationship banking: Evidence from small business financing in finland. *Journal of Financial Services Research*, 31(2-3):153–171.

Peng, C.-Y. J., Lee, K. L., and Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting. *The Journal of Educational Research*, 96(1):3–14.

Petersen, M. A. and Rajan, R. G. (1995). The effect of credit market competition on lending relationships. *The Quarterly Journal of Economics*, 110(2):407–443.

Petersen, M. A. and Rajan, R. G. (2002). Does distance still matter? the information revolution in small business lending. *The Journal of Finance*, 57(6):2533–2570.

Peón, D. and Guntín, X. (2021). Bank credit and trade credit after the financial crisis: Evidence from rural galicia. *Journal of Business Economics and Management*, 22(3):616–635.

Pietrovito, F. and Pozzolo, A. F. (2021). Credit constraints and exports of smes in emerging and developing countries. *Small Business Economics*, 56(1):311–332.

Preece, A. D., Harborne, D., Braines, D., Tomsett, R. J., and Chakraborty, S. (2018). Stakeholders in explainable ai. *ArXiv*, abs/1810.00184.

Presbitero, A. F. and Zazzaro, A. (2011). Competition and relationship lending: Friends or foes? *Journal of Financial Intermediation*, 20(3):387–413.

Qi, J., Yang, R., and Wang, P. (2021). Application of explainable machine learning based on catboost in credit scoring. In *Journal of Physics: Conference Series*, volume 1955.

Rabetti, D. (2022). Non-information asymmetry benefits of relationship lending. *SSRN Electronic Journal*.

Rajan, R. G. and Zingales, L. (1998). Financial dependence and growth. *The American Economic Review*, 88(3):559–586.

Ralcheva, A. and Roosenboom, P. (2020). Forecasting success in equity crowdfunding. *Small Business Economics*, 55(1):39–56.

Ribeiro, M. T., Singh, S., and Guestrin, C. (2016). ”why should i trust you?”: Explaining the predictions of any classifier. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.

Rupeika-Apoga, R. (2014). Access to finance: Baltic financial markets. *Procedia Economics and Finance*, 9:181–192. The Economies of Balkan and Eastern Europe Countries in the Changed World (EBEEC 2013).

Saarela, M. and Jauhiainen, S. (2021). Comparison of feature importance measures as explanations for classification models. *SN Applied Sciences*, 3(2).

Sapienza, P. (2004). The effects of government ownership on bank lending. *Journal of Financial Economics*, 72(2):357–384.

Savignac, F. (2007). The impact of financial constraints on innovation: What can be learned from a direct measure? *SSRN Electronic Journal*.

Schober, P., Boer, C., and Schwarte, L. A. (2018). Correlation coefficients: Appropriate use and interpretation. *Anesthesia & Analgesia*, 126(5):1763–1768.

Seiffer, C. and Gerling, A. (2021). Detection of concept drift in manufacturing data with shap values to improve error prediction. In *The Tenth International Conference on Data Analytics*, pages 51–60.

Senderovitz, M. (2009). How are smes defined in current research? In *Proceedings of AGSE*, Australia. AGSE.

Shi, S., Tse, R., Luo, W., D’Addona, S., and Pau, G. (2022). Machine learning-driven credit risk: a systemic review. *Neural Computing and Applications*, 34(17):14327–14339.

Shrikumar, A., Greenside, P., and Kundaje, A. (2017). Learning important features through propagating activation differences. In *34th International Conference on Machine Learning, ICML 2017*, volume 7, pages 4844–4866.

Shumway, T. (2001). Forecasting bankruptcy more accurately: A simple hazard model. *The Journal of Business*, 74(1):101–124.

Sikochi, A. (2020). Corporate legal structure and bank loan spread. *Journal of Corporate Finance*, 64:101656.

Silva, L., Silva, N. F., and Rosa, T. (2020). Success prediction of crowdfunding campaigns: a two-phase modeling. *International Journal of Web Information Systems*, 16(4):387–412.

Sundararajan, M., Taly, A., and Yan, Q. (2017). Axiomatic attribution for deep networks. *34th International Conference on Machine Learning, ICML 2017*, 7:5109–5118.

Trivedi, S. K. (2020). A study on credit scoring modeling with different feature selection and machine learning approaches. *Technology in Society*, 63:101413.

Uchida, H., Udell, G. F., and Watanabe, W. (2008). Bank size and lending relationships in japan. *Journal of the Japanese and International Economies*, 22(2):242–267.

USITC (2010). *Small and Medium Sized Enterprises: Characteristics and Performance*. U.S. International Trade Commission.

Uzzi, B. (1999). Embeddedness in the making of financial capital: How social relations and networks benefit firms seeking financing. *American Sociological Review*, 64(4):481.

Van Beveren, I. and Vandenbussche, H. (2010). Product and process innovation and firms' decision to export. *Journal of Economic Policy Reform*, 13(1):3–24.

Walsh, C. (2010). *Key Management Ratios*. Financial Times Series. Pearson Education Limited.

Wang, G., Hao, J., Ma, J., and Jiang, H. (2011). A comparative assessment of ensemble learning for credit scoring. *Expert Systems with Applications*, 38(1):223–230.

Wang, W., Lesner, C., Ran, A., Rukonic, M., Xue, J., and Shiu, E. (2020). Using small business banking data for explainable credit risk scoring. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(08):13396–13401.

Wang, Y. (2018). Fickle capital flows and retrenchment: Evidence from bilateral banking data. *Journal of International Money and Finance*, 87:1–21.

Wasiuzzaman, S., Lee, C. L., Boon, O. H., and Chelvam, H. P. (2021). Examination of the motivations for equity-based crowdfunding in an emerging market. *Journal of Theoretical and Applied Electronic Commerce Research*, 16(2):63–79.

Xu, Y., Saunders, A., Xiao, B., and Li, X. (2020). Bank relationship loss: The moderating effect of information opacity. *Journal of Banking & Finance*, 118:105872.

Zainol Abidin, J., Abdullah, N. A. H., and Khaw, K. L.-H. (2021). Bankruptcy prediction: Smes in the hospitality industry. *Vol. 16, Number 2, 2021*, 16(Number 2):51–80.

Zarutskie, R. (2006). Evidence on the effects of bank competition on firm borrowing and investment. *Journal of Financial Economics*, 81(3):503–537.

Zarutskie, R. (2013). Competition, financial innovation and commercial bank loan portfolios. *Journal of Financial Intermediation*, 22(3):373–396.

Zhao, Z., Xu, S., Kang, B. H., Kabir, M. M. J., Liu, Y., and Wasinger, R. (2015). Investigation and improvement of multi-layer perceptron neural networks for credit scoring. *Expert Systems with Applications*, 42(7):3508–3516.

Ćehajić, A. and Košak, M. (2022). Bank lending and small and medium-sized enterprises' access to finance – effects of macroprudential policies. *Journal of International Money and Finance*, 124:102612.

Činčikaitė, R. and Meidutė-Kavaliauskienė, I. (2023). Assessment of attractiveness of the baltic states for foreign direct investment: The topsis approach. *Journal of Risk and Financial Management*, 16(2):63.

CURRICULUM VITAE

Education:

2015 - 2017 Financial Economics, Master's degree, ISM University of Management and Economics.

2011 - 2015 Economics, Bachelor's degree, ISM University of Management and Economics.

Work experience:

2023 - now, Head of Transformation and Innovation, Swedbank Baltics, AS.
2019 - 2023, Head of Financing Transformation, Swedbank Baltics, AS.
2016 - 2019, Area Manager, Swedbank, AB.
2015 - 2016, Analyst, Swedbank, AB.

Contact email:

aidasmal@gmail.com

LIST OF SCIENTIFIC PUBLICATIONS

ARTICLES IN REVIEWED SCIENTIFIC PUBLICATIONS

Articles in conference proceedings in the Scopus database

1. **Malakauskas, A.**, Lakštutienė, A, The application of artificial intelligence tools in creditworthiness modelling for SME entities // 2021 IEEE International Conference on Technology and Entrepreneurship (ICTE), 24-27 August, 2021, Kaunas, Lithuania. Piscataway, NJ: IEEE, 2021.

2. Medianovskyi, K., **Malakauskas, A.**, Lakštutiene, A., Ben Yahia, S., Interpretable machine learning for heterogeneous treatment effect estimators with Double ML: a case of access to credit for SMEs, 27th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems, 2023, Athens, Greece. (Accepted)

Articles published in other peer reviewed journals

1. **Malakauskas, A.**, Lakštutiene, A., Malakauskienė, K., Credit accessibility evaluation model for small and medium enterprises // The 10th international scientific conference ”Finance, economics and tourism” (FET 2022), 22-24 September, 2022, Pula, Croatia. Pula : Juraj Dobrila University of Pula. 2022, p. 347-366.

Articles published in publications of other international databases

1. Medianovskyi, K., **Malakauskas, A.**, Lakštutienė, A., Ben Yahia, S., Interpretable machine learning for SME financial distress prediction // 12th international conference on information systems and advanced technologies “ICISAT 2022”: intelligent information, data science and decision support system. Springer, 2023. p. 454-464.

PRESENTATION OF SCIENTIFIC RESEARCH RESULTS AT SCIENTIFIC CONFERENCES

1. **Malakauskas, A.**, Lakštutienė, A., Modelling credit rating outlook for SME entities in the Baltic states. In Accounting and finance: innovative solutions for sustainable bioeconomy and rural development: 12th international scientific conference programme and abstracts, November 19-20, 2020 (pp. 28-28). Kaunas: Vytauto Didžiojo universitetas.

2. **Malakauskas, A.**, Lakštutienė, A., The application of artificial intelligence tools in creditworthiness modelling for SME entities. In 2021 IEEE International Conference on Technology and Entrepreneurship (ICTE) “Leading digital transformation in business and society”: book of abstracts (pp. 70-71). Kaunas: Technologija.

3. Medianovskyi, K., **Malakauskas, A.**, Lakštutiene, A., Ben Yahia, S., Interpretable Machine Learning for SME Financial Distress Prediction, 12th International Conference on Information Systems and Advanced Technologies “ICISAT 2022”, 26-27 August 2022.
4. **Malakauskas, A.**, Lakštutienė, A., Malakauskienė, K., Credit accessibility evaluation model for small and medium enterprises, 10th International Scientific Conference ‘Finance, Economics and Tourism - FET 2022‘, 22-24 September, 2022, Pula, Croatia.
5. Medianovskyi, K., **Malakauskas, A.**, Lakštutiene, A., Ben Yahia, S., Interpretable machine learning for heterogeneous treatment effect estimators with Double ML: a case of access to credit for SMEs, 27th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems, 6-8 September, 2023, Athens, Greece.

ACKNOWLEDGMENTS

'No one can whistle a symphony. It takes an orchestra to play it.'

Halford E. Luccock

I would like to express my sincere gratitude and appreciation to the many individuals who have supported and contributed to the completion of this endeavor. First and foremost, I am deeply grateful to my supervisor, Assoc. Prof. Aušrinė Lakštutienė, for her guidance, mentorship, and unwavering support throughout this research journey. Her expertise and encouragement have been invaluable.

I am also thankful to my colleagues in the sustainable economics research group, the reviewers, and the defense committee, who generously provided knowledge and expertise, thus contributing to the refinement of this work. Their insights and constructive feedback have been instrumental in its improvement.

I would like to offer special thanks to Kyrylo Medianovskyi for his cooperation and valuable discussions. His ideas have greatly enriched my research, and I am grateful for his support.

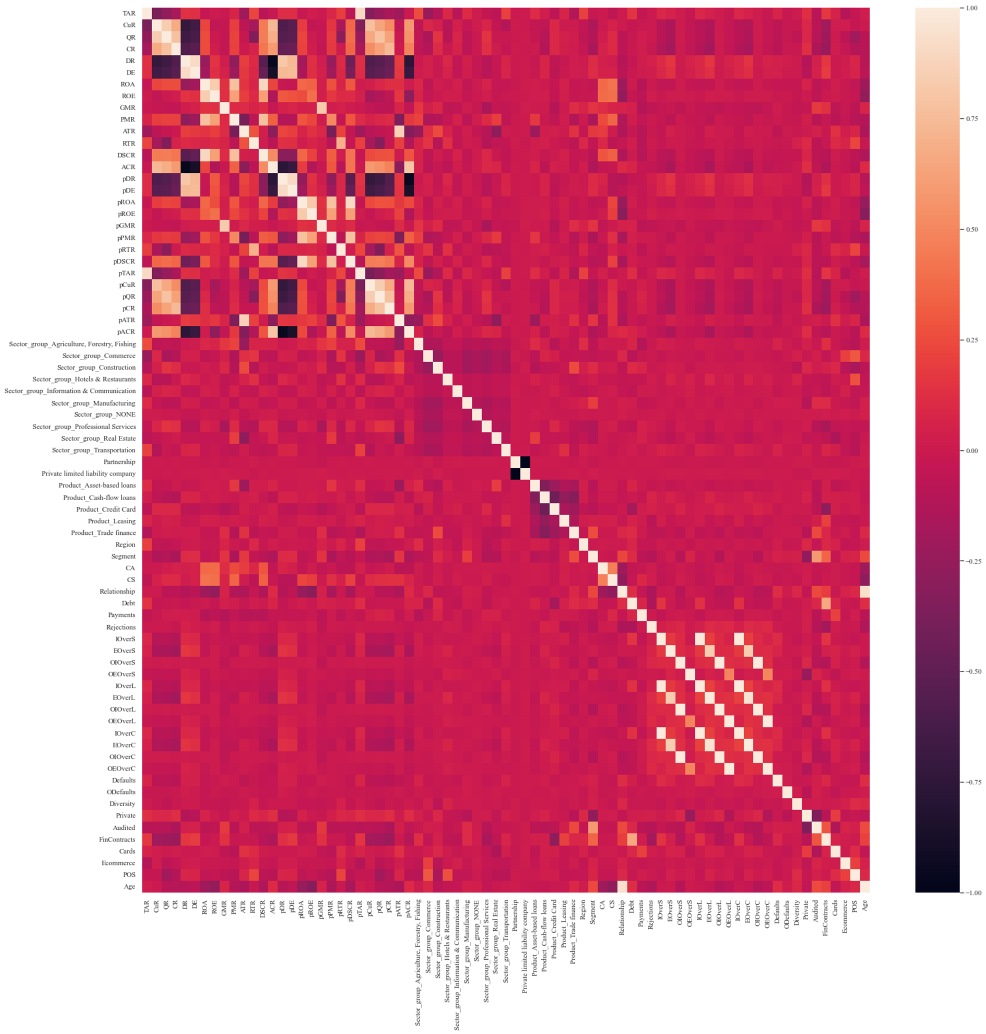
Furthermore, I could not have undertaken this journey without the unwavering support of my wonderful wife, Karolina, who, together with our children, Elena and Bernardas, has been by my side throughout this entire journey. Their love, understanding, and encouragement have been my anchor during challenging times.

Lastly, I would be remiss not to mention my entire family, especially my parents, whose belief in me has kept my spirits and motivation high during this difficult process. Their unwavering support and faith in my abilities have been a source of inspiration.

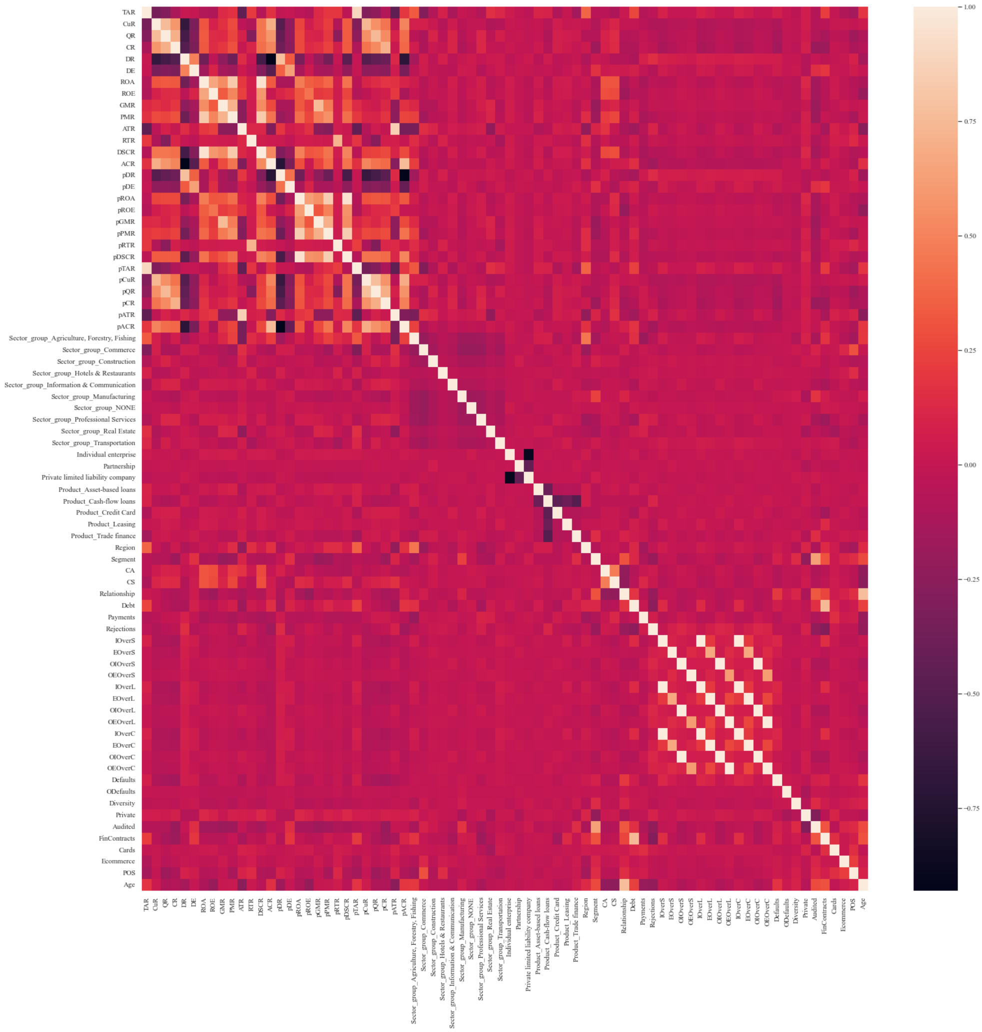
To all those mentioned above and anyone else who has contributed to this endeavor, I extend my heartfelt thanks. Your support, guidance, and belief in me have been indispensable, and I am deeply grateful for your presence in my life

ANNEXES

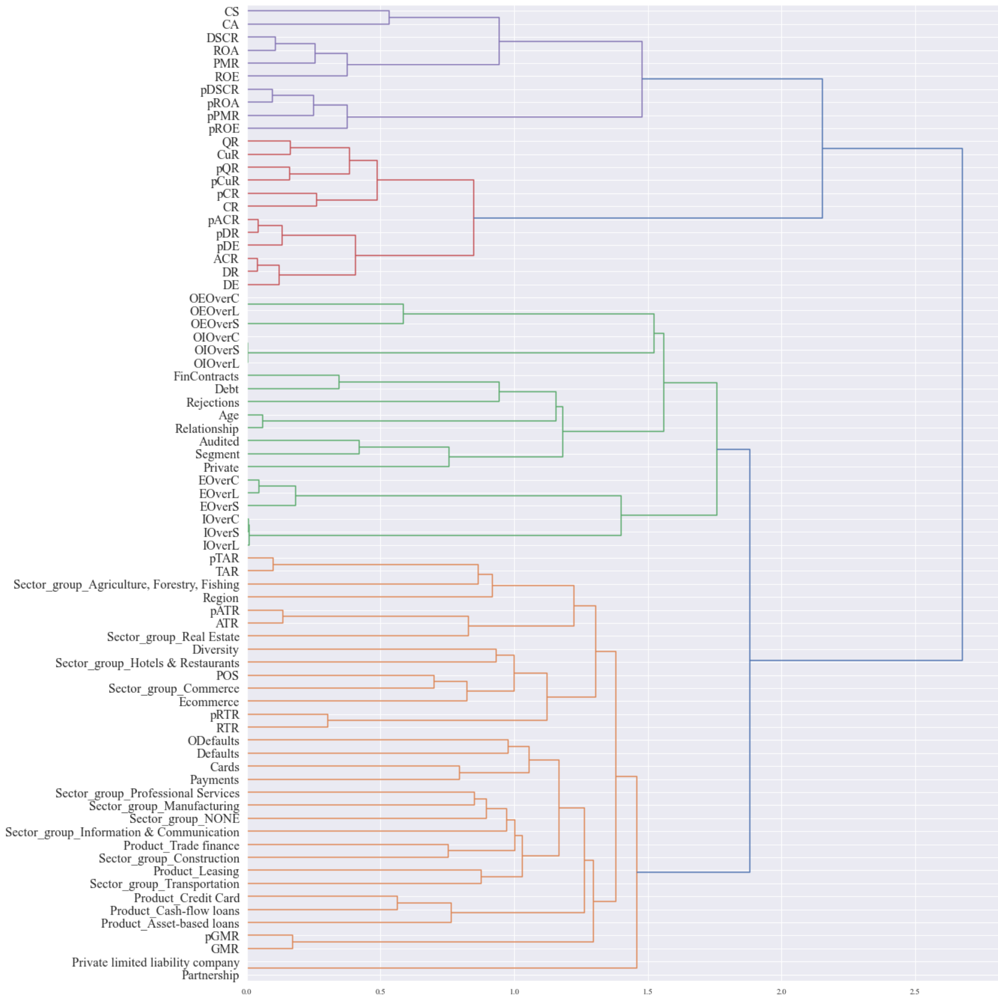
Annex 1. Feature correlation heatmap for Estonian dataset.



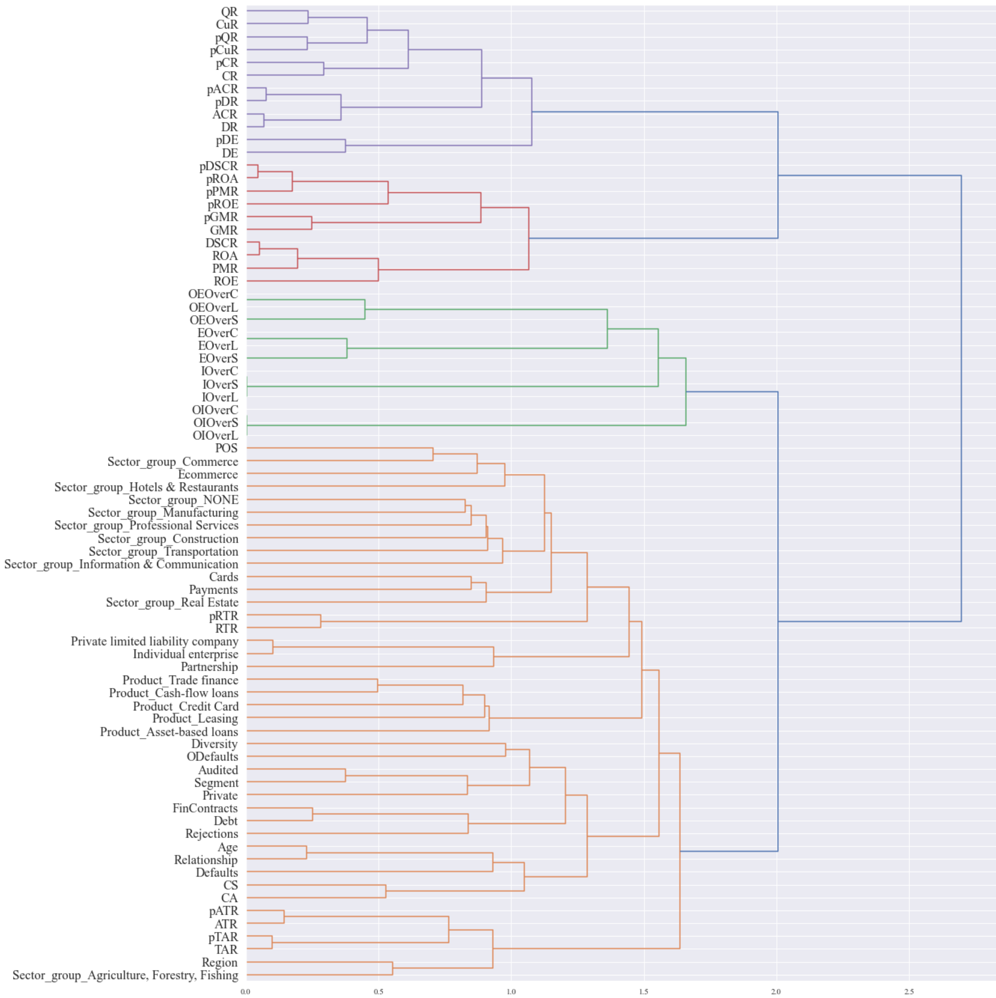
Annex 2. Feature correlation heatmap for Latvian dataset.



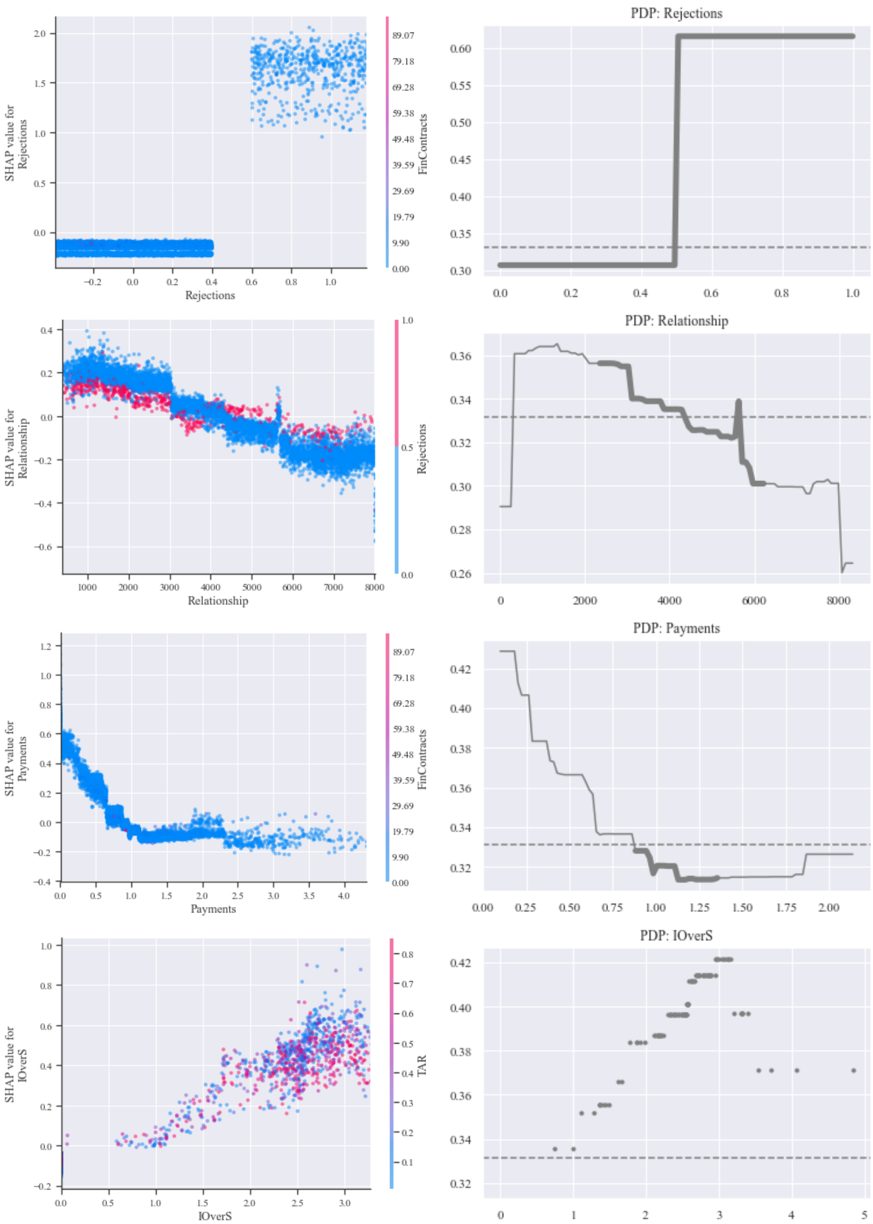
Annex 3. Hierarchical clustering dendrogram for Estonian dataset.



Annex 4. Hierarchical clustering dendrogram for Latvian dataset

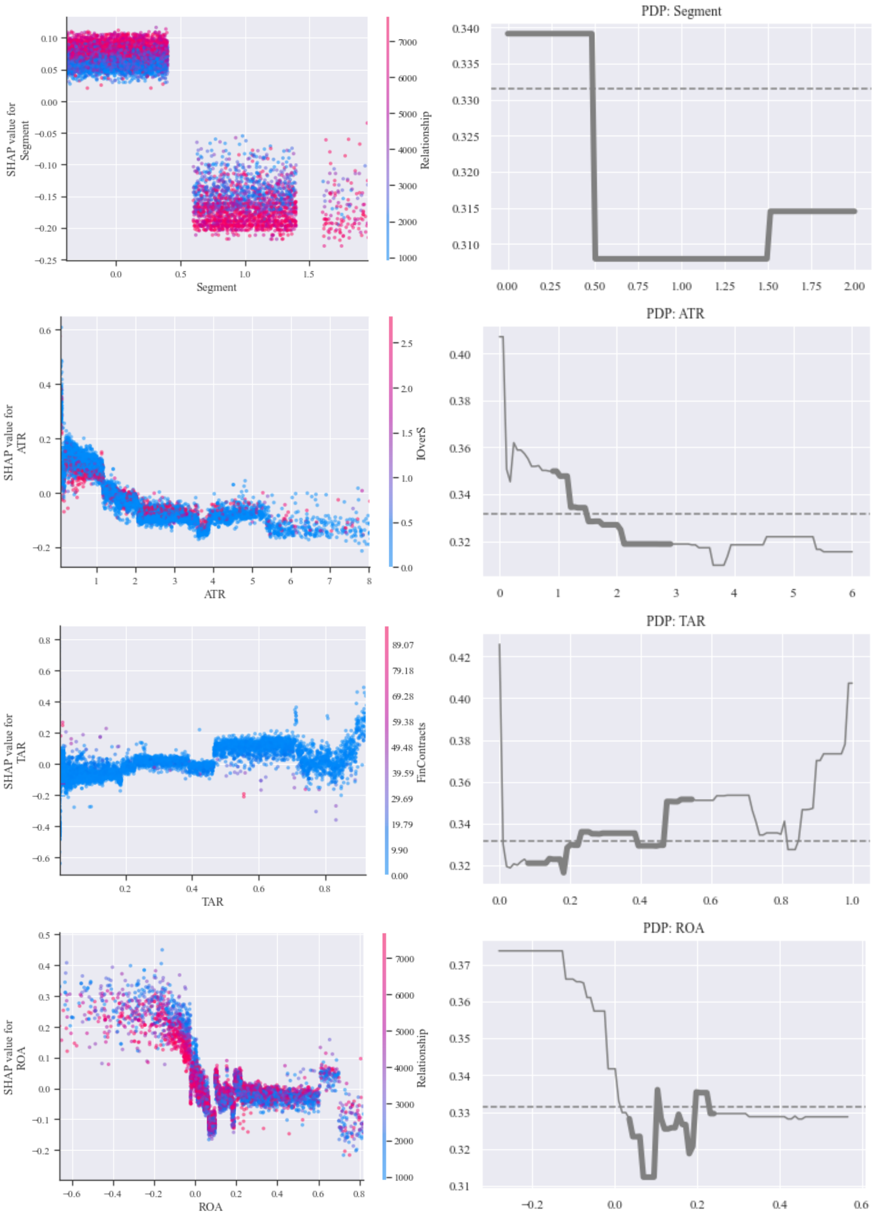


Annex 5. SHAP dependence and PDP plots for Estonian dataset.



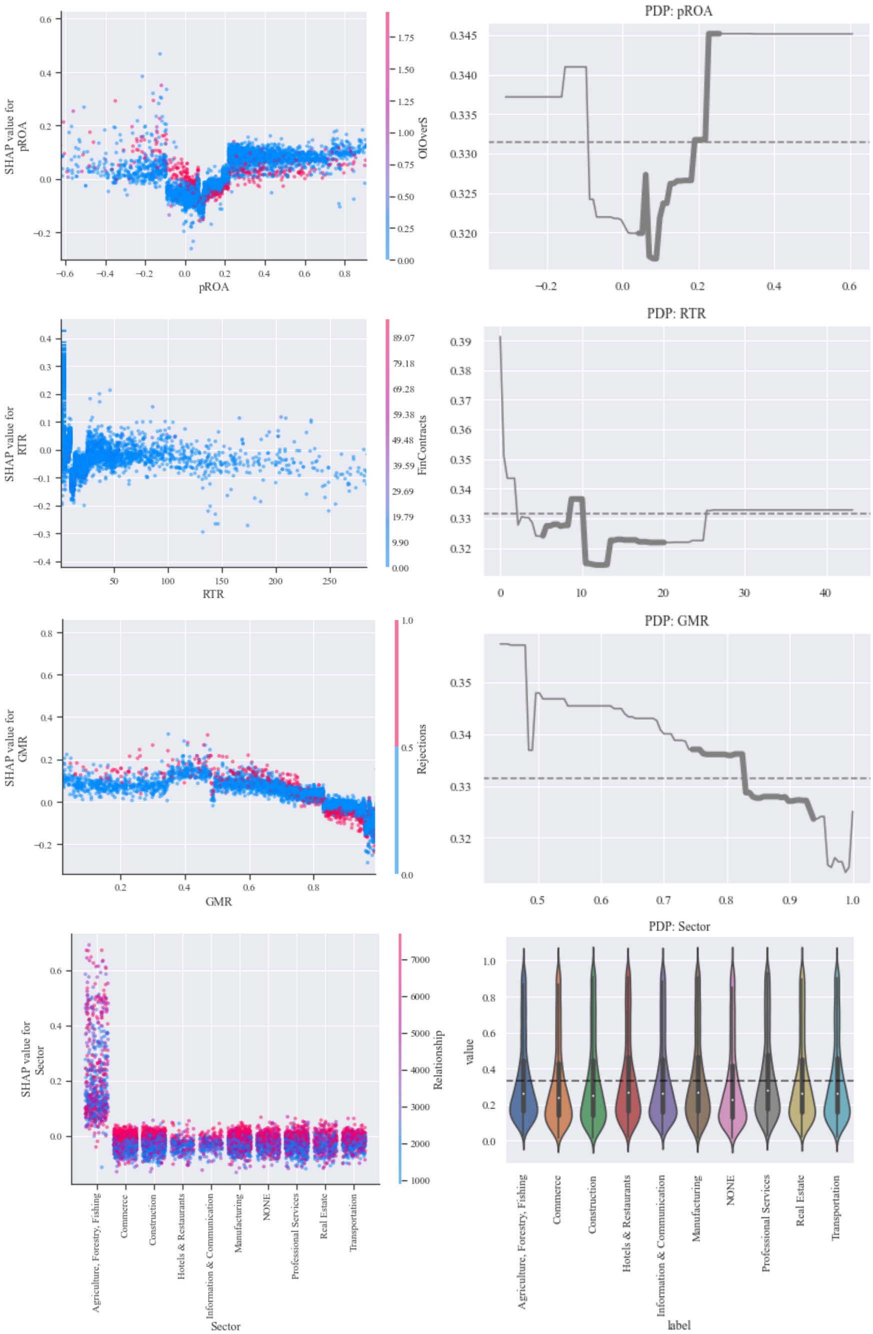
Continued on the next page.

Annex 5 continued

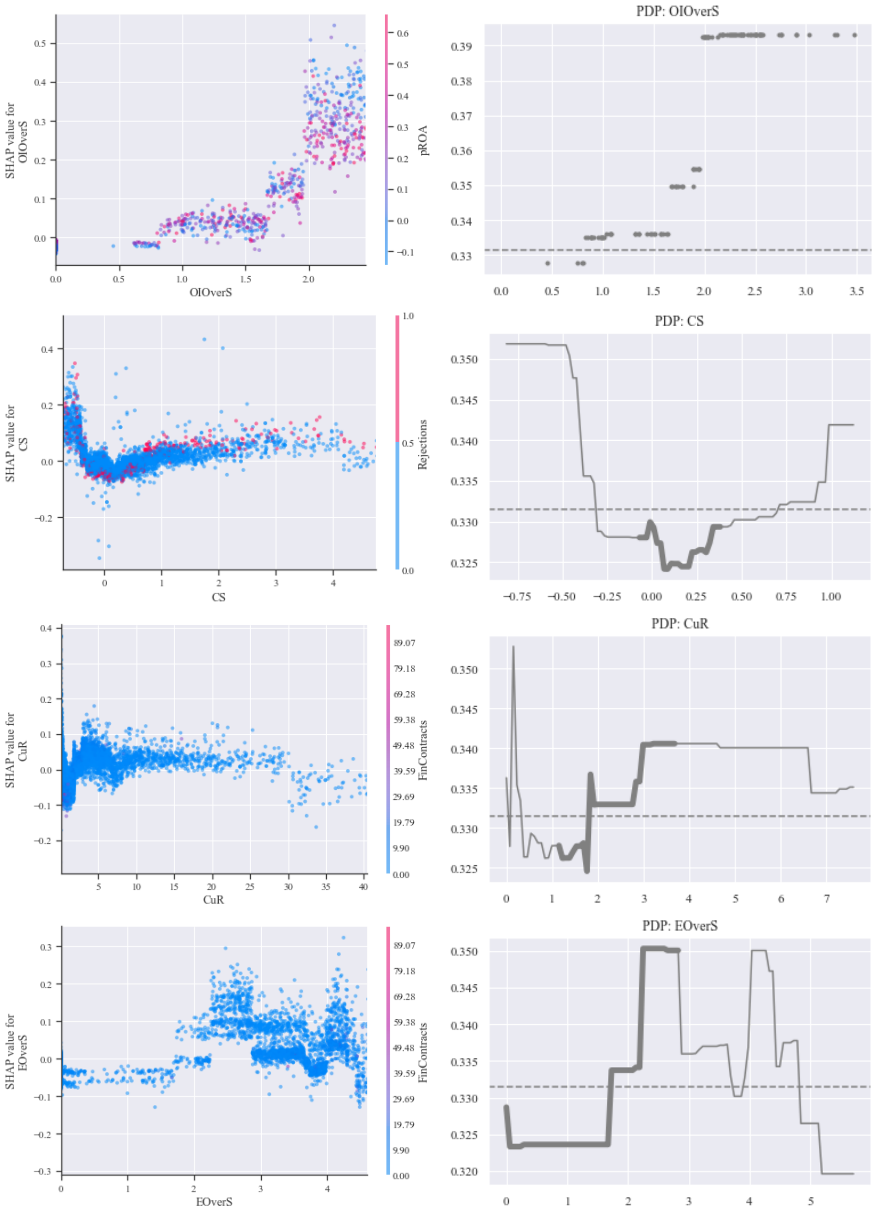


Continued on the next page.

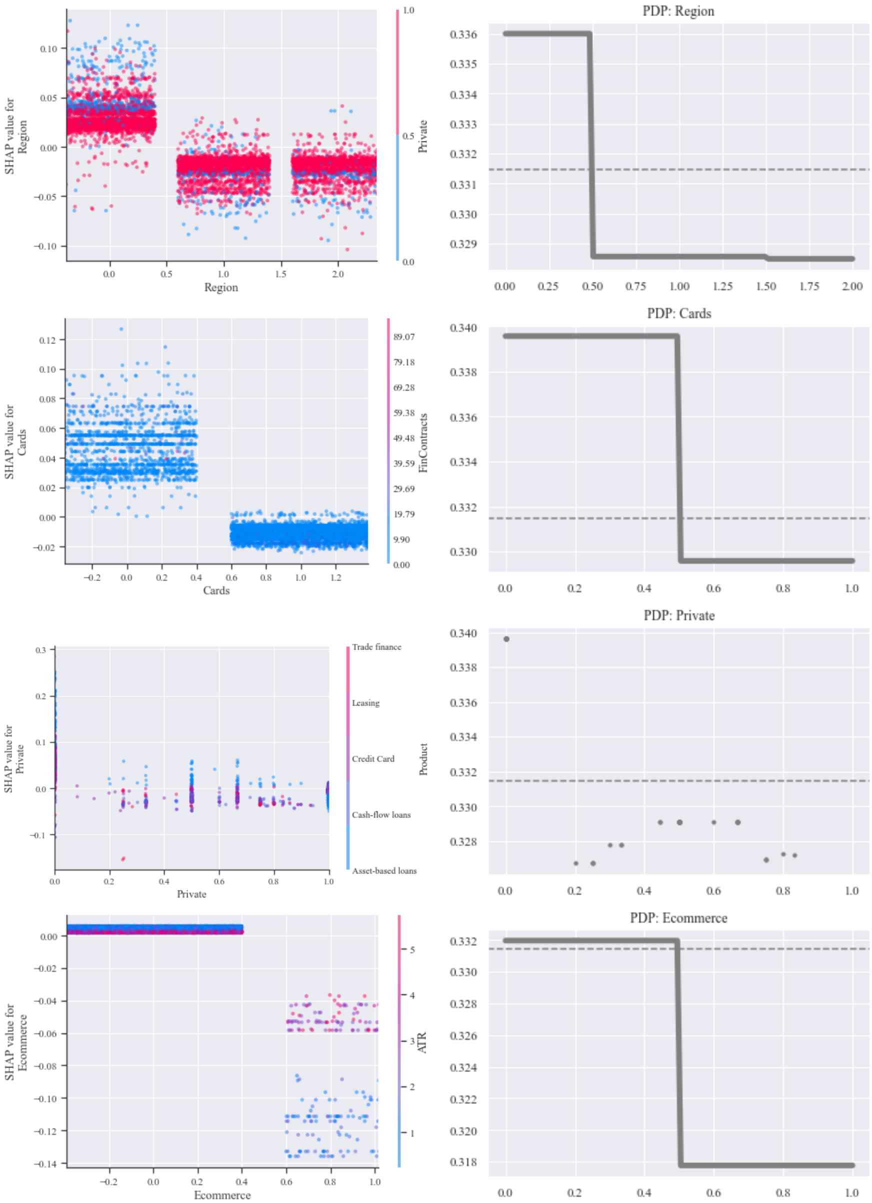
Annex 5 continued



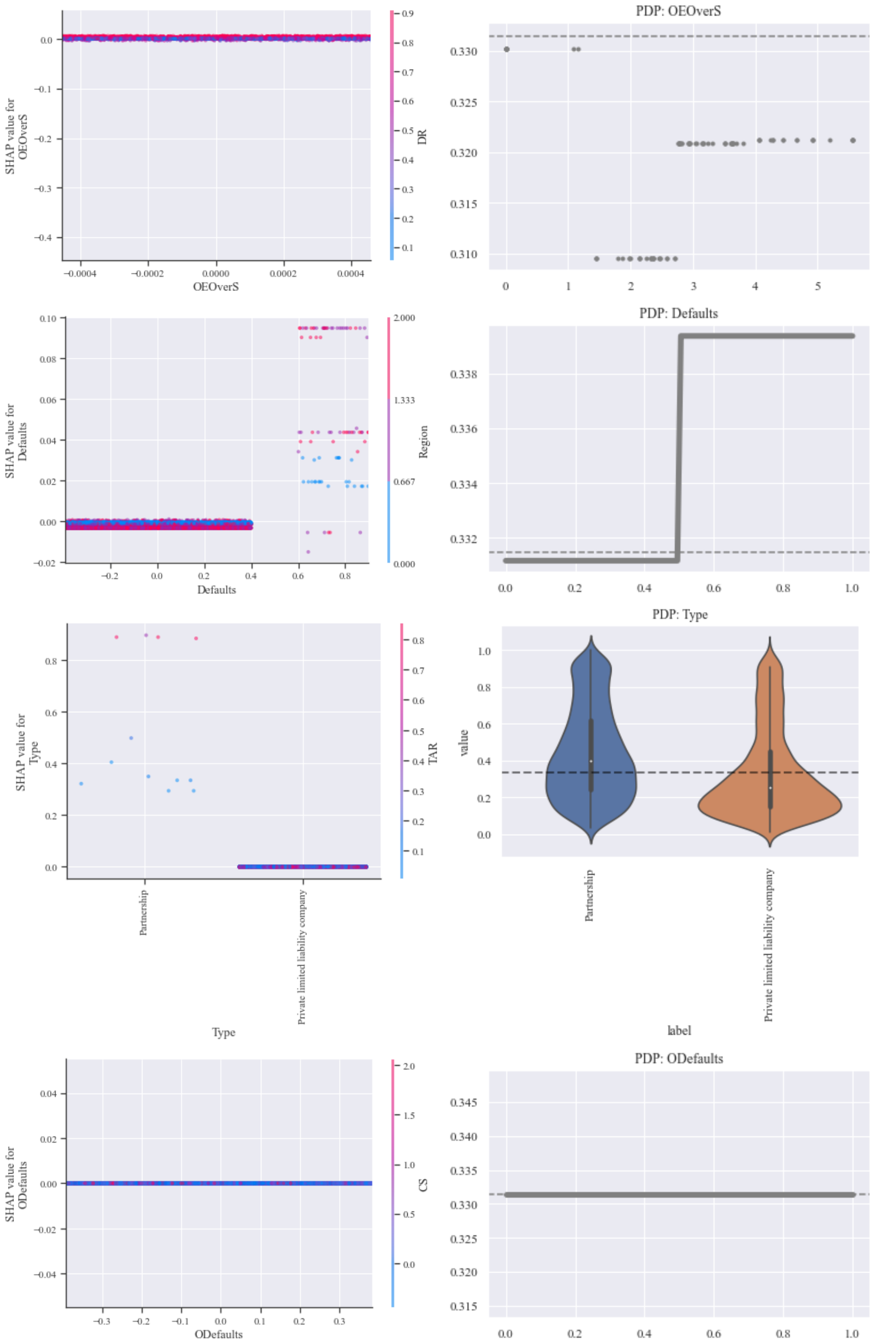
Continued on the next page.



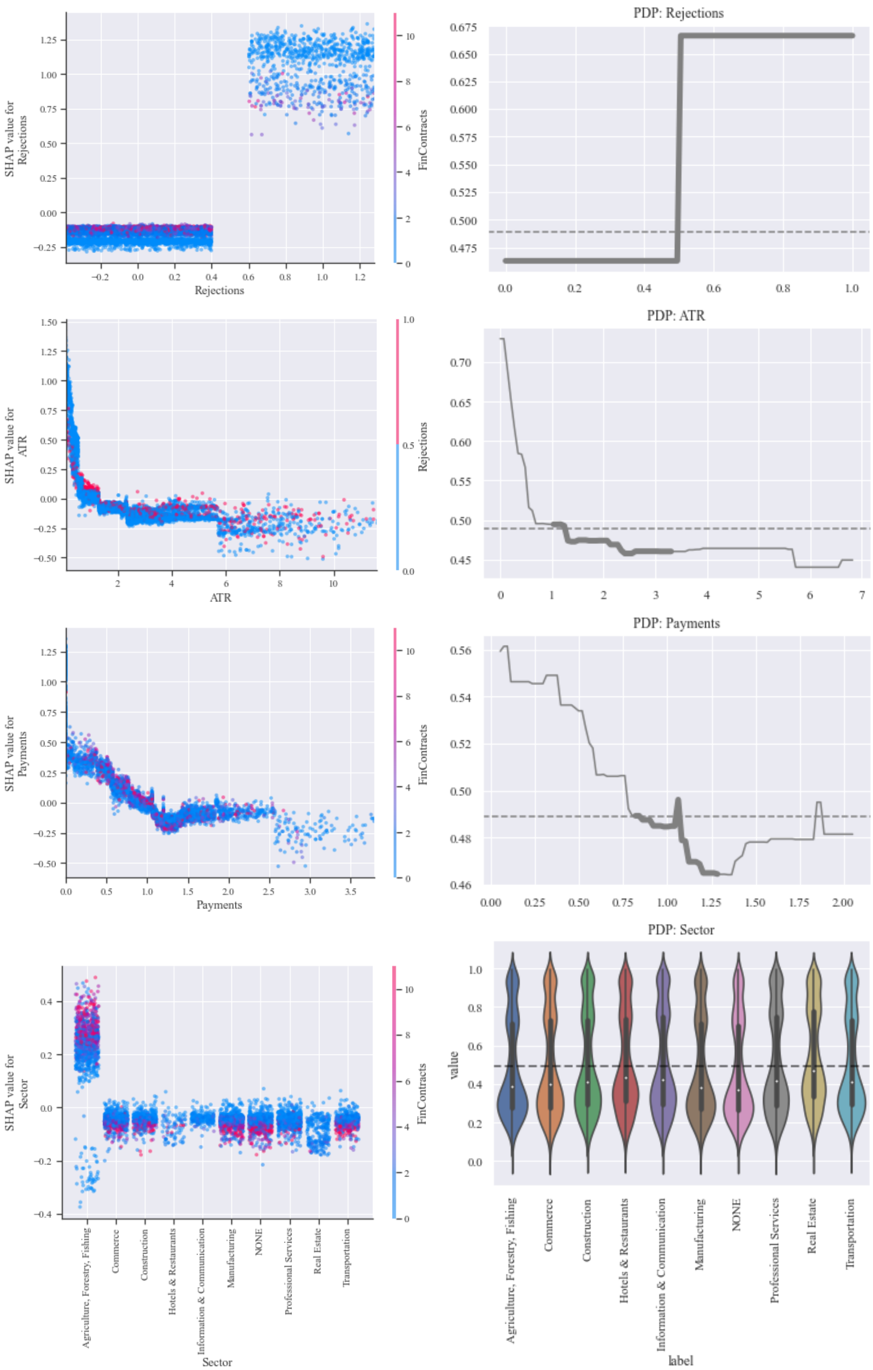
Continued on the next page.



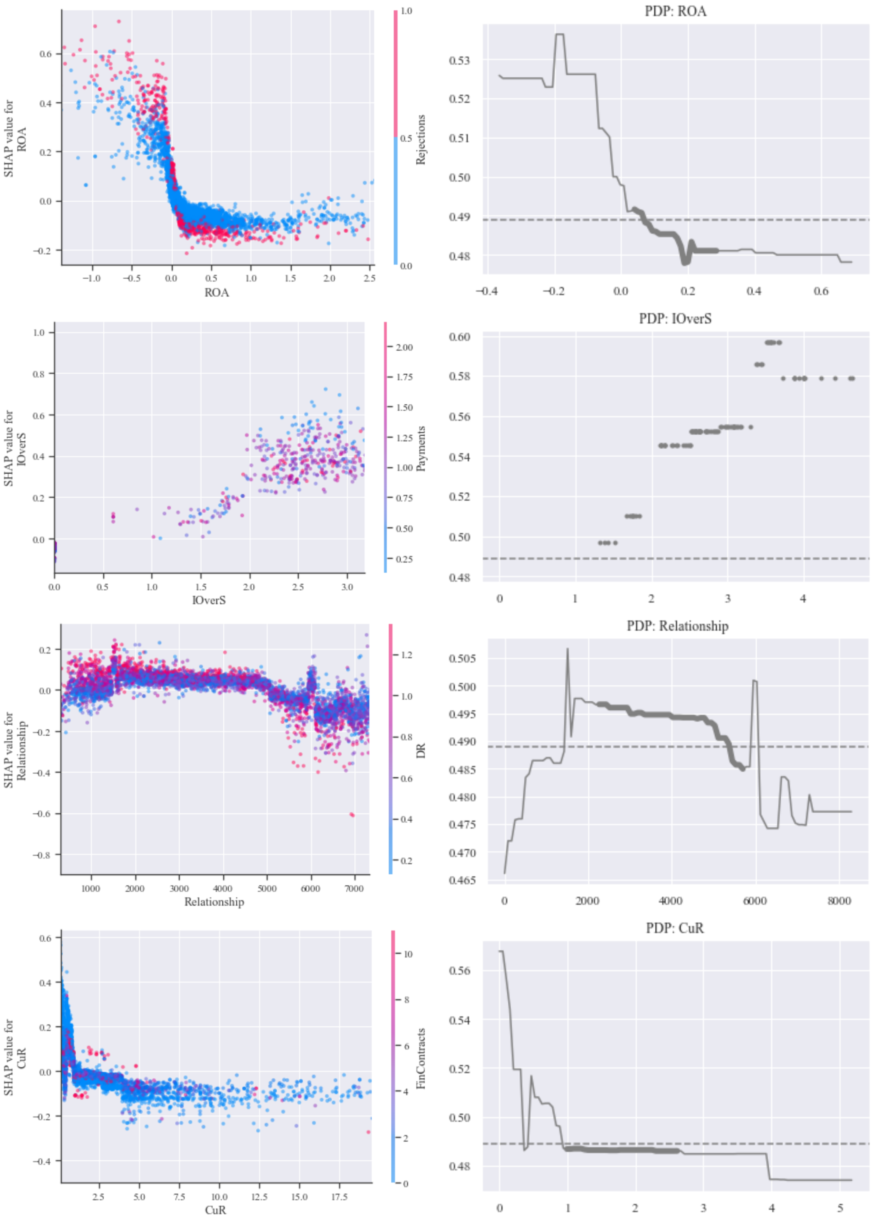
Continued on the next page.


PDP: OEOverS
SHAP value for OEOverS
OEOverS
DR
PDP: Defaults
SHAP value for Defaults
Defaults
Region
PDP: Type
SHAP value for Type
Type
TAR
value
label
Partnership
Private limited liability company
PDP: ODefaults
SHAP value for ODefaults
ODefaults
CS

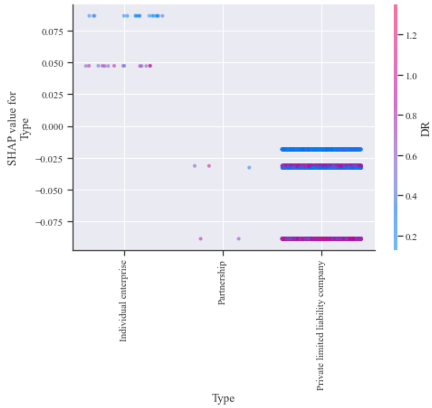
Annex 6. SHAP dependence and PDP plots for Latvian dataset.

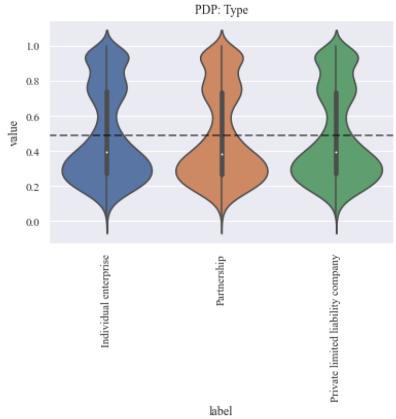


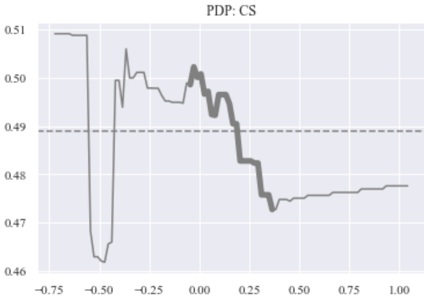
Continued on the next page.

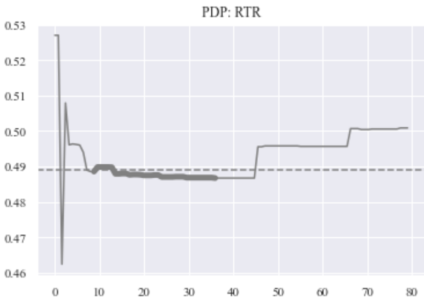


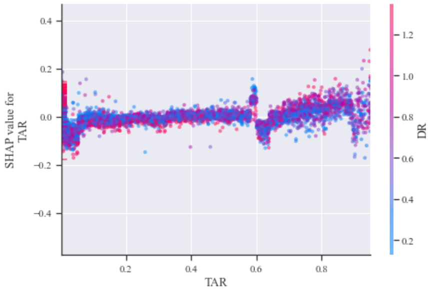
Continued on the next page.

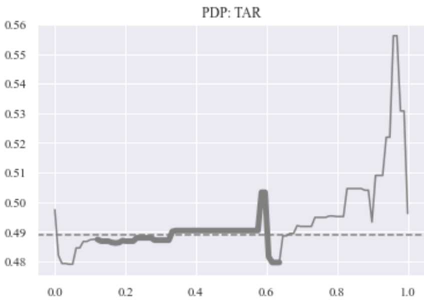




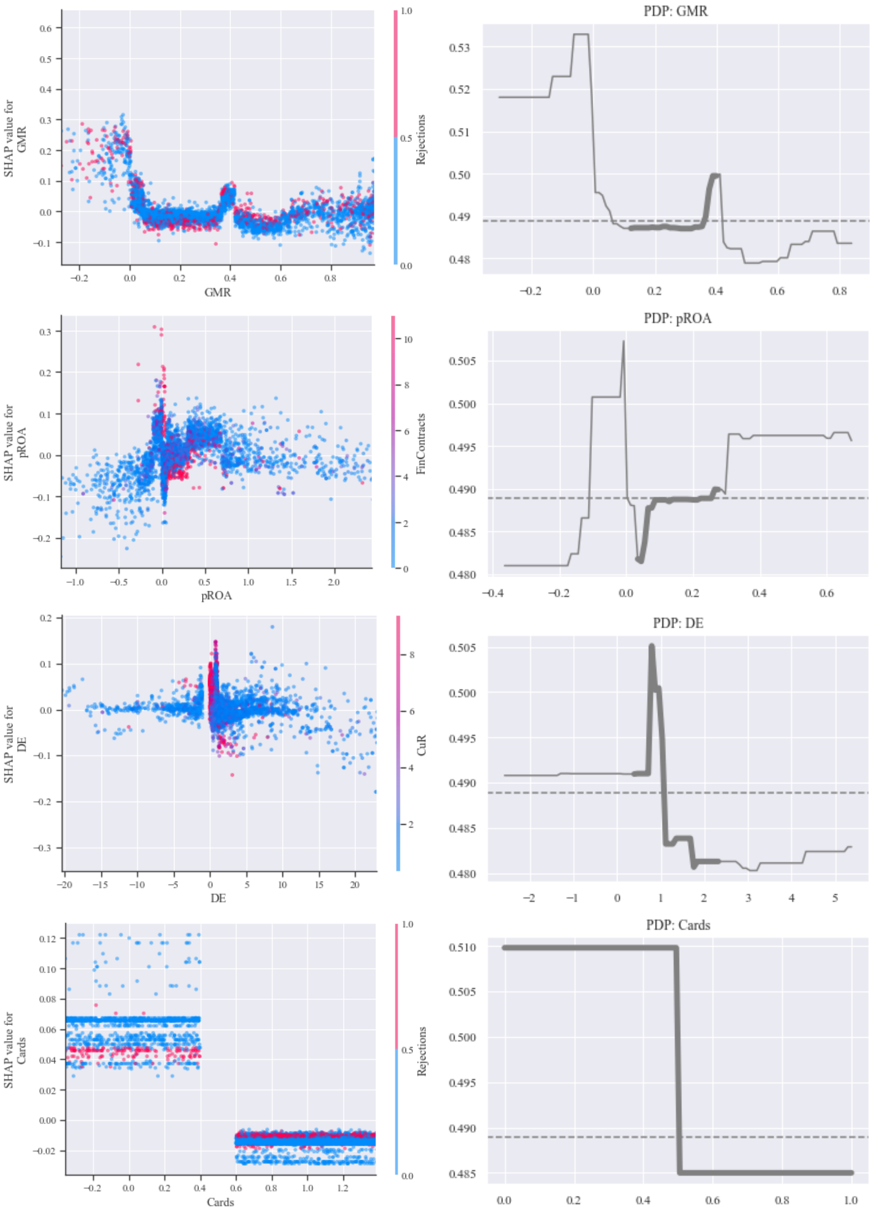




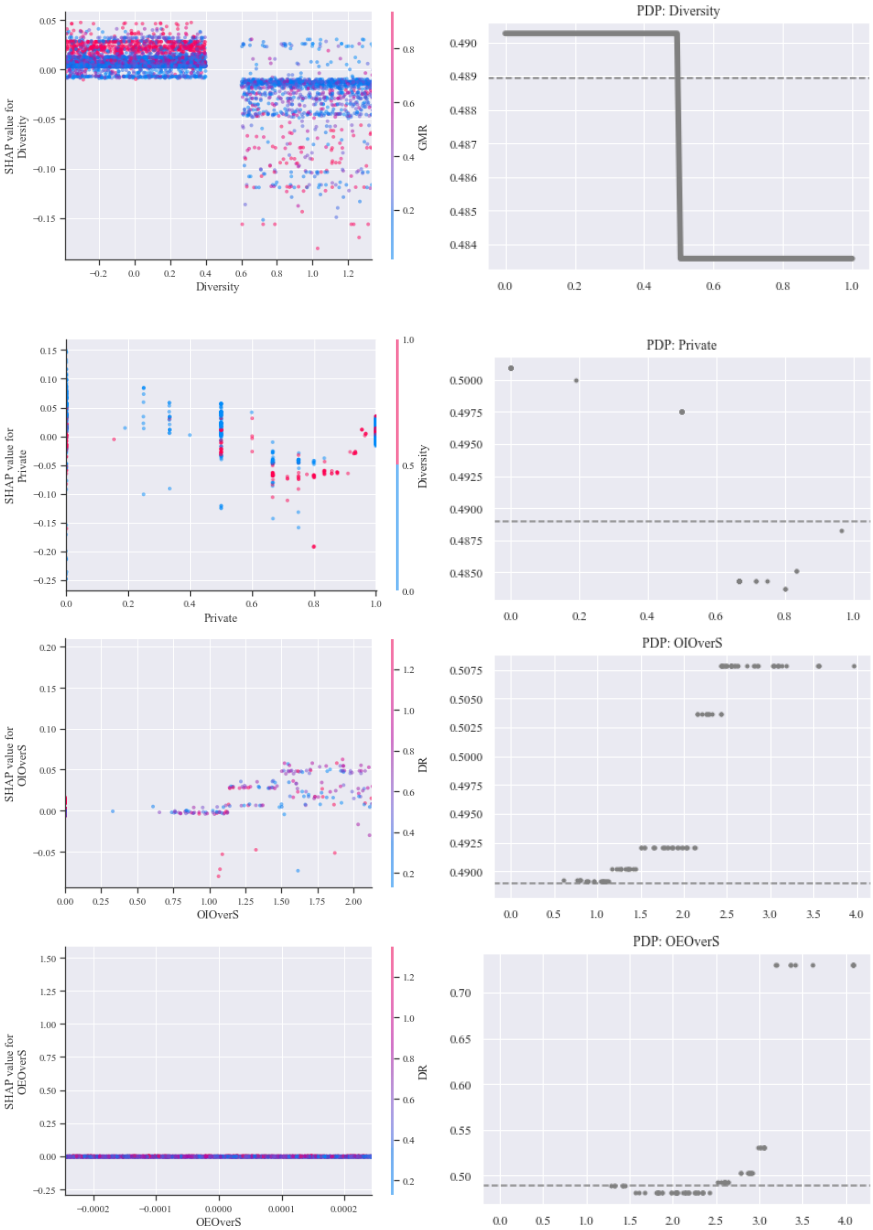




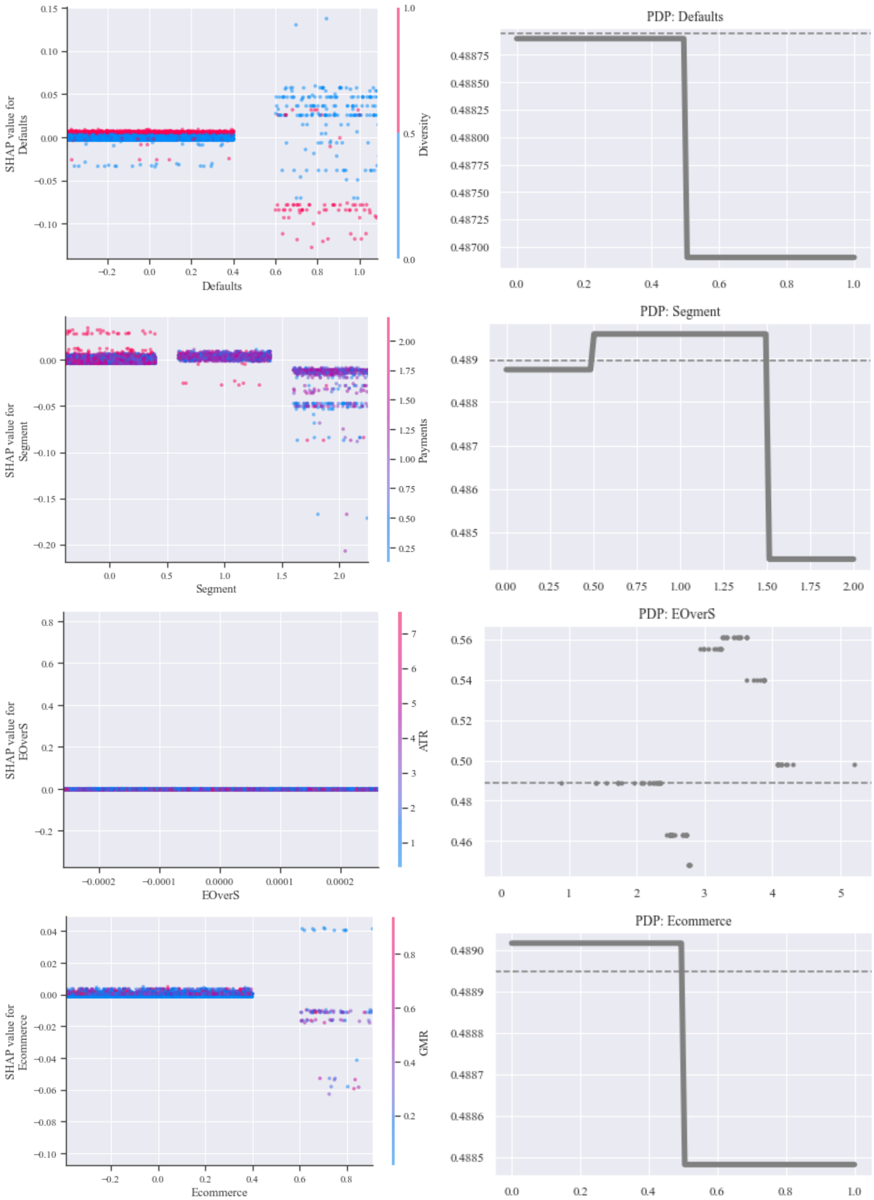
Continued on the next page.



Continued on the next page.

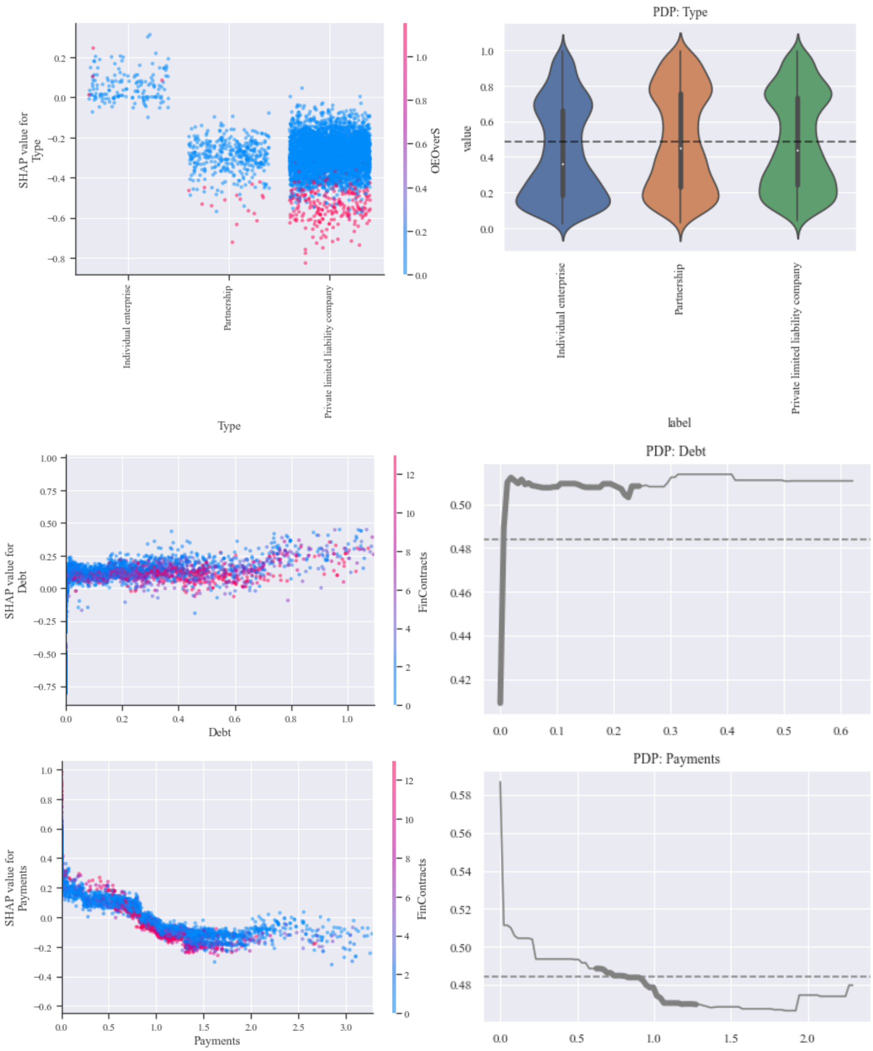


Continued on the next page.

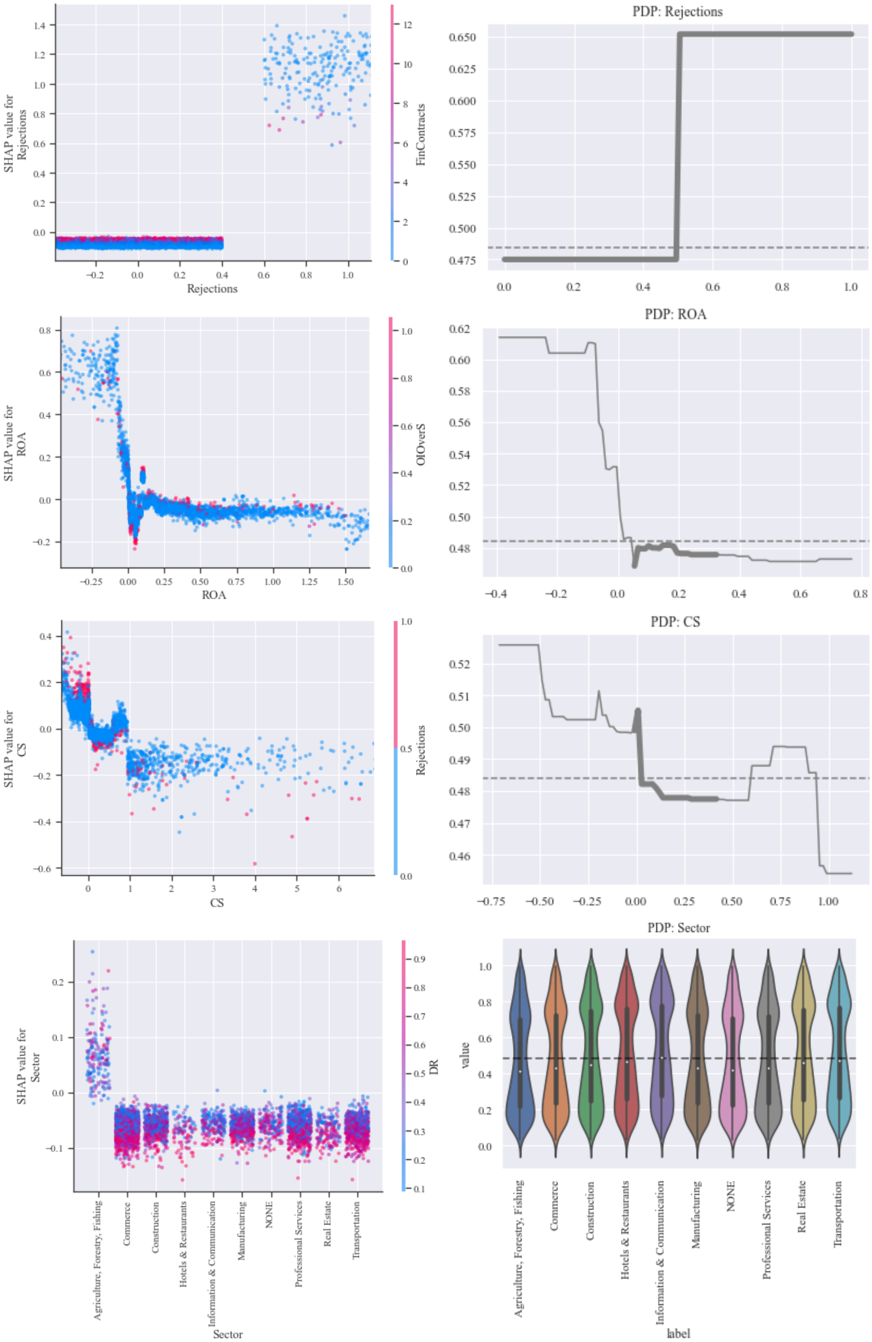


Continued on the next page.

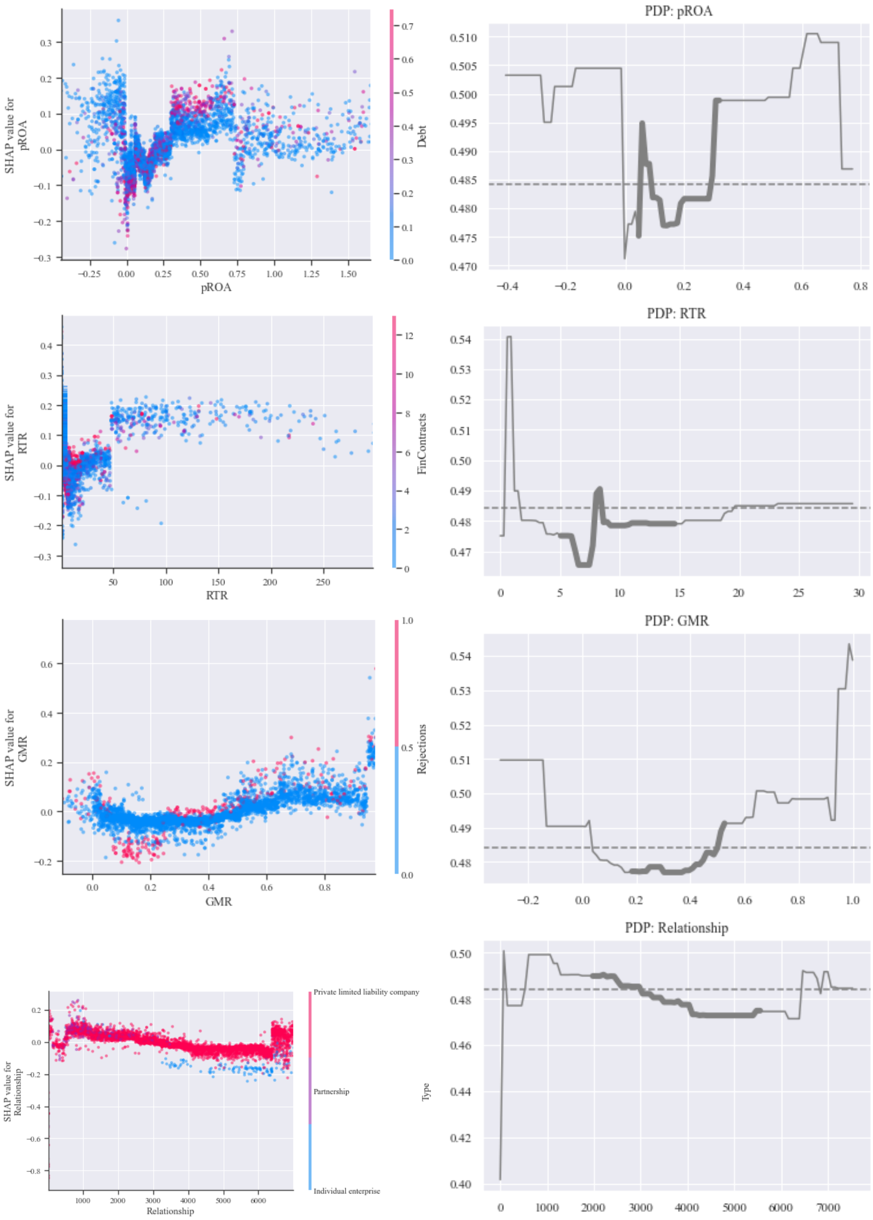
Annex 7. SHAP dependence and PDP plots for Lithuanian dataset.



Continued on the next page.

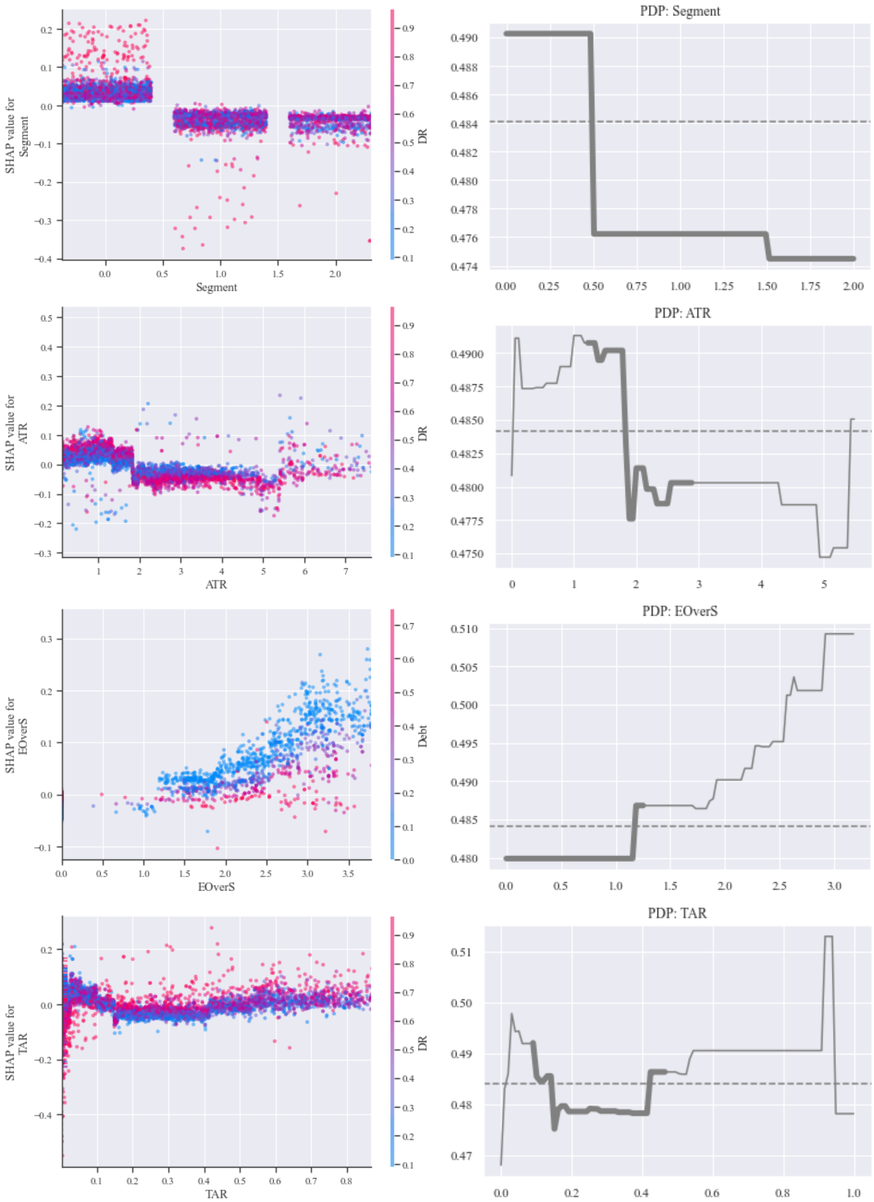


Continued on the next page.

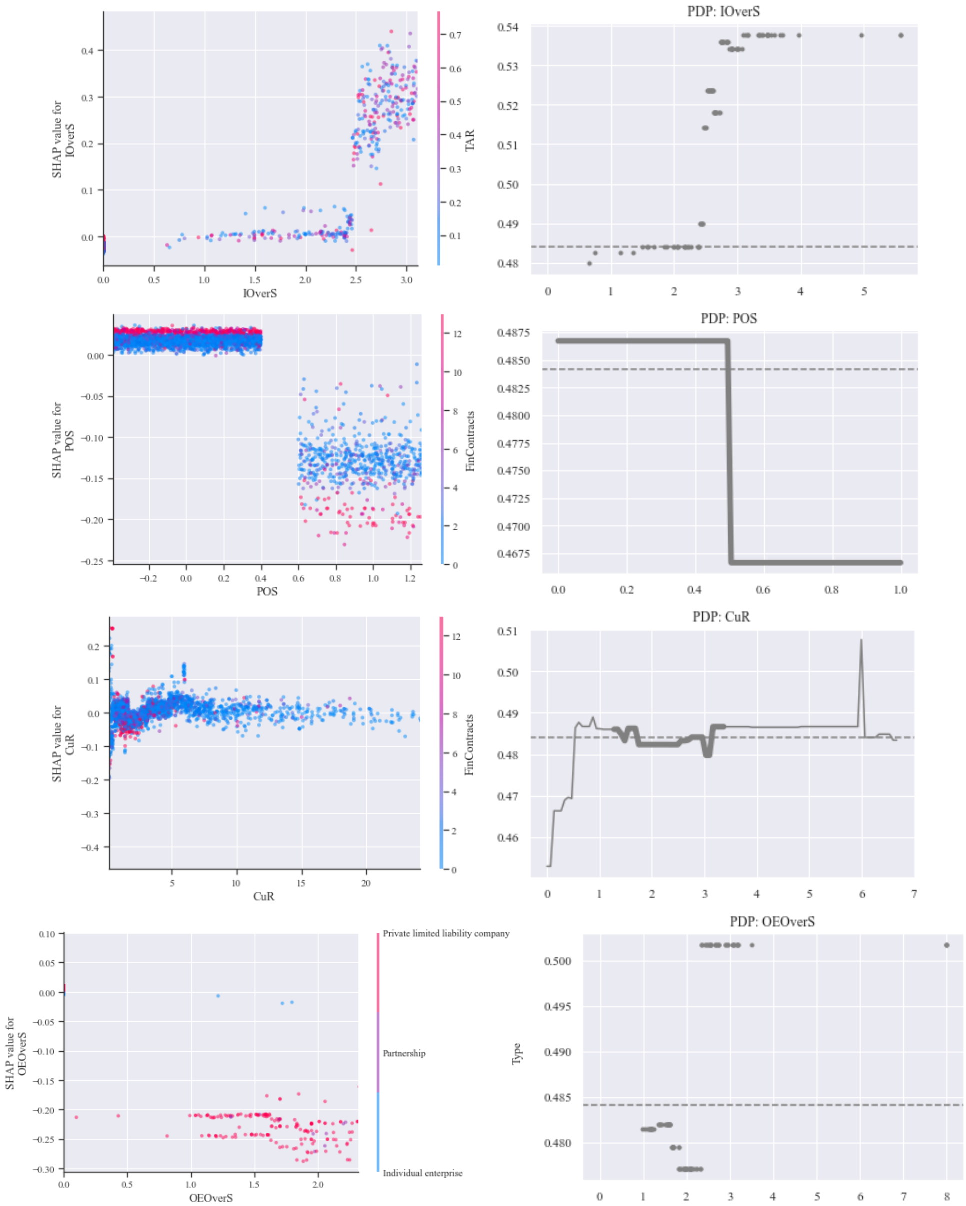


Continued on the next page.

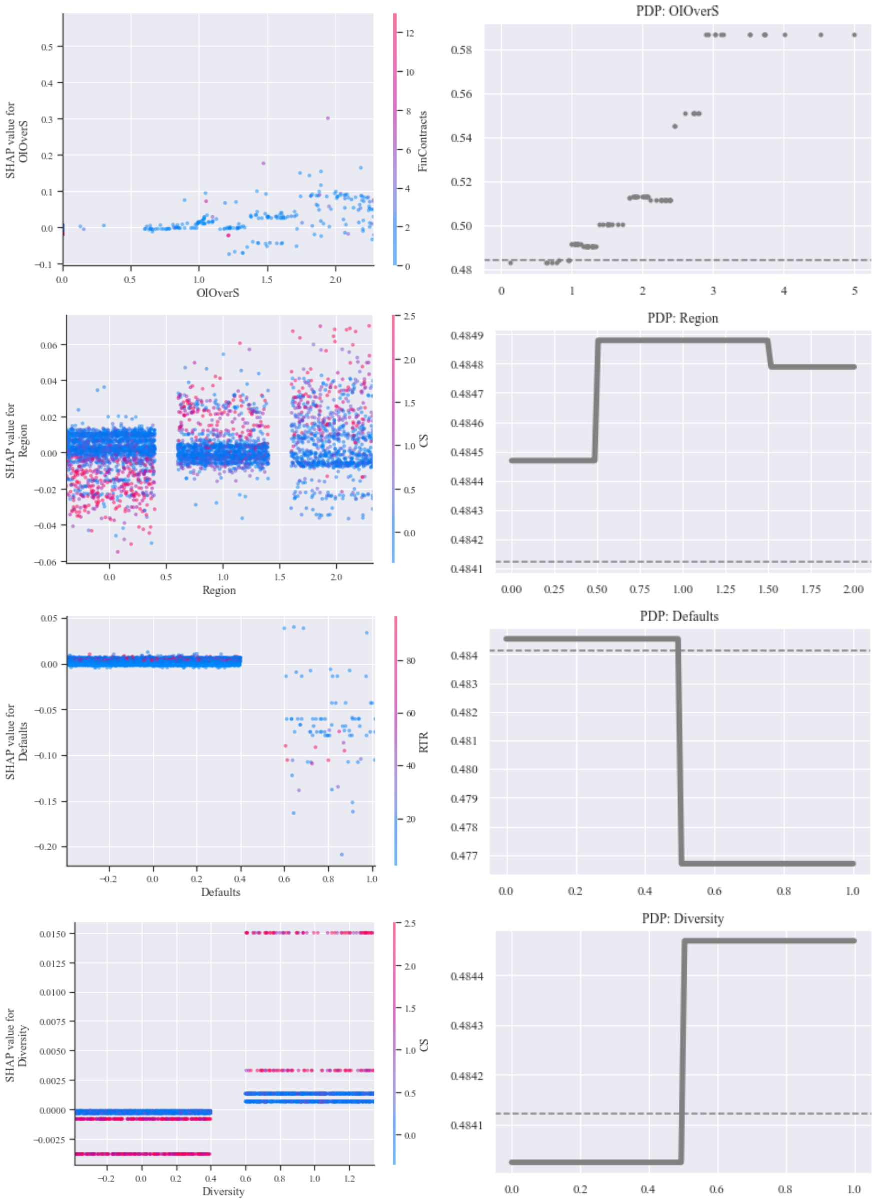
Annex 7 continued



Continued on the next page.

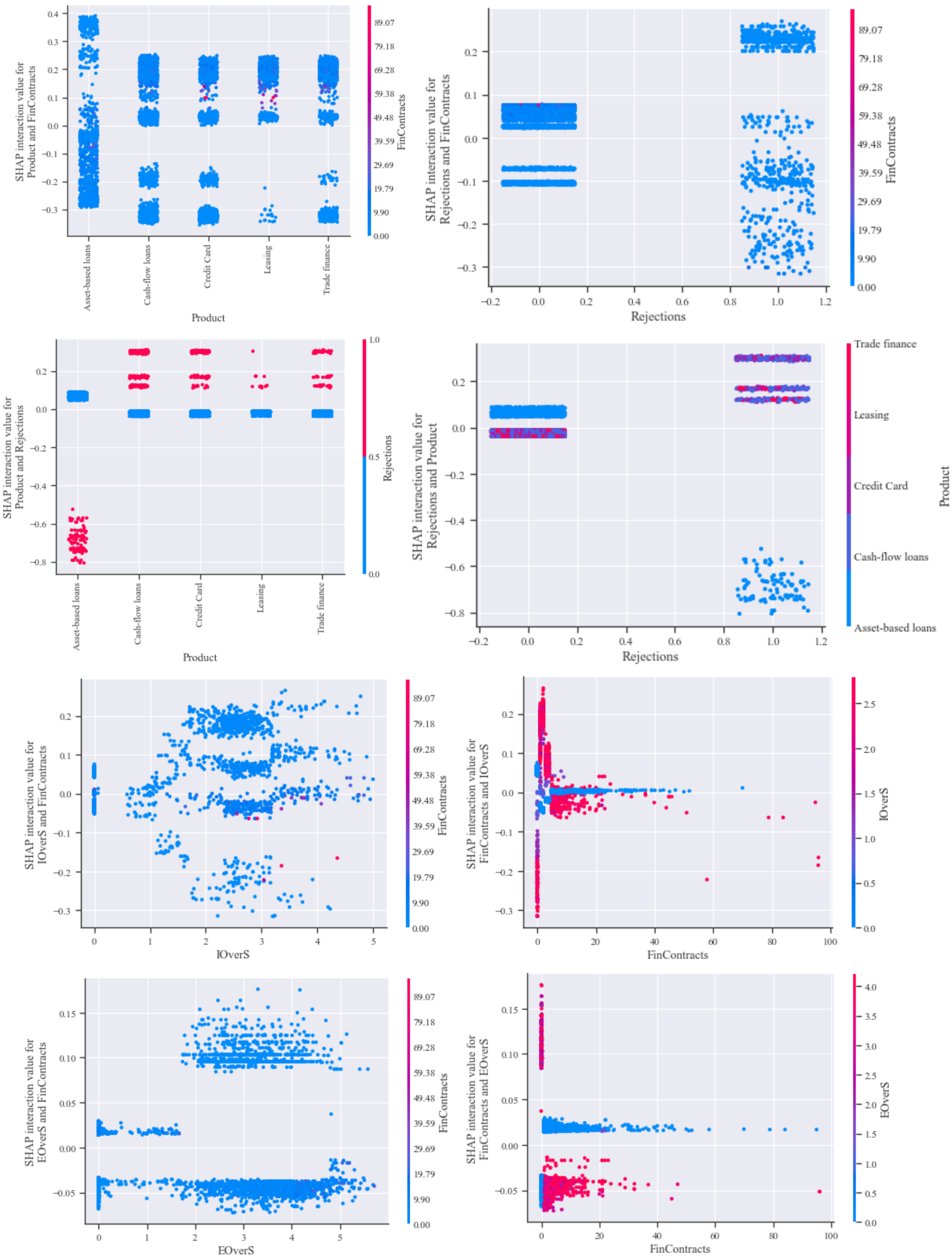


Continued on the next page.

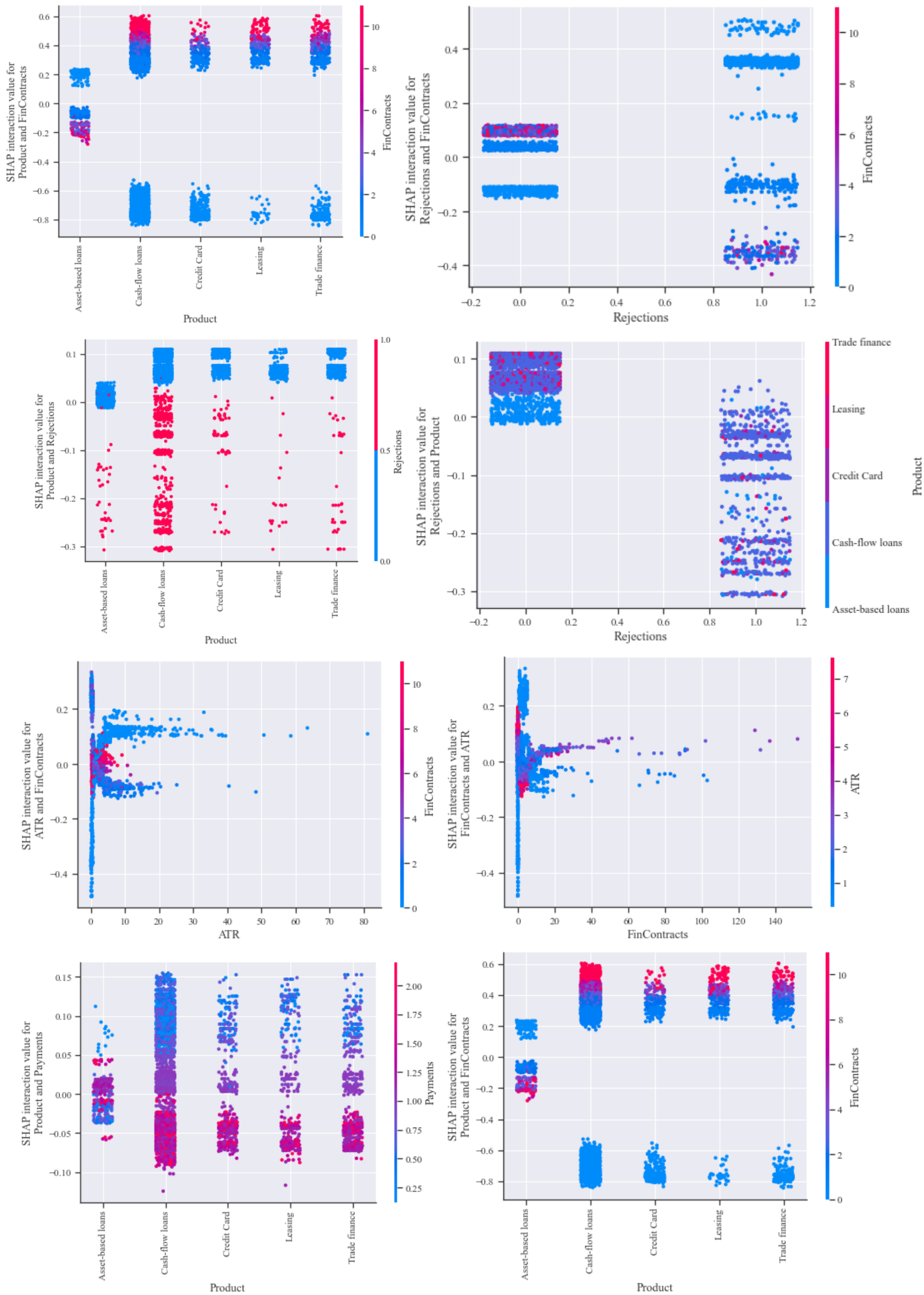


Continued on the next page.

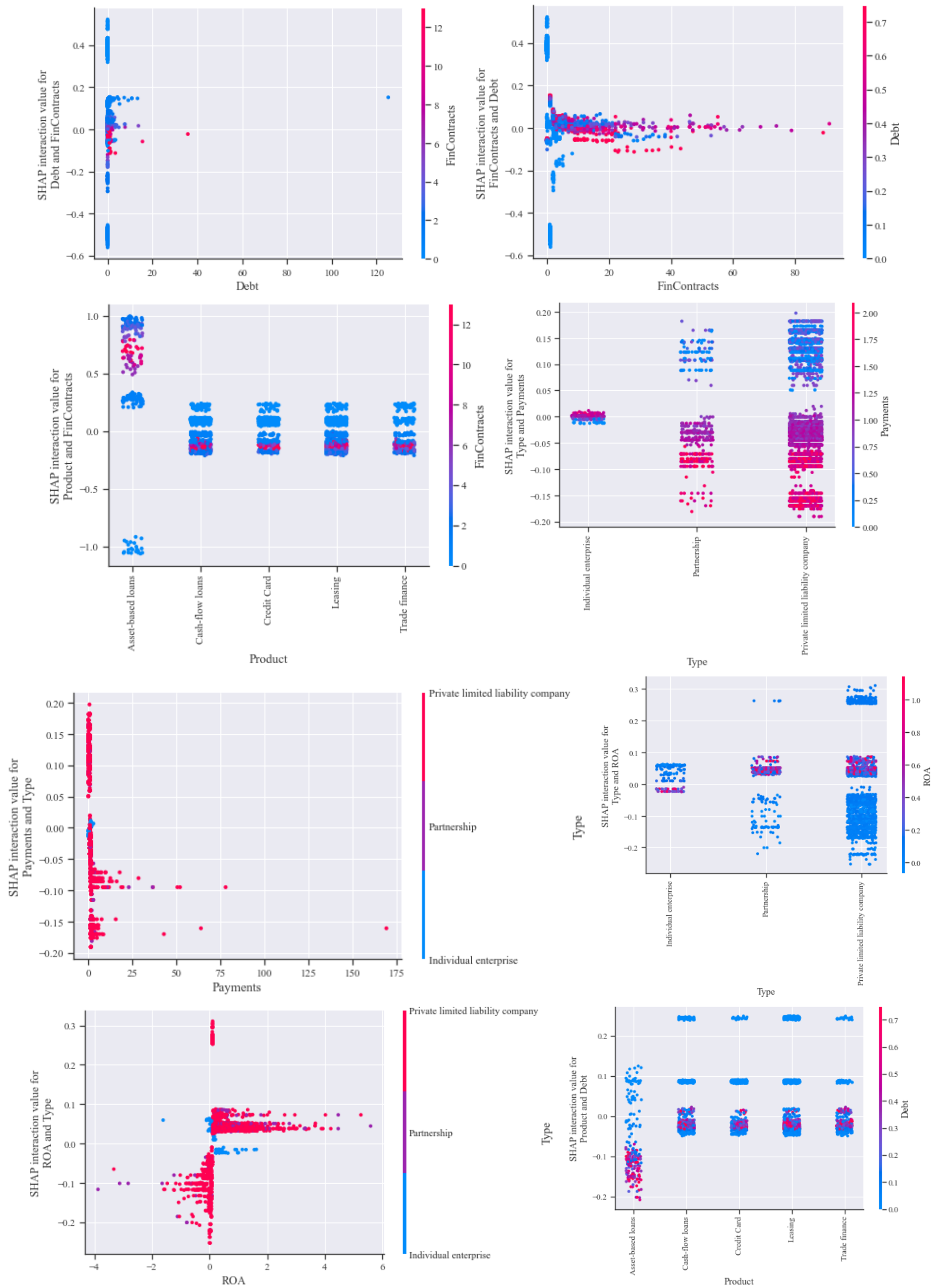
Annex 8. SHAP interaction plots for Estonian dataset.



Annex 9. SHAP interaction plots for Latvian dataset.



Annex 10. SHAP interaction plots for Lithuanian dataset.



UDK 336.77+334.72](043.3)

SL 344. 2023-09-25, 26,25 leidyb. apsk. l. Tiražas 14 egz. Užsakymas 163
Išleido Kauno technologijos universitetas, K. Donelaičio g. 73, 44249 Kaunas
Spausdino leidyklos „Technologija“ spaustuvė, Studentų g. 54, 51424 Kaunas