CHARLES UNIVERSITY
FACULTY OF SOCIAL SCIENCES
Institute of Economic Studies

Climate risk in financial markets

Master's thesis

Author: Bc. Simona Ivančová
Study program: Economics and Finance
Supervisor: prof. PhDr. Petr Teplý, Ph.D.
Year of defense: 2024

Declaration of Authorship

I hereby declare that I compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain any other academic title.

I grant to Charles University permission to reproduce and to distribute copies of this thesis in whole or in part and I agree with the thesis being used for study and scientific purposes.

Prague, April 29, 2024

Simona Ivančová

Abstract

This thesis investigates the relationship between climate transition risk and credit risk using a unique dataset provided by an anonymous Czech bank consisting of financial and carbon footprint information on corporate clients belonging to the SME category. Firstly, employing logistic regression, a standard credit scoring model was estimated using client-level financial predictors from 2022. Four significant financial drivers of credit default were identified based on the provided data. Second, a set of 11 variables on a client's carbon footprint was separately added to the standard credit scoring model. Results of the climate-stressed models imply that while direct emitters tend to default less, indirect emitters pose a higher threat to the bank in terms of credit risk. Finally, the predictive power of the climate-stressed models was compared to the standard model. Integrating Scope 2 carbon footprint into the credit scoring model enhances its discriminatory power both in terms of sensitivity and specificity.

Keywords climate risk, credit risk, credit default, ESG, carbon footprint, carbon intensity, probability of default, logistic regression

Title Climate risk in financial markets

Abstrakt

Tato práce zkoumá vztah mezi rizikem tranzice k uhlíkově čisté ekonomice a úvěrovým rizikem pomocí unikátní datové sady poskytnuté anonymní českou bankou, která obsahuje finanční ukazatele a uhlíkové stopy klientů patřících do kategorie malých a středních podniků. Nejprve byl pomocí logistické regrese odhadnut standardní model hodnocení úvěru s využitím finančních prediktorů klientů z roku 2022. Na základě poskytnutých dat byly identifikovány čtyři významné finanční faktory úvěrového selhání. Poté byla k standardnímu modelu hodnocení úvěru osobitě přidána sada 11 proměnných týkajících se uhlíkové stopy klienta. Výsledky klimaticky zatížených modelů naznačují, že zatímco přímí emitoři mají tendenci k nižšímu výskytu úvěrového selhání, nepřímí emitoři pro banku představují vyšší hrozbu z hlediska úvěrového rizika. Nakonec byla prediktivní síla klimaticky zatížených modelů porovnána se standardním modelem. Začlenění nepřímé uhlíkové stopy do modelu hodnocení úvěru zvyšuje jeho diskriminační schopnost jak z hlediska citlivosti, tak specificity.

Klíčová slova klimatické riziko, úvěrové riziko, úvěrové selhání, ESG, uhlíková stopa, uhlíková intenzita, pravděpodobnost defaultu, logistická regrese

Název práce Klimatické riziko na finančních trzích

Acknowledgments

I am grateful especially to prof. PhDr. Petr Teplý, Ph.D., who directed me to this interesting topic and, beyond, has been willing to share his knowledge and help in the creation of this thesis. I would also like to express my thanks to doc. PhDr. Jozef Baruník, Ph.D., whom I have consulted the method of this thesis. Furthermore, I would like to express my gratitude to my family, who have been always a support to me.

Bibliographic Record

Ivančová, Simona: *Climate risk in financial markets.* Master's thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies, Prague. 2024, pages 91. Advisor: prof. PhDr. Petr Teplý, Ph.D.

Contents

List of Tables

List of Figures

Acronyms

AIC Akaike Information Criterion

BCBS Basel Committee on Banking Supervision

BIS Bank for International Settlments

BP Brown Penalty

CBAM Carbon Border Adjustment Mechanism

CDS Credit default swap

DD Merton distance-to-default

EBA European Banking Authority

EC European Commission

ECB European Central Bank

EDF Expected Default Frequency

ESG Environmental Social and Governance

EU ETS European Union Emission Trading System

GHG Greenhouse gas

GSF Green Supporting Factor

IPCC Intergovernmental Panel on Climate Change

NGFS Network for Greening the Financial System

OECD Organisation for Economic Co-operation and Development

PD Probability of default

ROA Return on Assets

ROC Receiver operating characteristic

ROE Return on Equity

SME Small and medium-sized companies supporting factor

VIF Variance Inflation Factor

Master's Thesis Proposal

Author	Bc. Simona Ivančová
Supervisor	prof. PhDr. Petr Teplý, Ph.D.
Proposed topic	Climate risk in financial markets

Motivation The term climate risk has become increasingly relevant in recent years due to the growing recognition of the impact of climate change on global economies. In 2015, the Paris Agreement set the goal of mitigating the global temperature increase by 1.5 °C for the 21st century compared to the pre-industrial era, achieved mostly by cutting greenhouse gas emissions. A transition towards carbon neutrality is a supposed solution to reducing the impacts of climate change. Financial institutions play a crucial role in providing capital to support low-carbon transition projects and influencing corporate behavior skewed towards sustainability.

Climate risk in financial markets can be generally recognized in two forms, physical and transition risk (BIS, 2022). Physical risk materializes as a direct result of either one-off extreme natural event or long-term changes in climate. Transition risk stems from the implications of the transition to a carbon-neutral economy (Batten et al., 2016). The financial institutions are exposed to both climate risks indirectly, via their assets (Monasterolo et al., 2017). Supervisory bodies have increasingly pushed banks to develop internal climate stress testing and it is expected the practice will be implemented into the Basel IV agreement (ECB, 2021; BIS 2022). Given the straightening European climate policy, it is expected the transition risk impacts are going to be increasingly material for financial institutions.

There is a literature gap concerning the impact of the transition risk on the credit risk. One reason is that data on the carbon footprint of assets tends to be undisclosed or of questionable transparency (Monasterolo et al., 2017). Academics suggest the default risk increases with the level of carbon emissions. Existing research mostly points out the negative correlation between carbon footprint and Merton Distance-to-Default (Capasso et al. 2020; Carbone et al. 2021; Kabir 2021). Farrali & Ruggiero

(2022) suggested the expected default frequency rises with the carbon emissions of the borrower.

The purpose of this thesis is to analyze the transition risk in terms of credit risk. Given the access to unique data from an anonymous Czech bank on corporate loan data including client-level information on total and financed emissions, this analysis is set to quantify the impact of the client's carbon footprint on its probability of default.

Hypotheses

1. Client carbon footprint increases the odds of the default on corporate loan in the Czech Republic.
2. Client carbon-intensity increases the odds of the default on a corporate loan in the Czech Republic.
3. Inclusion of the client's carbon footprint increases the performance of the model for default prediction.
4. Inclusion of the client carbon intensity increases the performance of the model for default prediction.

Methodology The dataset provided by an anonymous Czech bank consists of time series data on corporate loans in the Czech Republic. The bank measures the carbon footprint and the carbon intensity of the borrowers internally. Control variables, such as financial and business sector information, are included in the dataset. Given the observed variable is the default on the loan, the logistic regression will be employed, based on its ability to handle binary outcomes such as default and non-default (Greene 2003; Jorion, 2010). The expected scope of the analysis is the following:

- Estimate a standard and a climate-stressed credit scoring model.
- Assess the estimates of the climate-stressed model.
- Compare the performance of the standard and climate-stressed models in terms of their ability to correctly predict a credit default.

Expected Contribution The expected contribution is two-fold. Firstly, the analysis will contribute to the scarce existing research on the presence of climate risk in financial markets by providing a specific case study of the transition risk in terms of credit risk of a Czech bank. Secondly, given the opportunity to access a unique dataset, the thesis proposes a new methodological approach to quantifying the transition climate risk, which was not yet introduced in the academic literature.

Outline

1. Theoretical background: We will introduce the main terms such as climate risk, credit risk and their measure. We will provide a specific definition of banking default. The current European climate policy will be described in more detail.
2. Literature review: We will comment on existing literature on climate risk in financial markets, and identify the literature gap, based on what we formulate the hypotheses.
3. Data: We will explain the choice of variables in the datasets, and then we will comment on descriptive statistics of the data.
4. Methodology: We will introduce the method and its assumptions.
5. Model: We will estimate standard credit scoring model.
6. Results: We will provide the baseline results and perform robustness checks for each hypothesis test.
7. Conclusion: We will summarize findings and implications for future research.

Core bibliography

1. Alogoskoufis, Spyros, et al. ECB economy-wide climate stress test: Methodology and results. No. 281. ECB Occasional Paper (2021).
2. Batten, Sandra, Rhiannon Sowerbutts, and Misa Tanaka. "Let's talk about the weather: the impact of climate change on central banks." (2016).
3. Battiston, Stefano, et al. "A climate stress-test of the financial system." Nature Climate Change 7.4 (2017): 283-288.
4. Battiston, Stefano, and Irene Monasterolo. "On the dependence of investor's probability of default on climate transition scenarios." Available at SSRN 3743647 (2020).
5. Battiston S, Dafermos Y, Monasterolo I. "Climate risks and financial stability." Journal of Financial Stability. (2021).
6. BIS (Bank for International Settlements: "Frequently asked questions on climate-related financial risks." (2022).
7. Capasso, Giusy, Gianfranco Gianfrate, and Marco Spinelli. "Climate change and credit risk." Journal of Cleaner Production 266 (2020).

8. Carbone, S., M. Giuzio, S. Kapadia, J. S. Kramer, K. Nyholm, K. Vozian (2021): "The low-carbon transition, climate commitments and firm credit risk.".

9. Faralli, M. & F. Ruggiero (2022): "The rise of climate risks: Evidence on firms' expected default frequencies." Available at SSRN 4144802.

10. Greene, William H. Econometric analysis. (2003).

11. Jorion, Philippe. Financial Risk Manager Handbook: FRM Part I/Part II. John Wiley & Sons, (2010).

12. Kabir, Md Nurul, et al. "Carbon emissions and default risk: International evidence from firm-level data." Economic Modelling 103 (2021).

13. Monasterolo, Irene, et al. "Vulnerable yet relevant: the two dimensions of climate-related financial disclosure." Climatic change 145 (2017): 495-507.

Chapter 1

Introduction

Climate risk represents a potential loss resulting from either extreme natural events or a transition towards a carbon-neutral economy. The European Green Deal is an agreement between the European Union member states upon the same goal: to cut 55 % of the carbon emissions by 2030 and reach carbon neutrality by 2050. To achieve such an ambitious goal, those who emit the most carbon emissions are set to be restrained financially, since money yet again proved to be the biggest motivation for saving the planet. Financial institutions do not directly generate any carbon footprint, but they provide financing to carbon-intensive firms, making them indirect polluters. European Commission once threatened to raise capital requirements of banks that do so (Dombrikovskis 2018). Looking at the business sector, the non-financial carbon-intensive firms are already being sanctioned for their carbon footprint by regulations such as carbon tax or emission trading systems setting a price per ton of CO_2 they emit. Climate policy might be an indirect credit default driver, directly impacting banks. The relationship between credit and climate risk has been subject to several academic papers, suggesting the level of carbon pollution might be an actual driver of a corporate client default (Capasso *et al.* 2020).

Given the access to a unique dataset provided by an anonymous Czech bank (further referred to as the Bank) consisting of client-level data on the financial situation and the level of carbon footprint (further referred to as transition risk variables), we were able to examine how the amount and the intensity of carbon emissions influences the probability of credit default. The relationship will be studied through a climate-stress test of a standard credit scoring model by transition risk variables.

The thesis is structured as follows. In Chapter 2 we define the most important terminology used throughout the thesis. In Chapter 3 we summarize the existing academic literature concerning climate risk in financial markets and we identify the literature gap based on which the hypotheses are formulated. Chapter 4 introduces the methodology used to test the hypotheses. In Chapter 5, the adjustments to the original dataset and the descriptive statistics of the data are discussed. Chapter 6 introduces a standard model for credit default prediction. Chapter 7 discusses the climate-stressed models. Finally, the findings are summarized in Chapter 8. Given the confidential nature of the banking data, the dataset is not disclosed but further details on data and the code used for the analysis can be provided upon request.

Chapter 2

Theoretical Background

This chapter elaborates on the terminology important to the thesis. Section 2.1 introduces two types of climate risk, i.e., physical risk and transition risk. Credit risk is discussed in Section 2.2.

2.1 Climate risks

There are two main drivers of climate risk in the context of financial markets: physical risk and transition risk. This section defines these risks and their relevance to the financial system. The focus of this research is solely placed on the transition risk, the concept of physical risk is included for completeness purposes. Unless stated otherwise, the definitions of climate risk discussed in this section are based on those presented by the European Central Bank (ECB 2020) and the Basel Committee on Banking Supervision (BCBS 2021).

2.1.1 Physical risk

Physical risk represents the financial losses stemming from physical damage caused by extreme climate change-induced events. The Intergovernmental Panel on Climate Change recognizes either acute events, i.e., sudden extreme weather events such as heatwaves, floods, droughts, storms, and wildfires, or chronic events, i.e., long-term shifts in climate patterns such as rising sea levels, melting glaciers and changing temperature patterns (IPCC 2021). Both acute and chronic events undermine the stability of the financial system. ECB recognizes the financial impact of physical risk in all four main risk categories: credit risk, market risk, operational risk, and liquidity risk.

In the context of credit risk, financial institutions are exposed via default risk driven by an increase in the probability of default and decrease in the recovery rates. For example, severe flood events tend to result in damaged residential property, causing both mortgage default and the destruction of physical collateral. Various forms of physical risk tend to be concentrated in distinct geographical areas. Banks operating in limited geographical areas are likely to suffer distress induced by the occurrence of a single extreme natural event.

2.1.2 Transition risk

Transition risk represents a risk of financial loss caused by the process of adjustment towards a low-carbon economy. The process of transition might involve carbon pricing policies or adjustments to market prices. For a carbon-intensive industry, transition risk translates to higher operating or financing expenses. The amount of carbon emissions produced by a firm can be interpreted by two concepts:

- **Carbon footprint:** the total amount of carbon emissions that are directly or indirectly associated with a firm, measured in units of mass of carbon dioxide.

- **Carbon intensity:** the amount of carbon emissions produced per unit of economic output. Countries usually use GDP-based carbon intensity, firms tend to use revenue-based or production-based carbon intensity.

The European Central Bank (ECB) recognizes three scopes for estimating the carbon emissions of a firm:

- **Scope 1:** Direct carbon emissions that result from sources that are owned or controlled by the firm, such as the combustion of fossil fuels in firm-owned facilities or vehicles.

- **Scope 2:** Indirect emissions associated with the firm's operations, such as electricity, steam, heating or cooling purchased from suppliers and used in production.

- **Scope 3:** All other indirect emissions that occur in the value chain of the firm. Financial institutions belong to Scope 3 since they provide financing to direct and indirect polluters.

The transition towards a low-carbon economy can consist of gradual interventions (Nordhaus 2007) or more restrictive policies (Stern 2009). The Network for Greening the Financial System published four main categories of future climate scenarios (NGFS 2021) :

- **Orderly transition:** Policies to reduce carbon emissions are implemented at an early stage and are gradually scaled up to achieve carbon neutrality by 2050. Physical and transition risks are effectively mitigated.
- **Disorderly transition:** Policies to mitigate carbon emissions are implemented at a later stage in a more abrupt manner, resulting in a failure of investors to anticipate the impact of climate policies on their business models (Battiston *et al.* 2017). The resulting stranded assets are those that have suffered from unanticipated or premature write-downs or devaluations (Caldecott *et al.* 2014). While physical risk might be mitigated, the transition risk is dominant in this scenario.
- **Hot house world:** Strict climate policies are implemented in some regions, but globally they remain uncoordinated and insufficient. Physical risk becomes the most dominant threat to the financial system.
- **Too little too late:** Transition policies are ineffective on an international level. There is elevated transition risk in some countries and high physical risk in all countries as critical temperature thresholds have been exceeded.

European climate policies

The European Green Deal is a package of policy initiatives by the European Union with the aim of making the EU carbon-neutral by 2050. The European climate law obliges the member states to reduce their net GHG emissions by at least 55 % compared to 1990 by 2030 and reach carbon neutrality by 2050. Fit for 55 is a set of proposals for the EU to revise and put in use in order to achieve the 55% reduction goal by 2030. This subsection discusses the most important adopted or planned climate measures by the EU that represent significant drivers of transition risk for the financial markets.

- **EU ETS:** The European Union Emission Trading System is a market-based mechanism that sets a yearly cap on total carbon emissions allowed to be emitted. One traded allowance equals one ton of carbon emitted. If the carbon emissions of a firm exceed the amount covered by permits, the firm is fined. The yearly cap is progressively lowered, and the carbon price is derived from the market. The measure has been in practice since 2005 and applies in all member states.

- **Carbon tax:** a governmental fiscal policy to price emissions. Tax is levied directly per ton of Scope 1 emissions. The tax rate and the tax scope are set by the member state government. Carbon tax applies in 14 member states. The Czech Republic is currently only subject to the EU ETS.
- **CBAM:** Carbon Border Adjustment Mechanism is a mechanism to prevent carbon leakage (EC 2023). Carbon leakage occurs when firms move their carbon-intensive production to countries with less stringent climate policies. The price set per ton of emissions is derived from the price under the EU ETS. The measure is expected to enter fully into force in 2026.
- **GSF:** The Green Supporting Factor is a proposed measure that would allow a reduction of the capital requirements for banks that finance green projects. It represents a downward shift in risk weights for green loans. Contrary to GSF, the Brown Penalty represents an upward shift in risk weights. The proposal is currently being reconsidered due to the concerns of weakening the financial sector (Dombrikovskis 2018).

2.2 Credit risk

Credit risk is defined as the potential financial loss of the lender due to the counterparty's unwillingness or inability to fulfill their contractual obligations (Jorion *et al.* 2010). Credit risk is the most significant financial risk for the banking sector since loans compose the majority of its assets (BIS 2023). In the context of credit risk, a usual distribution of returns is asymmetrical and left-skewed, as the upside return potential for the lender is limited by the contract. Negative skewness also implies a higher probability of extreme losses than extreme gains.

Expected loss (EL), a measure of credit risk, represents an estimate of the potential loss to the lender over a given time horizon. EL is calculated as a product of three components: the probability of default (PD), loss-given default

(LGD), and exposure at default (EAD).

$$EL = PD * LGD * EAD - PD * (1 - recovery\ rate) * EAD \tag{2.1}$$

PD is a statistical measure representing the likelihood that a borrower will fail to meet its debt obligations within a specified timeframe. PD is estimated by statistical models leveraging a range of quantitative and qualitative factors, including financial ratios, credit histories, and economic indicators, to assess the creditworthiness of individuals, companies, or entities. The most common methodologies for PD estimation encompass logistic regression or more complex machine learning algorithms.

LGD represents the extent of financial loss that a lender is likely to incur in the event of a borrower's default. It can be expressed as $(1 - recovery\ rate)$. The recovery rate represents the share of debt that can still be recovered. The recovery rate is estimated using historical data on past defaults, calculating the average or median recovery rate for a specific business segment or type of exposure. A more case-specific method for recovery rate estimation is using the Discounted Cash Flow model.

Credit risk of a firm can also be proxied by credit rating or credit default swap. A credit rating refers to a quantified evaluation of a borrower's creditworthiness concerning a particular debt. Expressed as a letter grade, it is issued by a credit rating agency or assigned to a borrower by a bank itself based on its internal metrics. A credit default swap is a financial derivative that enables an investor to offset the credit risk stemming from a defaultable loan or bond by transferring it to a swap seller (Jorion *et al.* 2010). The swap buyer pays an ongoing regular protection premium to the swap seller in return for a conditional payment in case of loan or bond default. The higher the premium paid for the swap, the more risky the debt.

As further mentioned in Chapter 3, researchers tend to proxy the credit risk by the Merton distance-to-default (DD) or the Expected Default Frequency (EDF) since these measures can be derived from available market data on publicly traded companies. The DD is a credit risk measure based on the structural model of default introduced by Merton (1974). The model assumes that the equity of a firm is a call option on the underlying value of the firm and that a firm defaults when its value falls below the face value of its debt. The DD is computed as the number of standard deviations between the expected asset value and the liability threshold. The EDF is a measure of the probability that a firm will default over a specified time period, typically one year. EDF is built on an assumption that firm default is driven by the market value of the assets being lower than the liabilities payable (Hamilton *et al.* 2011). EDF value is determined by the market value of assets, the level of the firm's obligations, and the volatility of the firm's market value.

Chapter 3

Literature Review

This chapter discusses the literature relevant to the thesis. The academic research relevant to the thesis can be classified into three main strains. First, climate risk impact on financial system resilience is discussed in Section 3.1. Second, climate risk impact on the pricing of securities is studied in Section 3.2. Third, climate risk implications for credit risk management are covered in Section 3.3. Finally, after the literature gap is identified, the hypotheses are formulated in Section 3.4. Given the research design of this thesis, the focus is solely placed on studies aimed at climate transition risk.

3.1 Climate risk impact on financial system

Research on the transition risk effect on the economy gained prominence mostly after the Paris Agreement in December 2015, when the majority of global economies agreed upon a coordinated effort to mitigate carbon emissions. The conference gave a clear signal of upcoming climate policies intended to financially restrain polluter firms and industries. Therefore, transition risk is most often proxied by the carbon footprint or the carbon intensity of the studied subject. When quantifying transition risk, the academics face inconsistency in the reporting of carbon emissions across the business sectors (Battiston *et al.* 2021).

Battiston *et al.* (2017) build a methodological base for conducting a climate-stress test of the financial system using data on shareholders of all EU and US publicly listed companies and loan portfolios of the 50 largest EU banks. The study revealed that investment and pension funds have the largest holdings of brown assets. Around 11.4 % of bank assets were exposed to transition risk, which implies the banking sector is resilient at the moment, but this figure is expected to arise with the intensifying reach of climate policy (Battiston *et al.* 2017). Monasterolo *et al.* (2017) extended the model by Battiston *et al.* (2017) by examining the weight of each subject's exposure to transition risk on the resilience of the financial system overall. The study suggests that governmental and individual investors are the most vulnerable to transition risk, yet the structure of their portfolios is not systematically important. Industry companies and investment funds emerge as the most exposed and important. Banks stand the middle ground both in terms of exposure and importance (Monasterolo *et al.* 2017).

Studies evaluating individual climate policies examine both the consequences of the policy on the financial system and their contribution to fostering a green economy transition. Dunz *et al.* (2019) examines the implications of GSF and carbon tax on the economy. Study suggest GSF might have only a short-term effect on increased green lending but may introduce potential trade-offs for financial stability. A carbon tax should be complemented with welfare measures to prevent unintended effects on non-performing loans and household budgets (Dunz *et al.* 2019). Dafermos & Nikolaidi (2021) also find only a limited effect of GSF and BP on the increase of green lending. Nevertheless, both studies underline the transition risk stemming from these policies.

3.2 Climate risk impact on pricing of securities

Green securities have been gaining popularity among investors during the past decade. Green bonds, i.e., bonds that are used for sustainable project financing, are the most relevant among green securities. The first green bond was issued in 2010, and the first corporate green bond in 2013. Growing demand for these bonds can be partially attributed to arousing social responsibility among the investors, not seeking a financial advantage but rather a contribution to the environmental cause (Maltais & Nykvist 2020). At the same time, numerous academic papers suggest potential evidence of the yield and pricing difference in comparison to peers. The greenium is the amount by which the yield on the green instrument is lower, implying a lower cost of debt financing (CBI 2017). The concept of the greenium has raised the academic debate as mixed empirical evidence has been uncovered. Although green securities specifically are not the subject of this thesis research, the academic evidence on greenium is a valid argument when examining the climate risk in financial markets.

Greenium was first mentioned in the study published by the Climate Bond Initiative (CBI 2017). This study did not find any clear difference in yields, however, suggested that green bonds are generally issued at yields above the yield curve. Karpf & Mandel (2018a) compared the green bonds with their peers over 6 years, from 2010 to 2016. There has been a significant positive yield premium on green bonds during the first 4 years of the observed period. During the last 2 years, the yield of conventional bonds significantly exceeded the yield of green bonds (Karpf & Mandel 2018a). Gianfrate & Peri (2019) analyze 121 European green bonds issued between 2013 and 2017 to find that green bonds are issued at significantly lower prices. Zerbib (2019) identifies a small but significant negative yield premium of green bonds, using data on the European bond market between 2013 and 2017. The study employed a coarsened matching method, which consisted of paring green and conventional bonds with the same characteristics except the green label to avoid selection bias. The results suggest an average yield difference of -2 basis points (Zerbib 2019).

Contrary to findings of the previous studies, Hachenberg & Schiereck (2018) did not find a clear difference in pricing, but suggested that on average, green bonds trade with tighter bid-ask spread than conventional bonds. A tighter bid-ask spread is an indicator of lower credit and liquidity risk for an investor. Fatica *et al.* (2021) uncover the presence of the greenium on bonds issued by supranational institutions and corporates. The result implies that issuers that have a positive environmental reputation are favored by investors indicating the investors are wary of the greenwashing tactics. Despite the mixed evidence, there is a visible trend in recent literature, that points out the incentives of many investors are rather turning in the direction of social responsibility. This supports the idea that green investments carry a lower cost of debt and, hence, less risk for the lender.

3.3 Climate risk implications for credit risk management

In terms of approximating transition risk, academic studies suggest several approaches for modeling the relationship between credit risk and climate risk. Earlier studies suggested employing the ESG rating as a proxy of transition risk, mostly due to lack of disclosed data on borrower emission amount. However, as climate policy impact on a firm is directly derived from its carbon footprint, carbon intensity, and carbon footprint represent the closest determinants of transition risk.

Concerning the studies using ESG rating as an observed variable, Oikonomou *et al.* (2014) observe the ESG effect on credit spread and credit bond rating. The study suggests that a better sustainability rating can materially reduce the risk premium associated with corporate bonds and thus decrease the cost of corporate debt. These findings appear to be fairly robust across sectors. Newer studies find evidence that a better ESG score is associated with better credit rating (Devalle 2017), lower CDS spreads (Blasberg *et al.* 2021), and a lower bond risk premium (Kotró & Márkus 2020). Höck *et al.* (2020) suggest that for firms with ex-ante high creditworthiness, the higher environmental score is associated with low leverage and higher market capitalization. Chava (2014) argues that firms with multiple environmental concerns must pay higher costs on their bank loans.

Employing direct carbon emissions as an independent observed variable, Jung *et al.* (2018) find the cost of bank debt of Australian firms increases with an increase in firms' carbon footprint, using a panel regression. Wang *et al.* (2021), using the Generalized Method of Moments, find firms with high carbon footprint facing more stringent loan covenants and tighter loan repayment schedules. Zhang *et al.* (2023), using pooled OLS regression, suggest higher Scope 1 emissions increase CDS spreads of the firms, whereas Scope 2 and 3 emissions are not yet priced by the credit market.

Several recent academic papers focus solely on direct drivers of credit risk, i.e., probability of default, Merton distance-to-default (DD) or Expected default frequency (EDF), and their dependence on the carbon footprint of the borrower. This strain of literature is the most relevant given the research design. Capasso *et al.* (2020) show that DD is negatively correlated with the firm's carbon footprint and carbon intensity. The result implies firms with a higher carbon footprint have a lower distance to default, i.e., are more likely to default. Market data on 458 companies between 2007 and 2017 was used to estimate DD. Then the DD was employed as a dependent variable in panel regression, with carbon intensity and carbon footprint used as observed independent variables. Capasso *et al.* (2020) identified COP21 as an exogenous policy shock, as DD was significantly lower for carbon-intensive firms post-2015.

Carbone *et al.* (2021) performed an analysis similar to Capasso *et al.* (2020), using stock data on 558 large US and EU companies from 2010 to 2019. The study came to the conclusion that high emissions tend to be associated with higher credit risk, but disclosing emissions and setting a forward-looking environmental commitment may mitigate this effect. Carbone *et al.* (2021) also capture the deterioration of credit ratings of the most climate risk-exposed firms post-COP21. Kabir *et al.* (2021) extends the research by Carbone *et al.* (2021) by studying over 2700 firms from 42 economies, to come to the same causal relationship between the firm DD and emissions. Kabir *et al.* (2021) find evidence that control variables ROA and cash flow volatility are the potential channels through which a firm carbon footprint affects the DD. The public environmental commitments of firms tend to reduce the negative impact of emissions on DD.

Faralli & Ruggiero (2022) employ EDF as a measure of credit risk along with carbon footprint to investigate their causal relationship. Data from 1841 firms from 2008 to 2019 were used in fixed-effect panel regression. The use of EDF as the credit risk proxy allowed to observe the effects of carbon footprint on the main drivers of default separately. The emission level is relevant to the probability of default mainly through the asset volatility channel. In line with Capasso *et al.* (2020) and Carbone *et al.* (2021), Faralli & Ruggiero (2022) also identified an effect of COP21 on the riskiness of polluter firms. Nguyen *et al.* (2023) perform a follow-up study to Capasso *et al.* (2020), observing S&P 500 firms during the 2010-2019 period, to find that the negative effect of transition risk on firms' distance to default is stronger for firms headquartered in states with carbon pricing.

Not employing direct carbon emission data but rather testing transition risk scenarios, Bell & van Vuuren (2022) simulate corporate equity price projections using Geometric Brownian Motion to approximate the effects of climate policy shocks. Inputs are then used to estimate a climate-stressed PD. As a result, there are negligible increases in PDs for highly rated credit, but the effect of policy shocks increases with worsened credit quality.

3.4 Formulation of the hypotheses

Based on the literature review, there is evidence of the possible relationship between transition and credit risk. Given the access to a unique dataset provided by the Bank, consisting of data on corporate loans and client-level financial and carbon footprint information, this thesis will closely examine the effect of client carbon emissions on the probability of loan default. The research questions are the following:

1. Client's carbon footprint increases the odds of default on a corporate loan in the Bank.

2. Client's financed carbon footprint increases the odds of default on a corporate loan in the Bank.

3. Client's carbon intensity increases the odds of default on the corporate loan in the Bank.

Chapter 4

Methodology

This chapter elaborates on the method chosen for the empirical analysis. First, Section 4.1 introduces the concept of logistic regression. Second, Section 4.2 is dedicated to the estimation and evaluation of logistic regression models.

4.1 Logistic Regression

Logistic regression is a statistical method used for the estimation of the dependent variable of a binary outcome. The linear form of the logistic regression with n features can be expressed as a natural logarithm of the odds ratio:

$$\ln\left(\frac{P}{1-P}\right) = \beta_0 + \sum_{i=1}^{n} \beta_i X_i \tag{4.1}$$

$$\frac{P}{1-P} = e^{\beta_0 + \sum_{i=1}^{n} \beta_i X_i} \tag{4.2}$$

$$P = e^{\beta_0 + \sum_{i=1}^{n} \beta_i X_i} - Pe^{\beta_0 + \sum_{i=1}^{n} \beta_i X_i} \tag{4.3}$$

$$P = \frac{e^{\beta_0 + \sum_{i=1}^{n} \beta_i X_i}}{1 + e^{\beta_0 + \sum_{i=1}^{n} \beta_i X_i}} \tag{4.4}$$

$$P = \frac{1}{1 + e^{-(\beta_0 + \sum_{i=1}^{n} \beta_i X_i)}} \tag{4.5}$$

where:

- $\frac{P}{1-P}$ represents the odds of an observed event occurring,
- β_i represents a vector of coefficients of the model,
- X_i represents a vector of independent variables.

If an estimator of the predictor is negative, the increase in this predictor decreases the odds of the event occurring and vice versa. Logistic regression requires several assumptions to be fulfilled. Dependent variables are of a binary outcome, either 1 or 0. These classes are exclusive and no observation can be part of both. The classes should be balanced, as the model estimates might be biased towards the majority class. The relationship between the dependent and independent variables is linear. There should be no perfect separation of the classes, i.e., estimates of odds should not be equal to 1 or 0. Logistic regression assumes no multicollinearity, i.e., high correlation between predictors. Multicollinearity can be either identified in the correlation matrix or by computing the Variance Inflation Factor (VIF) for each predictor. VIF is calculated as the ratio of the variance of a coefficient when fitting the full model to the variance of a coefficient when fitting that predictor variable in a single variable model. A value between 5 and 10 represents a moderate correlation. If the VIF of a predictor is higher than 10, it is likely to be severely correlated with another predictor in the model (Greene 2003).

Logistic regression is a standard method used for the credit risk assessment of a potential client (Jakubík *et al.* 2011). Compared to more sophisticated machine learning models such as neural networks or decision trees, logistic regression is favored because of its simplicity and interpretability of results (Dumitrescu *et al.* 2022). A trained logistic model represents a set of parameters on selected financial inputs, estimated using historical data. Financial information on a new potential client is plugged into the model which returns a probability of default, i.e., a credit score, based on which the decision whether a loan will be granted is taken.

4.2 Evaluation of the Logistic Regression

When the model is specified, its predictive power is tested on an unobserved testing set. The odds returned by the model are translated to binary predictions based on the threshold probability value. If this threshold is set to 0.5, all generated estimates carrying a value of 0.5 or higher would be evaluated as positive, hence the event is predicted to occur. Estimates below 0.5 would be evaluated as negative, hence the event is not predicted to occur. These binary predictions are compared to actual events in testing data and the model is evaluated based on the number of the correct matches.

True positive prediction (TP) means the model correctly predicted the occurrence of an event given the threshold. True negative (TN) prediction means the model correctly identified an event not happening. False positive prediction (FP), i.e., Type 1 error, occurs when a model predicted an event to happen but in fact, it did not. False negative prediction (FN), i.e., Type 2 error, occurs when the model fails to predict an event that occurred. A more sensitive model tends to create fewer Type 2 errors while a more specific model will aim for fewer Type 1 errors. Whether a more specific or more sensitive model is needed depends on the nature of the analysis. In the context of credit risk, a Type 1 error results in credit denial to creditworthy clients. A Type 2 error leads to classifying high-risk clients as creditworthy. A credit scoring model aims to reach as high sensitivity as possible while keeping specificity at high enough levels.

To find a balance between specificity and sensitivity, the Receiver Operating Characteristic (ROC) curve is employed in this thesis for the model assessment. The ROC curve plots the true positive rate (sensitivity) against the true negative rate (specificity) for all possible thresholds of the model's predicted probabilities. The size of the area under the ROC (AUC-ROC) curve, also referred to as the AUC score, indicates the discriminatory power of the model. AUC score value of 0.5 suggests the model is random guessing, with no predictive power. The closer the AUC score gets to a value of 1, the better the model is in the outcome prediction. An AUC score above 0.7 indicates that the model has reasonably good discriminatory power. The major drawback of AUC is that it ignores the class imbalance. Therefore, even a model unable to detect any defaults can achieve a high AUC score as the number of correctly identified negative events is prevalent. Models should be able to reach as high a magnitude of AUC as possible but it is important to look at rates of true positives and false positives separately.

Given this thesis employs logistic regression, three new hypotheses are formulated, as we uncovered new potential academic contributions that can be drawn from the analysis:

1. Inclusion of the client's carbon footprint increases the performance of the model for default prediction.

2. Inclusion of the client's financed carbon footprint increases the performance of the model for default prediction.

3. Inclusion of the client's carbon intensity increases the performance of the model for default prediction.

Chapter 5

Data

This chapter discusses the data used in the empirical analysis. First, the original dataset from the Bank and the adjustments made are introduced in Section 5.1. Second, descriptive statistics of the adjusted data are presented in Section 3.2. Finally, the correlation matrix of all variables is interpreted in Section 5.3.

5.1 Dataset

We were provided with unique data from the Bank consisting of three parts. First, the dataset of all active accounts on corporate loans in years 2022 and 2023, including the information on defaults in 2023. Secondly, a list of all active corporate clients and their financial information spanning from 2017 to 2022. Finally, a dataset consisting of total and loan-financed carbon footprint information on all corporate clients. One corporate client could have several open accounts on different loan products. The Bank stated that if a client defaults on one of his accounts, all his other accounts are flagged as defaulting too. Therefore, any information on the loan type does not contribute to further analysis as it is impossible to confidently assume which account of the particular client initiated the default.

The initial dataset consisting of 2,329 observations on 69 variables was created by merging data from three datasets based on the client ID. We keep only the financial data for 2022, due to the amount of missing values in earlier years. We add new predictor variables derived from the provided data. Finally, we are left with an initial dataset of 2,329 observations on 39 variables. Defaults represent 31 observations and non-defaults represent 2,289 observations.

5.1.1 Adjustments of the dataset

There are three major issues with the initial dataset, which required several adjustments to proceed with the analysis. First, the amount of missing values in the subset of defaulting clients. Second, a large number of extreme values. Finally, a standard challenge in credit risk research is a class imbalance, i.e., the count of default observations being very small (Jorion *et al.* 2010). The issues were handled using standard approaches used in empirical research.

Data imputation

To address the problem with the missing values, we opted for the data imputation technique consisting of the replacement of the missing observations with the value of selected sample statistics. Data imputation has to be done with caution as the original distribution of data may become altered due to excessive usage. Our main reason for data imputation is to prevent further loss of default observations, therefore, we impute only those. Missing values of non-default observations are omitted. As sample statistics, we select the sample median since the outliers do not have such an effect on it as they would on the sample mean (Jorion *et al.* 2010). We recorded 17 out of 31 default observations that lack financial data from 2022, mainly caused by the shifted fiscal years of these clients, hence the audited financial statements were not available at the time of data collection. Part of the missing financial data from 2022 can be replaced by data from 2021. For the remaining 6 observations, we impute the medians of the sample for every missing observation.

Data partition

Data have to be partitioned into the training set for model preparation and the testing set for model validation. We split the data proportionally, 55:45 in favor of the training set. We check whether the shape of the distribution of the sample and the defaulting subgroup corresponds between the training and testing sets. Random partition is repeated until the most corresponding distributions are achieved, although it is nearly impossible to split a default sample of 31 observations into two identical distributions as every single instance in such a small subset influences a distribution shape.

Winsorization

Winsorization, a technique to handle extreme values, consists of computing the upper and lower percentile of the data distribution and replacing all values that exceed the upper and lower threshold by the value of the upper and lower percentile respectively. This prevents further loss of data while conserving the original distribution. It has to be applied separately on the training and testing set to prevent data leakage from the testing set. The percentile threshold values are derived only from the training set. The extreme values above the upper and below the lower thresholds are replaced by the percentile values. The threshold values derived from the training set are also used to replace extreme values in the testing set. We set the lower percentile to be equal to 1 % of the distribution and the upper percentile to be at 99 % of the distribution (Leone *et al.* 2019).

Oversampling

Oversampling is a data augmentation technique used to address class imbalance. If one class is significantly underrepresented compared to another class, logistic regression model may be biased towards the majority class. Oversampling is applied only to the training data to prevent data leakage from the testing set. The classes should be balanced based on the natural proportions of the classes. Before adjustment, defaults represent 2.25 % of the training set. We set the desired proportion of defaults to 10 %, which is in line with the reality as the non-performing loan ratios in Europe were generally ranging from around 2 % to over 10 % in 2022 (Huljak *et al.* 2022).

We opt for the SMOTE algorithm (Synthetic Minority Over-sampling Technique), which generates synthetic samples in the minority class to balance the dataset. For each minority class observation, the algorithm finds its k-nearest neighbors given the variable. SMOTE randomly selects one of these neighbors and creates a synthetic example at a randomly selected point between the example and its chosen neighbor. This process continues until the desired balance between the minority and majority classes is achieved (Douzas *et al.* 2018). Given the small number of original observations, the original and oversampled distributions are not perfectly corresponding but the oversampled distribution of defaults mirrors the original data. Table 5.1 shows the number of observations after the adjustments were made.

Table 5.1: Observation count after adjustments

Data	0	1
Original data	2.289	31
Imputation		
Imputed data	1.375	31
Partition		
Train data	757	17
Test data	618	14
Winsorization		
Train data	757	17
Test data	618	14
Oversampling		
Train data	757	85
Test data	618	14

5.2 Descriptive statistics

This section contains a descriptive analysis of the adjusted data, as described in the previous section. Since the data was partitioned into the training and testing set so that the distributions of all predictors correspond, this section features descriptive statistics of the merged testing and training set prior to the oversampling, assuming both sub-samples carry similar patterns.

The default flag from the year 2023 serves as the dependent variable of the analysis. There are two qualitative variables in the dataset, the NACE code of an industry sector and the postal code of the headquarters of a client. The possible predictors that are part of the dataset can be classified into two main categories: transition risk variables and financial predictors.

Several important insights can already be taken from the first look at the data. Most of the clients in the dataset are headquartered in Prague and Northern Bohemia. The regions with the highest proportion of defaulting clients are Central Bohemia followed by Northern and Eastern Bohemia as featured in Table 5.2. The dataset also contains a small group of corporate clients headquartered in Slovakia, whose postal code begins with the number 8 or 9.

Table 5.2: Regional classification of clients

Regional	**N. of observations**	**Proportion of defaults**
Prague	364	2.20 %
Northern Moravia	240	2.08 %
Eastern Bohemia	214	2.34 %
Southern and Western Bohemia	180	2.22 %
Central Bohemia	163	3.07 %
Southern Moravia	160	1.25 %
Northern Bohemia	80	2.50 %
Slovakia	5	0.00 %

The NACE (Nomenclature of Economic Activities) code is a four-digit code that indicates business activity and is primarily used within the EU. The first digit indicates a high-level industry sector of a firm. Table 5.3 features the number of clients in the dataset per industry sector and the proportion of defaults in that sector. Table 5.4 features the average amount of emissions per industry sector observed in the dataset. Most of the sample belongs to category (4), characterized by Scope 3 emissions, representing roughly 75 % of the total average emissions. Category (2), second in default proportion, has around 80 % of the total average emissions reported as Scope 3.

The most polluting categories in terms of the total average emission are the categories (3) and (2). The least polluting industry sector is the category (6), the financial sector. The largest Scope 1 (direct, from production) emitters belong to categories (3) and (1). The largest Scope 2 emitters (indirect, from production) belong to categories (2) and (8). The largest Scope 3 emitters (indirect, from the supply chain) are part of the categories (2) and (3). Scope 3 emissions represent a majority of the average emissions across all categories, and Scope 1 emissions are usually the second.

Table 5.3: Industry sector of clients

Industry sector	**N. of observations**	**Proportion of defaults**
(1) Agriculture, forestry and fishing	241	2.90 %
(2) Manufacturing, mining and quarrying	250	3.60 %
(3) Construction	90	1.11 %
(4) Wholesale, retail, transportation, accommodation and food services	650	1.69 %
(5) Information and communication	29	0.00 %
(6) Finance and insurance	73	2.74 %
(7) Real estate	55	0.00 %
(8) Professional, scientific, technical and administration services	24	4.17 %
(9) Public administration, defense, education, human health and social work	24	0.00 %

Table 5.4: Average carbon footprint per industry sector (mil.m^3 $C0_2$)

Industry sector	**Scope 1**	**Scope 2**	**Scope 3**
(1) Agriculture, forestry and fishing	2,265	190	3,148
(2) Manufacturing, mining and quarrying	907	395	5,269
(3) Construction	3,271	209	3,684
(4) Wholesale, retail, transportation, accommodation and food services	573	43	1,836
(5) Information and communication	1,094	162	3,338
(6) Finance and insurance	383	77	785
(7) Real estate	523	61	1,622
(8) Professional, scientific, technical and administration services	606	265	2,893
(9) Public administration, defense, education, human health and social work	119	119	940

The banking sector is exposed to climate transition risk via its interest-earning assets, i.e., via the emissions financed through its loans. The financed carbon footprint of a client is important for the Bank in the context of climate policy. Table 5.4 illustrates that categories (2) and (3) have the highest average financed carbon footprint just as it is in the case of the average carbon footprint. Categories (9) and (7) are financing the least emissions through their loans.

Table 5.5: Average financed carbon footprint per industry sector (mil.m^3 $C0_2$)

Industry sector	Scope 1,2	Scope 3
(1) Agriculture, forestry and fishing	325	497
(2) Manufacturing, mining and quarrying	183	683
(3) Construction	380	417
(4) Wholesale, retail, transportation, accommodation and food services	98	272
(5) Information and communication	149	354
(6) Finance and insurance	300	273
(7) Real estate	72	204
(8) Professional, scientific, technical and administration services	120	275
(9) Public administration, defense, education, human health and social work	39	144

5.2.1 Transition risk variables

Table 5.6 features the descriptive statistics of the total and financed carbon footprint. Separate financed carbon footprint for Scope 1 and 2 could not be provided by the Bank. The largest proportion of emissions is attributed to Scope 3 for both total and financed carbon footprint. Scope 2 carbon emissions represent the smallest portion of the emissions in the sample. Distributions of all variables tend to be right-skewed, as the medians are significantly lower than the means.

Table 5.6: Descriptive statistics of carbon footprint (ths. m^3 $C0_2$)

Statistic	Mean	St. Dev.	Min	Median	Max
			Carbon footprint		
Scope 1	1,087,094	2,277,454	6,731	330,170	16,342,336
Scope 2	151,978	306,767	19,881	41,067	1,851,534
Scope 3	2,760,649	3,870,008	13,200	1,309,172	22,101,437
			Financed carbon footprint		
Scope 1 and 2	180,095	377,352	0	52,030	2,285,294
Scope 3	389,810	652,077	0	157,402	3,935,159

Looking at the distributions of financed carbon footprint for defaulting and non-defaulting subsets, one can notice a median of defaulters being significantly higher than that of non-defaulters, both in the case of the Scope 1,2 financed footprint in Figure 5.1 and of Scope 3 financed footprint in Figure 5.2. No other carbon footprint variable had a distinctively higher median of the distribution of defaulters.

Figure 5.1: Distribution of Scope 1 and 2 financed carbon footprint

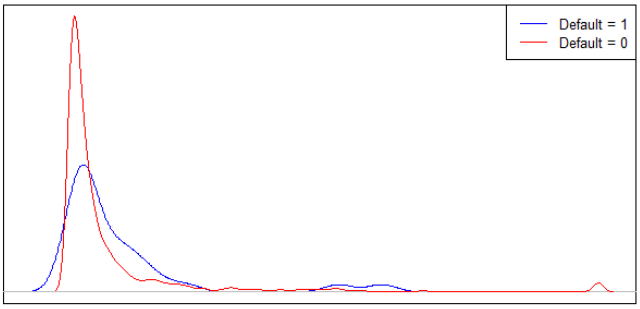
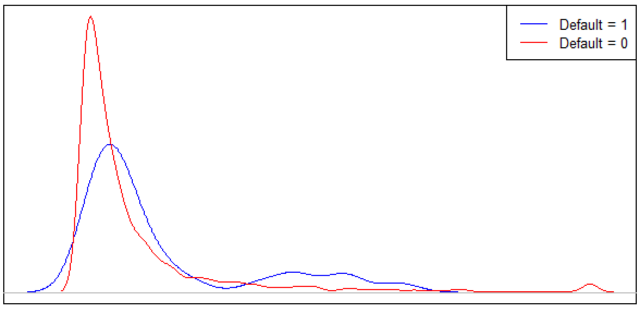


Figure 5.2: Distribution of Scope 3 financed carbon footprint



There are two types of variables representing the carbon intensity in this dataset. Carbon intensity per unit of sales, calculated as a fraction of the total carbon footprint and the gross sales of a particular client, describes how much of the emissions are linked to each unit of sales. The carbon intensity of assets is calculated as a fraction of the total carbon footprint and total assets. Both types of carbon intensity account for the size bias that might occur in the case of the total or financed carbon footprint, as bigger and more profitable firms produce more emissions, but usually tend to default less. Looking at the descriptive statistics featured in Table 5.7, carbon intensity per earning asset tends to be higher than per unit of sales. Distributions are generally right-skewed, as the mean values are significantly higher than the medians.

Table 5.7: Descriptive statistics of carbon intensity (m^3 $C0_2$/CZK)

Statistic	Mean	St. Dev.	Min	Median	Max
		per unit of Sales			
Scope 1	5.357	9.248	0.142	1.834	50.519
Scope 2	0.588	0.779	0.0003	0.383	4.768
Scope 3	10.446	6.850	0.263	10.138	44.247
		per unit of Assets			
Scope 1	6.265	11.151	0.031	3.369	82.942
Scope 2	0.776	1.039	0.0003	0.367	5.685
Scope 3	16.724	14.798	0.090	12.760	79.206

A similar trend in distributions of the Scope 2 and 3 carbon intensities can be observed. The median of the distribution of defaulters is distinctively pointed to the right in the case of both types of carbon intensities as seen in Figure 5.5, Figure 5.4, Figure 5.5 and Figure 5.6. This might suggest that Scope 2 and 3 carbon emissions might be the potential drivers of a credit default.

Figure 5.3: Distribution of Scope 2 carbon intensity per sales unit

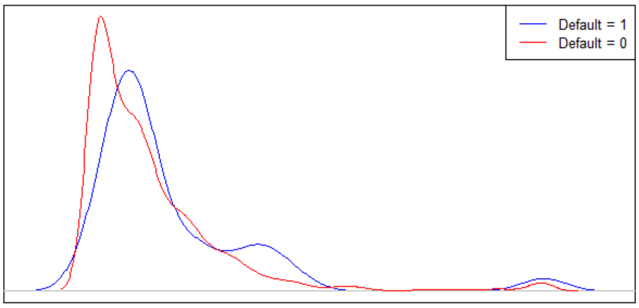


Figure 5.4: Distribution of Scope 3 carbon intensity per sales unit

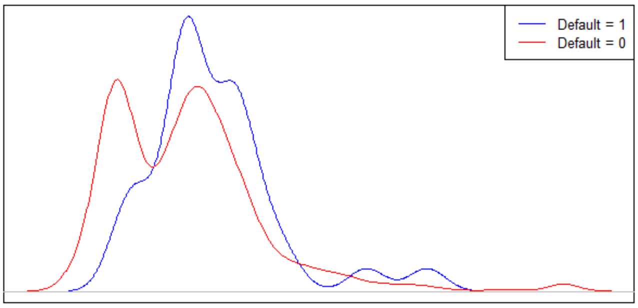


Figure 5.5: Distribution of Scope 2 carbon intensity per asset unit

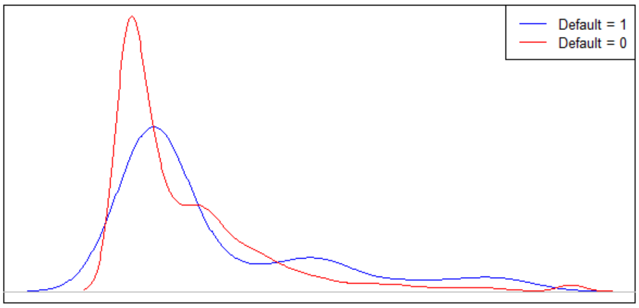
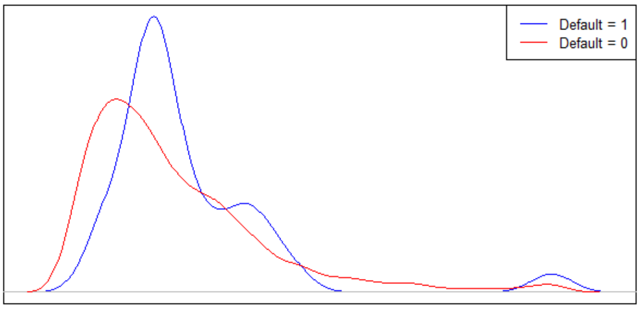


Figure 5.6: Distribution of Scope 3 carbon intensity per asset unit



5.2.2 Financial predictors

A set of 13 financial variables was provided by the Bank: Gross Sales, Total Turnover, EBITDA, Pre-tax Earnings, Non-current Assets, Current Assets, Total Assets, Net Profit, Total Debt, Equity, Total Liabilities, Cash, and Operating Cash Flow. Current Liabilities are derived as a difference between Total Liabilities and Non-current Liabilities. A set of 10 financial predictors, commonly used in the credit scoring models, was derived from original financial data (Jakubík *et al.* 2011). Table 5.8 shows the list of the derived financial predictors, their definitions, and optimal values. Ratios are organized into four categories, based on what aspect of the firm's financial health they account for. Table 5.9 features the descriptive statistics of the financial predictors.

Comparing the median or mean values to the optimal values, we observe that these values roughly correspond in the case of most of the ratios. Regarding the liquidity measures, the median of the Current ratio suggests the sample firms are liquid, while the Operating Cash Flow ratio median is quite below the liquidity threshold. Profitability-wise, we observe the sample to be slightly under-performing. Solvency and economic structure indicators are in line with the expectations.

Table 5.8: List of derived financial predictors

Predictor	Definition	Heathy value
Liquidity indicators		
Current ratio	$\frac{Current\ Assets}{Current\ Liabilities}$	1.5 or more
Cash Flow ratio	$\frac{Operating\ Cash\ Flow}{Current\ Liabilities}$	1 or more
Profitability indicators		
Return on Assets	$\frac{Net\ profit}{Total\ Assets}$	5 % or more
Return on Equity	$\frac{Net\ profit}{Total\ Equity}$	15 % or more
Profit margin	$\frac{Net\ profit}{Gross\ Sales}$	10 % or more
Cash Flow margin	$\frac{Operating\ Cash\ Flow}{Gross\ Sales}$	10 % or more
Solvency indicators		
Debt-to-Equity	$\frac{Total\ Debt}{Equity}$	0.5 or less
Debt-to-Assets	$\frac{Total\ Debt}{Total\ Assets}$	0.6 or less
Equity ratio	$\frac{Total\ Equity}{Total\ Assets}$	0.4 or more
Economic structure indicators		
Debt-to-Capital	$\frac{Total\ Debt}{Total\ Debt + Equity}$	0.4 or less

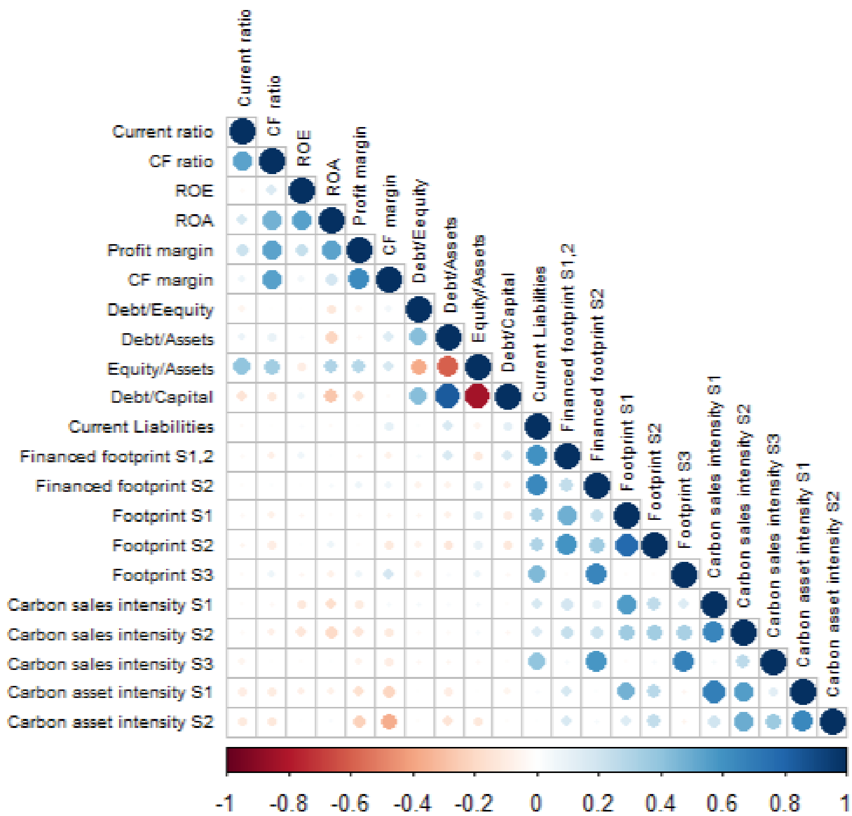
Table 5.9: Descriptive statistics of financial predictors

Statistic	Mean	St. Dev.	Min	Median	Max
Current ratio	3.004	2.455	0.185	2.275	12.955
Cash Flow ratio	0.760	0.950	−0.508	0.458	5.181
ROE	0.159	0.336	−1.147	0.122	1.727
ROA	0.057	0.084	−0.182	0.043	0.374
Profit margin	0.049	0.106	−0.400	0.031	0.564
Cash Flow margin	0.127	0.173	−0.164	0.076	0.886
Debt to Equity	1.456	3.375	−11.041	0.644	21.547
Debt to Assets	0.309	0.197	0.002	0.280	0.835
Equity ratio	0.403	0.226	−0.087	0.403	0.873
Debt to Capital	0.442	0.268	0.001	0.41	1.205

5.3 Correlation matrix

Figure 5.7 features the correlation matrix of all variables included in the dataset. Transition risk variables showcase light to moderate positive correlation between each other but they are not to be included in one model together. There is a moderate to strong correlation between the Equity ratio, Debt-to-Capital, Debt-to-Assets, and Debt-to-Capital, signaling potential multicollinearity in the model if these variables are included together. The correlation coefficient between the Equity ratio and the Debt-to-Capital is -0.836, for the Equity ratio and Debt-to-Assets is -0.599, and for Debt-to-Capital and Debt-to-Assets is 0.838. Debt-to-Equity is not strongly correlated with any of these predictors. There is a moderate correlation between Profit margin, Operating Cash Flow margin, and Operating Cash Flow ratio. There is a moderate correlation between Current Liabilities and Scope 3 financed carbon footprint as the correlation coefficient is equal to 0.642.

Figure 5.7: Correlation matrix



Chapter 6

Model

The last step in the preparation for the empirical analysis is to construct a standard model for the probability of default estimation. The standard model will be climate-stressed in Chapter 7 by transition risk variables to observe their effects. Therefore, the model should capture as much endogeneity in the data as possible and be able to correctly identify a default.

6.1 Model specification

A credit scoring model for corporate loans should be built on four financial dimensions of a firm (Bartual Sanfeliu *et al.* 2012). Liquidity, to capture the short-term financial viability of a firm. Solvency, to capture the long-term financial health of a firm. Profitability, to assess the ability of a firm to generate earnings relative to underlying factors and economic structure, to provide insight into the capital structure of a company. A set of 10 financial ratios and 4 financial inputs were selected as potential predictors in the model based on the literature on credit scoring models (Jakubík *et al.* 2011). Values of the 4 financial inputs were standardized, to avoid bias caused by different scales of variables.

All potential predictors were estimated in a simple model to examine the significance and the sign of the estimator. Table 6.1 features the estimators of the simple models. The effects of predictors, i.e., signs of estimators, are in line with the expectations of the variable effect on the odds of the default. Debt estimator has a negative sign, suggesting more debt lowers the chances of default, however, this estimate is of low weight and has no statistical significance.

Current Liabilities estimate is negative and significant, as increased value implies fewer long-term debt, i.e., lower interest expense. Operating Cash Flow margin and Operating Cash Flow ratio are both strongly significant and gain the lowest value of the Akaike Information Criterion, i.e., these variables might represent the best fit for the data.

Table 6.1: Simple models

Predictor	Coefficient	AIC	Significance
Operaing CF ratio	-2.575	477	***
Operating CF margin	-11.753	484	***
Equity ratio	-3.358	518	***
ROA	-9.177	522	***
ROE	-1.607	535	***
Profit margin	-4.494	539	***
Equity	-0.596	544	**
Current ratio	-0.199	545	**
Debt-to-Assets	1.638	547	**
Current Liabilities	-0.404	550	*
Assets	-0.385	550	*
Debt-to-Capital	0.769	552	
Debt-to-Equity	0.018	555	
Debt	-0.043	555	

The model was trained by incremental feature selection, i.e., by progressive adding of features one by one and observing their effect on the goodness of model fit and model performance. Two models were selected as the potential standard models for climate-stressing. Table 6.2 features the summary of the two models. The simpler model (1) consists of all pillars except for the solvency ratios which were not significant in any model. Model (2) is a more complex version of the first model containing a solvency indicator. Table 6.3 indicates there is no significant multicollinearity in the selected models as the values of VIF are well below 5. There are only minor differences between the estimates of the two models, and the levels of significance remained similar. Operating Cash Flow margin has the most profound effect on the probability of default, yet its statistical significance is the weakest.

Table 6.2: Summary of the logistic regression

	Dependent variable:	
	flag_default	
	(1)	(2)
Operating_CF_margin	-4.551^{*}	-4.082^{*}
	(2.504)	(2.428)
Operating_CF_ratio	-1.561^{***}	-1.686^{***}
	(0.561)	(0.561)
Current_liabilities	-0.630^{**}	-0.590^{**}
	(0.271)	(0.269)
Equity_ratio	-2.225^{***}	-1.893^{***}
	(0.621)	(0.696)
Debt_to_Assets		0.809
		(0.713)
Constant	-0.796^{***}	-1.150^{***}
	(0.215)	(0.385)
Observations	842	842
Log Likelihood	−221.914	−221.266
Akaike Inf. Crit.	453.828	454.532
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 6.3: Variance inflation factors

Model	CF margin	CF ratio	Equity ratio	Current liab	D/A
(1)	2.429	2.478	1.046	1.022	
(2)	2.411	2.480	1.272	1.039	1.260

Moving on to the evaluation of model performance, Table 6.4 features the number of correct and false predictions of the models on the testing set as well as the specificity and sensitivity rates. As expected, the severe class imbalance caused the model estimates to be biased toward the majority class as no defaults were correctly indicated using a threshold of 0.5. The defaults originally represented only 2.2 % of the training sample and the proportion was synthetically balanced to 10 % as discussed in Chapter 5.

Table 6.4: Performance evaluation

Model	TN	FN	TP	FP	Specificity	Sensitivity
Model (1)	614	14	0	4	0.9935	0
Model (2)	615	14	0	3	0.9951	0

To show the predictive power of the model not influenced by the class imbalance bias, we replicate all the default observations in the training set eight times to obtain a more balanced panel. The number of default observation is equal to 680 while the number of non-defaults is 757. The models are estimated using balanced training data as presented in Table 6.5. There is no major change in the values of the estimators, all the estimators are now of strong statistical significance. The value of the constant increased significantly as by balancing the classes, the occurrence of default became more frequent and so did the odds of a default happening.

Table 6.5: Summary of the logistic regression (balanced data)

	Dependent variable:	
	flag_default	
	(1)	(2)
Operating_CF_margin	-4.485^{***}	-4.201^{***}
	(1.177)	(1.142)
Operating_CF_ratio	-1.403^{***}	-1.621^{***}
	(0.251)	(0.256)
Current_liabilities	-0.811^{***}	-0.780^{***}
	(0.149)	(0.153)
Equity_ratio	-2.774^{***}	-2.108^{***}
	(0.365)	(0.381)
Debt_to_Assets		1.912^{***}
		(0.421)
Constant	1.379^{***}	0.566^{***}
	(0.132)	(0.213)
Observations	1,437	1,437
Log Likelihood	−740.828	−730.510
Akaike Inf. Crit.	1,491.656	1,473.019
Note:	*p<0.1; **p<0.05; ***p<0.01	

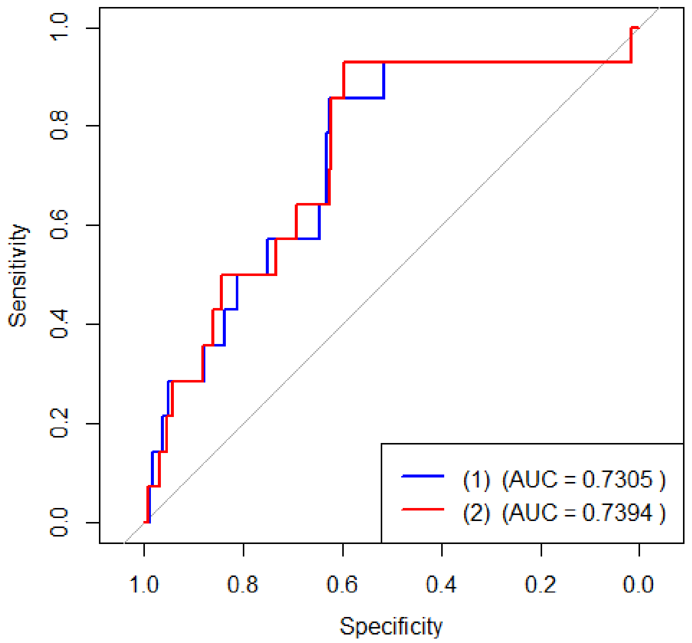
Looking at the performance of the models estimated on balanced data, Table 6.6 indicates the model (1) is better at identifying defaults than model

(2) since the number of true positive predictions is higher. The magnitude of the AUC, shown in Figure 6.1, is almost equal for the two models indicating both models have moderately good discriminatory power. Model (1) is slightly more sensitive, indicating it might be better in avoiding Type 2 errors, i.e., failing to predict a default.

Table 6.6: Performance evaluation (balanced data)

Model	TN	FN	TP	FP	Specificity	Sensitivity
Model (1)	456	6	8	162	0.738	0.571
Model (2)	455	7	7	163	0.736	0.500

Figure 6.1: ROC curve (balanced data)



6.2 Robustness check

The robustness of the models was assessed by applying the model on data adjusted by alternative parameters from those described in Subsection 5.1.1. Data was randomly partitioned at a ratio of 60:40, winosorization percentiles were set to 99.5 and 0.5 % for the upper and lower thresholds respectively. Defaults in the training set were oversampled so that they represent 8 % of the training sample. The distributions of the key predictors corresponded across training and testing samples. Training data were balanced in the similar manner as in the previous section.

Table 6.8 features the models (1) and (2) applied to the alternative data. We observe similar signs of the estimators with several differences such as the Operating Cash Flow margin not being significant in both models. The differences are caused by different random data partitions of defaults, which represent a very small subset of the data impossible to split into two perfectly corresponding partitions. The models estimated on imbalanced alternative data rendered similar results as featured in Appendix A. There is no sign of multicollinearity as featured in Table 6.7.

Table 6.7: Variance inflation factors (Robustness check)

Model	CF margin	CF ratio	Equity ratio	Current liab	D/A
(1)	1.572	1.704	1.065	1.046	
(2)	2.911	3.258	1.599	1.130	1.806

Table 6.8: Summary of the logistic regression (Robustness check)

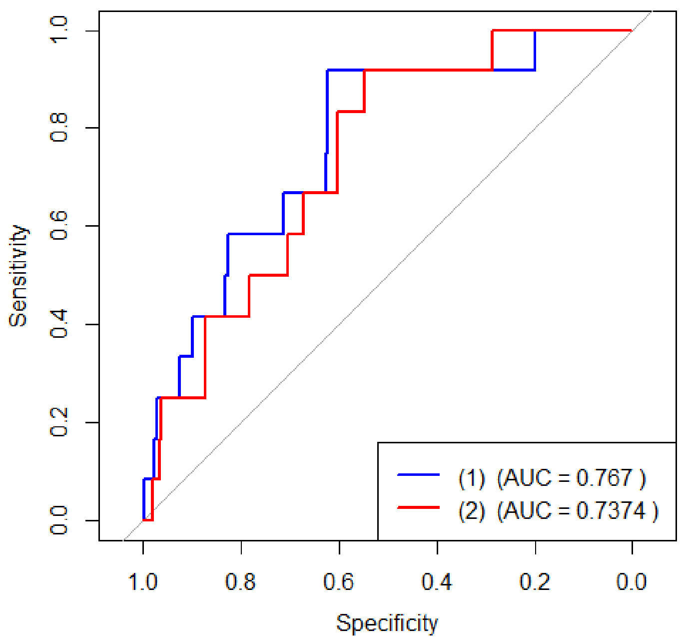
	Dependent variable:	
	flag_default	
	(1)	(2)
Operating_CF_margin	−0.508	−0.743
	(0.580)	(0.607)
Operating_CF_ratio	−0.483***	−0.697***
	(0.117)	(0.129)
Current_liabilities	−1.879***	−1.756***
	(0.212)	(0.217)
Equity_ratio	−3.098***	−1.141***
	(0.313)	(0.389)
Debt_to_Assets		3.444***
		(0.448)
Constant	0.860***	−0.912***
	(0.130)	(0.257)
Observations	1,585	1,585
Log Likelihood	−934.029	−903.462
Akaike Inf. Crit.	1,878.058	1,818.923
Note:	*p<0.1; **p<0.05; ***p<0.01	

Concerning the model performance Table 6.9 shows the two models are equally good in their ability to correctly identify a default given threshold of 0.5. In general, the AUC scores correspond with the baseline models, model (1) achieves a slightly higher magnitude of the AUC-ROC.

Table 6.9: Performance evaluation (Robustness check)

Model	TN	FN	TP	FP	Specificity	Sensitivity
Model (1)	358	4	8	192	0.651	0.667
Model (2)	368	4	8	182	0.669	0.667

Figure 6.2: ROC curve (Robustness check)



To conclude, the models were estimated on two different random partitions of data returning similar estimators and achieving similar effectivity in default detection. Model (1) is chosen to be climate-stressed in Chapter 7, since it managed to correctly predict more defaults and all of the variables are of statistical significance.

Chapter 7

Empirical analysis

The chapter is structured based on the respective hypotheses set in Chapter 3 and Chapter 4. The hypotheses on the carbon intensity variable will be split into two separate hypotheses, one for carbon intensity per unit of sales and one for carbon intensity per asset. The results are summarized in Section 7.9 and further research opportunities are presented in Section 7.10. The final list of the hypotheses is following:

1. Client's carbon footprint increases the odds of default on a corporate loan in the Bank.

2. Inclusion of the client's carbon footprint increases the performance of the model for default prediction.

3. Client's financed carbon footprint increases the odds of default on a corporate loan in the Bank.

4. Inclusion of the client's financed carbon footprint increases the performance of the model for default prediction.

5. Client's carbon intensity per unit of sales increases the odds of default on a corporate loan in the Bank.

6. Inclusion of the client's carbon intensity per unit of sales increases the performance of the model for default prediction.

7. Client's carbon intensity per earning asset increases the odds of default on a corporate loan in the Bank.

8. Inclusion of the client's carbon intensity per earning asset increases the performance of the model for default prediction.

To test the hypotheses, the standard model featured in Chapter 6 was climate-stressed separately by all types of transition risk variables. Values of carbon footprint and financed carbon footprint were standardized, to avoid bias induced by different scales of variables. For the robustness check, the analysis was repeated using the alternative data, also used for the standard model robustness check in Chapter 6. This Chapter features the models estimated on the balanced data so that their ability of correctly estimate defaults can be compared. Models estimated on imbalanced data generated similar values and significance of parameters as shown in Appendix A.

7.1 Hypothesis 1 testing

Table 7.2 features the summary of the standard and the climate-stressed models. Values of VIF are below the multicollinearity threshold for all models as shown in Table 7.1. Transition risk variables used for the Hypothesis 1 testing are the three scopes of carbon footprint further referred to as Scope 1,2 and 3. The standard model estimators were not changed significantly by adding a transition risk variable. Scope 1 carbon footprint estimator is not statistically significant. Scope 2 and 3 carbon footprint estimators are statistically significant in the model. An increase in Scope 2 carbon footprint results in an approximately 212 % increase in the odds of default. One unit increase in Scope 3 carbon footprint results in a roughly 276 % increase in the probability of default. On imbalanced data, the effects were a bit lower yet significant, i.e., 142% and 132% respectively.

Table 7.1: Variance inflation factors (H1)

Model	CF margin	CF ratio	Equity ratio	Current liab	Emission var
(1)	2.429	2.478	1.046	1.022	
(2)	2.438	2.479	1.056	1.098	1.095
(3)	2.782	2.859	1.174	1.639	1.875
(4)	2.563	2.653	1.122	1.824	1.969

Table 7.2: Summary of the logistic regression (H1)

	Dependent variable:			
	flag_default			
	(1)	(2)	(3)	(4)
Operating_CF_margin	-4.485^{***}	-4.544^{***}	-7.315^{***}	-3.949^{***}
	(1.177)	(1.188)	(1.579)	(1.332)
Operating_CF_ratio	-1.403^{***}	-1.399^{***}	-1.496^{***}	-1.819^{***}
	(0.251)	(0.251)	(0.316)	(0.289)
Equity_ratio	-2.774^{***}	-2.815^{***}	-4.183^{***}	-3.716^{***}
	(0.365)	(0.372)	(0.431)	(0.403)
Current_liabilities	-0.811^{***}	-0.836^{***}	-2.447^{***}	-2.441^{***}
	(0.149)	(0.157)	(0.284)	(0.263)
Scope_1_TotEm		0.064		
		(0.108)		
Scope_2_TotEm			1.140^{***}	
			(0.106)	
Scope_3_TotEm				1.325^{***}
				(0.132)
Constant	1.379^{***}	1.398^{***}	1.384^{***}	1.416^{***}
	(0.132)	(0.136)	(0.150)	(0.144)
Observations	1,437	1,437	1,437	1,437
Log Likelihood	−740.828	−740.655	−621.879	−673.215
Akaike Inf. Crit.	1,491.656	1,493.310	1,255.757	1,358.430
Note:				*p<0.1; **p<0.05; ***p<0.01

7.1.1 Robustness check

Applying the climate-stressed models to the alternative data, the values and the significance of the Scope 2 and Scope 3 estimators correspond to the baseline results as presented in the Table 7.4. There is no major change in the values or significance standard model estimators nor in the values of VIF shown in Table 7.3 that would suggest multicollinearity.

Table 7.3: Variance inflation factors (Robustness check H1)

Model	CF margin	CF ratio	Equity ratio	Current liab	Emission var
(1)	1.572	1.704	1.066	1.046	
(2)	1.569	1.709	1.072	1.131	1.093
(3)	1.628	1.743	1.124	1.475	1.511
(4)	1.658	1.776	1.121	2.306	2.378

Table 7.4: Robustness check (H1)

	Dependent variable:			
	flag_default			
	(1)	(2)	(3)	(4)
Operating_CF_margin	−0.508	−0.511	−0.276	0.242
	(0.580)	(0.576)	(0.634)	(0.635)
Operating_CF_ratio	−0.483***	−0.468***	−0.543***	−0.674***
	(0.117)	(0.116)	(0.128)	(0.132)
Current_liabilities	−1.879***	−1.703***	−3.038***	−3.236***
	(0.212)	(0.215)	(0.276)	(0.296)
Equity_ratio	−3.098***	−2.988***	−3.701***	−3.392***
	(0.313)	(0.315)	(0.333)	(0.325)
Scope_1_TotEm		−0.373***		
		(0.136)		
Scope_2_TotEm			1.066***	
			(0.110)	
Scope_3_TotEm				0.988***
				(0.114)
Constant	0.860***	0.791***	0.714***	0.707***
	(0.130)	(0.132)	(0.139)	(0.138)
Observations	1,585	1,585	1,585	1,585
Log Likelihood	−934.029	−929.103	−860.976	−891.470
Akaike Inf. Crit.	1,878.058	1,870.207	1,733.953	1,794.941
Note:				*p<0.1; **p<0.05; ***p<0.01

The inclusion of Scope 2 and Scope 3 carbon footprint into the standard model improved its goodness to fit and generated robust positive and statistically significant estimators suggesting that higher levels of Scope 2 and Scope 3 carbon emissions might be a driver of defaults in the Bank. Therefore, Hypothesis 1 could not be rejected.

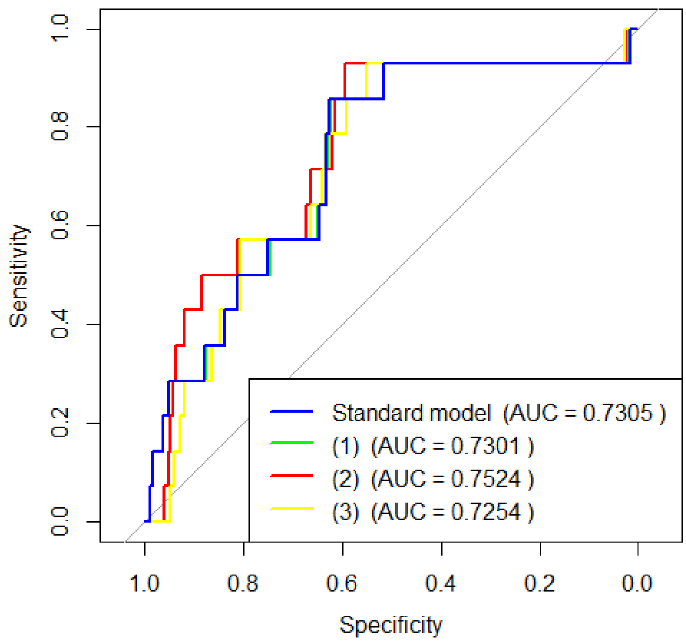
7.2 Hypothesis 2 testing

Standard and climate-stressed models are moderately good at ranking instances as presented in Figure 7.1. In terms of the magnitude of the AUC score, only the model (3) featuring Scope 2 emissions exceeded the standard model. As shown in Table 7.5 all models managed were equally good in default detection at a threshold equal to 0.5 but models (3) and (4) achieved a higher sensitivity score than the standard model.

Table 7.5: Performance evaluation (H2)

Model	TN	FN	TP	FP	Specificity	Sensitivity
Standard model	456	6	8	162	0.738	0.571
Model (2)	454	6	8	164	0.735	0.571
Model (3)	480	6	8	138	0.777	0.571
Model (4)	468	6	8	150	0.757	0.571

Figure 7.1: ROC curve (H2)



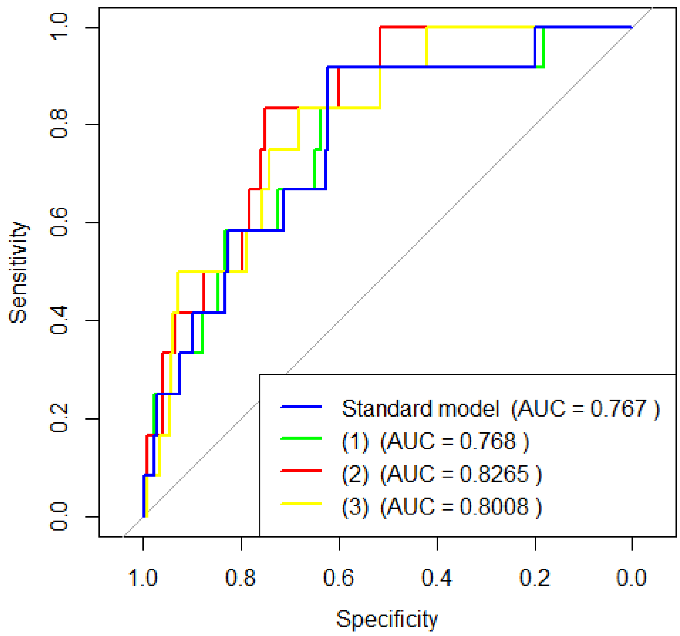
7.2.1 Robustness check

Looking at Figure 7.2 all climate-stressed models achieved a higher magnitude of the AUC and also performed better in terms of sensitivity given a threshold of 0.5 as shown in Table 7.6. Models (3) and (4) also achieved higher specificity rates than the standard model.

Table 7.6: Performance evaluation (Robustness check H2)

Model	TN	FN	TP	FP	Specificity	Sensitivity
Standard Model	358	4	8	192	0.651	0.667
Model (2)	353	3	9	197	0.642	0.750
Model (3)	374	2	10	176	0.680	0.830
Model (4)	370	2	10	180	0.673	0.833

Figure 7.2: ROC curve (Robustness check H2)



Model (3) stressed by the Scope 2 carbon footprint gained a higher magnitude of the AUC score on both data partitions. Moreover, it also repeatedly achieved higher sensitivity and specificity rates than the standard model. The inclusion of the Scope 2 carbon footprint can enhance the model's predictive power and therefore Hypothesis 2 can not be rejected.

7.3 Hypothesis 3 testing

Table 7.7 features the summary of the standard and the climate-stressed models. Values of VIF are below the multicollinearity threshold for all models as shown in Table 7.8. Transition risk variables used for the Hypothesis 1 testing are Scope 1,2 and Scope 3 financed carbon footprint. The standard model estimators were not changed significantly by adding a transition risk variable. Both Scope 1,2 and Scope 3 financed carbon footprint estimators are positive and statistically significant. One unit increase in Scope1,2 emissions results in an increase of 70 % in the odds of the default. The probability of default would be shifted by 177 % if the Scope 3 financed footprint increases by one unit. On imbalanced data, the effects were a bit lower yet significant, i.e., 49% and 133% respectively.

Table 7.7: Summary of the logistic regression (H3)

	Dependent variable:		
	flag_default		
	(1)	(2)	(3)
Operating_CF_margin	-4.485^{***}	-4.424^{***}	-2.360^{*}
	(1.177)	(1.217)	(1.219)
Operating_CF_ratio	-1.403^{***}	-1.530^{***}	-2.124^{***}
	(0.251)	(0.260)	(0.288)
Equity_ratio	-2.774^{***}	-2.855^{***}	-2.402^{***}
	(0.365)	(0.370)	(0.379)
Current_liabilities	-0.811^{***}	-1.021^{***}	-1.650^{***}
	(0.149)	(0.165)	(0.212)
Fin_Scope12		0.530^{***}	
		(0.089)	
Fin_Scope3			1.022^{***}
			(0.084)
Constant	1.379^{***}	1.372^{***}	0.857^{***}
	(0.132)	(0.134)	(0.141)
Observations	1,437	1,437	1,437
Log Likelihood	−740.828	−721.030	−635.961
Akaike Inf. Crit.	1,491.656	1,454.060	1,283.923
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 7.8: Variance inflation factors (H3)

Model	CF margin	CF ratio	Equity ratio	Current liab	Emission var
(1)	2.429	2.478	1.046	1.022	
(2)	2.377	2.441	1.049	1.191	1.188
(3)	2.397	2.562	1.043	1.422	1.560

7.3.1 Robustness check

Looking at the results of the robustness check featured in Table 7.9, both Scope 1,2 and Scope 3 financed carbon footprint estimators remained positive and statistically significant. The weight of both estimators decreased. An increase of one unit in Scope 1,2 financed footprint increases the odds of default by 34 % and Scope 3 financed footprint by 116 %. Multicollinearity is not present in the model as shown in Table 7.10.

Table 7.9: Robustness check (H3)

	Dependent variable:		
	flag_default		
	(1)	(2)	(3)
Operating_CF_margin	−0.508	−0.353	0.257
	(0.580)	(0.586)	(0.605)
Operating_CF_ratio	−0.483***	−0.514***	−0.605***
	(0.117)	(0.120)	(0.126)
Current_liabilities	−1.879***	−2.126***	−2.816***
	(0.212)	(0.233)	(0.277)
Equity_ratio	−3.098***	−3.070***	−2.532***
	(0.313)	(0.314)	(0.318)
Fin_Scope12		0.291***	
		(0.080)	
Fin_Scope3			0.773***
			(0.087)
Constant	0.860***	0.789***	0.296**
	(0.130)	(0.132)	(0.146)
Observations	1,585	1,585	1,585
Log Likelihood	−934.029	−927.327	−881.915
Akaike Inf. Crit.	1,878.058	1,866.654	1,775.829
Note:	*p<0.1; **p<0.05; ***p<0.01		

Table 7.10: Variance inflation factors (Robustness check H3)

Model	CF margin	CF ratio	Equity ratio	Current liab	Emission var
(1)	1.572	1.704	1.066	1.046	
(2)	1.564	1.722	1.065	1.231	1.183
(3)	1.601	1.737	1.071	1.530	1.502

Scope 1,2 and Scope 3 financed carbon footprints had positive and statistically significant estimators in the climate-stressed models, suggesting the level of loan-financed emissions increases the odds of default. Therefore, Hypothesis 3 could not be rejected.

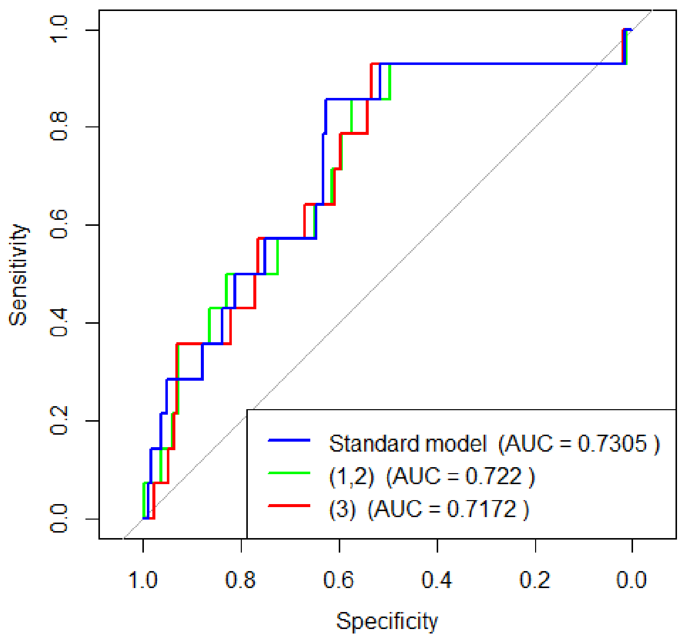
7.4 Hypothesis 4 testing

Standard and climate-stressed models perform moderately well in ranking instances, as presented in Figure 7.3. In terms of the magnitude of the AUC, none of the climate-stressed models exceeded the standard model. Given the threshold equal to 0.5, the models perform none of the climate-stressed models exceeded the standard model in terms of sensitivity rate as presented in Table 7.11.

Table 7.11: Performance evaluation (H4)

Model	TN	FN	TP	FP	Specificity	Sensitivity
Standard Model	456	6	8	162	0.738	0.571
Model (2)	457	7	7	161	0.739	0.500
Model (3)	488	8	6	130	0.790	0.429

Figure 7.3: ROC curve (H4)



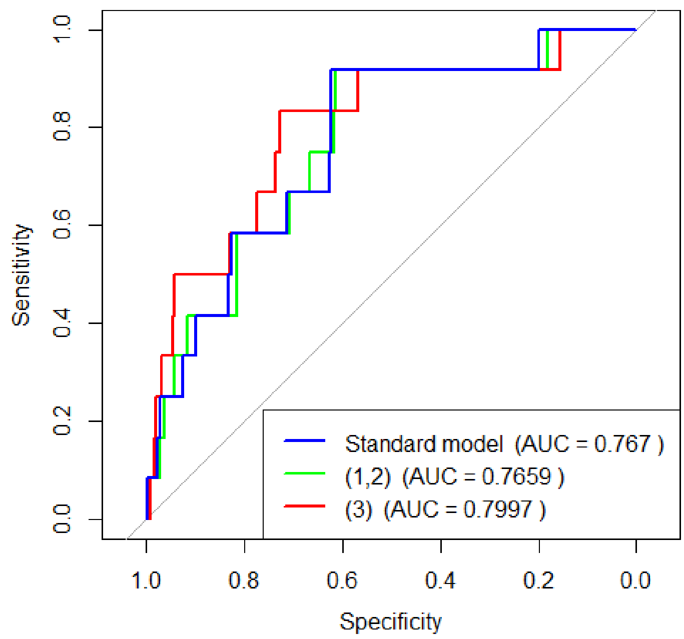
7.4.1 Robustness check

When assessing the performance of the climate-stressed models on the alternative data, the model stressed by the Scope 3 financed carbon footprint reaches a higher magnitude of AUC than the standard model as presented in Figure 7.4. Given a threshold equal to 0.5, both models predict more defaults correctly.

Table 7.12: Performance evaluation (Robustness check H4)

Model	TN	FN	TP	FP	Specificity	Sensitivity
Standard model	358	4	8	192	0.651	0.667
Model (2)	353	3	9	197	0.641	0.750
Model (3)	362	2	10	188	0.658	0.833

Figure 7.4: ROC curve (Robustness check H4)



Assessing the performance of the climate-stressed models on the two differently adjusted partitions of the data, the climate-stressed models did not prove to perform better in terms of default prediction than the standard model, therefore Hypothesis 4 can be rejected.

7.5 Hypothesis 5 testing

Table 7.13 features the summary of the standard and the climate-stressed models. Values of VIF are below the multicollinearity threshold for all models as shown in Table 7.14. Transition risk variables used for the Hypothesis 1 testing are Scope 1,2 and Scope 3 carbon intensities per unit of sales. The signs and significance of the standard model estimators remained similar in the standard and climate-stressed models. All transition risk variables are statistically significant. Scope 1 carbon intensity estimator is negative suggesting its increase result in a decrease in the probability of default by 8 %. Scope 2 carbon intensity increase translates to an upward shift of 133 % in odds of default and Scope 3 carbon intensity to a slight upward shift of 11 %. The effects correspond to those measured on imbalanced data.

Table 7.13: Summary of the logistic regression (H5)

	Dependent variable:			
	flag_default			
	(1)	(2)	(3)	(4)
Operating_CF_margin	-4.485^{***}	-2.920^{**}	-7.620^{***}	-6.783^{***}
	(1.177)	(1.137)	(1.434)	(1.387)
Operating_CF_ratio	-1.403^{***}	-1.673^{***}	-1.176^{***}	-1.188^{***}
	(0.251)	(0.255)	(0.274)	(0.264)
Equity_ratio	-2.774^{***}	-2.607^{***}	-3.363^{***}	-2.876^{***}
	(0.365)	(0.367)	(0.399)	(0.385)
Current_liabilities	-0.811^{***}	-0.820^{***}	-1.303^{***}	-0.973^{***}
	(0.149)	(0.148)	(0.211)	(0.175)
carbon_intensity_scope1_sales		-0.090^{***}		
		(0.021)		
carbon_intensity_scope2_sales			0.846^{***}	
			(0.089)	
carbon_intensity_scope3_sales				0.107^{***}
				(0.012)
Constant	1.379^{***}	1.580^{***}	0.900^{***}	0.146
	(0.132)	(0.141)	(0.152)	(0.189)
Observations	1,437	1,437	1,437	1,437
Log Likelihood	−740.828	−724.230	−681.648	−691.026
Akaike Inf. Crit.	1,491.656	1,460.459	1,375.296	1,394.052
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 7.14: Variance inflation factors (H5)

Model	CF margin	CF ratio	Equity ratio	Current liab	Emission var
(1)	2.429	2.478	1.046	1.022	
(2)	2.600	2.636	1.044	1.029	1.009
(3)	2.485	2.515	1.099	1.041	1.085
(4)	2.402	2.485	1.062	1.020	1.021

7.5.1 Robustness check

Table 7.16 features the summary of the logistic regression models applied to alternative data. There are no major changes in standard predictor estimators and their significance. Scope 2 and 3 carbon intensity estimators are both positive and statistically significant. These estimators narrowly correspond with the baseline results. There is no multicollinearity suggested as shown in Table 7.15.

Table 7.15: Variance inflation factors (Robustness check H5)

Model	CF margin	CF ratio	Equity ratio	Current liab	Emission var
(1)	1.572	1.704	1.066	1.046	
(2)	1.627	1.729	1.067	1.046	1.028
(3)	1.649	1.761	1.061	1.049	1.033
(4)	1.586	1.696	1.057	1.048	1.013

Table 7.16: Robustness check (H5)

	Dependent variable:			
	flag_default			
	(1)	(2)	(3)	(4)
Operating_CF_margin	−0.508	−0.352	−0.821	−0.348
	(0.580)	(0.588)	(0.631)	(0.636)
Operating_CF_ratio	−0.483***	−0.512***	−0.442***	−0.488***
	(0.117)	(0.120)	(0.125)	(0.127)
Current_liabilities	−1.879***	−1.903***	−1.963***	−1.752***
	(0.212)	(0.215)	(0.226)	(0.212)
Equity_ratio	−3.098***	−3.102***	−3.301***	−2.725***
	(0.313)	(0.316)	(0.326)	(0.318)
carbon_intensity_scope1_sales		−0.019***		
		(0.007)		
carbon_intensity_scope2_sales			0.967***	
			(0.096)	
carbon_intensity_scope3_sales				0.095***
				(0.010)
Constant	0.860***	0.946***	0.159	−0.421**
	(0.130)	(0.134)	(0.149)	(0.183)
Observations	1,585	1,585	1,585	1,585
Log Likelihood	−934.029	−926.096	−873.413	−880.906
Akaike Inf. Crit.	1,878.058	1,864.192	1,758.826	1,773.812
Note:				*p<0.1; **p<0.05; ***p<0.01

Two out of three variables on carbon intensity per unit of sales proved to have positive and statistically significant estimators in the climate-stressed models. The baseline results corresponded with the results from the robustness check analysis. Therefore, Hypothesis 5 could not be rejected.

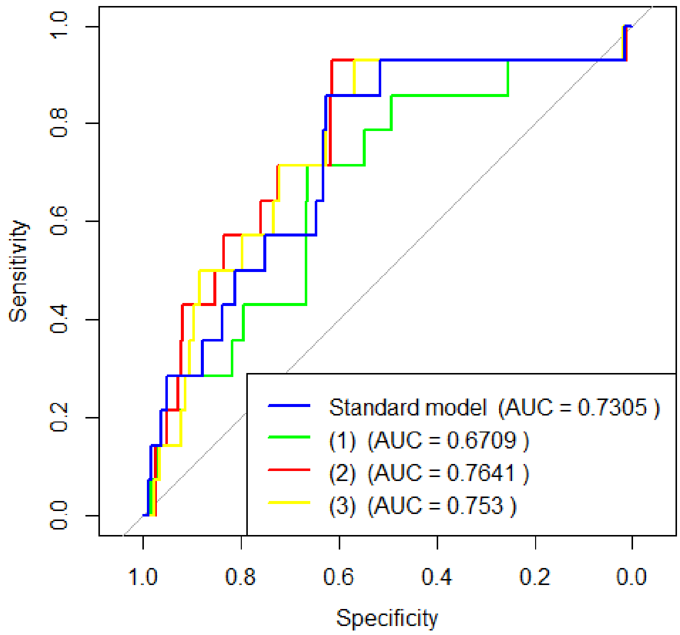
7.6 Hypothesis 6 testing

Standard and climate-stressed models perform moderately well in terms of ranking instances, as presented in Figure 7.5. The models climate-stressed by Scope 2 and Scope 3 carbon intensity per sales unit exceeded the standard model in terms of the AUC score. Given a threshold equal to 0.5, no model could exceed the standard model in terms of sensitivity, but models (2) and (3) were more specific as shown in 7.17.

Table 7.17: Performance evaluation (H6)

Model	TN	FN	TP	FP	Specificity	Sensitivity
Standard Model	456	6	8	162	0.738	0.571
Model (2)	463	8	6	155	0.749	0.429
Model (3)	471	6	8	147	0.762	0.571
Model (4)	471	6	8	147	0.762	0.571

Figure 7.5: ROC curve (H6)



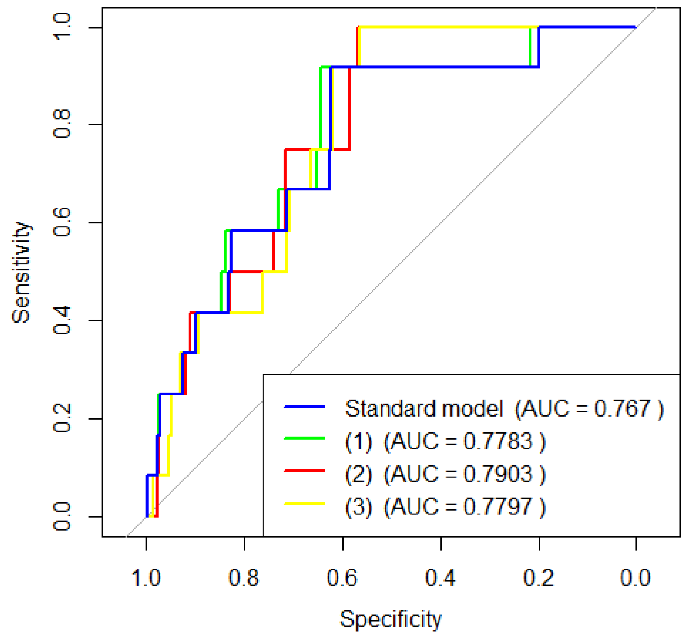
7.6.1 Robustness check

When assessing the performance of the climate-stressed models on the alternative data, all three models perform better than the standard model in terms of the magnitude of the AUC as presented in 7.6. Table 7.17 shows all models managed to correctly predict the same or higher number of defaults than the standard model given the threshold of 0.5. Models (3) and (4) achieved higher specificity than the standard model.

Table 7.18: Performance evaluation (Robustness check H6)

Model	TN	FN	TP	FP	Specificity	Sensitivity
Standard model	385	4	8	192	0.651	0.667
Model (2)	357	3	9	193	0.649	0.750
Model (3)	384	3	9	166	0.698	0.750
Model (4)	383	4	8	167	0.696	0.667

Figure 7.6: ROC curve (Robustness check H6)



Both Scope 2 and 3 carbon intensity per sales unit-stressed models achieved a higher magnitude of the AUC than the standard model in both analyses featuring different data partitions. Even though none of the climate-stressed models could perform better in terms of default detection in both analyses, these two models showed they are more specific while keeping the same sensitivity rate. Therefore, Hypothesis 6 could not be rejected.

7.7 Hypothesis 7 testing

Table 7.19 features the summary of the standard and the climate-stressed models. Values of VIF are below the multicollinearity threshold for all models as shown in Table 7.20. Transition risk variables used for the Hypothesis 1 testing are Scope 1,2 and Scope 3 carbon intensities per asset. The signs and significance of the standard model estimators remained similar. Scope 1 carbon intensity estimator is statistically significant and negative suggesting its increase results in a decrease in the probability of default by 17 %. Scope 2 carbon intensity increase by one unit results in an increase of 83 % in odds of default. The estimator of Scope 3 carbon intensity per asset is not significant when using imbalanced data. Using balanced data, the effect is significant, yet its magnitude is only 1 %.

Table 7.19: Summary of the logistic regression (H7)

	Dependent variable:			
	flag_default			
	(1)	(2)	(3)	(4)
Operating_CF_margin	-4.485^{***}	-4.674^{***}	-3.959^{***}	-3.840^{***}
	(1.177)	(1.221)	(1.288)	(1.177)
Operating_CF_ratio	-1.403^{***}	-1.470^{***}	-1.803^{***}	-1.525^{***}
	(0.251)	(0.262)	(0.283)	(0.254)
Equity_ratio	-2.774^{***}	-3.165^{***}	-2.509^{***}	-2.511^{***}
	(0.365)	(0.384)	(0.381)	(0.374)
Current_liabilities	-0.811^{***}	-1.032^{***}	-0.946^{***}	-0.768^{***}
	(0.149)	(0.166)	(0.170)	(0.148)
carbon_intensity_scope1_assets		-0.187^{***}		
		(0.028)		
carbon_intensity_scope2_assets			0.609^{***}	
			(0.068)	
carbon_intensity_scope3_assets				0.014^{***}
				(0.005)
Constant	1.379^{***}	2.229^{***}	0.711^{***}	1.029^{***}
	(0.132)	(0.187)	(0.153)	(0.174)
Observations	1,437	1,437	1,437	1,437
Log Likelihood	−740.828	−702.053	−692.767	−736.127
Akaike Inf. Crit.	1,491.656	1,416.107	1,397.534	1,484.253
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 7.20: Variance inflation factors (H7)

Model	CF margin	CF ratio	Equity ratio	Current liab	Emission var
(1)	2.429	2.478	1.046	1.022	
(2)	2.532	2.561	1.104	1.062	1.096
(3)	2.485	2.566	1.051	1.027	1.030
(4)	2.450	2.500	1.068	1.044	1.051

7.7.1 Robustness check

Table 7.22 features the summary of the models applied to alternative data. All three transition risk variables are statistically significant. An increase of one unit of Scope 2 carbon intensity per asset results in an increase of 84 % in the probability of default, which is similar to the baseline results. An increase of Scope 3 emissions per asset translates to an increase of 1 % in the odds of default. As presented in Table 7.21, the presence of multicollinearity is not suggested by the variance inflation factors.

Table 7.21: Variance inflation factors (Robustness check H7)

Model	CF margin	CF ratio	Equity ratio	Current liab	Emission var
(1)	1.572	1.704	1.066	1.046	
(2)	1.642	1.776	1.076	1.056	1.025
(3)	1.641	1.708	1.071	1.046	1.062
(4)	1.631	1.650	1.084	1.063	1.138

Table 7.22: Robustness check (H7)

	Dependent variable:			
	flag_default			
	(1)	(2)	(3)	(4)
Operating_CF_margin	−0.508	−0.694	0.155	0.185
	(0.580)	(0.592)	(0.623)	(0.600)
Operating_CF_ratio	−0.483***	−0.476***	−0.487***	−0.554***
	(0.117)	(0.120)	(0.128)	(0.123)
Current_liabilities	−1.879***	−2.040***	−1.689***	−1.692***
	(0.212)	(0.225)	(0.212)	(0.206)
Equity_ratio	−3.098***	−3.282***	−3.137***	−2.715***
	(0.313)	(0.322)	(0.328)	(0.317)
carbon_intensity_scope1_assets		−0.045***		
		(0.009)		
carbon_intensity_scope2_assets			0.613***	
			(0.054)	
carbon_intensity_scope3_assets				0.018***
				(0.003)
Constant	0.860***	1.125***	0.118	0.372**
	(0.130)	(0.141)	(0.145)	(0.153)
Observations	1,585	1,585	1,585	1,585
Log Likelihood	−934.029	−909.396	−854.134	−917.661
Akaike Inf. Crit.	1,878.058	1,830.792	1,720.268	1,847.321
Note:				*p<0.1; **p<0.05; ***p<0.01

Scope 1 and 2 carbon intensity per asset managed to remain statistically significant on both balanced and imbalanced data using two different data partitions. The estimators of the Scope 1 carbon intensity per asset were negative but the Scope 2 estimators were positive. Overall, the Hypothesis 7 could not be rejected.

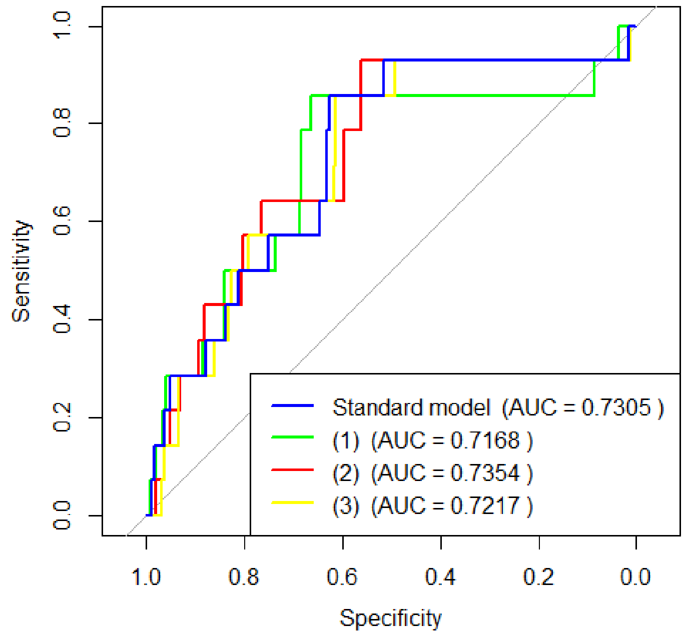
7.8 Hypothesis 8 testing

Standard and climate-stressed models perform moderately well, model (3) stressed by Scope 2 carbon intensity per asset reached a higher AUC score as shown in Figure 7.7. Moreover, model (3) reached higher specificity and sensitivity than the standard model as shown in Table 7.23.

Table 7.23: Performance evaluation (H8)

Model	TN	FN	TP	FP	Specificity	Sensitivity
Standard model	456	6	8	162	0.738	0.571
Model (2)	475	7	7	143	0.769	0.500
Model (3)	473	5	9	145	0.765	0.643
Model (4)	432	6	8	156	0.748	0.571

Figure 7.7: ROC curve (H8)



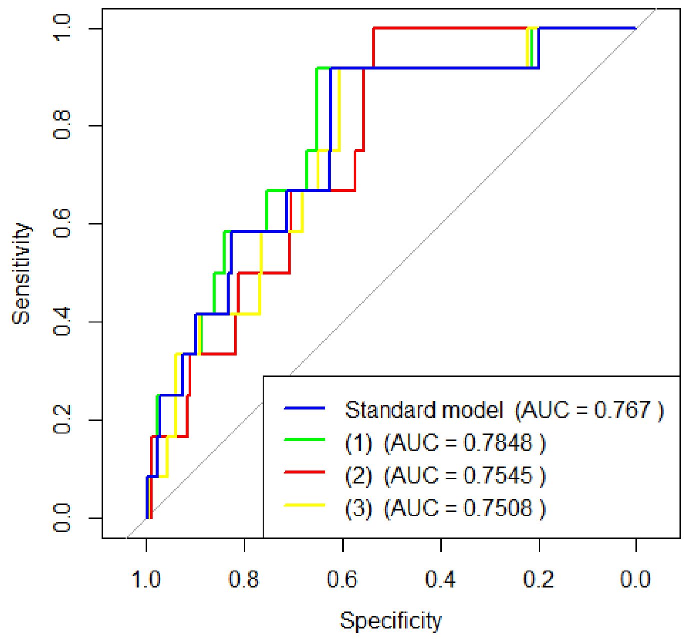
7.8.1 Robustness check

When assessing the performance of the climate-stressed models on the alternative data, only model (2), stressed by the Scope 1 carbon intensity per asset gained a magnitude of the AUC higher than the standard model as presented in Figure 7.8. Model (2) also gained distinctively higher sensitivity and specificity rates as shown in Table 7.24.

Table 7.24: Performance evaluation (Robustness check H8)

Model	TN	FN	TP	FP	Specificity	Sensitivity
Standard Model	358	4	8	192	0.651	0.667
Model (2)	350	1	11	200	0.917	0.917
Model (3)	387	4	8	163	0.667	0.667
Model (4)	362	4	8	188	0.667	0.667

Figure 7.8: ROC curve (Robustness check H8)



None of the climate-stressed models managed to gain a higher magnitude of the AUC or predict more defaults correctly on both sets of data. Therefore, Hypothesis 8 is rejected.

7.9 Summary of results

Overall, 8 hypotheses were tested, out of which two were rejected and six could not be rejected. There were two types of hypotheses, on the nature of the relationship between the transition risk variables and the probability of default and on the contribution of transition variables to the model performance.

Table 7.25: Overview of the results

Independent variable	Results
Hypothesis 1: Not rejected	
Scope 1	Rejected
Scope 2	Not rejected
Scope 3	Not rejected
Hypothesis 3: Not rejected	
Scope 1,2	Not rejected
Scope 3	Not rejected
Hypothesis 5: Not rejected	
Scope 1	Rejected
Scope 2	Not rejected
Scope 3	Not rejected
Hypothesis 7: Not rejected	
Scope 1	Rejected
Scope 2	Not rejected
Scope 3	Rejected

Table 7.25 summarizes the partial hypotheses on each transition risk variable for hypotheses 1,3,5 and 7. When Scope 1 emissions were studied separately in climate-stressed models, the estimator was negative and statistically significant in almost all cases. This means the firms, that are direct polluters tend to default less in the Bank. This result leads us to the conclusion that the most directly polluting firms are more stable given the Bank data.

An increase in Scope 2 and Scope 3 carbon footprint almost always resulted in a significant upward shift in the odds of default given the data from the Bank. The increase in odds tends to be higher for Scope 2 emissions when studied separately in climate-stressed models. This leads us to the conclusion that firms that are indirect polluters tend to be more prone to default on the loan granted by the Bank. Moreover, firms that use indirect emissions in their productions are more likely to default than the firms that only gain a carbon footprint through their supply chain.

As financial and emission data from 2022 were used, these findings correspond with the fact that the cost of electricity and gas sky-rocketed in the Czech Republic in 2022, mainly due to post-covid demand and geopolitical situation. Gas and electricity are the main sources of carbon emissions. Scope 1 polluters were not hit as hard by prices as they either sold these commodities or had the prices fixed. Scope 2 polluters were hit the hardest as they were buying these commodities at the market price. Scope 3 emitters might have been hit with some time lag as the high energy prices did not reflected instantly to other production inputs along their supply chain.

Although this situation may have been seen as a potential bias to the analysis, we rather interpreted it as an approximation of a climate policy shock on the economy which is likely to see the energy prices rise even further. Results are in line with the academic literature (Capasso *et al.* 2020; Carbone *et al.* 2021).

Table 7.26: Overview of the results

Independent variable	Results
Hypothesis 2: Not rejected	
Scope 1	Rejected
Scope 2	Not rejected
Scope 3	Rejected
Hypothesis 4: Rejected	
Scope 1,2	Rejected
Scope 3	Rejected
Hypothesis 6: Not rejected	
Scope 1	Rejected
Scope 2	Not rejected
Scope 3	Not rejected
Hypothesis 8: Rejected	
Scope 1	Rejected
Scope 2	Rejected
Scope 3	Rejected

Table 7.25 summarizes the partial hypotheses on each transition risk variable for hypotheses 2,4,6 and 8. Although these hypotheses are more of a complementary character, there are several findings to be presented. All climate-stressed models performed moderately well in terms of models' discriminatory power and reached a specificity rate higher than 0.5 meaning the majority of defaults were predicted well by the model given a threshold of 0.5. Using two different random data partitions, we were able to observe the inclusion of the Scope 2 carbon footprint significantly improved the performance of the model in terms of specificity, sensitivity, and the magnitude of the AUC-ROC curve.

The inclusion of Scope 2 and Scope 3 carbon intensity per sale to the standard model improved its specificity rate and the magnitude of the AUC-ROC curve.

7.10 Further research opportunities

We recognize three further research opportunities linked to the subject of the thesis. First, using the same dataset as in this thesis, a climate-stress test of the Bank could be done. This would consist of analyzing and quantifying different scenarios of the future European climate policies and their direct impact on the level of capital requirements of the Bank.

Second, given the pressure on banks to foster more green finance, it is expected several other banks might start or already collecting similar emission data on their clients. This would allow researchers to compare the results of this thesis to a case study featuring the data of a different bank. Moreover, the availability of time-series data on client carbon footprint might allow for more complex analysis.

Finally, this thesis employs logistic regression which is a simple and widely used method for credit risk estimation. However, more complex machine learning models tend to yield more precise results, therefore employing different models to answer similar questions might represent a research-worthy opportunity.

Chapter 8

Conclusion

The thesis builds on a unique opportunity to analyze internal data collected by the Bank which, except for the standard financial information on the client, contains the information on carbon footprint of all corporate clients belonging to SME category.

The main contribution of this thesis is the empirical evidence on the relationship between climate transition risk and credit risk. The analysis was structured as a stress test of a standard credit scoring model to which a set of transition risk variables was added separately. Logistic regression was employed for the climate-stress test. Transition risk variables consisted of three scopes of carbon footprint, financed carbon footprint, and carbon intensity of the client. We uncovered that Scope 1 emitters are generally less prone to credit default in the Bank. Scope 2 and Scope 3 indirect emitters tend to default more. Given the data provided by the Bank, the level of Scope 2 and Scope 3 carbon footprint and carbon intensity was recognized as a potential driver of default.

The second contribution of this thesis is methodological since the performance of the climate-stressed logistic regression models could be observed and compared to the standard model. We found the inclusion of the Scope 2 carbon footprint into the credit scoring model improved its discriminatory power and ability to correctly predict a credit default. Integration of Scope 2 and Scope 3 carbon intensity per unit of sales improved discriminatory power and ability to correctly predict non-defaults while correctly predicting the same number of defaults.

The third contribution of the thesis is the identification of 4 financial predictors that are drivers of credit default since these indicators proved to be statistically significant forming a standard model capable of detecting the majority of defaults in testing data. The standard model predictors thus represent an empirical cross-check on the data provided by the Bank. Finally, the thesis provides a comprehensive summary of academic literature on climate transition risk in financial markets, effectively identifying a wide literature gap for further research opportunities.

Bibliography

Bartual Sanfeliu, C., F. García García, F. Guijarro, & A. Romero Civera (2012): "Probability of deafault using the logit model: The impact of explanatory variable and data base selection." In "International Scientific Conference: Whither our Economics," pp. 118–124. Mykolas Romeris University.

Battiston, S., Y. Dafermos, & I. Monasterolo (2021): "Climate risks and financial stability."

Battiston, S., A. Mandel, I. Monasterolo, F. Schütze, & G. Visentin (2017): "A climate stress-test of the financial system." *Nature Climate Change* **7(4)**: pp. 283–288.

BCBS (2021): "Climate-related risk drivers and their transmission channels."

Bell, F. & G. van Vuuren (2022): "The impact of climate risk on corporate credit risk." *Cogent Economics & Finance* **10(1)**: p. 2148362.

BIS (2023): "Credit risk." .

Blasberg, A., R. Kiesel, & L. Taschini (2021): "Climate default swap: Disentangling the exposure to transition risk through cds." *Available at SSRN* .

Caldecott, B., J. Tilbury, & C. Carey (2014): "Stranded assets and scenarios." .

Capasso, G., G. Gianfrate, & M. Spinelli (2020): "Climate change and credit risk." *Journal of Cleaner Production* **266**: p. 121634.

Carbone, S., M. Giuzio, S. Kapadia, J. S. Krämer, K. Nyholm, & K. Vozian (2021): "The low-carbon transition, climate commitments and firm credit risk." .

CBI (2017): "Climate bond initiative report." .

Chava, S. (2014): "Environmental externalities and cost of capital." *Management science* **60(9)**: pp. 2223–2247.

Dafermos, Y. & M. Nikolaidi (2021): "How can green differentiated capital requirements affect climate risks? a dynamic macrofinancial analysis." *Journal of Financial Stability* **54**: p. 100871.

Devalle, A. (2017): "The linkage between esg performance and credit ratings: A firm-level perspective analysis." .

Dombrikovskis (2018): "Speech by vice-president dombrovskis at the high-level conference on financing sustainable growth." .

Douzas, G., F. Bacao, & F. Last (2018): "Improving imbalanced learning through a heuristic oversampling method based on k-means and smote." *Information sciences* **465**: pp. 1–20.

Dumitrescu, E., S. Hué, C. Hurlin, & S. Tokpavi (2022): "Machine learning for credit scoring: Improving logistic regression with non-linear decision-tree effects." *European Journal of Operational Research* **297(3)**: pp. 1178–1192.

Dunz, N., A. Naqvi, & I. Monasterolo (2019): "Climate transition risk, climate sentiments, and financial stability in a stock-flow consistent approach." *Climate Sentiments, and Financial Stability in a Stock-Flow Consistent Approach (April 1, 2019)* .

EC (2023): "Carbon border adjustment mechanism factsheet." .

ECB (2020): "Guide on climate-related and environmental risks, supervisory expectations relating to risk management and disclosure." *Guide on climate-related and environmental risks. Supervisory expectations relating to risk management and disclosure* .

Faralli, M. & F. Ruggiero (2022): "The rise of climate risks: Evidence on firms' expected default frequencies." *Available at SSRN 4144802* .

Fatica, S., R. Panzica, & M. Rancan (2021): "The pricing of green bonds: are financial institutions special?" *Journal of Financial Stability* **54**: p. 100873.

Gianfrate, G. & M. Peri (2019): "The green advantage: Exploring the convenience of issuing green bonds." *Journal of cleaner production* **219**: pp. 127–135.

Greene, W. H. (2003): *Econometric analysis.* Pearson Education India.

Hachenberg, B. & D. Schiereck (2018): "Are green bonds priced differently from conventional bonds?" *Journal of Asset Management* **19**: pp. 371–383.

Hamilton, D. T., Z. Sun, & M. Ding (2011): "Through-the-cycle edf credit measures." *Moody's Analytics, August* .

Höck, A., C. Klein, A. Landau, & B. Zwergel (2020): "The effect of environmental sustainability on credit risk." *Journal of Asset Management* **21(2)**: pp. 85–93.

Huljak, I., R. Martin, D. Moccero, & C. Pancaro (2022): "Do non-performing loans matter for bank lending and the business cycle in euro area countries?" *Journal of applied economics* **25(1)**: pp. 1050–1080.

IPCC (2021): "Weather and climate extreme events in a changing climate (chapter 11)." .

Jakubík, P., P. Teplỳ *et al.* (2011): "The jt index as an indicator of financial stability of corporate sector." *Prague Economic Papers* **20(2)**: pp. 157–176.

Jorion, P. *et al.* (2010): *Financial Risk Manager Handbook: FRM Part I/Part II.* John Wiley & Sons.

Jung, J., K. Herbohn, & P. Clarkson (2018): "Carbon risk, carbon risk awareness and the cost of debt financing." *Journal of Business Ethics* **150**: pp. 1151–1171.

Kabir, M. N., S. Rahman, M. A. Rahman, & M. Anwar (2021): "Carbon emissions and default risk: International evidence from firm-level data." *Economic Modelling* **103**: p. 105617.

Karpf, A. & A. Mandel (2018a): "The changing value of the green label on the us municipal bond market." *Nature Climate Change* **8(2)**: pp. 161–165.

Kotró, B. & M. Márkus (2020): "The impact of esg rating on corporate bond yields." *Economy and Finance: English-Language Edition of Gazdaság és Pénzügy* **7(4)**: pp. 474–488.

Leone, A. J., M. Minutti-Meza, & C. E. Wasley (2019): "Influential observations and inference in accounting research." *The Accounting Review* **94(6)**: pp. 337–364.

Maltais, A. & B. Nykvist (2020): "Understanding the role of green bonds in advancing sustainability." *Journal of sustainable finance & investment* pp. 1–20.

Merton, R. C. (1974): "On the pricing of corporate debt: The risk structure of interest rates." *The Journal of finance* **29(2)**: pp. 449–470.

Monasterolo, I., S. Battiston, A. C. Janetos, & Z. Zheng (2017): "Vulnerable yet relevant: the two dimensions of climate-related financial disclosure." *Climatic change* **145**: pp. 495–507.

NGFS (2021): "Ngfs climate scenarios for central banks and supervisors." .

Nguyen, Q., I. Diaz-Rainey, & D. Kuruppuarachchi (2023): "In search of climate distress risk." *International Review of Financial Analysis* **85**: p. 102444.

Nordhaus, W. D. (2007): *The challenge of global warming: economic models and environmental policy*, volume 4. Citeseer.

Oikonomou, I., C. Brooks, & S. Pavelin (2014): "The effects of corporate social performance on the cost of corporate debt and credit ratings." *Financial Review* **49(1)**: pp. 49–75.

Stern, N. (2009): "Climate change, internationalism and india in the 21 st century." *CCCEP Policy Paper* .

Wang, H.-J., Y. Geng, X.-Q. Xia, & Q.-J. Wang (2021): "Impact of economic policy uncertainty on carbon emissions: evidence from 137 multinational countries." *International Journal of Environmental Research and Public Health* **19(1)**: p. 4.

Zerbib, O. D. (2019): "The effect of pro-environmental preferences on bond prices: Evidence from green bonds." *Journal of Banking & Finance* **98**: pp. 39–60.

Zhang, Y., Y. Liu, & H. Wang (2023): "How credit default swap market measures carbon risk." *Environmental Science and Pollution Research* pp. 1–21.

Appendix A

Figure A.1: Distribution of Scope 1 carbon footprint

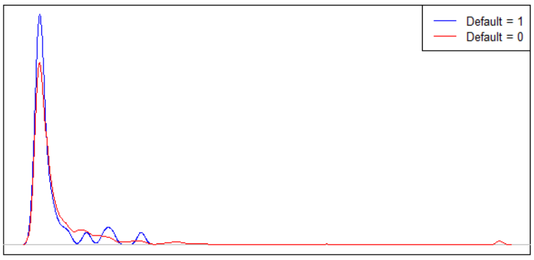


Figure A.2: Distribution of Scope 2 carbon footprint

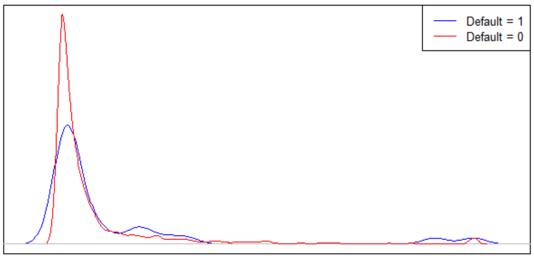


Figure A.3: Distribution of Scope 3 carbon footprint

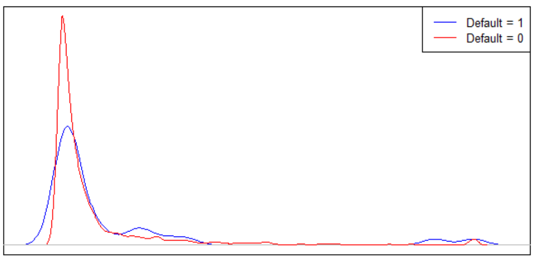


Figure A.4: Distribution of Scope 1 carbon intensity per sales unit

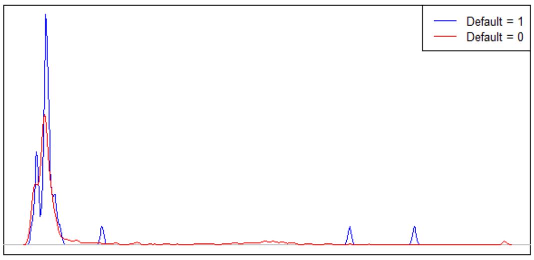


Figure A.5: Distribution of Scope 1 carbon intensity per asset unit

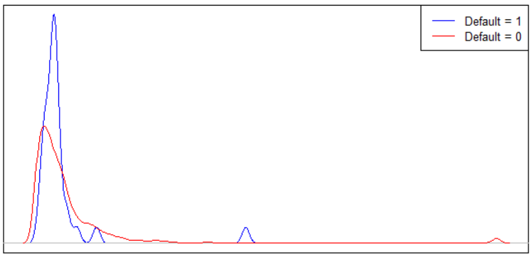


Table A.1: Summary of logistic regression - Robustness check on imbalanced data

	Dependent variable:	
	flag_default	
	(1)	(2)
Operating_CF_margin	−0.913	−0.743
	(1.304)	(0.607)
Operating_CF_ratio	−0.457*	−0.697***
	(0.267)	(0.129)
Current_liabilities	−1.221***	−1.756***
	(0.387)	(0.217)
Equity_ratio	−1.913***	−1.141***
	(0.586)	(0.389)
Debt_to_Assets		3.444***
		(0.448)
Constant	−1.642***	−0.912***
	(0.245)	(0.257)
Observations	901	1,585
Log Likelihood	−238.580	−903.462
Akaike Inf. Crit.	487.161	1,818.923
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table A.2: Summary of logistic regression (H1) - imblanced data

	Dependent variable:			
	flag_default			
	(1)	(2)	(3)	(4)
Operating_CF_margin	-4.551^{*}	-4.562^{*}	-5.290^{*}	-3.381
	(2.504)	(2.510)	(2.988)	(2.702)
Operating_CF_ratio	-1.561^{***}	-1.561^{***}	-1.847^{***}	-1.958^{***}
	(0.561)	(0.562)	(0.650)	(0.616)
Equity_ratio	-2.225^{***}	-2.230^{***}	-3.332^{***}	-2.838^{***}
	(0.621)	(0.624)	(0.692)	(0.657)
Current_liabilities	-0.630^{**}	-0.636^{**}	-1.686^{***}	-1.670^{***}
	(0.271)	(0.281)	(0.403)	(0.420)
Scope_1_TotEm		0.016		
		(0.199)		
Scope_2_TotEm			0.891^{***}	
			(0.139)	
Scope_3_TotEm				0.841^{***}
				(0.172)
Constant	-0.796^{***}	-0.792^{***}	-0.723^{***}	-0.758^{***}
	(0.215)	(0.219)	(0.233)	(0.230)
Observations	842	842	842	842
Log Likelihood	-221.914	-221.911	-197.283	-210.058
Akaike Inf. Crit.	453.828	455.822	406.565	432.117
Note:				*p<0.1; **p<0.05; ***p<0.01

Table A.3: Robustness check (H1) - imbalanced data

	Dependent variable:			
	flag_default			
	(1)	(2)	(3)	(4)
Operating_CF_margin	−0.913	−0.891	−0.384	0.116
	(1.304)	(1.289)	(1.329)	(1.319)
Operating_CF_ratio	-0.457^{*}	-0.438^{*}	-0.515^{*}	-0.656^{**}
	(0.267)	(0.265)	(0.280)	(0.289)
Current_liabilities	-1.221^{***}	-1.085^{***}	-2.135^{***}	-2.728^{***}
	(0.387)	(0.394)	(0.497)	(0.632)
Equity_ratio	-1.913^{***}	-1.848^{***}	-2.452^{***}	-2.331^{***}
	(0.586)	(0.587)	(0.616)	(0.609)
Scope_1_TotEm		−0.309		
		(0.312)		
Scope_2_TotEm			0.719^{***}	
			(0.143)	
Scope_3_TotEm				0.908^{***}
				(0.202)
Constant	-1.642^{***}	-1.697^{***}	-1.743^{***}	-1.825^{***}
	(0.245)	(0.252)	(0.267)	(0.281)
Observations	901	901	901	901
Log Likelihood	−238.580	−237.875	−225.556	−229.199
Akaike Inf. Crit.	487.161	487.749	463.113	470.398
Note:				*p<0.1; **p<0.05; ***p<0.01

Table A.4: Summary of logistic regression (H3) - imbalanced data

	Dependent variable:		
	flag_default		
	(1)	(2)	(3)
Operating_CF_margin	-4.551^{*}	-4.367^{*}	-1.972
	(2.504)	(2.444)	(2.474)
Operating_CF_ratio	-1.561^{***}	-1.679^{***}	-2.367^{***}
	(0.561)	(0.556)	(0.609)
Equity_ratio	-2.225^{***}	-2.230^{***}	-2.389^{***}
	(0.621)	(0.632)	(0.687)
Current_liabilities	-0.630^{**}	-0.882^{***}	-1.569^{***}
	(0.271)	(0.320)	(0.398)
Fin_Scope12		0.400^{***}	
		(0.133)	
Fin_Scope3			0.860^{***}
			(0.124)
Constant	-0.796^{***}	-0.823^{***}	-1.094^{***}
	(0.215)	(0.222)	(0.250)
Observations	842	842	842
Log Likelihood	-221.914	-218.041	-195.454
Akaike Inf. Crit.	453.828	448.081	402.907
Note:			*p<0.1; **p<0.05; ***p<0.01

Table A.5: Robustness check (H3) - imbalanced data

	Dependent variable:		
	flag_default		
	(1)	(2)	(3)
Operating_CF_margin	-0.913	-0.819	-0.270
	(1.304)	(1.303)	(1.322)
Operating_CF_ratio	-0.457^{*}	-0.511^{*}	-0.631^{**}
	(0.267)	(0.272)	(0.291)
Current_liabilities	-1.221^{***}	-1.445^{***}	-2.084^{***}
	(0.387)	(0.432)	(0.523)
Equity_ratio	-1.913^{***}	-1.879^{***}	-1.663^{***}
	(0.586)	(0.591)	(0.612)
Fin_Scope12		0.244^{*}	
		(0.129)	
Fin_Scope3			0.530^{***}
			(0.114)
Constant	-1.642^{***}	-1.688^{***}	-1.998^{***}
	(0.245)	(0.252)	(0.290)
Observations	901	901	901
Log Likelihood	-238.580	-237.133	-228.526
Akaike Inf. Crit.	487.161	486.266	469.052
Note:			*p<0.1; **p<0.05; ***p<0.01

Table A.6: Summary of logistic regression (H5) - imbalanced data

	Dependent variable:			
	flag_default			
	(1)	(2)	(3)	(4)
Operating_CF_margin	-4.551^{*}	-3.919	-5.249^{**}	-4.823^{*}
	(2.504)	(2.744)	(2.559)	(2.530)
Operating_CF_ratio	-1.561^{***}	-1.718^{***}	-1.387^{**}	-1.380^{**}
	(0.561)	(0.594)	(0.566)	(0.561)
Equity_ratio	-2.225^{***}	-2.267^{***}	-2.580^{***}	-2.364^{***}
	(0.621)	(0.628)	(0.642)	(0.631)
Current_liabilities	-0.630^{**}	-0.643^{**}	-0.706^{**}	-0.615^{**}
	(0.271)	(0.272)	(0.289)	(0.273)
carbon_intensity_S1_sales		-0.084^{*}		
		(0.045)		
carbon_intensity_S2_sales			0.449^{***}	
			(0.127)	
carbon_in tensity_S3_sales				0.045^{***}
				(0.015)
Constant	-0.796^{***}	-0.542^{**}	-1.033^{***}	-1.314^{***}
	(0.215)	(0.244)	(0.229)	(0.278)
Observations	842	842	842	842
Log Likelihood	−221.914	−217.985	−216.129	−217.442
Akaike Inf. Crit.	453.828	447.970	444.257	446.883
Note:				*p<0.1; **p<0.05; ***p<0.01

Table A.7: Robustness check (H5) - imbalanced data

	Dependent variable:			
	flag_default			
	(1)	(2)	(3)	(4)
Operating_CF_margin	−0.913	−0.698	−1.254	−0.663
	(1.304)	(1.300)	(1.304)	(1.315)
Operating_CF_ratio	−0.457*	−0.477*	−0.346	−0.405
	(0.267)	(0.272)	(0.247)	(0.256)
Current_liabilities	−1.221***	−1.229***	−1.244***	−1.145***
	(0.387)	(0.389)	(0.388)	(0.381)
Equity_ratio	−1.913***	−1.895***	−2.139***	−2.022***
	(0.586)	(0.588)	(0.588)	(0.594)
carbon_intensity_S1_sales		−0.014		
		(0.017)		
carbon_intensity_S2_sales			0.474***	
			(0.124)	
carbon_intensity_S3_sales				0.059***
				(0.014)
Constant	−1.642***	−1.590***	−1.950***	−2.368***
	(0.245)	(0.251)	(0.263)	(0.306)
Observations	901	901	901	901
Log Likelihood	−238.580	−237.712	−231.810	−229.780
Akaike Inf. Crit.	487.161	487.423	475.620	471.559
Note:				*p<0.1; **p<0.05; ***p<0.01

Table A.8: Summary of logistic regression (H7) - imbalanced data

	Dependent variable:			
	flag_default			
	(1)	(2)	(3)	(4)
Operating_CF_margin	-4.551^{*}	-4.624^{*}	-4.300	-4.174
	(2.504)	(2.518)	(2.644)	(2.561)
Operating_CF_ratio	-1.561^{***}	-1.477^{***}	-1.672^{***}	-1.627^{***}
	(0.561)	(0.563)	(0.581)	(0.569)
Equity_ratio	-2.225^{***}	-2.616^{***}	-2.279^{***}	-2.080^{***}
	(0.621)	(0.651)	(0.632)	(0.628)
Current_liabilities	-0.630^{**}	-0.779^{***}	-0.566^{**}	-0.571^{**}
	(0.271)	(0.291)	(0.268)	(0.268)
carbon_intensity_S1_assets		-0.118^{***}		
		(0.046)		
carbon_intensity_S2_assets			0.393^{***}	
			(0.095)	
carbon_intensity_S3_assets				0.012
				(0.007)
Constant	-0.796^{***}	-0.224	-1.176^{***}	-1.074^{***}
	(0.215)	(0.297)	(0.239)	(0.279)
Observations	842	842	842	842
Log Likelihood	-221.914	-216.097	-213.920	-220.680
Akaike Inf. Crit.	453.828	444.193	439.840	453.360
Note:				*p<0.1; **p<0.05; ***p<0.01

Table A.9: Robustness check (H7) - imbalanced data

	Dependent variable:			
	flag_default			
	(1)	(2)	(3)	(4)
Operating_CF_margin	-0.913	-1.179	0.100	0.132
	(1.304)	(1.313)	(1.302)	(1.283)
Operating_CF_ratio	-0.457^{*}	-0.424	-0.449^{*}	-0.549^{**}
	(0.267)	(0.271)	(0.273)	(0.269)
Current_liabilities	-1.221^{***}	-1.306^{***}	-1.098^{***}	-1.080^{***}
	(0.387)	(0.399)	(0.384)	(0.378)
Equity_ratio	-1.913^{***}	-2.038^{***}	-2.126^{***}	-1.713^{***}
	(0.586)	(0.593)	(0.616)	(0.604)
carbon_intensity_S1_assets		-0.051^{*}		
		(0.027)		
carbon_intensity_S2_assets			0.435^{***}	
			(0.081)	
carbon_intensity_S3_assets				0.017^{***}
				(0.006)
Constant	-1.642^{***}	-1.354^{***}	-2.154^{***}	-2.102^{***}
	(0.245)	(0.278)	(0.274)	(0.292)
Observations	901	901	901	901
Log Likelihood	-238.580	-234.644	-224.886	-233.955
Akaike Inf. Crit.	487.161	481.287	461.771	479.909
Note:		*p<0.1; **p<0.05; ***p<0.01		