Egeli, Martin and Grimsmo, Anniken Skeisvoll

Disentangling determinants of CDS spreads: A machine learning approach

Master's thesis in Industrial Economy and Technology Management
Supervisor: Lavrutich, Maria
Co-supervisor: Pimentel, Rita
June 2022

Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Economics and Management
Dept. of Industrial Economics and Technology Management

NTNU
Kunnskap for en bedre verden

Egeli, Martin and Grimsmo, Anniken Skeisvoll

Disentangling determinants of CDS spreads: A machine learning approach

Master's thesis in Industrial Economy and Technology Management
Supervisor: Lavrutich, Maria
Co-supervisor: Pimentel, Rita
June 2022

Norwegian University of Science and Technology
Faculty of Economics and Management
Dept. of Industrial Economics and Technology Management

NTNU
Norwegian University of
Science and Technology

Department of Industrial Economics and Technology Management

Master Thesis

Disentangling determinants of CDS spreads: A machine learning approach

Authors:
Egeli, Martin
Grimsmo, Anniken Skeisvoll

Supervisors:
Lavrutich, Maria
Pimentel, Rita

June, 2022

Abstract

This thesis analyzes the determinants of corporate credit default swaps (CDS) by using a global sample of 240 unique CDS's from 2010 until 2020. We apply a systematic approach by using the feature selection method LASSO to find new variables and identify potential redundancies in the existing literature. Furthermore, we use fixed-effect panel regression to provide significance levels and the models R-squared, and quantile panel regression to provide a more nuanced understanding of what determines CDS spreads. The analysis is performed on level data and first-differenced data, using a broad range of variables.

First, our results provide a concise yet comprehensive model with only seven and ten variables needed to explain the CDS spread and changes in the CDS spread, respectively. Second, we provide several novel insights related to the variables. The ESG variables social pillar and governance pillar has a positive sign, thus supporting the over-investment view related to ESG. Quantitative Easing (QE) has previously only been demonstrated to have a risk-reducing effect for bank CDS spreads, but we show the same results apply to corporate CDS spreads. Furthermore, as shown by our quantile panel regression, QE has a larger risk-reducing effect for the more risky firms. For Fama-French variables, we find the variables market excess return (MKT) and small-minus-big (SMB) significant for describing the CDS spread for first-differenced data. Lastly, our results show that the variables needed to explain the variations in the CDS spread is not the same variables needed to explain the variations in the changes in CDS spreads.

Sammendrag

I denne avhandlingen analyseres determinantene til bedrifters *credit default swaps* (CDS), ved å bruke et globalt utvalg av 240 unike CDS-er fra 2010 til 2020. Vi bruker en systematisk tilnærming med variabelseleksjons-metoden LASSO for å verifisere nye variabler og identifisere potensiell redundans i den eksisterende literaturen. Deretter bruker vi fikset-effekt- og kvantil panelregresjon for å vise signifikans og gi en mer nyansert forståelse av hvordan variablene bestemmer CDS-spreaden. Analysen utfører vi på level data og data som er differensiert en gang, og vi inkluderer et bredt spekter av variabler.

For det første gir resultatene våre en konsis, men likevel omfattende modell der kun syv og ti variabler nødvendige for å forklare henholdsvis CDS-spreaden og endringer i CDS-spreaden. For det andre, viser resultatene våre flere nye innsikter knyttet til CDS determinantene. ESG-variablene *social pillar* og *governance pillar* har en positiv effekt på CDS spreaden, og støtter dermed overinvesterings hypotesen knyttet til ESG. Videre har Quantitative easing (QE) tidligere kun vist seg og ha en risikoreduserende effekt for bank-CDS-spreader, men vi viser at de samme resultatene gjelder for bedrifts-CDS-spreader. Kvantil panelregresjon viser at QE har en større risikoreduserende effekt for de mer risikofylte firmaene. For Fama-french variabler finner vi at variablene market excess return (MKT) og small-minus-big (SMB) er signifikante for å beskrive CDS-spreaden for differensert data, noe som ikke har blitt demonstrert i tidligere litteratur. Til slutt viser vi at kombinasjonen av variabler som forklarer variasjonene i CDS-spreader ikke er den samme som kombinasjoner av variabler som viser endringer i CDS-spreader

Preface

This thesis completes our Master of Science (MSc) degree in Industrial Economics and Technology Management at the Norwegian University of Science and Technology (NTNU). It was written in the period from January 2022 to June 2022. The thesis is written independently by Martin Egeli and Anniken Skeisvoll Grimsmo. Our motivation for writing the thesis is based on academic and professional interests. Academic in terms of using theory, as we have learned empirical finance and statistics to understand real-world financial data. Professional in terms of a broader understanding of credit risk and its determinants, which give valuable insight and understanding for our future professional careers. We are grateful for the possibility of working on developing new research on such an important topic.

We want to thank our primary supervisor Maria Lavrutich for her academic support. We have received excellent guidance in academic writing, valuable insights through her expertise in empirical finance, and extensive feedback throughout the entire project period. We would also like to thank Rita Pimentel for her valuable insights, interpretation of results and methodology, and showing a thorough interest in our work throughout the project period.

Table of Contents

List of Figures

List of Tables

Abbreviations

AIC	Akaike Information Criteria
BIC	Bayesian Information Criteria
CAPM	Capital Asset Pricing Model
CDS	Credit Default Swap
CSR	Corporate Social Responsibility
FE	Fixed Effects
GOV	Governance Pillar Score
HML	High-Minus-Low
LASSO	Least Absolute Shrinkage and Selection Operator
MKT	Market excess return
QE	Quantitative Easing
ROA	Return On Assets
ROCE	Return on Capital Employed
SMB	Small-Minus-Big
SOC	Social Pillar Score
VIX	Chicago Board Options Exchanges CBOE Volatility Index

Chapter 1

Introduction

During the last 20 years, several important events have altered the market's perception of credit risk. Among them is the great financial crisis in 2008, which caused severe damage worldwide. The crisis was largely driven by derivative credit risk instruments that lacked transparency, as they allowed banks to reduce the capital required against their loans and thus enabled them to issue even more loans (Acharya and Richardson, 2009). By observing the credit derivatives throughout the crisis, the movements in the derivatives reflect the changing market perception of what constitutes credit risk (Cesare and Guazzarotti, 2010). For instance, firm-specific leverage became more important for describing credit risk, which may reflect that investors became more aware of individual risk factors (Cesare and Guazzarotti, 2010). In the wake of the crisis, the European sovereign debt crisis arose when Iceland's bank system collapsed in 2008. The crisis caused fear in the financial markets and peaked in 2012 when Greece defaulted on its debt. As a result of each crisis, the need for a more regulated financial environment for credit instruments became apparent. Authorities responded to this challenge by enforcing various regulations such as standardization of contracts, expansion of reporting requirements, requirements of margins, and mandatory central clearing (FSB, 2017).

Central in the financial crisis and European sovereign debt crisis was the Credit Default Swap (CDS). The swap was introduced by Morgan Stanley in 1994 and was originally designed to transfer credit risk. The credit instrument quickly became popular, and in 2007 the value of the CDS market reached \$61.2 trillion, which was more than double of the value invested in the stock market (BIS, 2021). Since then, CDS spreads have been widely used by the academic community as a metric for credit risk. There was a need for a more liquid metric that responded quicker to changes in credit risk than bonds and credit ratings previously used (Finnerty et al., 2013, Zhu, 2006, Ederington et al., 2015, Blanco et al., 2005). Additionally, the CDS made it easier to compare credit risk across firms as it is a more standardized instrument that is less dependent on individual features and primary market issuance (B. Y. Zhang et al., 2009, Norden and Weber, 2009).

In this thesis, we employ a systematic approach to find the most important determinants of CDS spread and identify redundancies in the literature in order to facilitate a better understanding of the drivers of credit risk. Various studies have sought to uncover what determines the CDS spread by examining a broad range of variables. Most studies focus on finding new variables in order to increase the model's ability to explain the variation in the CDS spread. As a result, a model using all CDS determinants that were found to be important in the existing literature is complex and possibly suffers from redundancies. However, no study has performed a structural approach for identifying potential redundancies in the variables considered by the literature. Thus, it is not easy for investors and financial institutions to disentangle what affects credit risk in the current market environment. Our study is a response to the need for a concise yet comprehensive model of credit risk.

Early academic literature on CDS determinants primarily focused on whether accounting-based or market-based variables are better at explaining the variations in CDS spreads. For instance, Das et al. (2009) concludes that a model using accounting-based variables explains the spread at least as well as a market-based model, however, a model using both categories are needed to capture most of the variation. More recently, Tang and Yan (2017) demonstrates that market-wide and firm-specific variables together only account for 40% of the variation in the CDS spread. Literature thus evolved to focus on increasing how much of the variation in the CDS spread the model explains, leading to an uptake in studies searching for new determinants. The uptake was also driven by global events, and as a result, several new determinants of CDS spreads were uncovered.

For example, among the latest events impacting the perception of credit risk is the Paris Agreement. The legally binding international treaty was adopted by 196 parties and entered into force in 2016 (UN, 2015). The agreement increased awareness of credit risk related to companies' sustainable performance. In addition, it led to a stream of literature demonstrating the impact different metrics of sustainability have on CDS spreads (Delis et al., 2018, and Kölbel, Leippold et al., 2020). For example, Blasberg et al. (2021) finds that higher yearly emissions levels increase the CDS spread. Related to the communication of sustainability, Kölbel, Leippold et al. (2020) finds that disclosing transitional risk increases the CDS spread. Moreover, Naumer and Yurtoglu (2019) demonstrates that companies with larger volumes of positive ESG-related news face lower CDS spreads. The debate on ESG variables has sparked two opposing views debated in literature; the risk-mitigation view and the over-investment view (Goss and Roberts, 2011). While the risk-mitigation view argues that investments in ESG-related activities reduce risk, the over-investment view argues that these activities are costly diversions of a firm's resources and thus increase risk.

Intuitively, recent and historical impactful events shape the market's perception of what constitutes credit risk. This implies that determinants might change over time, and that determinants found to affect credit risk in the past become obsolete in the current market environment. For instance, return on assets (ROA) has been demonstrated to be negatively correlated to the CDS spread by Das et al. (2009) and Hu et al. (2017) and not significant by Tang and Yan (2017) and Kölbel, Leippold et al. (2020). Moreover, other variables like term-structure slope and firm size were found to be both positively-, negatively-, and not significantly correlated to the CDS spread in different studies (Galil et al., 2013, Hu et al., 2017 and Barth et al., 2021). The lack of consensus regarding CDS determinants may also be caused by different study designs that include different methodologies and data types. The most common methodologies used, employed by studies such as Pereira et al. (2018), Kölbel, Leippold et al. (2020) and Barth et al. (2021), is the fixed-effect panel regression that accounts for the unobserved heterogeneity between CDS's. However, earlier studies like Ericsson et al. (2009) used average regression coefficients obtained by running a series of time-series regressions, one for each firm. Moreover, Pires et al. (2013) uses quantile panel regression to provide more insights regarding quantiles and not the mean. In addition to methodological differences, studies diverge in data types. For example, some use CDS spreads as the dependent variable, whereas others focus on explaining the variation in changes in the CDS spreads. Although this might impact the results, no study has discussed whether the driving factors of the variability in CDS spreads and changes in CDS spreads differ.

We contribute to the existing literature in several ways. First, we employ a systematic approach that allows us to identify the most important determinants for describing the CDS spread among both new variables and variables previously demonstrated as determinants. Using the feature selection methodology Least Absolute Shrinkage and Selection Operator (LASSO), we identify redundancies in previous studies, which allows us to build a concise model of the important determinants of CDS spreads. Further, using quantile regression, we examine which variables explain each quantile and not only the expected value, which gives a broader understanding of the variables affecting credit risk. We apply the LASSO methodology on both level data and first-differenced data. The LASSO method results in seven variables showing significance for level data and ten for first-differenced data, which implies that different variables are important for explaining the variations in CDS spreads and changes in CDS spreads. Our results show that adding more variables does not necessarily improve the model's ability to explain the CDS spread. Thus, we find a concise yet comprehensive model for explaining credit risk that financial institutions can easily employ.

Methodologically, Chalamandaris and Vlachogiannakis (2016) is the first, and to the best of our knowledge, the only study using a feature selection method to find the most important determinants of corporate CDS spreads. However, the study focuses solely on financial ratios as determinants. While they conclude that 5 out of 43 financial ratios are important, they do not apply feature selection methods to the pool of determinants found in the literature. In contrast to Chalamandaris and Vlachogiannakis (2016), our study examines all categories of determinants on a more recent dataset.

Our thesis provides several novel insights. By employing a fixed-effects LASSO panel regression on level data, we find the new variables social pillar score (SOC), governance pillar score (GOV), and quantitative easing (QE) to be significant in explaining the CDS spread. Additionally, the traditional variables stock volatility, firm size, spot rate, and yield slope are significant. Previously commonly used variables in literature, such as stock return, VIX, and leverage ratio (Tang and Yan, 2017), do not give further explanation of CDS spread in our model. In terms of ESG, our results imply that the ESG variables GOV and SOC support the over-investment view. The implication is that the market sees these investments as wasteful. Moreover, we find the novel result that QE is significant for explaining corporate CDS spread, which has previously only been found to be significant for bank CDS spread (Ajer and Bohler, 2021). Further, we use the chosen variables in a quantile regression, where a higher quantile corresponds to more risky firms. This gives an understanding of the conditional distribution of the determinants in relation to the riskiness of the firm. Our quantile regression shows that the largest risk-reducing effect applies to the riskiest firms, implying that the probability of default is reduced with increased QE. This is interesting in light of the planned quantitative tightening by the FED in mid 2022, which might cause credit risk to increase when the QE level is reduced (Councile, 2022).

By employing a fixed-effects LASSO panel on first-differenced data, our results show that different variables explain the changes in the CDS spread than on levels. The only significant variables for regressions on both data types are stock volatility, firm size, and yield slope. A total of six firm-specific variables are significant for regressions on first-differenced data, which also includes stock return, leverage ratio, ROA and return on capital employed (ROCE). The signs of the first three variables are in line with previous literature. ROCE is a new metric for profitability not previously used, where our results show negative sign, same as ROA. While spot rate and QE are no longer significant, the model includes VIX from the macroeconomic variables. Interestingly, the Fama-French variables market excess return (MKT) and small-minus-big (SMB) are also significant in our model, in contrast to the findings of Galil et al. (2013). To the best of our knowledge, we are the first study to demonstrate the effect of Fama-French variables on CDS spread changes. Furthermore, by employing a quantile regression on first-differenced data, we can understand how the determinants impact credit risk in different regimes, as the higher quantiles represent regimes with large positive movements in credit risk. In line with previous research, we show that both volatility and leverage ratio has a more substantial effect in regimes with large positive movements (Pires et al., 2013). Moreover, we find that stock return has a non-linear effect on changes in CDS spreads, consistent with the result of Blasberg et al. (2021). Finally, the quantile regression shows how the determinants have a more substantial impact in times of large positive movements, such as in periods of crisis.

The remainder of the thesis is organized as follows. Chapter 2 presents the relevant literature. Chapter 3 contains an overview of the variables, data cleaning process, and descriptive statistics of our final dataset. Chapter 4 explains the methodology used to examine the determinants of CDS spreads. Chapter 5 present and discusses the results. Finally, Chapter 6 concludes the thesis.

Chapter 2

Literature review

The CDS was first introduced in 1994 as a means to allow an investor to offset or "swap" his or her credit risk with another investor. Since then, several studies have examined the impact and importance of both traditional and non-traditional variables on CDS spreads. The increased global attention to CDS after the great financial crisis in 2008 and the European sovereign debt crisis in 2012 resulted in an uptake in CDS literature, motivated by the need to better understand the financial product. Although many years of research have passed, there is still a lack of consensus regarding the impact of several determinants.

To the best of our knowledge, Blanco et al. (2005) is the first study to investigate the determinants of CDS spreads. The motivation behind the study is to understand whether theoretical determinants impact CDS spreads and bond prices differently. The variables included in the study are the interest rate, the slope of the yield curve, and market- and firm-specific volatility and returns. However, Blanco et al. (2005) gets inconsistent results for coefficients across the models found. Additionally, the small sample of investment-grade corporate and bank CDS spreads from 2001 to 2002 leaves different avenues open for further research. Since Blanco et al. (2005), several studies have explored determinants for CDS spreads, with diverging results. Table 2.1 presents an overview of studies.

Ericsson et al. (2009) expanded on the research of Blanco et al. (2005) by including all graded corporate CDS and a more updated dataset, and is one of the first contributions to verify that commonly used theoretical variables explain the CDS spread. The study was motivated by the fact that variables that in theory should explain credit risk had shown limited explanatory power on corporate bond data in previous research. The study finds that firm leverage and asset volatility increase the CDS spread, while term-structure slope and spot rate reduce it. In a similar study of theoretical CDS determinants, B. Y. Zhang et al. (2009) finds that equity volatility and equity jumps explain 32% in level data of credit spreads. In contrast to Ericsson et al. (2009), the study concludes that these variables explain most of the variation in the CDS spreads.

In an attempt to bring more structure to the pool of determinants discovered by early research, Das et al. (2009) is the first study to examine whether one category of variables explains more of the variation in the CDS spread. The study divides the theoretical variables explaining CDS spreads into accounting-based and market-based. While accounting variables are firm-specific, market-based variables contain both firm-specific (in terms of stock values) and macroeconomic variables. Das et al. (2009) concludes that a model using accounting-based variables explains the

	No.	#CDSs	Period	Frequency	Geography	Levels/Returns	Model
Blanco et al. (2005)	[1]	119	2001-2002	Daily	US/Europe	Returns	Panel regression
Ericsson et al. (2009)	[2]	NA	1999-2002	Daily	US	Levels/Returns	Series of regressions
B. Y. Zhang et al. (2009)	[3]	307	2001-2003	Weekly	US	Levels/Returns	Panel regression
Das et al. (2009)	[4]	506	2001-2005	Quarterly	US	Levels	Panel regression
Pires et al. (2013)	[5]	260	2002-2007	Monthly	US/Europe	Levels	Quantile regression
Galil et al. (2013)	[6]	695	2002-2013	Monthly	US	Returns	Panel regression
Chalamandaris and Vlachogiannakis (2016)	[7]	467	2005-2014	Daily	Worldwide	Returns	LASSO/Panel regression
Hu et al. (2017)	[8]	231	2005-2011	Yearly	US	Levels	Panel regression
Tang and Yan (2017)	[9]	861	2002-2009	Daily	US	Returns	Panel regression
Kim et al. (2017)	[10]	384	2004-2012	Monthly	US	Returns	Panel regression
Fonseca and Gottschalk (2018)	[11]	85	2007-2010	Weekly	Asia/Pacific	Returns	Panel regression
Lecce et al. (2018)	[12]	330	2006-2008	Weekly	US	Levels/Returns	Panel regression
Pereira et al. (2018)	[13]	704	2005-2012	Quarterly	US/UK/EU	Levels	Panel regression
Xinjie Wang et al. (2018)	[14]	448	2001-2016	Monthly	US	Levels	Panel regression
Blasberg et al. (2021)	[15]	212	2013-2018	Daily	US/Canada	Returns	Quantile regression
Barth et al. (2021)	[16]	470	2007-2019	Monthly	US/Europe	Levels	Panel regression
Kölbel, Leippold et al. (2020)	[17]	323	2010-2018	Monthly	US	Returns	Panel regression
Our study		240	2010-2019	Monthly	US/Europe	Levels/Returns	LASSO/Panel/Quantile regression

Table 2.1: Overview of literature on CDS determinants.

spread at least as well as a market-based model. On the other hand, the paper also concludes that a model using both categories of variables performs best in explaining the CDS spread. As a result, subsequent CDS literature generally uses both market-based and accounting-based variables when analyzing determinants.

Bringing the discussion to firm-specific variables, several studies build on the work of Das et al. (2009) and Ericsson et al. (2009) and verify that leverage ratio and stock volatility have a positive significant relationship with the CDS spread (Galil et al., 2013, Tang and Yan, 2017, Kim et al., 2017, Xinjie Wang et al., 2018 and Barth et al., 2021). Leverage ratio is defined as debt divided by the sum of debt and equity, and the positive sign can be explained by the fact that a firm defaults when its leverage ratio reaches one. For stock volatility, firms with high stock volatility are perceived as having a higher credit risk due to the increased probability for the firm value of going below the default barrier (Han and Zhou, 2015). However, Blasberg et al. (2021) finds stock volatility to be negatively correlated with the spread in lower quantiles and positively correlated in higher quantiles. As Blasberg et al. (2021) uses differenced data, this indicates that in credit regimes with large negative movements in CDS spread, stock volatility is negatively related to the CDS spread. Furthermore, stock returns are demonstrated to be negatively correlated with the CDS spread in most studies (Das et al. (2009), Pires et al. (2013), Galil et al. (2013), Fonseca and Gottschalk (2018), Tang and Yan (2017), Blasberg et al. (2021)). However, Barth et al. (2021) is one of the more recent contributions from literature, and finds stock return to be not significant.

Despite, to a large extent, consistent findings for the aforementioned firm-specific variables, the results regarding the firm size and ROA diverge substantially. While Das et al. (2009), Pires et al. (2013), Xinjie Wang et al. (2018) and Barth et al. (2021) demonstrate a negative sign for firm size, Hu et al. (2017), Tang and Yan (2017) and Pereira et al. (2018) find a positive sign. The difference in results could be explained by the proxy used for firm size. The three studies finding a positive sign consider total assets, while the majority of the studies finding a negative sign use market capitalization. The only study that also uses total assets and finds a negative sign is Das et al. (2009). This result could, however, be explained by the fact that Das et al. (2009) is not using the log-transform of the proxy, which most other studies have done. Another study finding a negative sign, such as Tang and Yan (2017), uses daily frequency and first-differenced data for total assets, which is only reported quarterly. Linearly interpolation in this case leads to very small changes in the variable, which might affect the validity of the results. Total assets on the balance sheet may not necessarily show the full value of the firm due to large intangible assets, which makes market capitalization a more accurate proxy for firm size. Moreover, literature disagrees on whether ROA is a significant variable explaining the CDS spread. While Tang and Yan (2017) and Kölbel, Leippold et al. (2020) find it not significant, Das et al. (2009), Hu et al. (2017), Pereira et al. (2018) and Barth et al. (2021) find ROA to have a significant negative impact on the CDS spread.

Bringing the discussion to macroeconomic variables, following Blanco et al. (2005), studies throughout the years have verified the negative sign of the spot rate using different spot rates from 3 year to 10 year (Das et al., 2009, Galil et al., 2013, Hu et al., 2017, Pereira et al., 2018, Barth et al., 2021). Kölbel, Leippold et al. (2020) finds a positive spot rate but argues that the relationship should be non-linear and demonstrate a negative sign when using the square of spot rate. On the other hand, B. Y. Zhang et al. (2009) argues that the effect of the spot rate is ambiguous. A negative sign might come from the theoretical standpoint that a higher spot rate increases the risk-neutral drift of the firm value process, which reduces the probability of default (Longstaff and Schwartz, 1995). On the other side, an increased spot rate might also reflect a tightened monetary policy stance which increases the probability of default (B. Y. Zhang et al., 2009).

Continuing the discussion with macroeconomic variables, VIX is a well-known metric for market volatility and is primarily consistent with having a positive sign on the CDS spread. A higher VIX is viewed as an increased risk in the market (B. Y. Zhang et al., 2009 and Bai and Wu, 2016). However, Galil et al. (2013) finds it not to be significant and Kölbel, Leippold et al. (2020) gets the counter-intuitive result that the VIX has a negative impact on the CDS spread, and thus leaves it out of the regression. In contrast to other studies, Kölbel, Leippold et al. (2020) uses one proxy for asset volatility and includes either stock volatility or the VIX. In our study, we include both variables as the stock volatility and market volatility may capture different effects. For example, stock volatility can show differences between firms in different sectors and countries, where VIX is a global variable equal to all firms.

The macroeconomic variable with the most diverging results is the term-structure slope. While Hu et al. (2017), Ericsson et al. (2009), B. Y. Zhang et al. (2009), Tang and Yan (2017), Kim et al. (2017), Fonseca and Gottschalk (2018) and Lecce et al. (2018) find a negative sign, Galil et al. (2013) finds it not to be significant, and Barth et al. (2021), Blanco et al. (2005), Chalamandaris and Vlachogiannakis (2016) find a positive sign. In these studies, the negative sign is explained by an increase in the yield curve anticipating improved economic growth, while the positive sign comes from the steepening of the slope may reduce the number of projects with a positive net present value (Galil et al., 2013). While the studies finding a negative sign only include US or Asian companies, the studies demonstrating a positive sign also include European companies, which might impact the results.

In contrast to fundamental macroeconomic variables, to the best of our knowledge, quantitative easing (QE) has never been used in studies on determinants of corporate CDS spreads. FED introduced QE in 2008 as a response to fight the great financial crisis. Thus the data on QE, which historically has been increased during periods of crisis, was unavailable for the first studies on CDS determinants. Among the few studies examining this variable is Albu et al. (2013) that finds that ECB's QE affects sovereign CDS spreads by using an ARMA-GARCH model for abnormal returns on CDS spreads. However, they did not find a significant directional effect. Moreover, Ajer and Bohler (2021) finds QE to reduce the CDS spread for banks because high liquidity in the banking sector intuitively reduces default risk for banks. However, Ajer and Bohler (2021) uses a rolling regression on the credit spread mean, which does not account for potential unobserved heterogeneity between each CDS as opposed to, for example, a fixed-effect panel regression. Another new variable that was considered a new determinant is inflation. In a study using corporate bond yield as a metric for credit risk, Kang and Pflueger (2015) finds the uncertainty about the long-run price level and the changing relationship of inflation with the business cycle to be determinants of credit risk. To the best of our knowledge, our study is the first to include QE and inflation in a study on corporate CDS.

Although the majority of literature on CDS determinants is centered around the early established categories firm-specific and macroeconomic variables, some studies have expanded the search beyond traditional fundamental variables to find other variables explaining the CDS spread. Motivating the search, Tang and Yan (2017) demonstrates that market-based and firm-specific variables altogether only account for 40% variations in CDS spread changes and concludes that much of the variation can be found in previously omitted variables. For example, the Fama-French variables extend the traditional CAPM as risk-return asset pricing, considered to be a richer model (Womack and Y. Zhang, 2003). Theory suggests that the Fama-French factors should be determinants of credit risk, as higher factors indicate better economic conditions and, therefore, lower credit spreads (Galil et al., 2013). Galil et al. (2013) is the first to examine the impact of the Fama-French factors on CDS spreads. By only using the factors in a fixed-effects panel regression, the study finds them to explain 16% of the CDS spread changes. However, when adding other control variables, the factors become not significant. In contrast, Barth et al. (2021) finds the excess return on the market to be significant with a negative sign, thus in line with the theoretical expectations. An interesting aspect of the Fama-French variables is that they are reported on a monthly basis by subtracting the value at the end of the month from the value at the beginning of the month. In this way, the Fama-French factors resemble first-differenced data. Previous studies vary between using level data and first-differenced data, which might impact the results from regressions.

The search for new determinants has also been impact by historical events. Among the influential events that had important implications for the pricing of credit risk was the Paris Agreement in 2015 (Delis et al., 2018, Kölbel, Leippold et al., 2020). The agreement increased the awareness of the risk associated with a company's sustainability profile and led to an uptake in the literature regarding the link between sustainability and credit risk. Here, literature has found two explanations; over-investment and risk-mitigation. The over-investment view argues that acting sustainable requires substantial investments that are costly diversions of firm resources. Thus, investments will cause lower or more volatile cash flows and higher credit risk for the firm (Goss and Roberts, 2011). On the contrary, the risk mitigation view argues that acting sustainable reduces company risk (Goss and Roberts, 2011). In other words, the over-investment view would imply a positive sign, while the risk-mitigation view would imply a negative sign.

Supporting the over-investment view, Flammer (2015) finds *decreasing marginal returns* on increased investments in Corporate Social Responsibility (CSR) related activities. When a firm has reached a level of CSR, adding more will not necessarily give the same returns as previously encountered. In contrast, supporting the risk-mitigation view Barth et al. (2021) demonstrates a negative link between Environmental, Social, and Governance (ESG) factors and CDS spreads. The study finds ESG to be a more important determinant for European companies than US companies, justified by European financial benefits linked to sustainability such as higher sustainability focus, investor protection, and transaction costs (Barth et al., 2021). Focusing on environmental risk solely, Blasberg et al. (2021) demonstrates that higher yearly emission levels increase the CDS spread of a company. In addition, the study provides evidence that firms with higher emissions have a higher cost of capital and support the view that many investors require compensation for holding "brown" companies instead of green companies. Bringing the discussion to the communication of sustainability, Kölbel, Leippold et al. (2020) investigates whether disclosing transitional and physical climate risk in 10-K reports impacts the CDS spread of a company. The study reveals that disclosing transitional risk increases the spread. In contrast, disclosure of physical climate risk decreases the spread, as it removes some of the unobserved risks in the credit risk premium. Focusing more on sustainability controversies, Kölbel, Busch et al. (2017) investigates the relationship between negative media items regarding sustainability and CDS spread. The study proves that negative media items significantly increase credit risk for companies, the level depending on severity and media coverage. On the other hand, Naumer and Yurtoglu (2019) demonstrates that companies with larger volumes of positive ESG-related news face lower CDS spreads.

With the growth of the literature expanding the search for CDS determinants, there have also been developments in the methodologies applied. Early studies, such as Ericsson et al. (2009) used average regression coefficients from a series of time-series regressions, one for each company.

However, this approach did not fully exploit the information of a panel data structure. Thus the literature shifted towards using the fixed-effects panel regression, which allows for unobserved heterogeneity in the dataset, thus a unique intercept for each CDS (Barth et al., 2021). The methodology is used by the majority of studies discussed and presented in Table 2.1. Moreover, quantile regression is used by Pires et al. (2013), as quantile regressions provide more nuanced insights. Pires et al. (2013) shows that the coefficient for implied volatility is increased for higher quantiles. Following Pires et al. (2013), Blasberg et al. (2021) also employs a quantile regression framework, showing that stock volatility has a negative sign for lower quantiles and a positive sign for higher quantiles.

To summarize, there is a large body that investigates the importance of new determinants of CDS spreads. However, there is a lack of studies that address the issue of identifying a relevant set of input variables among those already found to be important in previous studies. One study close to do this, Chalamandaris and Vlachogiannakis (2016), uses LASSO and panel-consistent estimations to investigate whether financial ratios are used in the decision-making process of CDS traders. However, the study does not bring in the pool of determinants but focuses solely on financial ratios. While LASSO selects 18 of the 43 ratios, the panel-consistent estimation models only result in 5 being significant. Thus, although several financial ratios are found significant in explaining the CDS spread in previous literature, one does not need more than 5 out of 43 to describe the CDS spread.

Apart from Chalamandaris and Vlachogiannakis (2016), no studies have looked at reflecting on the findings in the literature systematically and identifying potential redundancies. Our goal is to include all variables found in the literature (see Table 2.2) in a more comprehensive setting to examine whether some determinants are more important than others. Additionally, different variables may have similar explanatory power, enabling us to identify any redundancies in previous literature findings. Therefore, we thoroughly analyze the determinants through an extensive literature review and update variables, proxies, and datasets.

Variables	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]	[17]
Sustainability-related																	
ESG Controversies																	
Environmental risk															-		-
Social risk																	
Governance risk																	
Firm-specific																	
Stock return				-	-	-			-		-				-	N.S	
Stock volatility				+		+			+	+	+			+	+	+	+
Debt to EBITDA							+										
ROA				-				-	N.S				-			-	N.S
Leverage ratio		+	+		+	+		+	+	+	+	+	+	+		+	+
Firm size				-	-			+	+				+	-		-	
Macroeconomic																	
VIX			+			N.S			+	+		+		+			
Inflation																	
Spot rate	-	-	-	-		-	-	-	-	-	-	-	-			-	+
Term-structure slope	+	-	-			N.S	+	-	-	-	-	-				+	
Quantitative easing																	
Fama-French																	
HML						N.S											
SMB						N.S											
MKT						N.S										-	

Table 2.2: Overview of variables in previous literature. "+" represent a significant positive relationship between the variable and CDS spread, "-" a significant negative relationship, and "N.S" a not significant relationship.

Chapter 3

Data

In order to identify redundancies in existing literature, we use an extensive dataset consisting of firm-specific, sustainability-linked, macroeconomic, and Fama-French variables. This chapter begins with Section 3.1 describing the dependent variable in our data set, while the independent variables are presented in Section 3.2. Further, Section 3.3 describes the data cleaning process and shows descriptive statistics of the final dataset.

3.1 Dependent variable

The monthly corporate CDS spread is the dependent variable in this study and is gathered from Refinitiv Eikon. We use CDS with 5-year maturity as it is the most liquid (Pereira et al., 2018). Each CDS is denominated in U.S. Dollars and follows senior unsecured issue seniority. For the restructuring event specified in the CDS, we use the Modified Restructuring (MR) and Modified-Modified Restructuring (MM) from the 2014 ISDA protocol (ISDA, 2020). The two clauses are viewed to be similar in terms of structuring a CDS, with the most notable difference being geography. We exclude banks and firms categorized as financial firms, as these are highly regulated.

3.2 Independent variables

The first group of independent variables is the sustainability-linked. We use Refinitiv Eikon for the ESG pillars, namely the Environmental, Governance, and Social pillar. What constitutes the different pillars are described in Appendix A. Furthermore, we also include the ESG controversy score, which is calculated based on 23 ESG controversy topics and addresses the fact that large companies attract more media attention than smaller companies (Refinitiv, 2021).

The second group of independent variables is the firm-specific, namely stock return, stock volatility, debt to ebitda, price to cash flow per share, ROA, ROCE, leverage ratio, firm size, cash to asset and revenue growth. To calculate stock return and stock volatility, we extract the daily closing price for each firm. We use net income over total assets as a proxy for ROA, consistent with previous studies (Das et al., 2009). We define the other profitability metric, ROCE, as EBIT divided by capital employed, where capital employed is total assets minus current liabilities. Cash to assets represents the amount of cash divided by total assets. Revenue growth is the year-over-year change in total revenue. Leverage ratio follows the most common proxy used in the literature, calculating total debt over total debt plus market capitalization. It gives the percentage of debt in the capital structure, where a higher ratio would imply higher leverage. Debt to EBITDA is not often included but is another metric for leverage (Chalamandaris and Vlachogiannakis, 2016). It is calculated as total debt over the last twelve months' EBITDA. The last ratio, price to cash flow per share, is a stock valuation indicator. The proxy used is the stock price divided by firm cash flow per share.

The third group are the macroeconomic variables, namely VIX, inflation, spot rate, term structure slope, and QE. The most commonly used macroeconomic variable in literature is the risk-free rate. We follow Ericsson et al. (2009) and Blasberg et al. (2021) and proxy the rate by the 10-year maturity. We use the 10-year Constant Maturity Treasury rate for U.S. firms and the ECB Government 10-year rate for European firms, consistent with Blasberg et al. (2021). Kölbel, Leippold et al. (2020) argues for a non-linear relationship for the spot rate, and thus we include the squared spot rate. Furthermore, we proxy the yield curve's slope as the 10-year rate minus the 2-year rate, using the rates of the respective geographies as described above. We also use the inflation rates from the CDS' respective geography. The U.S. data mentioned above is collected from St. Louis Federal Reserve (FRED)[1], while European data has been collected from the statistical data warehouse by the European Central Bank (ECB)[2]. For market volatility, we use the VIX, defined as the CBOE Volatility Index. Due to the significant increase in quantitative easing (QE) in the last ten years, we also include QE as a variable. The variable is defined as total assets on FED's balance sheet.

Finally, the last category of variables is the Fama-French variables, known for expanding on the traditional Capital Asset Pricing Model (CAPM) (Fama and French, 1989). There are three variables, called market excess return (MKT), Small-Minus-Big (SMB) and High-Minus-Low (HML). All data has been downloaded from Kenneth R. French's homepage[3], which is collected from CSRP.

Table 2.2 presents a summary of the variables presented in the chapter, including a variable description, signs in literature, and the source of the data.

[1]https://fred.stlouisfed.org/

[2]https://sdw.ecb.europa.eu/

[3]https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data$_library.html$

Name in model	Variable	Variable description	Sign (literature)	Source
Sustainability				
ESGC	ESG Controversies	Refinitiv Eikon ESG Controversy Score from 0 to 100		Refinitiv Eikon
ENV	Environmental risk	Refinitiv Eikon Environmental Score from 0 to 100	-	Refinitiv Eikon
SOC	Social risk	Refinitiv Eikon Social Score from 0 to 100		Refinitiv Eikon
GOV	Governance risk	Refinitiv Eikon Governance Score from 0 to 100		Refinitiv Eikon
Firm-specific				
RET	Stock return	Average monthly stock return	-/N.S.	Refinitiv Eikon
VOL	Stock volatility	Monthly mean of annualized 90-day rolling window	+	Refinitiv Eikon
DTE	Debt to EBITDA	Debt over EBITDA	+	Refinitiv Eikon
PCFS	Price To Cash Flow Per Share	Share price divided by cash flow per share		Refinitiv Eikon
ROA	Return on assets	Net income over total assets	-/N.S	Refinitiv Eikon
ROCE	Return on capital employed	EBIT over capital employed		Refinitiv Eikon
LEV	Leverage ratio	Total debt over total debt plus market capitalization	+	Refinitiv Eikon
SIZE	Firm size	log(market capitalization)	-/+	Refinitiv Eikon
CASHTA	Cash to Assets	The sum of cash and cash equivalents divided by current liabilities		Refinitiv Eikon
REVGROW	Revenue Growth	Percentage growth in revenue		Refinitiv Eikon
Macroeconomic				
VIX	VIX	CBOE Volatility Index	+/N.S	FRED
INF	Inflation	FRED Inflation rate for US and ECB inflation rate for Europe		FRED and ECB
RATE	Spot rate	10Y Constant Maturity Treasury rate for US and Government 10-Year rate for Europe	+/-	FRED and ECB
SLOPE	Term-structure slope	10-year rate minus the 2-year rate using FRED and ECB respectively	+/-/N.S	FRED and ECB
QE	Quantitative easing	Total assets on FEDs balance sheet		FRED
Fama-French				
HML	HML	Small Minus Big	N.S	Kenneth R. French
SMB	SMB	High Minus Low	N.S	Kenneth R. French
MKT	MKT	Excess return of the market	N.S/-	Kenneth R. French

Table 3.1: Overview of independent variables used in our model. "+" represent a significant positive relationship between the variable and CDS spread, "-" a significant negative relationship, and "N.S" a not significant relationship.

3.3 Data preparation and final dataset

After downloading all CDS's which match our search described in Section 3.1, we employed a cleaning and validation process to ensure our final results are reliable. The initial dataset consisted of 539 CDS's, with daily data from January 2010 to December 2019. We only include CDS with over 2000 data points in the extraction, thus having a sufficient time period and liquidity for our analysis. Further, to ensure our results are not affected by outliers or data errors, we remove CDS which for one moment in time trades above 4000 bps. This is consistent with the approach of Barth et al. (2021) and B. Y. Zhang et al. (2009). Furthermore, as firm-specific data are necessary for our analysis, we filter out the CDS's which do not have a primary equity identifier connected to them. Reasons for the lack of an identifier are either that the firm related to the CDS is not listed or that Refinitiv Eikon does not have the data. We end up with 353 CDS's after this initial cleaning and organize them into a panel structure based on the monthly mean of each CDS spread.

The ESG data is updated yearly, and the firm-specific variables from company reporting are quarterly. Thus, we interpolate the values missing between the quarters as we apply a monthly regression. We use linear interpolation, in line with previous studies such as Galil et al. (2013) and Ericsson et al. (2009). After merging the ESG data and firm-specific data with the CDS spread panel data, we are left with 240 CDS's for which we have no missing data. Macroeconomic and Fama-French variables have no missing values, and most of the data are reported monthly, consistent with our panel data structure. The interest rate and yield curve for U.S. data and the VIX are reported daily, thus we use the monthly mean of these variables.

In our regression, we do not use the stock price but the average stock return for the month. As stock price itself is highly correlated with market capitalization, and given the common proxies for stock return used by, e.g. Das et al. (2009), we view this as the most suitable way to include stock return. More technically, we take a monthly mean of daily returns for the stock. For first-differenced data, we use the percentage change of month-to-month stock price. Stock volatility is modeled directly from the daily stock prices, where we annualize 30-day rolling window volatility from past stock prices for each day. To align it with our monthly panel structure, we take the monthly mean. Our proxy is consistent with previous studies such as Pereira et al. (2018), but we use a shorter time horizon for the rolling window to better reflect short-term changes.

First-differenced data requires more data preparation than level data. We take the month-to-month percentage change in value for the CDS spread, stock return, and market capitalization variables. For the rest of the variables, except the Fama-French variables, we use absolute difference, consistent with studies such as Barth et al. (2021) and Collin-Dufresne et al. (2001). The Fama-French variables are equal for both data types, due to the nature of the factors resembling first-differenced data when downloaded.

Geography	**#CDS's**
North America	164
Europe	71
Other Countries	5

Table 3.2: Overview of CDS Geography

An overview of the geographical segmentation of our observations is presented in Table 3.2. As can be seen, our final dataset consists of 240 CDS's with a broad geographical cover. The majority of our observations are from North America, including the United States and Canada. The CDS's from Europe come from 13 countries, where most of the CDS's are from Germany. The remaining five CDS's are from South Korea, South Africa, Brazil, and Israel.

Statistic	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
CDS	4.473	0.796	2.417	3.937	4.931	8.224
ESGC	74.298	33.232	0.781	50.000	100.000	100.000
ENV	64.131	23.124	0.000	50.978	81.645	98.546
GOV	63.544	19.734	3.117	50.227	79.265	98.507
SOC	67.147	20.256	4.422	55.010	83.101	98.470
RET	0.007	0.064	−0.429	−0.027	0.042	1.262
VOL	0.228	0.111	0.000	0.154	0.272	2.163
DTE	2.608	14.790	0.00002	1.069	3.094	1,601.357
PCFS	12.266	61.678	0.057	5.595	13.274	5,712.146
SIZE	23.768	1.549	18.052	22.837	24.689	31.801
ROA	1.233	2.015	−32.194	0.545	1.945	37.832
LEV	0.317	0.192	0.000	0.167	0.436	0.977
ROCE	0.030	0.028	−0.409	0.015	0.040	0.497
CASHTA	0.068	0.058	−0.319	0.025	0.093	0.534
REVGROW	0.139	12.235	−14.846	−0.040	0.064	1,269
RATE	2.316	0.904	0.048	1.707	2.889	4.661
SLOPE	1.331	1.240	−9.813	0.946	1.948	2.834
INF	1.620	0.894	−0.600	1.100	2.200	3.900
VIX	16.861	5.153	10.125	13.486	19.048	36.530
QE	3,738,939	794,793	2,256,424	2,870,242	4,462,752	4,507,150
MKT	1.091	3.729	−9.550	−0.917	3.400	11.350
SMB	−0.012	2.244	−4.820	−1.862	1.195	5.670
HML	−0.198	2.282	−4.950	−1.865	1.105	8.210

Table 3.3: Descriptive statistics of dataset on level data before scaling

The descriptive statistics of our dataset are described in Table 3.3 and the correlation between the variables in the dataset is described in Appendix B. The descriptive statistics are calculated after applying the logarithms on CDS spread and market capitalization but before scaling all the variables. As we can see from market capitalization, the lower percentile has a factor of 22.8 or a market cap of 8 billion dollars. The value implies that our dataset is somewhat skewed towards larger firms. The scaling is done before we run the regressions in order to compare the coefficients. The correlation matrix shows that none of the explanatory variables are highly correlated.

Chapter 4

Methodology

This chapter describes our modeling techniques. First, Section 4.1 introduces panel data and the fixed-effects panel regression. Further, we describe our primary methodology by introducing the LASSO panel regression and how we use the Bayesian Information Criteria (BIC) to choose the most suitable penalization effect in the LASSO. Finally, Section 4.2 describes how we used a quantile regression on the variables chosen by the LASSO estimator.

4.1 LASSO Panel Data Regression

Our model uses monthly data from January 2010 to December 2019, resulting in a time dimension of 144, and 240 CDS's make the space dimension. The main methodological novelty of our application is that we combine LASSO as a variable selection method with a fixed-effect panel regression.

In our model, we employ a fixed-effects (FE) model, which allows the intercept to be firm-specific. In our case, that is an intercept for each CDS employed. In the most simple form, the FE panel regression equation can be expressed as described in Equation (4.1).

$$ln(CDS_{i,t}) = \boldsymbol{\beta}\mathbf{X_{i,t}^T} + \mu_i + \epsilon_{i,t} \tag{4.1}$$

$CDS_{i,t}$ represents the CDS spread of CDS i at time t. The $\boldsymbol{\beta}$ is a vector of all independent variables' coefficients, as described in Section 3.2. $\mathbf{X_{i,t}}$ is a vector representing all observations related to a given CDS i at time t. μ_i is the intercept for CDS_i, which can vary across each CDS and thus represents the unobserved heterogeneity in the dataset. The differences are caused by variables such as industry, country and management. $\epsilon_{i,t}$ is the error term for CDS i at time t.

We view LASSO as a suitable data-driven approach to select the determinants that best capture the variability of our dependent variable. There are several reasons for this choice; (1) by penalizing the sum of the variates, we can shrink the determinants where LASSO will put some of the coefficients to zero, and (2) it deals effectively with possible multicollinearity issues (Tibshirani et al., 2013). Chalamandaris and Vlachogiannakis (2016) employs LASSO as it allows "to start from a very large number of explanatory variables and look for the optimal linear specification among those defined by their combination". LASSO can be expressed to minimize the sum of squared errors (SSE) as in the equation below

$$SSE = \sum_{i=1}^{N}\sum_{t=1}^{T}\hat{\epsilon}_{i,t}^2 + \lambda\sum_{j=1}^{K}|\beta_j| \tag{4.2}$$

where N is the number of CDS's, T is the time dimension, and K is the total number of variables. β_j is the coefficient of the jth independent variable and λ the chosen penalty parameter. When increasing λ, the sum of the objective function will increase. $\hat{\epsilon}_{i,t}$ denotes the estimated error term. In a minimization problem, the penalty will shrink the coefficients. At the same time, the sum of squared errors will increase; thus, the LASSO makes a trade-off between increasing the error term and decreasing the coefficients.

We use information criteria to select the most suitable penalty parameter, λ. We choose the Bayesian Information Criterion (BIC), which is expressed as

$$BIC = kln(n) - 2ln(\hat{L}) \tag{4.3}$$

where k is the number of parameters estimated, n is the number of observations, and $\hat{L}$ is the maximized value of the likelihood function of the model. The likelihood function is determined as $\hat{L} = p(x|\boldsymbol{\beta})$, where $\boldsymbol{\beta}$ is the parameter values that maximize the likelihood function. Among the specifications with different values of λ (in our case 100), the one with the lowest BIC is the preferred. Another popular information criteria is the Akaike Information Criterion (AIC). AIC might have a tendency to choose a more complex model, whereas BIC will favor a more concise model. In our study, we therefore find it most relevant to use BIC.

For our results, we first employ panel LASSO to shrink non-significant variables to zero. In this way, we can say that LASSO "choose" variables shown to have a significant effect on the dependent variable. Then, in order to report R^2 and significance level, we use a regular fixed-effects model for the variables selected by LASSO, using Equation (4.1). We report the R^2 and not adjusted R^2, as the penalization for increasing the amount of variables are very small due to the relation between our large panel data set and relatively smaller amount of predictors. Comparing models on adjusted R^2 would thus not make any significant difference from comparing R^2.

4.2 Quantile Panel Regression

To provide further nuance to our result, we include a quantile panel regression for the variables selected by LASSO. Mandelbrot and Hudson (2004) illustrates how the tails of the distributions can have an extraordinary effect on the entire financial system. Compared to OLS, quantile regression will give the conditional median and quantiles of the distribution instead of the mean. In addition, quantile regression makes the method more robust to outliers, and it does not assume any distribution for the target variable.

Quantile regression does not divide the data into subsets but rather adds different weights to the data based on the quantile for the dependent variable. The model used is a quantile panel regression with fixed effects as first proposed by Koenker (2004). The conditional quantile function for the CDS i at time t can be written as

$$Q_{ln(CDS_{i,t})}(\tau|x_{i,t}) = \boldsymbol{\beta}_{\tau}\mathbf{X_{i,t}^T} + \mu_{\tau,i} + \epsilon_{i,t} \tag{4.4}$$

where i represents CDS i and t is time-period t, for $i = 1, ...N, t = 1, ..., T$, respectively. τ is the chosen quantile, thus a number between 0 or 1, where we use {0.1, 0.25, 0.5, 0.75, 0.9} as quantiles. The quantile regression will produce a series of quantile coefficients, in our model shown as β_τ, one for each quantile. This will give the conditional distribution of CDS for each control variable. Also, the firm fixed-effect μ will be different throughout the quantiles, where the firm-specific intercept will be constant but with a changing global intercept. That is, we write $\mu_{\tau,i} = \alpha_\tau + \mu_i$ for the firm-specific intercept.

The quantile regression have a different interpretation when employed on level data or first-differenced data. Intuitively, higher quantiles on the level data correspond to the CDS's with high CDS spread, or in other words, those CDS's which are deemed riskier by the market. We thus get the quantiles based on the riskiness of the firm. On the other hand, when we look at the change for the first-differenced data, the higher quantiles would be when the change in the CDS spread is the highest. Higher quantiles represent either when a firm has been riskier or if the whole market is in a high credit risk period. Using this interpretation, we can determine whether variables have a different effect in periods of increasing changes in credit risk than in decreasing changes in credit risk. The quantiles will then correspond to different economic periods, more specifically, upper quantiles denote crisis periods with high credit risk, while lower quantiles denote periods of economic tranquillity with low credit risk (Koutmos, 2019).

Chapter 5

Results and Discussion

This chapter presents and discusses our results. In Section 5.1 we present the result of a fixed-effect panel regression and quantile panel regression on the variables chosen by LASSO for level data. Our models investigate the effect variables might have on the CDS spread overall and for companies with different levels of credit risk. In comparison to level data, using regression on first-differenced data investigates how the changes in the variables affect the change in the CDS spread, which is explored in Section 5.2. The interpretation for the quantile regression is also different, as the upper quantile corresponds to a period in which the market experiences significant increases in credit risk, called credit regimes with positive movements. On the other side, lower quantiles respond to credit regimes with negative movements. Finally, Section 5.3 compares the results of the models on level and difference data.

5.1 Level data

To investigate variables affecting CDS spreads on level data, we perform LASSO regressions as specified in (4.2) on all variables presented in Table 3.1. The algorithm chooses the optimal λ by minimizing the BIC, where the λ penalizes the coefficients by shrinking them. When the λ is non-zero, LASSO ensures that the coefficients of the variables not important will be shrunk to zero and not "chosen". Of the 23 variables included, the algorithm chooses only seven variables to be considered. The included variables are governance pillar (GOV) and social pillar (SOC) from ESG variables, stock volatility and firm size from firm-specific variables, and interest rate, yield slope, and QE from macroeconomic variables. None of the Fama-French variables is chosen. We include the chosen variables in a fixed-effect panel regression as shown in (4.1). The result of this regression is presented in Table 5.1.

With only seven variables chosen in the full model, our study finds a parsimonious model that reduces the complexity of previously used models. The model combines traditional fundamental variables such as volatility and market capitalization but also includes new variables such as GOV, SOC, and QE. This combination of variables for explaining the CDS spreads has never been used in the literature before. The R^2 obtained is 0.475, which indicates that these variables account for 47.5% of the variation in the data. In comparison, when we run a regression using all 23 variables the R^2 is 0.490, which is only slightly higher. Our results thus indicate that for investors and banks that might value a concise model, our model is favorable as it explains close to the same amount of the CDS spreads.

	CDS
GOV	0.029***
	(0.003)
SOC	0.026***
	(0.003)
VOL	0.167***
	(0.003)
SIZE	−0.833***
	(0.011)
RATE	0.045***
	(0.003)
SLOPE	−0.043***
	(0.002)
QE	−0.116***
	(0.003)
Observations	26,418
R^2	0.475
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 5.1: Results of the fixed-effect model in (4.1) on level data. Variables used in regression are the ones previously chosen by LASSO.

The two ESG variables, GOV and SOC, are significantly different from zero. Both have a positive sign, which means a higher level of ESG indicates a higher credit spread. The use of ESG variables is one of the most recent additions to the pool of determinants suggested by the literature (Barth et al., 2021, Kölbel, Leippold et al., 2020). The fact that the governance pillar and social pillar have positive signs would imply that firms with higher scores on these pillars have a higher credit risk. Contrary to previous studies such as Kölbel, Leippold et al. (2020), we do not find a risk-mitigating effect for the ESG variables. Barth et al. (2021) did also investigate ESG, without dividing it into the separate pillars but did not get a significant result for the 5-year CDS spread in a complete model. Our results are, however, consistent with the over-investment hypothesis that investments in ESG are a drain on a company's sparse resources which should be redirected to other projects (Goss and Roberts, 2011). At the same time, one should notice the corresponding coefficients as they are relatively small compared to the other variables presented. The small coefficients indicate that while the level of SOC or GOV impacts credit risk, they do not impact the credit risk as much as the other variables. Moreover, the environmental pillar is not chosen by LASSO as an important variable for explaining the CDS spread. This result could indicate that the effect of environmental investments is captured in some of the other variables, or that the degree of environmental investments does not impact credit risk substantially.

The chosen firm-specific variables, stock volatility and firm size, are consistent with literature on sign and significance. Similar to Barth et al. (2021), in our model stock return is not significant on level data. Firm size is the most important determinant given the size of its coefficient, which implies a firm could reduce its credit risk by increasing its size. Literature has revealed diverging signs for firm size. However, our results indicate an apparent risk mitigation effect from this determinant.

For macroeconomic variables, the variables chosen are spot rate, yield slope, and QE. The positive sign for spot rate is the opposite of most previous studies. According to Collin-Dufresne et al. (2001), an increase in the spot rate reduces the probability of default. For a firm taking on projects, a higher rate should increase the internal rate of return (IRR) needed, thus reducing the riskiness of projects. However, in the current low-rate environment, an increase in the spot rate will increase the financing costs more than in high-rate environments, which might increase the risk. Also, higher rates tend to lower asset prices and reduce consumer spending, leading to a downturn in the business cycle and increasing credit risk. Given the sign in our model, although the coefficient is small, the latter effect seems to dominate.

The yield slope has a negative sign in our result, consistent with theory. The variable is often included as a business cycle predictor. For example, we would have an inverted yield curve when the yield slope is negative, meaning the 2-year rate is higher than the 10-year rate. An inverted yield curve is by many considered a recession sign, and a higher spread is thus warranted, meaning more soundness in the cycle.

QE is the last variable included in our model and, to our knowledge, only has been studied in the context of bank CDS (Ajer and Bohler, 2021). Consistent with Ajer and Bohler (2021) we have a negative sign for the variable. Interestingly, our study is the first to demonstrate that QE is a determinant for corporate CDS spreads and one of seven variables chosen in a feature selection method. The overall effect of QE is that it increases the supply of bank reserves in the financial system, with the hope that the lenders go on to pass that liquidity along as credit to companies and households, spurring growth. Another effect QE has is that it drives borrowing rates down, effectuated by increasing supply in the supply-demand interest rate calculation. Our negative sign can thus be explained by the increased credit for households and lower interest rates, which stimulate to more spending and thus larger demand in the economy.

Table 5.2 show the results of separate LASSO regressions for each group of the variables. First, we only use the four ESG variables in the LASSO regression before doing a fixed-effects panel regression on the variables chosen by LASSO. We then take the same approach with each group of variables. According to Galil et al. (2013), a comparison between the model with all regressors (Table 5.1) and models with subsets of explanatory variables (Table 5.2) reveals the potential collinearity among the explanatory variables. As we can see, none of the Fama-French variables are included, meaning this group of variables does not affect the CDS spread.

For ESG variables, the LASSO estimator chooses all variables. Although the R^2 is low, the variables are still shown to have a significant effect. While GOV and SOC have the same sign as the full model, it now includes the environmental pillar score (ENV) with a negative sign and the ESG controversy score (ESGC) with a small positive sign. A negative sign on the environmental pillar implies a lower credit risk for firms with a higher score on the environmental pillar. The result underscores the risk-mitigating effect of investments in factors related to the environment, such as less pollution and carbon intensity. It is interesting to see such a result even though the CDS's are only 5-year maturity, implying investors also see environmental risk in the short-term. The variable is significant in this model but not in the full model, meaning that other variables capture this effect. From Appendix B, ENV has a minor correlation on 0.12 with firm size. As previously research, there may exist an advantage to larger firms with more resources to communicate ESG practices and thus getting a higher ESG score (Drempetic et al., 2017). Since both ENV and SIZE has a negative sign, the environmental effect may be captured by the size of the firm.

	ESG	Firm	Macroeconomic	Fama-French
ESGC	0.020***			
	(0.004)			
ENV	−0.072***			
	(0.008)			
GOV	0.089***			
	(0.006)			
SOC	0.078***			
	(0.006)			
RET		–		
VOL		0.179***		
		(0.003)		
DTE		–		
PCFS		–		
SIZE		−1.036***		
		(0.011)		
ROA		–		
LEV		–		
ROCE		–		
CASHTA		–		
REVGROW		–		
RATE			0.426***	
			(0.015)	
$RATE^2$			−0.408***	
			(0.015)	
SLOPE			−0.062***	
			(0.003)	
INF			−0.011***	
			(0.003)	
VIX			0.095***	
			(0.003)	
QE			−0.208***	
			(0.004)	
MKT				–
SMB				–
HML				–
Observations	26,418	26,418	26,418	26,418
R^2	0.019	0.404	0.266	0
Note:			*p<0.1; **p<0.05; ***p<0.01	

Table 5.2: Results from fixed-effect panel regression as presented in (4.1) on level data for each group of variables separately. The variables included are the ones chosen by LASSO. Variables not chosen by LASSO have an ”-” in the table as they are not included in the regression.

ESGC has a significant positive sign, which might seem counter-intuitive. The ESGC variable gives a score of 100 if no media issues related to the ESG have appeared in the last year. Thus, a score of 100 would be deemed good. On the other hand, if a firm has media attention related to its ESG practice, the score could go down to zero, implying a low score would not be good for a firm. Intuitively, a score of 100 should reduce credit risk, as a higher value on levels should be better than a lower value. However, the ESGC variable has few changes, and most of the time, it has a score of 100 for each firm. Therefore, given the fact that the coefficient is relatively small and not significant when all variables are included, we do not emphasize the importance of this variable.

Comparing the R-squared for firm-specific and macroeconomic variables, we see that the model for firm-specific has better explainability power. This is consistent with Galil et al. (2013) and Ericsson et al. (2009) who concludes that firm-specific variables captured more of the variation. However, a full model is needed to capture most of the variation. This is consistent with findings from Das et al. (2009), which reports that models that make use of both accounting and marked-based variables explain a substantially larger portion of CDS spreads than one of them individually.

In this model, we can see that both the rate and the squared rate are significant. In other words, we give the rate a more prominent effect when it is higher, and we can now see a significant negative sign. This would say that the spot rate reduces the credit risk on level data in high-interest rate environments. Overall, we have seen a reduction in the spot rate from the beginning of 2010 until 2019. The reduction implies that the rate had a risk-reducing effect at the start of the decade when the rate was higher, but a more negligible risk-increasing effect at the end of the decade, with lower rates. An increase from the current low levels can further change the interpretation of this variable, as this is not something we have seen before. However, despite RATE^2 being significant in the macroeconomic model, LASSO does not choose the variable in the whole model.

To get a deeper understanding of the impact of the most important determinants, we run a quantile regression on the determinants chosen by the LASSO. The results are presented in Table 5.3. Interestingly, GOV varies in significance between the quantiles, and SOC is not significant in all quantiles. This shows how the result can be different when the regression uses the distribution of the conditional mean or the conditional quantile. Moreover, GOV is significantly positive for most quantiles and thus supports the over-investment view throughout the distribution. However, the coefficient is largest for the 10th and 90th quantile, thus indicating a non-linear relationship. This implies that the largest support for the over-investment is found in the most risky and the least risky firms. An explanation for this could be that companies facing a large amount of credit risk should not use money on investing in ESG, as they are likely to struggle with issues more crucial for credit risk such as volatility. On the other side, the least risky firms might redistribute money to shareholders or reduce debt instead of using scarce resources to improve governance.

Stock volatility has a positive sign, and the coefficient increases with higher quantiles, which is consistent with Pires et al. (2013). For example, while VOL had a coefficient of 0.167 using the conditional mean, we can see VOL changes from 0.112 in the lower quantile to 0.224 in the upper quantile, doubling the value. According to Pires et al. (2013), this shows that the conditional mean does not fully characterize a typical panel of heterogeneous CDS data. In other words, volatility leads to more credit risk for already risky firms.

QE has a significant negative sign on all quantiles, but as we can see, the coefficient is further decreasing in higher quantiles. This means that QE thus has a larger risk-reducing effect for the

	Quantiles				
	0.1	0.25	0.5	0.75	0.9
GOV	0.048*	0.025	0.036**	0.036**	0.039*
	(0.031)	(0.026)	(0.022)	(0.020)	(0.025)
SOC	0.007	0.017	0.019	0.026	0.032
	(0.031)	(0.025)	(0.023)	(0.023)	(0.032)
VOL	0.112***	0.131***	0.156***	0.185***	0.224***
	(0.014)	(0.012)	(0.012)	(0.015)	(0.019)
SIZE	−0.845***	−0.855***	−0.865***	−0.863***	−0.850***
	(0.068)	(0.068)	(0.070)	(0.070)	(0.073)
RATE	0.068***	0.050***	0.042***	0.039***	0.028*
	(0.019)	(0.014)	(0.013)	(0.013)	(0.016)
SLOPE	−0.011	−0.036***	−0.048***	−0.054***	−0.070***
	(0.022)	(0.011)	(0.007)	(0.011)	(0.018)
QE	−0.069***	−0.102***	−0.121***	−0.129***	−0.151***
	(0.017)	(0.015)	(0.016)	(0.018)	(0.021)

Note: *p<0.1; **p<0.05; ***p<0.01

Table 5.3: Results for quantile regression as presented in (4.2), on level data for the variables chosen by LASSO and included in model Table 5.1.

already more risky firms. The introduction and further increase of QE might have given the risky firms an artificial low credit risk. An important point here is that the implications of a reversal of QE might come sooner than expected, as the FED is now planning for quantitative tightening in mid 2022, thereby reducing the total level of FED's balance sheet (Councile, 2022). Thus, our results highlight the importance of investors and policymakers to understand that reversing QE will, ceteris paribus, increase the credit risk for firms in the period going forward.

5.2 Differenced data

In order to further examine the effect of determinants proven significant in literature, we perform the same analysis as in Section 5.1, but now on first-differenced data. The regressions on first-differenced data provide an understanding of the factors driving the variation in CDS spreads changes. We begin by employing the LASSO model presented in Equation (4.2) on all variables presented in Table 3.1. The variables chosen by LASSO, which are stock return, stock volatility, ROA, leverage ratio, ROCE, yield slope, VIX, MKT, and SMB, are then employed in a fixed-effect model, which results are presented in Table 5.4.

The combination of variables chosen by LASSO has not been used in literature before. Our study thus also contributes to the literature by suggesting a benchmark model. Interestingly, the R^2 for the model is 0.287, while the R^2 when running a model using all 23 variables is also 0.287. The concise model presented in Table 5.4 is thus able to explain the same variation as a considerably more complex model.

Not surprisingly, the ESG variables are not chosen by LASSO and thus not included in the fixed-effect model presented in Table 5.4. Although they are significant on levels, their changes are small and rare, which makes them unfit to explain the variation in the differenced data.

For the firm-specific variables, both stock return and stock volatility have signs similar to

	CDS
RET	−0.348***
	(0.006)
VOL	0.121***
	(0.005)
SIZE	−0.040***
	(0.008)
ROA	−0.013**
	(0.005)
LEV	0.052***
	(0.007)
ROCE	−0.018***
	(0.005)
SLOPE	−0.021***
	(0.006)
VIX	0.104***
	(0.007)
MKT	−0.109***
	(0.007)
SMB	0.065***
	(0.006)
Observations	28,560
R^2	0.287
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 5.4: Results for fixed-effect panel regression presented in Equation (4.1) on variables chosen by LASSO.

the literature (Galil et al., 2013, Tang and Yan, 2017, Fonseca and Gottschalk, 2018). The stock return has a substantially larger coefficient than the other variables. This indicates that stock return reduces the credit spread most compared to the other determinants. Stock volatility is significant, though smaller than stock return. However, its coefficient is also large, which indicates that stock volatility is an important determinant for CDS spread changes. The sign is positive, because higher volatility increases the probability for the firm value of going below the default barrier.

Despite firm size having both positive and negative signs in studies using level data, as far as we know, only Tang and Yan (2017) has included this variable on regression models for differenced data with diverging results. The study finds a positive relationship for investment-grade firms but a negative for high-yield firms. However, the study uses the quarterly reported total asset as a proxy for firm size and interpolates to get a daily frequency of the data, which might impact their results. A negative relationship in our model can be explained as larger firms better withstand negative shocks to their cash flow, which make them less likely to default (Goss and Roberts, 2011).

The coefficients of profitability ratios ROA and ROCE are both negative. Although ROA has been demonstrated to have a negative sign in studies on level data, studies on differenced data have found ROA not to be significant (Tang and Yan, 2017 and Kölbel, Leippold et al., 2020). Our findings indicate that increased profitability decreases the CDS spread, which aligns with theoretical expectations that more profitable firms have better financial conditions and are thus less likely to default. Continuing the discussion on firm-specific ratios, our results show that the leverage ratio is positively correlated to the changes in CDS spread. Higher leverage indicates a shorter distance to the default barrier and a higher probability of default (Merton, 1974), thus explaining a positive sign.

For the macroeconomic variables, both the term-structure slope and VIX have a positive sign. Although the term-structure slope has diverging results in previous literature, as discussed in Section 5.1, our result is consistent with the theoretical expectations. Furthermore, the positive sign of the VIX is also consistent with the literature, as a more volatile market indicates times of uncertainty in the economy, thus increasing the risk of all companies in the market.

In contrast to level data, the Fama-French factors MKT and SMB are chosen by LASSO and significant in the panel regression on differenced data. However, the Fama-French factors have, as far as we know, only been included in Galil et al. (2013) among the studies using differenced data. The study finds all three Fama-French factors not significant, despite the variables being theoretically acknowledged as important economic factors and determinants of credit risk. MKT is negatively correlated with the CDS spread, which can be explained by a high market excess return often being related to good economic times. The SMB factor is positively correlated with the CDS spread changes. This indicates that when the returns of small companies outperform large companies, the credit risk in the market increases. The result might seem counter-intuitive since SMB is usually an indicator of a positive business cycle (Galil et al., 2013). However, our sample consists of companies skewed towards large capitalization. The firms included in our dataset will thus see increased risk in periods when small companies outperform large companies, explained by investors moving their money towards smaller companies.

We continue by running LASSO on each of the four groups of variables as we did for level data. The variables chosen by LASSO are then used in a fixed effect regression in each category. The results are presented in Table 5.5. The model using only firm-specific variables has an R^2 of 0.259, which is rather high compared to the other models. For example, the regressions in Table 5.4 has an R^2 on 0.287 which is only slightly higher. This indicates that the firm-specific variables capture substantial variation for changes in the CDS spread, although a model using all categories performs slightly better. Thus, we contribute to the literature by complementing the results of Das et al. (2009), demonstrating that the study's result holds on first-differenced data as well.

For the macroeconomic variables, the R^2 is lower than firm-specific, but both QE and RATE^2 become significant in explaining the CDS spreads. The significant RATE^2 indicates that there may be a non-linear relationship between changes in spot rate and changes in CDS spreads, as LASSO does not choose the RATE itself. However, our result indicates that while the squared rate is deemed important when only using macroeconomic variables, it is not important when including all categories of variables as presented in Table 5.4. QE is also significant and negatively correlated in the model using only macroeconomic variables, although LASSO shrinks it to zero when including all variables.

The Fama-French variables are all chosen by LASSO when only including this category of variables, and all are significant in describing the CDS spread changes. The result is consistent with Galil et al. (2013), which got all factors significant when only including the Fama-French variables in the regression. However, Galil et al. (2013) has a negative sign on all variables, whereas we have a positive sign on SMB. The study uses a sample of US companies from 2002 until 2013, thus their sample differs from ours. As discussed in Section 5.1, Galil et al. (2013) states that a comparison between a model with all regressors (Table 5.4) and a subset of regressors (Table 5.5) reveals collinearity among the explanatory variables. Following this reasoning, our results indicate that the effect of QE, HML, and RATE^2 are captured by other variables in our model.

	ESG	Firm-specifc	Macroeconomic	Fama-French
ESGC	–			
ENV	–			
GOV	–			
SOC	–			
RET		−0.400***		
		(0.006)		
VOL		0.150***		
		(0.005)		
DTE		–		
PCFS		–		
SIZE		−0.054***		
		(0.008)		
ROA		−0.012**		
		(0.005)		
LEV		0.049***		
		(0.007)		
ROCE		−0.017***		
		(0.005)		
CASHTA		–		
REVGROW		–		
RATE			–	
$RATE^2$			−0.017***	
			(0.006)	
SLOPE			−0.027***	
			(0.006)	
INF			–	
VIX			0.347***	
			(0.006)	
QE			−0.035***	
			(0.006)	
MKT				−0.327***
				(0.006)
SMB				0.063***
				(0.006)
HML				−0.032***
				(0.006)
Observations	28,560	28,560	28,560	28,560
R^2	0	0.259	0.121	0.093
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 5.5: Results from a fixed-effect panel regression on first-differenced data. The variables used in the model are the ones chosen by LASSO. Variables not chosen by LASSO have an "-" in the table as they are not included in the regression

In order to further understand the impact of the variables chosen by LASSO on the CDS spread, we run a quantile regression on the variable chosen by LASSO. The results are presented in Table 5.4.

	Quantiles				
	0.1	0.25	0.5	0.75	0.9
RET	-0.309^{***}	-0.293^{***}	-0.262^{***}	-0.315^{***}	-0.387^{***}
	(0.013)	(0.014)	(0.015)	(0.015)	(0.015)
VOL	0.050^{***}	0.064^{***}	0.080^{***}	0.119^{***}	0.157^{***}
	(0.012)	(0.008)	(0.008)	(0.008)	(0.016)
SIZE	-0.027^{***}	-0.027^{**}	-0.032^{***}	-0.042^{***}	-0.038^{**}
	(0.015)	(0.013)	(0.008)	(0.011)	(0.017)
ROA	-0.008	-0.005	-0.004	-0.010	-0.010
	(0.008)	(0.007)	(0.005)	(0.007)	(0.008)
LEV	0.017	0.031^{**}	0.045^{***}	0.059^{***}	0.098^{***}
	(0.013)	(0.012)	(0.006)	(0.011)	(0.021)
ROCE	0.007	0.003	-0.006	-0.001	-0.017
	(0.008)	(0.006)	(0.005)	(0.006)	(0.013)
SLOPE	0.004	-0.015^{**}	-0.017^{***}	-0.023^{***}	-0.038^{***}
	(0.004)	(0.008)	(0.006)	(0.005)	(0.007)
VIX	0.095^{***}	0.091^{***}	0.101^{***}	0.113^{***}	0.145^{***}
	(0.012)	(0.011)	(0.008)	(0.009)	(0.015)
MKT	-0.111^{***}	-0.091^{***}	-0.070^{***}	-0.094^{***}	-0.128^{***}
	(0.011)	(0.007)	(0.008)	(0.010)	(0.014)
SMB	0.030^{***}	0.043^{***}	0.046^{***}	0.063^{***}	0.111^{***}
	(0.009)	(0.007)	(0.005)	(0.006)	(0.011)
Note:			*p<0.1; **p<0.05; ***p<0.01		

Table 5.6: Results for quantile regression on variables chosen by LASSO

We see that stock return is negatively correlated with the CDS spread changes and distinct as the largest coefficient. The coefficient has a non-linear effect on the CDS spread, consistent with the result of Blasberg et al. (2021). The coefficient is lowest in the median and has a larger effect in the outer quantiles. The results indicate that in regimes with large movements in the CDS spread, especially in regimes with increasing risk, stock returns are a stronger determinant than in regimes with small movements. In other words, the stock return has a larger relation to the CDS spread changes in times of considerable uncertainty.

In contrast to stock return, stock volatility is positively associated with the CDS spread changes and its impact becomes more prominent in higher quantiles. The result is in contrast to Blasberg et al. (2021) that finds stock volatility to be negative in lower quantiles and positive in higher quantiles. However, theoretically, higher volatility of a firm's assets resulting from more asset value uncertainty leads to higher credit spreads (Merton, 1974). Following the pattern of stock volatility, the leverage, yield slope and SMB also increase their absolute value of coefficients across quantiles. The findings indicate that the relation between CDS spread changes and the variables are larger in regimes with large positive movements.

The coefficients of firm size, VIX and MKT are slightly higher in absolute value for the 90th quantile than the lower quantiles. This indicates less heterogeneity in the explanatory power of these variables. Moreover, ROA and ROCE are not statistically significant in any quantile, which can result from quantile regression using the distribution of conditional median and not conditional mean.

In general, the increase in the absolute value of coefficients in higher quantiles for most variables indicates that the determinants have a more substantial impact in times of large positive movements. On the other hand, the coefficients are smaller in times of large negative movements. Thus, our result indicates that in times of increasing credit risk, the determinants become more important in explaining the firm's risk. This result has implications in different economic environments. For example, in periods of crisis such as COVID-19, companies aiming to reduce credit risk should be more aware of the variables included in our model. Moreover, in the context of the current economic environment, one might expect a regime with higher credit risk going forward, as interest rates are expected to increase, which might lower economic activity and growth. Based on our result, a company can then seek to reduce the leverage ratio and improve stock return to mitigate some of the increase in credit risk going forward.

5.3 A comparison of the use of levels and first-differenced data

A comparison and discussion are warranted after presenting the regression results for both level and first-differenced data. As we can see, three variables are consistently included with the same sign: VOL, SIZE, and SLOPE. On the other hand, the remaining variables included vary between the models, which shows that the choice of data type answers to different purposes. In this section, we discuss what variations we see and give weight to the implications this might have for future research.

Overall, the R^2 is much higher for regressions on level data than on first-differenced data. The R^2 is 0.475 and 0.287, respectively. This is consistent with previous studies such as Ericsson et al. (2009) and B. Y. Zhang et al. (2009) which also used both data structures, and indicate that the determinants found are able to describe the variation in CDS spread better than the CDS spread changes.

An obvious difference between regressions on level and first-differenced data can be seen on the ESG variables. On levels we have two ESG variables included in the full model (Table 5.1), as well as all four significant in the ESG model run (Table 5.2). Regressions on first-differenced data do not have any significant ESG variables. This is not surprising, as the ESG variables are reported yearly in the dataset, and firms mostly experience minor changes in ESG scores from year to year. Even though we do not get any significant variables on first-differenced data, we emphasize that the ESG score is an important factor for credit risk on levels. While our results show that GOV and SOC might be related to the over-investment view, ENV supports risk mitigation when only including ESG variables. However, since ENV is not significant in the complete model, our study supports the over-investment view in a concise model for credit risk.

Another finding is related to stock return and firm size. While firm size is an essential predictor of the CDS spread on levels, stock return is most important for the first-differenced data. The proxy for the variables has an influence on these results. For first-differenced data, we have the month-to-month return for the stock return, which is almost equal to the change in market capitalization. As market capitalization is the total amount of shares times the stock price, these two variables represent the same effect in first-differenced data. This would mean that market capitalization is vital for regressions on level data and that the change in market capitalization is an essential determinant for regressions on first-differenced data. If the stock price has taken a hit and the credit spread is still, investors should react.

We also note that stock volatility is a consistent determinant for CDS spreads, always showing a positive sign and increasing in higher quantiles on both level and first-differenced data. Still, the interpretation from regressions on level and first-differenced data are different, especially for quantile regression. Moreover, it shows that stock volatility has an increasing coefficient both for high-risk firms compared to low-risk firms and an increasing coefficient in periods of increased risk regimes compared to reduced-risk regimes. Finally, we note that stock volatility might be higher for high-risk firms and that volatility is also increased in periods of increased risk regimes.

LASSO regression on first-differenced data chooses six firm-specific variables, i.e., it chooses four more variables than it does for regressions on level data. The profitability metrics ROA and ROCE are not chosen on first-differenced data but not on level data. Leverage ratio is also chosen for LASSO on first-differenced data, implying that it is the change of higher leverage that increases the risk, not high leverage in itself. However, a firm can then reduce credit risk by reducing the leverage. Also, Figlewski and Xiaozu Wang (2000) notes that increasing leverage leads to an increase in stock volatility, showing how leverage might have an indirect effect on credit risk.

The QE variable falls out of the model when employing LASSO on first-differenced data. A reason for this is that the increase in QE does not lead to an immediate reduction in credit risk but will do so over time when the full effects are incorporated into the market. Also, as we are using data from January 2010 to December 2019, the changes in the variable are minor through this period. A reason can be that QE is heavily increased in crisis periods, and our time horizon has not seen any major crises in the U.S.

The Fama-French variables are only significant for first-differenced data. A reason can be that the Fama-French variables in themselves are differenced. However, Barth et al. (2021) uses level data and finds a significant negative relationship for market excess return. It is not clear if Barth et al. (2021) did change the proxy when using it in a regression on level data. Our results show that the Fama-French variables are not needed for explaining CDS spreads but are needed for explaining changes in CDS spreads.

Chapter 6

Conclusion

In a complex and constantly changing world, substantial research and understanding are required to make vital decisions. Policymakers, investors, and firms alike need a better understanding of the effects choices have, both now and for future generations. Using CDS spreads to measure credit risk, we can understand what factors the market view as crucial to explain credit risk. This study responds to the need for a thorough review of new and existing determinants of credit risk to provide a concise yet comprehensive model. By including 23 different variables from several categories, and an extensive dataset of 240 CDS's over ten years, we contribute to the growing literature on determinants for CDS spreads and provide several novel insights.

Our study provides a systematic approach to identify redundancies in the existing models for explaining CDS spreads and their changes. By using a large pool of both traditionally included and new variables, we show that only seven and ten variables are needed to explain the CDS spread and changes in the CDS spread, respectively. The new variables, social pillar, and governance pillar from ESG factors and QE from macroeconomic variables are important for describing credit risk. Our results for ESG-variables support the over-investment, contrary to the opposing risk-mitigation view. Moreover, QE has a risk-reducing effect on level data. Further, a quantile regression shows that the level of QE has a larger risk-reducing effect for the riskiest firms. In this way, we contribute to the literature by providing new insights into both data types. For example, our results show that it is the change in leverage ratio and not the total level that increases credit risk. We thus highlight that variables have different effects used with the CDS spread or changes in the CDS spread.

We emphasize several promising directions for further research in what follows. First, the same methodology and variables could be employed but split the data into different time periods. Then, one could see how the variables affecting the CDS spread change over time, giving a broader understanding of how the world has developed and what the market perceives as risk in the various environments.

Second, Asia could be included as an area, and a geographical study could be performed. Our study does not focus on the differences in geographical areas, where some new findings might be found, with Asia included. For instance, China opened for CDS trading in 2016, which may contribute to exciting results in the Asian market.

Lastly, as the new variables ESG and QE have been shown to have a significant effect, a deeper analysis of these variables can be explored. In terms of QE, one suggestion is to include lagged variables for regressions on first-differenced data to see if and when QE affects the CDS spread. For ESG, the research can explore different sectors and geographies. A deeper analysis can be done to find investments in ESG-related activities to map the direct investments together with credit risk. While we acknowledge that the data needed might not be available yet, we consider it may be possible in the future.

Bibliography

Acharya, Viral V. and Matthew Richardson (2009). 'CAUSES OF THE FINANCIAL CRISIS'. In: *Critical Review*.

Ajer, A. E. and J. H. Bohler (2021). 'How have quantitative easing influenced the credit risk of commercial banks in the US? Evidence from the CDS market'. In: *NTNU*.

Albu, Lucian Liviu et al. (2013). 'The effect of ECB's quantitative easing on credit default swap instruments in Central and Eastern Europe'. In: *Procedia Economics and Finance*.

Bai, J. and L Wu (2016). 'Anchoring Credit Default Swap Spreads to Firm Fundamentals'. In: *Journal of Financial and Quantitative Analysis*.

Barth, Florian, Benjamin Hübel and Hendrik Scholz (2021). 'ESG and Corporate Credit Spreads'. In: *SSRN*.

BIS (2021). *BIS statistic explorer*. URL: https://stats.bis.org/statx/toc/LBS.html (visited on 26th Oct. 2021).

Blanco, Robert, Simon Brennan and Ian W. Marsh (2005). 'An Empirical Analysis of the Dynamic Relation between Investment-Grade Bonds and Credit Default Swaps'. In: *The journal of finance*.

Blasberg, Alexander, Rüdiger Kiesel and Luca Taschini (2021). 'Climate Default Swap – Disentangling the Exposure to Transition Risk Through CDS'. In: *SSRN*.

Cesare, Antonio Di and Giovanni Guazzarotti (2010). 'An Analysis of the Determinants of Credit Default Swap Spread Changes Before and During the Subprime Financial Turmoil'. In: *SSRN*.

Chalamandaris, George and Nikos E. Vlachogiannakis (2016). 'Are financial ratios relevant for trading credit risk? Evidence from the CDS market'. In: *Annals of Operations Research*.

Collin-Dufresne, Pierre, Robert S. Goldstein and J. Spencer Martin (2001). 'The Determinants of Credit Spread Changes'. In: *The journal of Finance*.

Councile, Atlantic (2022). *Deploying QT – The Fed readies its new tool to fight inflation*. URL: https://www.atlanticcouncil.org/blogs/econographics/deploying-qt-the-fed-readies-its-new-tool-to-fight-inflation/ (visited on 9th May 2022).

Das, Sanjiv R., Paul Hanouna and Atulya Sarin (2009). 'Accounting-bassed versus market-based cross-sectional models of CDS spreads'. In: *Journal of Banking & Finance*.

Delis, M., K. de Greiff and S. Ongena (2018). 'Being Stranded the Carbon Bubble? Climate Policy Risk and the Pricing of Bank Loans'. In: *Swiss Finance Institute*.

Drempetic, Samuel, Christian Klein and Bernhard Zwergel (2017). 'The Infuence of Firm Size on the ESG Score: Corporate Sustainability Ratings Under Review'. In: *Journal of Business Ethics*.

Ederington, L., W. Guan and L Yang (2015). 'Bond market event study methods'. In: *Journal of Banking & Finance*.

Ericsson, Jan, Kris Jacobs and Rodoldo Oviedo (2009). 'The Determinants of Credit Default Swap Premia'. In: *Journal of Finance and Quantitative Analysis*.

Fama, Eugene F. and Kenneth R. French (1989). 'Business conditions and expected returns on stocks and bonds'. In: *Journal of Financial Economics*.

Figlewski, Stephen and Xiaozu Wang (2000). 'Is the 'Leverage Effect' a Leverage Effect?' In: *NYU Stern School of Business*.

Finnerty, J.D, C. D Miller and R.R Chen (2013). 'The impact of credit rating announcements on credit default swap spreads'. In: *Journal of Banking & Finance*.

Flammer, Caroline (2015). 'Does Corporate Social Responsibility Lead to Superior Financial Performance? A Regression Discontinuity Approach'. In: *INFORMS*.

Fonseca, José Da and Katrin Gottschalk (2018). 'The Co-Movement of Credit Default Swap Spreads, Equity Returns and Volatility: Evidence from Asia-Pacific Markets'. In: *International review of finance*.

FSB (2017). *OTC Derivatives Market Reforms*. URL: https://www.fsb.org/wp-content/uploads/P290617-2.pdf (visited on 29th Oct. 2021).

Galil, Koresh et al. (2013). 'The determinants of CDS spreads'. In: *Journal of Banking & Finance*.

Goss, Allen and Gordon S. Roberts (2011). 'The impact of corporate social responsibility on the cost of bank loans'. In: *Journal of Banking & Finance*.

Han, Bing and Yi Zhou (2015). 'Understanding the term structure of credit default swap spreads'. In: *Journal of Empirical Finance*.

Hu, Nan, Ling Liu and Lu Zhu (2017). 'Credit default swap spreads and annual report readability'. In: *Review of Quantitative Finance and Accounting*.

ISDA (2020). *ISDA Legal Guidelines for Smart Derivatives Contracts: Credit Derivatives*. URL: https://www.isda.org/a/ur4TE/Guidelines-for-Smart-Contracts-CDS.pdf (visited on 12th Oct. 2021).

Kang, Johnny and Carolin E. Pflueger (2015). 'Inflation Risk in Corporate Bonds'. In: *Journal of Finance*.

Kim, Tong Suk, Jae Won Park and Yuen Jung Park (2017). 'Macroeconomic Conditions and Credit Default Swap Spread Changes'. In: *The Journal of Futures Markets*.

Koenker, Roger (2004). 'Qunatile regression for longitudinal data'. In: *Department of Economics, University of Illnois at Urbana-Champaign*.

Kölbel, Julian, Timo Busch and Leonhardt Jancso (2017). 'How media coverage of corporate social irresponsibility increases financial risk'. In: *Strategic management journal*.

Kölbel, Julian, Markus Leippold et al. (2020). 'Ask BERT: How Regulatory Disclosure of Transition and Physical Climate Risks affects the CDS Term Structure'. In: *SSRN*.

Koutmos, Dimitrios (2019). 'Asset pricing factors and bank CDS spreads'. In: *Journal of International Financial Markets, Institutions & Money*.

Lecce, Steven et al. (2018). 'Short-selling and credit default swap spreads—Where do informed traders trade?' In: *The Journal of Futures Markets*.

Longstaff, F. A. and S. S. Schwartz (1995). 'A Simple Approach to Valuing Risky Fixed and Floating Rate Debt'. In: *The Journal of Finance*.

Mandelbrot, B. and R. Hudson (2004). *The (Mis)Behavior of Markets: A Fractal View of Risk, Ruin, and Reward*. Basic Books.

Merton, R.C (1974). 'On the pricing of Corporate Debt: The Risk Structure of Interest Rates'. In: *The Journal of Finance*.

Naumer, Hans-Jörg and Burcin Yurtoglu (2019). 'It is not only what you say, but how you say it: ESG, corporate news, and the impact on CDS spreads'. In: *Global Finance Journal*.

Norden, Lars and Martin Weber (2009). 'The Co-movement of Credit Default Swap, Bond and Stock Markets: an Empirical Analysis'. In: *European Financial Management*.

Pereira, John, Ghulam Sorwar and Mohamed Nurullah (2018). 'What drives corporate CDS spreads? A comparison across US, UK and EU firms'. In: *Journal of International Financial Markets, Institutions and Money*.

Pires, Pedro, João Pedro Pereira and Luís Filipe Martins (2013). 'The Empirical Determinants of Credit Default Swap Spreads: a Quantile Regression Approach'. In: *European Financial management*.

Refinitiv (2021). *Environmental, Social and Governance (ESG) Scores from Refinitiv*. URL: https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/refinitiv-esg-scores-methodology.pdf (visited on 1st Nov. 2021).

Tang, Dragon Y. and Hong Yan (2017). 'Understanding transactions prices in the credit default swaps market'. In: *Journal of Financial Markets*.

Tibshirani, Robert et al. (2013). *An Introduction to Statistical Learning: With Applications in R*. Springer.

UN, The (2015). *The Paris Agreement*. URL: https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement (visited on 1st Nov. 2021).

Wang, Xinjie, Weike Xu and Zhaodong Zhong (2018). 'Economic policy uncertainty, CDS spreads, and CDS liquidity provision'. In: *The Journal of Futures Markets*.

Womack, L. Kent and Ying Zhang (2003). 'Understanding Risk and Return, the CAPM, and the Fama-French Three-Factor Model'. In: *SSRN*.

Zhang, B. Y., H. Zhou and H. Zhy (2009). 'Explaining Credit Default Swap Spreads with the Equity Volatility and Jump Risks of Individual Firms'. In: *Review of Financial Studies*.

Zhu, H. (2006). 'An Empirical Comparison of Credit Spreads between the Bond Market and the Credit Default Swap Market'. In: *Journal of Financial Services Research*.

Appendix A

Refinitiv ESG Scores

ESG Pillar	Score	Definition
Environmental	Resource use	The resource use score reflects a company's performance and capacity to reduce the use of materials, energy or water, and to find more eco-efficient solutions by improving supply chain management.
	Emissions reduction	The emission reduction score measures a company's commitment and effectiveness towards reducing environmental emissions in its production and operational processes.
	Innovation	The innovation score reflects a company's capacity to reduce the environmental costs and burdens for its customers, thereby creating new market opportunities through new environmental technologies and processes, or eco-designed products.
Social	Workforce	The workforce score measures a company's effectiveness in terms of providing job satisfaction, a healthy and safe workplace, maintaining diversity and equal opportunities and development opportunities for its workforce.
	Human rights	The human rights score measures a company's effectiveness in terms of respecting fundamental human rights conventions.
	Community	The community score measures the company's commitment to being a good citizen, protecting public health and respecting business ethics.
	Product responsibility	The product responsibility score reflects a company's capacity to produce quality goods and services, integrating the customer's health and safety, integrity and data privacy.
Governance	Management	The management score measures a company's commitment and effectiveness towards following best practice corporate governance principles.
	Shareholders	The shareholders score measures a company's effectiveness towards equal treatment of shareholders and the use of anti-takeover devices.
	CSR strategy	The CSR strategy score reflects a company's practices to communicate that it integrates economic (financial), social and environmental dimensions into its day-to-day decision-making processes

Table A.1: Table of the category scores and their definitions from Refinitiv.
(Refinitiv, 2021)

Appendix B

Correlation matrices

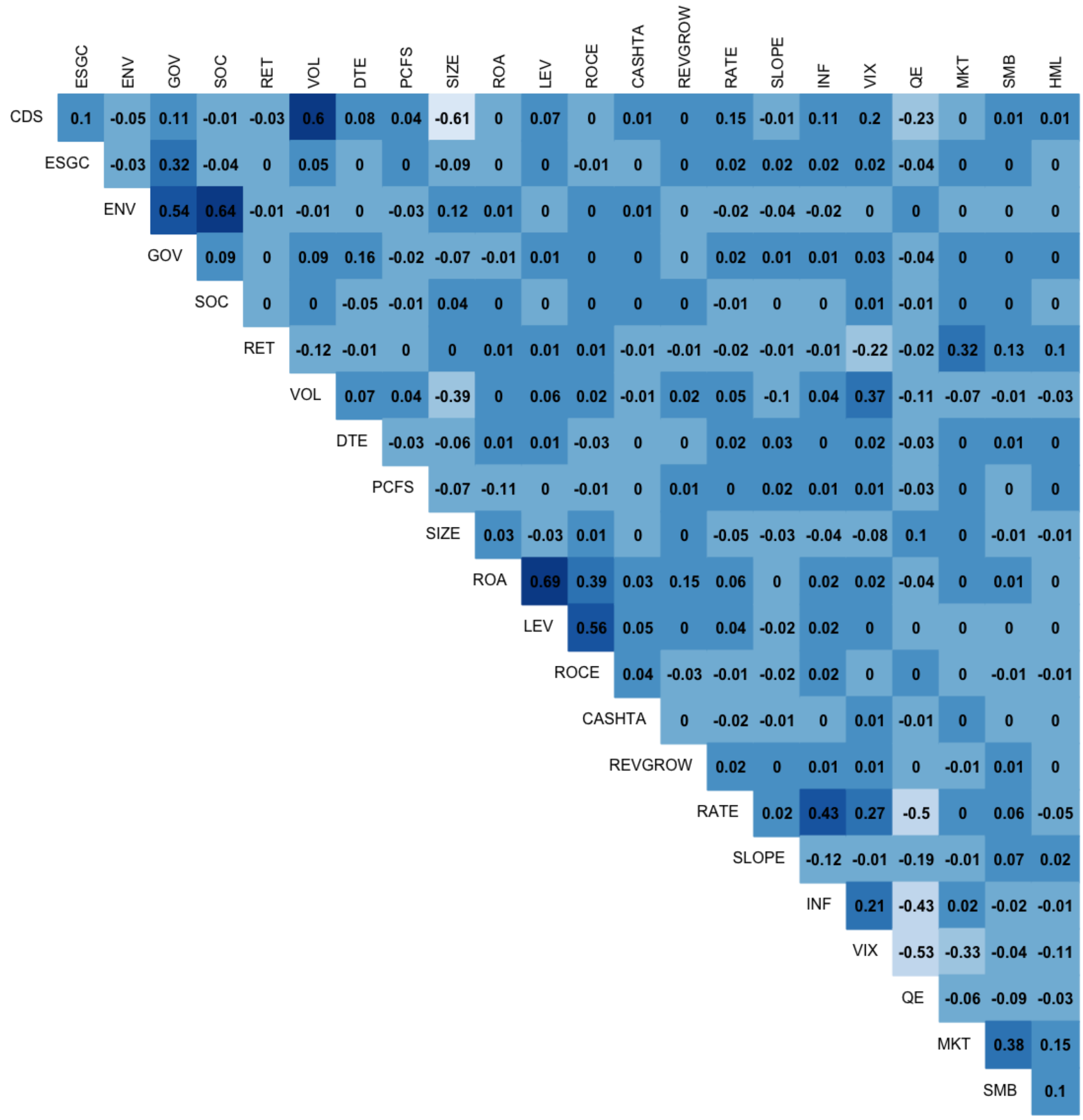

Figure B.1: Correlation plot between variables on level data.

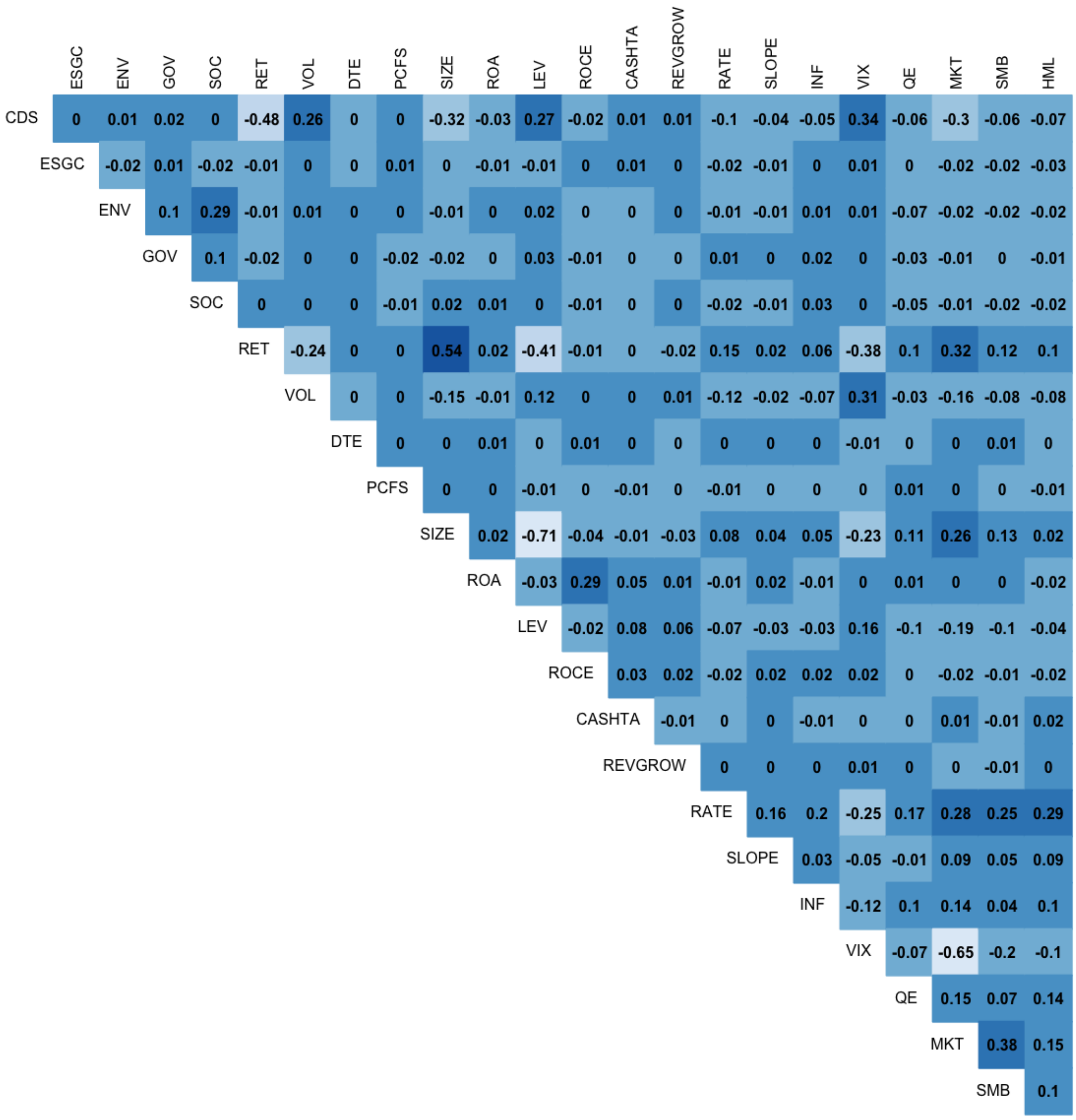

Figure B.2: Correlation plot between variables on first-differenced data.

NTNU

Kunnskap for en bedre verden