New Funding and Pricing Mechanism in Alternative and Sustainable Finance: The Role of Non-Financial Factors

Dissertation zur Erlangung des Grades eines Doktors der Wirtschaftswissenschaft

eingereicht an der Fakultät für Wirtschaftswissenschaften der Universität Regensburg

vorgelegt von: Rongxin Zhang

Berichterstatter:
Prof. Dr. Gregor Dorfleitner
Prof. Dr. Klaus Röder

Tag der Disputation: 01.12.2021

Rongxin Zhang

New Funding and Pricing Mechanism in Alternative and Sustainable Finance:

The Role of Non-Financial Factors

A dissertation submitted to the Faculty of Business, Economics, and Management Information Systems in partial fulfillment of the requirements for the degree of Doktor der Wirtschaftswissenschaft (Dr. rer. pol.)

Advisors:

Prof. Dr. Gregor Dorfleitner

Prof. Dr. Klaus Röder

Universität Regensburg

Regensburg, Germany

Acknowledgements

I would like first to thank Prof. Dr. Dorfleitner for giving me the opportunity to pursue a doctorate degree under his supervisory. I am grateful for his trust and confidence in me from the very beginning. He always inspired and guided me, sharing his knowledge and experience throughout the whole process. My thanks also go to my second supervisor Prof. Dr. Röder, for his insightful suggestions and comments.

Moreover, I am also grateful to Prof. Dr. Dr. h. c. Keller, for his effort in helping me find an opportunity to complete this dissertation.

At last, I would like to thank my family members, especially my wife, for their understanding and support. Without their encouragement and support, this dissertation would not have been possible.

List of Figures

List of Tables

Contents

Chapter 1

Introduction

1.1 Non-financial factors in alternative and sustainable finance

According to the survey report of the World Bank (Demirguc-Kunt et al., 2018), 1.7 billion people are excluded from the formal financial markets as it is too risky and costly to grant loans to them (Armendáriz and Morduch, 2010). Even though Microfinance in developing countries provide a very promising prospect in terms of poverty alleviation and financial inclusion (Morduch, 1999; Khandker, 2005), it depends heavily on donations and subsidies and is thought to be insufficient from financial perspective (Hudon and Traca, 2006). There is evidence that MFIs have to choose between social outreach and financial sufficiency (Morduch, 2000). In response, many new financing alternatives, such as crowdfunding, have been introduced and undergone substantial development in the past few decades thanks to the breakthrough in internet technology (Bruton et al., 2015). P2P lending, often referred to as crowdlending, enjoys incredible fast growth in the past ten years and is viewed as a powerful competitor to banks in retail finance. Since the investment decisions are made by the crowd instead of professional investors, the funding and pricing mechanism of these financing activities could also change. Recent studies reveal that individual investors have quite different mindset and considerations compared with institutional investors (Lee and Lee, 2012; Liu et al., 2015).

Besides the innovation in alternative finance, another important trend on financial markets is the focus on sustainability in financing activities. Investors, especially institutional investors, pay more and more attention to sustainability aspects of financing such as environmental, social and governance (ESG) issues (Schröder, 2004; Camilleri et al., 2017). The concept of Socially Responsible Investment (SRI) has become very popular and the demand for SRI mutual funds is also increasing due to investors' growing non-financial considerations (Bialkowski and Starks, 2016; Sparkes and Cowton, 2004). However, some other studies argue that the alignment of the socially responsible principles is just for the reason of risk management and economic motives rather than social values (e.g. Nofsiger et al., 2016) and thus raise doubt whether sustainability is valued by the market.

Despite these new developments mentioned above in the finance sector, our understanding regarding the role of non-financial factors in the funding and pricing mechanism in alternative and sustainable finance is still very limited. The research question that what kind of non-financial motivations may play a role in investment decisions has been seldom touched, not to mention how to investigate the influence of these special considerations of investors. Therefore, a systematic and comprehensive review of investors' non-financial considerations in the funding and pricing mechanism of alternative and sustainable finance, needs to be carried out to better understand their investment behavior.

A key obstacle to address this topic is the extraction and quantification of these non-financial factors. Non-financial factors, unlike financial ones, are often very difficult to process and quantify. In alternative finance, most non-financial information items are often stored implicitly in unstructured forms, such as pictures, audios and texts. Due to this reason, very few studies on alternative finance manage to investigate the role of non-financial signals in unstructured data. By contrast, the amount of unstructured data increases exponentially in the past few years (Dhar, 2013). Recently, some studies attempt to extract soft factors, including non-financial considerations, from unstructured data in alternative finance and explore their potential influence on the investors' investment behavior. For instance, the existence of racial disparities and facial attractiveness bias in online microfinance lending is confirmed by Pope and Sydnor (2011) and Jenq et al. (2015). In addition, some studies suggest that soft factors in the description texts are key funding and pricing determinants in crowdlending (Allison et al., 2013, 2015; Moss et al., 2015; Berns et al., 2018). These studies provide some interesting insights into the role of non-financial factors in alternative finance.

Also, sustainability as soft information cannot be measured with ease in the research of sustainable finance. Even though sustainability is often perceived to be of great importance for the general public, whether and how investors on financial markets appreciate it are left unanswered. For instance, it has been hotly debated that whether green bonds enjoy lower financing costs and no consensus has been reached in earlier studies (Climate Bond Initiative, 2019b; Larcker and Watts, 2020; Bachelet et al., 2019; Baker et al., 2019).

It still remains a challenge as regards how to measure the authenticity and true greenness of green bonds. Moreover, earlier studies on sustainable finance rely heavily on performance proxies from external data providers. Thus, the effectiveness and accuracy of these studies are subject to how data providers process soft information. Some recent studies on sustainable finance turn directly to text data to avoid such dependency (Krüger, 2015; Capelle-Blancard and Petit, 2019).

While these above findings indicate promising prospects in the research of non- financial factors in alternative and sustainable finance, they only reveal the tips of the iceberg of their role on financial markets. First of all, it is unclear whether and to what extent non-financial factors are valued and which of them are crucial to investors. Moreover, how investors optimize the tradeoff between financial and non-financial considerations is not yet well understood. In other words, it remains unanswered how much they are willing to sacrifice financial benefit for non-financial considerations. In addition, the way earlier studies extract non-financial factors is far from satisfactory (Guo et al., 2016). The necessity to measure non-financial factors presents an imminent challenge for finance researchers. As the analysis of unstructured data in alternative and sustainable finance is still at its early stage, more powerful and comprehensive analysis tools are needed to accurately extract those non-financial factors in order to investigate how they influence the investors' decision-making.

1.2 Research objectives

Overall, the major aim of this doctoral dissertation, is to identify and quantify various non-financial factors in alternative and sustainable finance, and evaluate their influence on the funding and pricing mechanism. To this end, non-financial factors should be measured in an accurate way in the first place. In this dissertation, the main focus is to quantify non-financial factors embedded in descriptive texts. For instance, loan applications in P2P lending are usually written by borrowers and thus provide extra signals, especially non-financial signals to investors. In sustainable finance, sustainability or ESG performance is often gauged in the form of third-party evaluation reports and related news. This dissertation intends to define and capture these signals in alternative and sustainable finance to facilitate further inspections.

After the quantification of non-financial factors, the first direct objective of this dissertation is to investigate the influence of non-financial factors such as empowerment and vulnerability in alternative finance. It tries to shed some light on what non-financial factors are important for different types of investors when they make investment decisions and to what degree their investment decisions can be influenced. Another associated question to be answered is that how investors may strike a balance between financial and non-financial considerations when making investment decisions.

This dissertation also seeks to examine the impact of non-financial considerations such as environmental protection and social care in sustainable finance. To be specific, it investigates the pricing mechanism of sustainability-related financial instruments and information. It adds to the debate of whether and how much investors are willing to pay higher prices for sustainability. Additionally, it studies the question that under which conditions investors may value sustainability. Overall, it aims to deepen understanding regarding investors' sustainability preferences by evaluating sustainability performance in an alternative way instead of relying on external performance proxies from agencies.

1.3 Research methodologies

The preliminary requirement of this dissertation is to extract non-financial signals from descriptive texts such as loan applications, research reports and instant news in a reliable manner. These texts have different lengths, formats and styles, and thus are very difficult to process. This dissertation shows how interdisciplinary methodologies may facilitate the investigation of non-financial factors in related studies. In particular, various innovative linguistic analysis methods are conducted whenever possible to quantify non-financial factors by extracting useful information from unstructured data. Throughout the whole dissertation, text data items are transformed into uniform and meaningful variables which can be seen as proxies of non-financial factors. Apart from the traditional lexicon-based approach, several other linguistic analysis techniques are applied to extract soft information in texts. To be specific, keywords analysis, sentiment analysis and other more advanced nature language processing (NLP) techniques developed in recent years will be adopted.

After the quantification of non-financial factors, various empirical methodologies such as univariate and multivariate analysis are applied to understand their impact on the funding and pricing mechanism. In alternative and sustainable finance, the empirical research settings are different from that of classical finance research and econometric models often need to be modified and adapted in order to more precisely dissect the influence of non-financial factors. For instance, hybrid models, instead of regular panel regressions such as fixed-effects and random-effects models, are applied to extract time-variant green bond premiums in the second paper. Another example is the adoption of a correlation robust t-statistic in the significance test of event study in the last paper. Besides econometric tools, comparison analysis (e.g. Zerbib, 2019) is carried out to more accurately identify and quantify the influence of specific non-financial factors in the second paper.

1.4 Possible contributions

In general, this dissertation contributes to the research of non-financial factors in alternative and sustainable finance. It aims not only to quantify various non-financial factors in alternative and sustainable finance, but also to further evaluate their impact on financial markets. Even though some recent studies start to focus on non-financial factors, it is still unclear what these factors are and how important they could be in the funding and pricing mechanism of alternative and sustainable finance. This dissertation provides interesting insights into these questions and may promote further research on related topics. In addition, the innovation in processing unstructured information by integrating new technique development from other disciplines may pave the way for new discovery in related research. Recent development in machine learning and NLP has shown very promising progress and provides powerful tools in analyzing unstructured information. Nevertheless, little effort has been made to adopt these new techniques in exploring the role of non-financial factors. In particular, by applying various linguistic analysis tools in processing text data in different ways, this dissertation shows promising potential for further research of non-financial factors and serves as an example of how soft factors can be properly measured.

This dissertation will also expand the understanding of non-financial drivers in both alternative and sustainable finance, and could be regarded as a basic theoretical guidance for market participants. By referring to findings of this study, companies will be in a better position to decide how to be engaged in financing activities related to environmental-friendly or ethical aspects. Furthermore, investors can learn how to target ideal investment objects and achieve good financial performance while taking environmental or social considerations into account. At last but not least, the dissertation could possibly contribute to the regulation development in alternative and sustainable finance. Regulators can only develop suitable policies if they understand how investors behave, especially regarding these non-financial factors that would have a significant impact on their decision-making.

1.5 Summary of research papers

This dissertation consists of three independent but closely connected research papers. The first paper investigates the funding determinants in interest-free P2P lending, while the second and third paper focus on the pricing mechanism of sustainable finance on financial markets. All of three papers pay special attention to the examination of the role of non-financial factors.

From credit risk to social impact: on the funding determinants in interest-free peer-to-peer lending This paper studies the funding determinants of US interest-free direct microloans on the famous crowdlending platform Kiva. We extract not only financial credibility signals but also non-financial factors from loan applications written by borrowers. Logistic regressions on the funding success and tobit regressions on the reversed funding time show interesting empirical results. Investors prefer applications with a social underwriting and clear signals in the description texts which help build trust in borrowers' repayment. Interestingly, the possibility to empower women and groups of borrowers appears to be attractive to investors. Regarding borrowers' vulnerability, we find evidence that there is significant difference in preference amongst investors.

The pricing of green bonds: external reviews and the shades of green In this paper, we revisit the question whether green bonds enjoy a premium based on a comprehensive green bond database. We apply a very strict matching process to find ideal conventional counterparts for every single green bond. By further removing the influence of liquidity difference in a hybrid model, we estimate a green bond premium for each green bond and for each trading day. We find that overall green bonds have a significant but very small premium. Moreover, we measure the authenticity of greenness by the existence of four types of external reviews and the greenness level by integrating the shade of green methodologies of different second party opinion providers. There is clear evidence that investors are willing to pay more for green bonds with specific external reviews and those with better shade of green evaluation. Lastly, this external validation effect decreases with increasing age of green bonds.

The pricing of ESG news: a comprehensive investigation via BERT We investigate the pricing implication of ESG news based on a large sample of ESG news constructed by ourselves instead of acquiring proprietary datasets. We show how the recent development in NLP, i.e. the BERT model, can be applied in several ways to build a ESG news dataset, and how news sentiment can be extracted in a consistent way. With such a comprehensive and unique ESG news dataset, we are able to investigate how ESG news are perceived by the market. We find that positive (negative) ESG news are associated with positive (negative) abnormal returns, and the market reactions to negative ESG news are stronger. More interestingly, past ESG records may serve as a buffer to the impact of ESG news. The negative impact of negative ESG news is smaller for companies with a good ESG profile, while the positive impact of positive ESG news is more pronounced for companies with a bad ESG image.

Chapter 2

From Credit Risk to Social Impact: On the Funding Determinants in Interest-Free Peer-to-Peer Lending

This research paper is joint work with Gregor Dorfleitner and Eva-Maria Oswald. The paper has been published as: Dorfleitner, G., Oswald, EM. & Zhang, R. From Credit Risk to Social Impact: On the Funding Determinants in Interest-Free Peer-to-Peer Lending. J Bus Ethics 170, 375–400.

Abstract: Based on a unique data set on US direct microloans, we study the funding determinants of interest-free peer-to-peer crowdlending aimed at borrowers in the US. By performing logistic regressions on funding success and Tobit regressions on the reversed funding time, the existence of a social underwriting by a third-party trustee as well as information in the description texts fostering the investors' trust are shown to be the main predictors of successful funding. Regarding social impact, the possibility to empower women and groups of borrowers appeals to the investors, whereas empowerment of the family or community beyond the borrowers themselves appears to remain unappreciated. When examining the vulnerability of the borrowers as a predictor, the results manifest differences amongst the attitudes of the investors towards social impact. In the subsample of non-endorsed loans the investors appear to prefer to support borrowers with an immigration background. In contrast, this is not the case with endorsed loans.

Keywords: Text analysis, Crowdlending, Microfinance, Funding probability, Funding time

JEL Classification: D64 D91 O16

2.1 Introduction

In this paper, we study the determinants of funding in interest-free peer-to-peer (P2P) lending. The interest rate is typically the most crucial parameter in P2P lending, as it usually reflects the repayment risk of a loan. Setting this parameter equal to zero changes the economic basis of the lending, as the investors who are willing to accept such conditions must derive some utility from sources other than the financial return. Therefore, the lenders in this context can be assumed to be socially-oriented or ethical investors. We study the question of the funding determinants in this context with a novel data set stemming from the online microfinance platform Kiva.

While crowdfunding enjoys rapid growth in the past decade, its application in microfinance has just recently drawn attention from scholars and is relatively under-researched (Berns et al., 2018). Traditionally, microfinance institutions (MFIs) grant microcredit to the poor who are excluded from the normal financial market. With the emergence of crowdfunding technique, altruistic individuals from all over the world can support more directly the unbanked population (Ly and Mason, 2012a). A few studies investigate the investors' investment behavior in prosocial crowdfunding and indicate the importance of both investors' financial and ethical considerations (e.g. Ly and Mason, 2012b; Burtch et al., 2014) . Nevertheless, it is noteworthy that prior studies focus exclusively on a specific type of prosocial crowdfunding, in which MFIs act as an intermediary between borrowers and investors (see e.g. Allison et al., 2013; Burtch et al., 2014; Allison et al., 2015; Moss et al., 2015; Dorfleitner et al., 2020). In this intermediary-based crowdfunding model, MFIs play a significant role throughout the loan life cycle (e.g. screening loan applicants, preparing loan applications, monitor loan repayment). Therefore, this kind of prosocial crowdfunding cannot be seen as pure P2P lending and the investors' investment behavior is influenced by the presence of MFIs (Allison et al., 2015; Berns et al., 2018). As a result, the investors' real attitude and preferences regarding the properties that make an applicant supportable could be obscured and not be well understood. The question arises how the investors in interest-free P2P lending can make investment decisions without mediating MFIs. However, no study has yet been conducted to answer this question and our knowledge regarding the complex motives of the prosocial investors is still very limited. This study seeks to fill this gap by investigating the investors' investment behavior, especially their ethical motives in a pure P2P setting.

Our investigation is related to business ethics in several ways. First, it touches the question of the fair interest rate in microcredit (Hudon and Ashta, 2013), which has been disputed for a long time. In our setting, the interest rate is zero and therefore can be regarded as fair to the borrower in any case. Second, as the lenders sacrifice the complete interest to let the borrowers profit, the transactions are also a matter of altruism and more concrete of philanthropic giving (Obaidullah and Shirazi, 2014). Third, in microcredit the responsibility of the lender for the borrower is an important problem, as providing microcredit has led to cases of over-indebtedness (Schicks, 2014). However, this issue is solved in our context because if the borrower is not able to repay the loan, the only penalty he or she faces is receiving no further loan. Thus, it is very unlikely that over-indebtedness emerges from a Kiva direct loan. Fourth, the honesty on side of the borrower is a relevant ethical dimension in our setting, as no one verifies the authenticity of the information given in the self-written description texts.

Our study follows the framework of prior studies analyzing the investors' dual motives in prosocial crowdfunding (e.g. Allison et al., 2015; Dorfleitner et al., 2020; Berns et al., 2018). Under this framework, the investors' financial and non-financial considerations can be examined at the same time. In general, we apply signaling theory (Spence, 1973, 2002) to understand the direct communication between borrowers and investors. In particular, special attention is paid to signals in the self-written description texts as recent studies show the informativeness of the unverified texts (see e.g. Allison et al., 2015; Berns et al., 2018).

To investigate the funding determinants of interest-free P2P lending, we examine more than 6,000 US direct loan applications on the online microfinance platform Kiva. Unlike prior studies that focus exclusively on Kiva's intermediary-based model in developing countries (e.g. Burtch et al., 2014; Moss et al., 2015), we utilize a unique data set of direct loans in the USA. The dataset is very unique as it includes not only the basic information about US direct loans from Kiva's official API but also other crucial information derived from original campaign webpages such as the description texts and endorsement details. The empirical examinations provide very interesting insights regarding the investors' investment behavior in interest-free P2P lending. First, a third-party endorsement is found to be crucial to funding success and funding speed, even if the so-called 'trustee' has no financial responsibility. Second, there is evidence that signals related to trust between investors and borrowers in the self-written description texts can influence the fundraising result. Third, the investors do appear to empower women and groups, but not others beyond the borrowers themselves. Last but not least, the investors appear to care about the borrowers' vulnerability, but to a varying extent.

With these findings, our study makes the following two contributions. First, to our knowledge, this is the first study that sheds some light on the financial and prosocial considerations of the investors funding interest-free P2P loans. While the two motivational dimensions that we investigate on the lenders' side, namely avoiding repayment risk and seeking social impact, are the same as in earlier research on the intermediary-based model, it should be noted that we do not expect to find the same well-known results now in a different setting. Rather one can say that while the two dimensions as such are canonical, we aim to study whether and how they are perceived and appreciated in a new and even purer ethical context. In the end, the details of the findings are important, as from these one can draw conclusions on the functionality of the platform and the real preferences of the investors involved in such interest-free P2P lending.

Second, our study contributes to the research of microfinance in developed countries as Kiva's direct loan model is only available in the USA. Despite growing interest in microfinance in developed countries, there is still limited research on this topic (Pedrini et al., 2016; Forcella and Hudon, 2016). Most studies on microfinance in developed countries are surveys or qualitative analysis (e.g. Kraemer-Eis and Conforti, 2009; Carboni et al., 2010; Bruhn-Leon et al., 2012; Diriker et al., 2018), and very few of them conduct empirical investigations (e.g. Cozarenco et al., 2014; Bourlès and Cozarenco, 2018; Cozarenco and Szafarz, 2018). We empirically investigate how altruistic investors make lending decisions to help the minority in developed countries who are less likely to attract attention from the public compared with their counterparts in developing countries, thus providing the opportunity to understand microfinance in different contexts.

The reminder of the paper is organized as follows: Kiva's funding model for direct loans is introduced in Section 2.2. In Section 2.3, four hypotheses are derived from theoretical considerations and existing studies. Section 2.4 describes the data and the employed methodology. The results of regressions and robustness checks are displayed in Section 2.5. Sections 2.6 concludes.

2.2 Kiva's funding model for direct loans

Kiva is well-known as an online crowdfunding platform that enables microlending to the poor by mobilizing debt capital from the worldwide crowd of altruistic investors. The standard lending model on Kiva is devoted to the crowdfunding of loans that are mediated through MFIs in developing countries. Under this intermediary-based microfinancing model, the investors refinance microloans which have already been granted to applicants by MFIs.

Apart from the intermediary-based microfinancing model, Kiva also facilitates a direct P2P lending model in which micro borrowers and socially-oriented lenders interact directly without any financial intermediation. Kiva direct loans, focusing on US inhabitants who wish to develop a promising business idea but struggle with access to capital, provide interest-free debt capital of up to 10,000 USD. The borrowers do not pay and the investors do not receive any interest on the loan. The investors fully bear the credit default risk. To minimize the risk of fraud, Kiva carries out an internal due diligence process[1]. Additionally, Kiva requires the loan applicant to successfully pass the process of so-called 'social underwriting'. During a private fundraising period, the applicant's network (family, friends) is asked to fund the loan application to further affirm the applicant's creditworthiness. Therefore a small portion of the loan amount has always been collected before the application is posted online[2]. Moreover, the loan applicant can be endorsed by an entity (an organization or an individual) that is in a relationship with the loan applicant. Even though the entity does not have any financial liability (Kiva, 2019a), Kiva calls it trustee and expects that the entity helps to strengthen the borrower's commitment to the repayment obligation. After the 3-stage screening process of the applicant's creditworthiness, the direct loan application is posted publicly and available to the crowd of socially-oriented investors. After the loan is granted, Kiva monitors the repayment behavior of the borrower. When the borrower fails to repay the loan in time, Kiva will remind the borrower via phone call or email. Kiva adjusts the trustee's ability to further endorse borrowers based on the repayment rate of the loans endorsed by the trustee. When the borrower defaults, the borrower can no longer apply for loans on Kiva. According to the official statistics (Kiva, 2019b), the repayment rate for US direct loans on Kiva is 78%, which is evidently lower than 97.5%, the repayment rate for MFIs facilitated loans. Kiva's direct P2P model is summarized in Figure 2.1.

It should be noted that Kiva's direct model is, to a large degree, unique in the practice of microfinance as well as in the field of P2P lending. From the microfinance perspective, this model is special as there is no MFI involved. From the standpoint of classical P2P lending, the fact that the borrowers do not need to pay any interest and that the investors, therefore, do not receive any financial compensation for the credit risk they take is very unusual. Therefore, Kiva's direct loan model combines the concepts of microfinance and P2P lending.

[1]The internal due diligence process includes a review of the financial history, a verification of the identity and a validation of the business. Also, all applicants are screened through the Office of Foreign Assets Control terrorism database due to national security reasons.

[2]Note that for our analysis the private fundraising does not play a significant role because every loan application fulfills this requirement (typically approximately 10% to 15% of the loan amount is prefunded).

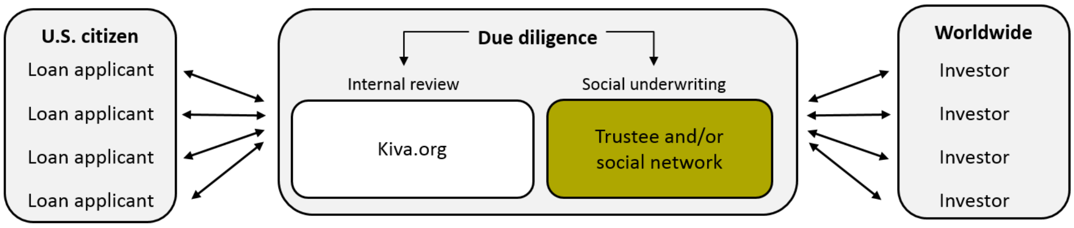


Figure 2.1: Kiva's direct P2P model for direct loans in the United States

2.3 Theory and hypotheses development

2.3.1 Theoretical basics

Findings from related fields While interest-free P2P lending is a relatively new phenomenon, its origin can be traced back to microfinance, as its underlying and fundamental objective is to help the poor population realize their economic potential (Kiva, 2018a). To better understand the investors' behavior in interest-free P2P lending, we first discuss multiple motivations of MFIs and their funders in the field of microfinance.

Traditionally, MFIs rely mainly on governmental subsidies or philanthropic donations (Hudon and Traca, 2011; Ghosh and Van Tassel, 2013). Accordingly, many MFIs focus mainly on the social outreach and impact of their business. Studies find that microfinance programs in developing countries can reduce poverty (Robinson, 2001; Khandker, 2005; Imai et al., 2010) and especially empower women (Cheston and Kuhn, 2002; Swain and Wallentin, 2009). As the microfinance industry has grown exponentially in the past few decades (Beatriz and Marc, 2011), it attracts a much broader range of funders including different public and private investors. Many non-governmental organizations (NGOs) that provide funding to MFIs are often very active in areas such as health, women's empowerment, and children's issues (Ledgerwood et al., 2013). Institutional investors like pension funds or insurance companies also fund MFIs as they seek for 'impact investing' (Ledgerwood et al., 2013). However, institutional investors could also be attracted to fund MFIs because investing in MFIs can be financially attractive (Krauss and Walter, 2009; Galema et al., 2011). There is a tendency that more and more MFIs in developing countries become for-profit organizations (Battilana and Dorado, 2010; Khavul, 2010), despite some criticism that MFIs experience 'mission drift' (Dichter and Harper, 2007). In addition, various funders or participants in the microfinance industry claim that MFIs should go beyond financial efficiency and social impact and be engaged in environmental issues as well (Hammill et al., 2008; Allet et al., 2011). Allet (2014) find that MFIs in developing countries for which social responsibility is the key motivation are more likely to promote an environmentally friendly practice.

In recent years, microfinance has also spread to Western economies. As the economic and social context in developed countries is different, microfinance in developed countries has slightly different focuses. According to Bendig et al. (2012, 2014) and Diriker et al. (2018), job creation, poverty reduction, and microenterprise development are the most important missions for MFIs in Western European countries. Although women's empowerment is also an objective of MFIs in developed countries, it plays a less prominent role (Bendig et al., 2012, 2014). MFIs in developed countries are niche institutions (Kraemer-Eis and Conforti, 2009; Cozarenco et al., 2014) and still rely heavily on government subsidies and support (Kraemer-Eis and Conforti, 2009; Bruhn-Leon et al., 2012). As a result, they focus particularly on encouraging entrepreneurial activities (Carboni et al., 2010; Cozarenco et al., 2014), as governments expect to create more employment opportunities and reduce the financial burden of social welfare (Underwood, 2006; Barinaga, 2014; Pedrini et al., 2016). Besides governments, an increasing number of commercial banks in developed countries fund or support MFIs to realize their socially responsible investment policies (Pedrini et al., 2016). However, while the microfinance sector in developing countries starts to experiment with a commercialization process, MFIs in developed countries are less profit-oriented (Kraemer-Eis and Conforti, 2009; Jayo et al., 2010). Moreover, environmental responsibility is also a concern of MFIs in developed countries (Forcella and Hudon, 2016). Forcella and Hudon (2016) find that investors' concern for environmental issues is an important determinant of MFI's environmental performance.

Despite the great achievement gained by microfinance in the past few decades, the problem of financial exclusion still prevails. According to a recent estimate of the World Bank (Demirguc-Kunt et al., 2018), 1.7 billion people do not have a bank account and can be defined as the unbanked population. Therefore, microfinance in developing and developed countries has a long way to go. Due to the development of internet technology in the recent decade, new financing alternatives, such as crowdfunding, provide the unbanked group new financing opportunity. P2P lending, sometimes also referred to as 'crowdlending', is the most important type of crowdfunding (Ziegler et al., 2017). Numerous studies (e.g. Freedman and Jin, 2008; Yum et al., 2012; Lin et al., 2013) investigate the investment behavior of individual investors in P2P lending. Some of them suggest that individual investors have a quite different mindset and show several biases when making lending decisions (e.g. Pope and Sydnor, 2011; Lee and Lee, 2012; Duarte et al., 2012). For instance, Lee and Lee (2012) observe investors' herding behavior in P2P lending. Duarte et al. (2012) and Pope and Sydnor (2011) suggest that P2P lending investors respond to signals of characteristics in attached pictures. Recent studies pay more attention to soft facts in the descriptive texts of loan applications (e.g. Herzenstein et al., 2011; Dorfleitner et al., 2016).

As a crowdfunding platform dedicated to promoting microfinance, Kiva has achieved huge success via its intermediary-based lending model (Kiva, 2018a). Many studies examine the behavior of individual investors under this model (see e.g. Burtch et al., 2014; Allison et al., 2015; Moss et al., 2015). Burtch et al. (2014) find that cultural differences and geography have a significant influence on the fundraising outcome of Kiva intermediated loans. Dorfleitner et al. (2020) observe that MFIs who have a better level of social performance in terms of lending to women, lending responsibly and charging low interest, are more likely to be refinanced through Kiva. Jenq et al. (2015) examine behavioral biases of the investors supporting Kiva's intermediated loans and find that the investors favor those borrowers who appear to be more attractive. Allison et al. (2015) assess the effect of linguistic cues on the funding result for Kiva intermediated loans and find evidence that the investors prefer to support loan applicants who position their ventures as an opportunity to help others.

Differences in the considered setting While some of the above findings on Kiva's intermediary-based model are important to our considerations, we argue that the interest-rate free P2P lending setting is very different. As Johnson et al. (2010) point out, most so-called P2P microlending models actually do not facilitate the direct interaction of borrowers and investors and thus can not be seen as real P2P lending. This fundamental difference between the intermediary-based model and real P2P model would probably lead to different investor behavior.

First, the repayment rate in the P2P setting is (with 78%) rather low when compared with that in Kiva's intermediary-based model (97.5%)[3]. This implies that the investors in interest-free P2P lending assume much higher credit risk. The credit risk in Kiva's intermediary-based model is less of a problem, and the corresponding investors may spend less effort in identifying trustworthy borrowers as the expected loss rate is only 2.5%. The fact that the funding probability in the interest-free P2P lending is less than 67%[4], which is much lower than 99% in the intermediary-based model (Berns et al., 2018), also implies the investors' serious concern about the default risk in the new setting. Second, as there is no financial compensation for the considerably higher potential credit risk of direct loans, the investors may take non-financial considerations more seriously. One could argue that the money spent on financing direct loans is 'play money'. But even if this were the case,

[3]See Kiva (2019b). It's even lower than that of usual P2P lending. As an example, the average repayment rates for the German P2P lending platforms, Auxmoney and Smava, are 88% and 86.2% (Dorfleitner et al., 2016).

[4]See descriptive statistics in section 2.4.

there still must be a reason that one loan application is preferred over another. Third, the participants of direct loans interact directly without any intermediation. The borrowers of direct loans have the chance to promote their campaigns by deciding what information they want to deliver to the investors as they write their description texts themselves. At the same time, direct loan investors have more autonomy and responsibility in screening loan applications as they can no longer utilize the information of credit profile and social performance of MFIs[5]. Taking the above together, we expect that the investors in the interest-free P2P lending are more likely to reveal their real attitude and preferences from both the financial and non-financial perspectives.

Signaling in interest-free P2P lending While the information asymmetry prevails in every lending situation, the problem is even more serious for P2P lending investors since they are not professionals like banks or other institutional investors (Yum et al., 2012; Lee and Lee, 2012). In the case of Kiva direct loans, the investors only have very limited information to evaluate loan applications. A typical US direct loan application on the Kiva website only includes very basic personal, geographical information, a brief loan description, and trustee information, while the repayment history of the borrower is difficult to obtain due to Kiva's effort to protect the borrowers' privacy. What makes the situation worse is the fact that there is even no interest rate for these direct loans, which usually serves as a signal of the credit risk of the loan[6]. Therefore, the investors of Kiva direct loans have to overcome adverse selection and the risk of moral hazard (Bruton et al., 2011).

According to signaling theory (Spence, 1973, 2002), high-quality insiders can intentionally send positive signals about themselves to influence the decision-making of outsiders (Connelly et al., 2011). Signaling theory is often applied in the entrepreneurship literature to explain how the entrepreneurs attract potential investors (Lester et al., 2006; Alsos and Ljunggren, 2017). Moss et al. (2015) and Jancenelle et al. (2018) argue that signaling theory is also applicable in the case of crowdfunding as the entrepreneurs are insiders and signals in crowdfunding are observable and costly. Several studies in crowdfunding literature adopt explicitly or implicitly signaling theory to investigate the investor's investment behavior (Allison et al., 2013, 2015; Moss et al., 2015; Jancenelle et al., 2018; Berns et al., 2018). In the context of interest-free P2P lending, the borrowers can send signals indi-

[5]In Kiva's intermediary-based model, the investors can see credit profiles of the MFIs, including default rate, delinquency rate, loans at risk rate, etc. Moreover, they can also see whether a special social performance badge is assigned to the MFI (Kiva, 2019c).

[6]The interest rate a potential borrower is willing to accept can signal the creditworthiness of the borrower in the sense that high interest rates are only accepted by borrowers with low creditworthiness, which corresponds to the idea of lemon markets (Akerlof, 1970).

cating their worthiness of being supported to reduce the severe information asymmetry. At the same time, the investors respond to these signals based on their financial and non-financial assessment. Even though signals sent by the borrowers in crowdfunding cannot be verified, Moss et al. (2015) argue that dishonest signals may not be in the best interest of the borrowers and they should strategically choose what signals to send. Michels (2012) also demonstrate that unverified information on the P2P lending platform Prosper can influence individuals' decisions and reduce the cost of debt.

Theoretical basis: A special type of investor reacting to signals From the fact that no interest rate is charged and therefore the expected financial return is negative, we conclude that the backers of campaigns in the direct loan model must have some other source of felicity when investing. As Ly and Mason (2012a) or Allison et al. (2013) show, the investors in the intermediary-based model appear to be socially oriented. There is no reason for the assumption that in the direct model totally different investors are active. However, due to the discussed differences, the investors surely are not identical either, especially because the expected repayment in the direct loan setting is much lower than in the intermediary case. Still, following Dorfleitner et al. (2020), we model an investor's utility as comprising the financial return r and the social return s weighted with the factor $\alpha > 0$:

$$r + \alpha \cdot s \tag{2.1}$$

Even if $E(r) < 0$, empirical evidence from the intermediary-based model shows that the investors still stress credit risk to be closest to zero (Dorfleitner and Oswald, 2016; Jenq et al., 2015). In contrast to kinship groups, the investors are not acquainted with the borrowers personally and face even greater information disadvantages due to the distance to the borrower and the limited information provided in the loan application. It is evident that the investors are willing to provide capital only under the condition of a positive personal utility. Consequently, the expected social return $E(s)$ should overcompensate for the expected negative financial return.

Combining signaling theory and the above theoretical considerations, we develop several concrete hypotheses to investigate where the investors might induce a positive $E(s)$ or an $E(r)$ close to zero.

2.3.2 Hypotheses development

To help investors evaluate the credit risk of borrowers, P2P platforms usually adopt several identifiable or quantifiable mechanisms such as the assignment of credit ratings and cooperation with partners. Several studies show that borrowers' credit ratings assigned by P2P platforms or external agencies are important to the investors' investment decisions (Freedman and Jin, 2008; Barasinska and Schäfer, 2014). Risk ratings of the MFIs in the intermediary-based microfinancing model could also be informative for the investors (Berns et al., 2018). However, the Kiva direct loan applicants do not have such a credit rating which may facilitate the investors' decision-making. Instead, the direct loan applications on Kiva can have trustees who endorse the borrowers.

Existence of an endorsement One of the most objective and obvious differences among direct loans is whether or not they are endorsed by a trustee. Some studies show that social networks of the borrowers are very important in the reduction of information asymmetry for online P2P lending (Liu et al., 2015; Lin et al., 2013; Freedman and Jin, 2017). Kiva direct loans with an endorsement from trustees could be perceived as being safer because trustees have to evaluate the creditworthiness of borrowers beforehand and also to monitor borrowers' repayment behavior to minimize reputation risk. Indeed, Berger and Gleisner (2009) and Collier and Hampshire (2010) document that a community endorsement on the P2P platform Prosper leads to favorable funding result, even though the endorsing lending-group leaders resume no financial responsibility. By considering the above, we expect that Kiva direct loans with a trustee endorsement are more likely to be funded.

> **H1 (Trustee endorsement):** The existence of a trustee is positively related to funding success.

Apart from the potential existence of a trustee endorsement, the investors require more information to help them evaluate the borrowers' creditworthiness. Since the hard facts are limited in interest-free P2P lending, the investors' attention could be drawn to soft facts regarding the borrowers' creditworthiness in the description texts, which constitute the main part of the campaign webpages.

Creditworthiness signals in the description texts A significant amount of studies investigate soft factors in the description texts on P2P lending platforms (e.g. Allison et al., 2015; Moss et al., 2017; Jiang et al., 2018). For instance, the empirical fact that the descriptive texts can reduce information asymmetries and thus contribute to fundraising has been documented several times (e.g. Larrimore et al., 2011; Michels, 2012). Although the description texts cannot be validated, they appear to contain some information (Michels, 2012). However, generally the lenders should take into account that with a certain probability the information given in these texts is not completely correct, i.e. some applicants may cheat about their true motives and circumstances. Yet, this rationally only makes sense if the potential borrowers know which factors influence the funding probability.

As the description texts are written by different individual borrowers in Kiva's direct loan model, the text lengths differ. Several studies find that the length of the description text is a crucial driver of funding success in P2P lending (Larrimore et al., 2011; Michels, 2012; Moss et al., 2017). Larrimore et al. (2011) argue that a lengthier text can provide more information about the borrower and thus build up trust between the borrower and investors in commercial P2P lending. Similarly, we expect that a longer and more detailed description text in interest-free P2P lending can also serve as a quality signal concerning the borrower's willingness to offer more information to the investors. However, as a too long description text could be troublesome for the non-professional investors to evaluate, we also expect that the positive effect of a longer description text to be dampened when the number of words exceeds a certain amount (e.g. Dorfleitner et al., 2016).

Besides the text length, linguistic signals which may indicate the project quality in the description texts could also affect the investors' decision-making. For instance, since microenterprise development is an important mission for microfinance in developed countries (Bendig et al., 2012, 2014; Diriker et al., 2018), the investors may pay special attention to the description of the loan usage. If there is little description related to business, the investors have no information to evaluate the feasibility of the underlying business and may be skeptical about the real intention of the borrower as Kiva direct loans are exclusively intended for entrepreneurial purposes (Kiva, 2018b). Dorfleitner et al. (2016) also suggest that the mentioning of a business purpose in the loan application is related to a higher funding probability in P2P lending because business activities are more likely to create additional positive cash flows and help repay loans. Thus, we anticipate that a clear signal of the willingness to do business with the loan proceeds in the description texts contributes to funding success.

In addition, many studies in the entrepreneurship literature indicate that human capital is very important for the success of entrepreneurial activities (Robinson and Sexton, 1994; Unger et al., 2011; Doms et al., 2010). A good education background has a strong and positive impact on entrepreneurship success, especially in a self-employment entrepreneurship setting (Robinson and Sexton, 1994). With appropriate education, the borrowers are more likely to succeed in their entrepreneurial activities as they may gain the knowledge needed to manage the business. Indeed, Dorfleitner et al. (2016) find empirical evidence that the borrowers on a German P2P platform who mention their education background in the descriptive texts have a lower probability of default. Therefore, we anticipate that the investors may look for signals in the description texts which can indicate higher human capital, such as the borrowers' education background.

Based on the above considerations, we expect signals in the descriptive texts that build trust between direct loan investors and borrowers to play an important and positive role for funding success.

> **H2 (Trust):** Signals in the descriptive texts that emphasize trustworthiness regarding the repayment are positively associated with funding success.

The theoretical considerations regarding the investors' utility imply that the investors are more likely to support loans with greater social impact to maximize their utility. The investors on prosocial P2P platforms are expected to help other people to alleviate impoverishment as they do not receive any interest from loans. Even return-oriented investors on commercial P2P lending platforms are occasionally motivated by social contributions (Pietraszkiewicz et al., 2017). Therefore, the concept of social impact is of large significance, especially for socially-oriented investors, as Allison et al. (2013), Moss et al. (2017) and Jancenelle et al. (2018) prove for the intermediary-based model on Kiva. If we, therefore, assume an ethical dimension of philanthropy in the investors' perspective, the question of interest then is, which social aspects and corresponding signals they are appealed to.

To develop our hypotheses, we adhere to two major fields in which a social contribution can be made in microfinance, namely empowerment and vulnerability, following Gaiha and Thapa (2006). At the same time, we assume that the investors can perceive signals indicating the possibility of creating social impact in the description texts wherever applicable since there is no simple and quantifiable indicator of potential social impact like the social performance badge in the intermediary-based model on Kiva.

Empowerment Empowerment is a process of change by which individuals or groups with little or no power (e.g. women, poor communities), gain in their power and ability to make choices that can change their lives (Cheston and Kuhn, 2002). Based on the conceptual framework from Schulz et al. (1995), empowerment can be viewed at the individual, organizational and community levels. Accordingly, we discuss empowerment possibility in interest P2P lending at these three levels.

Women's empowerment, particularly women's economic empowerment, is the core mission of the United Nations Industrial Development Organization (UNIDO, 2018). It is intensively investigated in the microfinance literature and many studies agree that women's empowerment is a very important objective for microfinance (e.g. Kabeer, 2001; Cheston and Kuhn, 2002; Kabeer, 2005; Swain and Wallentin, 2009). Kiva offers a special loan category, exclusively to female borrowers, and prioritizes it on the loan requests list. As of October 2017, 81% of borrowers supported through Kiva have been female (Kiva, 2018a). Heller and Badding (2012) find that female borrowers on Kiva in the intermediary-based model are funded 40% faster than their male counterparts. Ly and Mason (2012b) also find that it takes female borrowers of Kiva intermediated loans less time to gain funding. Therefore, we expect female borrowers of Kiva direct loans to receive more support from the investors.

Compared with individual direct loans, group direct loans are expected to attract more attention from direct loan investors as lending money to a group may increase the possibility of empowerment. As Thorp et al. (2005) argue, group formation can be an important way for the poor people to be empowered. Stewart (2005) also agrees that the poor people within an organization can achieve more by taking collective actions since it is often too difficult for them to escape poverty through their own efforts. In Kiva's intermediary-based model, group loans are more likely to raise funds (Berns et al., 2018). Ly and Mason (2012b) argue that if the group size is relatively large, group loans are preferred because more beneficiaries profit from these loans.

Moreover, beyond the borrowers themselves, the investors could also empower others who have a close relationship with the borrowers such as their family members and communities. When the borrowers mention their family members and communities in the description texts, the investors may perceive this as an opportunity to empower more unprivileged people, rather than just the borrowers. Freedman and Jin (2008) find evidence that loan requests on Prosper that mention family members are more likely to be funded. Allison et al. (2015) also find that words related to family members in the description texts, written by MFIs, can reduce time to funding for Kiva intermediated loans. Calic and Mosakowski (2016) suggest that social entrepreneurs who focus on the preservation of nature, life support, and community are more likely to be funded on the donation-based crowdfunding platform Kickstarter. By supporting prosocial borrowers, the investors of direct loans do not only help the borrowers to fulfill their personal goals but also help more people indirectly. Therefore, we expect these prosocial loan applications to be preferred by direct loan investors.

H3 (Empowerment): A description text indicating empowerment possibilities is positively related to funding success.

Vulnerability Besides empowerment possibility, direct loan investors can also look for the chance to help those who are in a very vulnerable position to make a social contribution. Vulnerability reduction is often seen as a desirable outcome of microfinance and closely examined in the microfinance literature (Zaman, 1999; Tchouassi, 2011; Swain and Floro, 2012). Jenq et al. (2015) find that perceived neediness in the attached picture is positively related to the funding speed of Kiva intermediated loans. According to Dorfleitner et al. (2016), P2P loan applications with negative keywords in the descriptive texts have a higher funding probability. Thus, we expect that the needy borrowers in the interest-free P2P lending can possibly attract more attention by expressing their misery in the descriptive texts.

Among the needy and vulnerable borrowers, the direct loan applicants with an immigration background are of special interest to us in this study as immigrants in developed countries often suffer from a lack of resources and financing support in the new environment. For instance, Aldén and Hammarstedt (2016) find that non-European immigrants in Sweden report upon more discrimination by traditional finance institutions. Pedrini et al. (2016) point out that the immigrant population is one of the most important targets for MFIs in developed countries. According to Jayo et al. (2010), more than 40% of MFIs in Europe identify the ethnic minorities and immigrants as their target clients. Therefore, the borrowers with an immigration background can be expected to be a target group of direct loan investors. In summary, we expect that direct loan applicants that appear to be vulnerable are more likely to be funded.

H4 (Vulnerability): If the description text indicates that a borrower is more vulnerable, the probability of funding is higher.

2.4 Data and methodology

2.4.1 Data description

Our analysis is based on interest-free direct loans which are requested by US inhabitants using the direct P2P model on Kiva. The data set is derived from Kiva's public API and includes loan applications posted on Kiva between 2011 and 2017 which can either be categorized as 'successfully funded' or 'unsuccessfully funded'. The data set is extended through additional information extracted from the original campaign webpages. Loan applications include information on loan conditions and the trustee endorsement if a trustee is provided. The applicant's personality and the purpose of the loan request are described in a descriptive text. The data set is cleared by removing 8 observations with unrealistic loan amounts of more than 10,000 USD (strict limit defined by Kiva) and 20 unsound loan applications without a description text and therefore lacking information both on the applicant and the purpose of the loan. The final data set comprises 6,121 observations. Therein, 4,077 loans are successfully funded and 2,044 loans have expired. All variables relevant to our analysis are explained in detail in Table 2.2.

Two dependent variables are observable. The first one is *Funding success*, being defined as a binary variable with a value of one if the loan request is successfully funded by the crowd of investors and zero otherwise. Additionally, the funding time for funded loans is observable. The funding time in days measures how long it has taken loan applicants to receive successful funding via the crowd. The second dependent variable, *Reversed funding time*, is defined by calculating 1,000 divided by the funding time in days. Thereby, the reversed funding time of non-funded loans is set to be zero as their funding time is infinite. This *Reversed funding time* can serve as a proxy for funding speed as it measures how fast a loan can be funded. Values are logarithmized.

All four hypotheses stated above are tested through several explanatory variables. Regarding H1, whether or not a trustee is given, is considered in the dummy variable *Trustee*. The variable *Type* distinguishes among different types of trustees: individuals, non-profit organizations and others. For loan applications with a trustee endorsement, we can calculate *Trustee's experience* in days on Kiva at the point of time the respective, new loan application is posted publicly. Furthermore, we include a dummy variable, *Trustee's proximity*, to indicate whether the trustee and the applicant are located in the same US state. The proximity of trustees and loan applicants located in the same US state is perceived as being higher.

Second, to test whether the applicant's effort to build trust helps to attract potential investors, signals within the description texts are considered. The extent of the description texts is often considered when examining the determinants of funding success in the crowdfunding literature (e.g. Parhankangas and Renko, 2017; Pietraszkiewicz et al., 2017). The variable *# of words*, calculated by counting the number of words in the description texts, is a measure of trustworthiness and could reflect the applicant's willingness to share information with potential investors (Dorfleitner et al., 2016). To capture the possible u-shape relationship between the text length and the funding result, we also include the quadratic term of the text length *# of words*^{2} following Dorfleitner et al. (2016). In addition to the text length, we extract more signals from the description texts by searching for keywords that could provide more insights into the applicant's creditworthiness and the possibility of making a social contribution (see e.g. Berns et al., 2018; Jancenelle et al., 2018). All keywords are defined and reported in Table 2.1.

Table 2.1: Categorical variables depicting possible keywords in the description texts

Hypothesis	Variable	Keywords
H2 Trust	Keyword_Business	business, career, client, company, customer, employment[1], entrepreneur[1], expand, financial stability, invest, job, network, profession[1], profitability[1], skills[1].
	Keyword_Education	academic, Bachelor, college, degree, education, exam, graduation[1], Master, PhD, (high- / home-) school, student, study, undergraduate, university.
H3 Empowerment	Keyword_Family	aunt, boy, brother, (grand-) child, dad, (grand-) daughter, family, (grand-) father, husband, kid, marriage[1], mom, (grand-) mother, (grand-) parents, partner, pregnant, siblings, sister, (grand-) son, uncle, wife.
	Keyword_Community	community, friend, help[1], serving others, support[1].
H4 Vulnerability	Keyword_Negative	abuse, addiction[1], cancer, civil war, death, defeat me, destiny, difficulty[1], disruption[1] drug, enemy, hard work, incarceration, insane, pain[1], passed away, poverty, prison, sick, ups and downs, victim.
Controls	Keyword_Positive	enjoy, fun, happiness[1], greatness[1], love[1], pleasure, smile[1], thankful, thank you.
	Keyword_Purpose	believe, better future, better life, chance, dream, fascination[1], motivation, passion[1], purpose, vision.

The keywords are obtained by analyzing the description text of loan applicants using the computerized text analysis software LIWC2015. All keywords are stated as being singular. The respective plural words are also taken into account.
[1] indicates that all respective verbs, adjectives, and adverbs are also taken into account as keywords.

The dummy variable *Keyword_Business* indicates whether the applicant's intention of entrepreneurship can be detected in the text, while *Keyword_Education* clarifies whether the applicant mentions an appropriate educational background to enable the successful management of the entrepreneurial activity.

In the context of social lending, the empowerment attained through the granted credit is highly valuable to the investors, being the subject of H3. The dummy variable *Individual* indicates whether the loan supports only one individual borrower or more people, as is the case with a group of borrowers. The applicant's gender as one of the most discussed aspects of microfinance and crowdlending is considered in the categorical variable *Gender*.

Female/male individuals or groups of only female/only male borrowers are defined as being female/male, respectively. Groups consisting of male and female individuals are categorized as being mixed. Furthermore, empowerment of the applicant's family and community is measured by *Keyword_Family* and *Keyword_Community*, which indicate the mentioning of family members and the community to which the loan applicant belongs respectively.

Last but not least, the applicant's vulnerability is measured by the dummy variables *Immigration* and *Keyword_Negative*.The immigration background of the applicant and/or his family is considered if this aspect is explicitly mentioned in the loan application. Otherwise, the applicant is assumed to be a native US inhabitant with no immigration background. Furthermore, the description text usually includes information about the applicant's social and emotional constitution. Negative keywords are associated with the applicant's vulnerability as the applicant appears to have faced severe difficulties and social abuse, such as serious illness, drug addiction, and incarceration.

The following control variables are considered in the analysis. Loan conditions like the loan amount in USD and the loan length in months are included through the variable *Principal per month*. Furthermore, the intended usage of the loan is categorized into one of 14 activity sectors, such as services and food, represented by *Activity sector*. In contract to *Keyword_Negative*, *Keyword_Positive* indicates whether the applicant has a more balanced social constitution and expresses a positive emotion in the description texts. The applicant's expectation associated with the loan is represented by the dummy variable *Keyword_Purpose*. While all loan applicants are visualized in a photograph, only a few loan applicants use a video to further emphasize their personality (dummy variable *Video*). Additionally, *US state* and *Year index* indicate where the loan applicant is located and when the loan application was posted, respectively. As a last control variable, we include *Expiration* to measure how much time the loan applicant has for fundraising on Kiva. All loan applications have a defined period during which the loan must be fully funded; otherwise, the loan application is removed from Kiva's webpage.

2.4.2 Methodology

The main determinants of successful debt funding through Kiva by socially-oriented investors are expected to be located in the areas of credit risk and social impact. All the variables related to H1 and H2 are summarized by the vector R_i in our models. The variables corresponding to H3 and H4 are represented by the vector S_i. The vector C_i

Table 2.2: Definition of variables

Variable	Expected effect	Description
Dependent variables		
Funding success		Binary variable with the value of one if a loan application meets its funding goal, zero otherwise.
Reversed funding time		Metric variable calculated as 1000 divided by the funding time (in days). The funding time indicates how long it takes loan applicants to meet funding goals. Values are logrithmized.
Cox survival time		Metric variable measuring the survival time of loan applications. The original survival time is multiplied by 100 and logarithmized. For none-funded loans, the time until expiration is employed as the original survival time.
H1 Trustee endorsement		
Trustee	+	Dummy variable with the value of one if the loan application has a trustee, zero otherwise.
Type		Trustees are categorized into individuals, non-profit organizations, others, and no trustee endorsement. Reference category: Individuals.
Trustee's experience	+	Time period in days the trustee has had experience on Kiva.
Trustee's proximity	+	Dummy variable with the value of one if the trustee and the applicant are located in the same US state, zero otherwise.
H2 Trust		
# of words	+	Length of the narrative description of the business idea and the applicant's background measured in 100 words.
Keyword_Business	+	Dummy variable with the value of one if the applicant's planned entrepreneurship is explained, zero otherwise.
Keyword_Education	+	Dummy variable with the value of one if the applicant's educational background is stated, zero otherwise.
H3 Empowerment		
Gender	+	Categorical variable for female individual/groups, male individual/group, and mixed group consisting of female and male borrowers. Reference category: Male individual/groups.
Individual	−	Dummy variable with the value of one if the loan is a individual loan, zero otherwise.
Keyword_Family	+	Empowerment in terms of family members being positively affected by the loan. Dummy variable with the value of one if family empowerment is stated, zero otherwise.
Keyword_Community	+	Empowerment in terms of the applicant's intention to benefit his or her community. Dummy variable with the value of one if community empowerment is stated, zero otherwise.
H4 Vulnerability		
Immigration	+	Dummy variable with the value of one if an immigration background of the applicant is given, zero otherwise.
Keyword_Negative	+	Dummy variable with the value of one if social dislocation of the loan applicant is mentioned, zero otherwise.
Controls		
Principal per month		Metric variable calculated as loan amount (in USD) divided by loan length (in months, the duration between the disbursal date, and the due date of the last repayment obligation).
Keyword_Positive		Dummy variable with the value of one if the applicant's positivity experienced is stated, zero otherwise.
Keyword_Purpose		Dummy variable with the value of one if the applicant's expectation with the help of loan proceeds is stated, zero otherwise.
Video		Dummy variable with the value of one if a video is available, zero otherwise.
Expiration		Metric variable (in months) calculated based on the duration between the posting date on Kiva and the planned expiration date.
Year index		Index variable for each year in which the loan application is posted in an ascending order (e.g. 1 for 2011 and 7 for 2017).
Activity sector		Activity sectors are categorized into agriculture, arts, clothing, construction, education, entertainment, food, health, housing, manufacturing, retail, service, transportation, and wholesale. Reference category: Agriculture
US state		US state in which the loan applicant is located.

represents the loan-specific controls and *Year Index*. The loan-specific error term is ϵ_i. The latent variable Y_i^* is determined through

$$Y_i^* = \beta_0 + \beta_1 R_i + \beta_2 S_i + \beta_3 C_i + \epsilon_i \, ,$$

which is fed into respective link functions according to the logistic and Tobit estimations. Primarily, funding success, being defined as a binary variable, is subject to our research. We use logistic regression models with Eicker-Huber-White robust standard errors to estimate the probability of successful debt funding. Furthermore, we are interested in the funding time which is only observable as a positive time interval for successfully funded loans but not for non-funded loans. In order not to lose the observations of non-funded loans, we investigate the reversed funding time instead of the funding time. Thus, the total data sample consists of censored (reversed funding time = 0) and uncensored (reversed funding time > 0) observations. Due to the left-censoring of the data set, Tobit regression models are chosen to estimate the linear relationship between variables. Alternatively, the Cox proportional hazard model can be applied to estimate the time until the event of successful funding without losing the observations of non-funded loans. Therefore, the Cox proportional hazard model is run as a robustness check to verify the Tobit regression results.

2.4.3 Descriptive statistics

Table 2.3 and Table 2.4 present the descriptive statistics for metric and categorical variables that contribute to testing our hypotheses, while the descriptive statistics for our control variables are displayed in Table 2.5.

The requested loan amount ranges from 100 USD to 10,000 USD, which is set as the upper credit limit by Kiva. On average, the loan duration is 25.2 months. The calculated principal per month is defined with a minimum value of 10.4 USD/month and a maximum value of 1,333.3 USD/month. Both extreme scenarios correspond to the subsample of non-funded loans. Only a small portion of the loans is requested by groups of at least two individuals as 98% are requested by individuals. The majority of loan applicants is female, compromising 57% of the entire sample. More than 60% of the successfully funded loans are given to female borrowers.

A trustee is available for less than half of the loan applications on Kiva. In the subsample of funded loans, 55% of the loans are endorsed by a trustee, whereas in the subsample of non-funded loans, only 16% of the observations are endorsed by a trustee. On average, a trustee has experience of almost 15 months, which is a factor that does not differ greatly between the subsamples. The negative minimum value of −119 days is reasonable in the case of a trustee being acquired after the public posting of a loan and the commencement of fundraising. Most of the trustees are categorized as being of the type other, followed by non-profit organizations and lastly by individuals. In more than 90% of the cases, the trustee and the loan applicant are located in the same US state.

The description text comprises an average of 545 individual words. The text description is more comprehensive in the subsample of successfully funded loans compared with the subsample of non-funded loans. Loan applications that do not state the entrepreneurial activity are seldom. The educational background is frequently stated. A share of 84% of the loan applicants provides insights into their family situation and 96% about their community. In 19% of the cases, an immigration background is explicitly mentioned in the description texts. The share of immigrants significantly differs by 7.5% between the subsamples of funded loans and non-funded loans. In less than 32% of all cases, the description texts include negative aspects.

Regarding our controls, more than 72% of the loan applications contain keywords indicating positive aspects. 80% of the loan applicants describe their expectations related to the loan. A video is not commonly available. The loans are widely distributed among the activity sectors with an emphasis on services, followed by food and retail.

Table 2.3: Descriptive statistics for metric variables

Total sample						
Variable	Obs.	Mean	S.D.	Min	Median	Max
Loan amount	6,121	4,914.41	3,036.05	100.00	5,000.00	10,000.00
Loan duration (in months)	6,121	25.24	8.14	1.00	24.00	51.00
Principal per month	6,121	183.80	86.74	10.42	208.33	1333.33
# of words (in 100 words)	6,121	5.45	2.27	0.66	5.25	26.25
Trustee's experience (in days)	2,588	442.34	472.86	-119.00	280.00	2073.00
Expiration (in days)	6,121	67.74	125.63	15.00	52.50	1,682.01
Funded loans						
Variable	Obs.	Mean	S.D.	Min	Median	Max
Funding time (in days)	4,077	44.15	30.09	0.10	39.04	300.55
Loan amount	4,077	5,206.48	2,994.86	100.00	5,000.00	10,000.00
Loan duration (in months)	4,077	25.87	8.10	1.00	24.00	51.00
Principal per month	4,077	191.92	82.07	12.50	208.33	1111.11
# of words (in 100 words)	4,077	5.70	2.22	0.84	5.56	26.25
Trustee's experience (in days)	2,255	440.50	472.36	-119.00	273.00	1986.00
Expiration (in days)	4,077	79.89	150.76	15.01	58.05	1682.01
Non-funded loans						
Variable	Obs.	Mean	S.D.	Min	Median	Max
Loan amount	2,044	4,331.85	3,034.47	125.00	5,000.00	10,000.00
Loan duration (in months)	2,044	23.98	8.09	6.00	24.00	42.00
Principal per month	2,044	167.62	93.30	10.42	166.67	1333.33
# of words (in 100 words)	2,044	4.95	2.28	0.66	4.65	21.39
Trustee's experience (in days)	333	454.76	476.72	-62.00	336.00	2073.00
Expiration (in days)	2,044	43.52	32.48	15.00	34.59	462.76

The entire data sample contains 6,121 observations. The variables are defined in Table 2.2.

Bravais-Pearson correlation coefficients for dependent and all explanatory variables are shown in Table 2.6. We do not expect any multicollinearity issues as all pairwise correlations for explanatory variables are far below 0.8, which is the critical value according to Kennedy (2008). The correlation between the two dependent variables *Funding success* and *Reversed funding time* is as high as 0.8566 which encourages us to examine both variables in separate regressions and an additional joint regression.

Table 2.4: Descriptive statistics for main categorical variables

	Total sample N=6,121		Funded loans N=4,077		Non-funded loans N=2,044	
Variable	Obs.	Relative	Obs.	Relative	Obs.	Relative
Funding success						
Yes	4,077	66.61	4,077	100.00	0	0.00
No	2,044	33.39	0	0.00	2,044	100.00
Trustee						
Yes	2,588	42.28	2,255	55.31	333	16.29
No	3,533	57.72	1,822	44.69	1,711	83.71
Type						
Individual	478	7.81	405	9.93	73	3.57
Non-Profit	899	14.69	804	19.72	95	4.65
Others	1,211	19.78	1,046	25.66	165	8.07
No endorsement	3,533	57.72	1,822	44.69	1,711	83.71
Trustee's proximity						
Yes	2,358	91.15	2,070	91.84	288	86.49
No	229	8.85	184	8.16	45	13.51
Keyword_Business						
Yes	6,053	98.89	4,031	98.87	2,022	98.92
No	68	1.11	46	1.13	22	1.08
Keyword_Education						
Yes	3,873	63.27	2,638	64.70	1,235	60.42
No	2,248	36.73	1,439	35.30	809	39.58
Individual						
Yes	6,020	98.35	3,993	97.94	2,027	99.17
No	101	1.65	84	2.06	17	0.83
Gender						
Male	2,521	41.19	1,532	37.58	989	48.39
Female	3,530	57.67	2,488	61.03	1,402	50.98
Mixed	70	1.14	57	1.40	13	0.64
Keyword_Family						
Yes	5,180	84.63	3,500	85.85	1,680	82.19
No	941	15.37	577	14.15	364	17.81
Keyword_Community						
Yes	5,897	96.34	3,937	96.57	1,960	95.89
No	224	3.66	140	3.43	84	4.11
Immigration						
Yes	1,183	19.33	889	21.81	294	14.38
No	4,938	80.67	3,188	78.19	1,750	85.62
Keyword_Negative						
Yes	1,954	31.92	1,334	32.72	620	30.33
No	4,167	68.08	2,743	67.28	1,424	69.67

The entire data sample contains 6,121 observations. Absolute values and relative values of the categorical variables are displayed. The variables are defined in Table 2.2.

Table 2.5: Descriptive statistics for categorical variables - controls

Variable	Total sample N=6,121 Obs.	Relative	Funded loans N=4,077 Obs.	Relative	Non-funded loans N=2,044 Obs.	Relative
Keyword_Positive						
Yes	4,450	72.70	2,992	73.39	1,458	71.33
No	1,671	27.30	1,085	26.61	586	28.67
Keyword_Purpose						
Yes	5,018	81.98	3,416	83.79	1,602	78.38
No	1,103	18.02	661	16.21	442	21.62
Video						
Yes	69	1.13	44	1.08	25	1.22
No	6,052	98.87	4,033	98.92	2,019	98.78
Year Index						
2011	4	0.07	4	0.10	0	0.00
2012	107	1.75	101	2.48	6	0.29
2013	361	5.90	337	8.27	24	1.17
2014	708	11.57	545	13.37	163	7.97
2015	1,163	19.00	733	17.98	430	21.04
2016	1,766	28.85	1,049	25.73	717	35.08
2017	2,012	32.87	1,308	32.08	704	34.44
Activity sector						
Agriculture	439	7.17	377	9.25	62	3.03
Arts	326	5.33	236	5.79	90	4.40
Clothing	445	7.27	288	7.06	157	7.68
Construction	95	1.55	56	1.37	39	1.91
Education	181	2.96	109	2.67	72	3.52
Entertainment	199	3.25	96	2.35	103	5.04
Food	1,361	22.23	1,071	26.27	290	14.19
Health	67	1.09	40	0.98	27	1.32
Housing	42	0.69	20	0.49	22	1.08
Manufacturing	26	0.42	20	0.49	6	0.29
Retail	974	15.91	611	14.99	363	17.76
Services	1,862	30.42	1,103	27.05	759	37.13
Transportation	92	1.50	41	1.01	51	2.50
Wholesale	12	0.20	9	0.22	3	0.15

The entire data sample contains 6,121 observations. Absolute values and relative values of the categorical variables are displayed. The variables are defined in Table 2.2.

Table 2.6: Bravais-Pearson correlation coefficients between all pairs of variables

		1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.
1.	Funding success	1.000															
2.	Reversed funding time	0.8566*	1.000														
3.	Trustee endorsement	0.3725*	0.2377*	1.000													
4.	Type_Individual	0.1118*	0.0718*	0.3401*	1.000												
5.	Type_Non_Profit	0.2008*	0.1362*	0.4848*	−0.1117*	1.000											
6.	Type_Others	0.2082*	0.1253*	0.5803*	−0.1334*	−0.1896*	1.000										
7.	Type_No endorsement	−0.3725*	−0.2377*	−1.000	−0.3029*	−0.4305*	−0.5140*	1.000									
8.	Trustee's experience	0.2120*	0.1894*	0.5794*	0.0102	0.3172*	0.4297*	−0.5794*	1.000								
9.	Trustee's proximity	−0.3426*	−0.2147*	−0.9455*	−0.3407*	−0.4663*	−0.5286*	0.9455*	−0.5381*	1.000							
10.	# of words	0.1538*	0.0928*	0.1936*	0.0442*	0.0787*	0.1404*	−0.1936*	0.0799*	−0.1896*	1.000						
11.	Keyword_Business	−0.0023	0.0044	−0.0292*	−0.0331*	−0.0001	−0.0139	0.0292*	0.0060	0.0266*	0.0659*	1.000					
12.	Keyword_Education	0.0419*	−0.0034	0.1053*	0.0323*	0.0279*	0.0840*	−0.1053*	0.0365*	−0.1050*	0.2803*	0.0130	1.000				
13.	Gender_male	−0.1036*	−0.1386*	0.0303*	0.0293*	0.0570*	0.0239	0.0329*	−0.0233	−0.0346*	−0.0717*	−0.0348*	−0.0173	1.000			
14.	Gender_female	0.0959*	0.1339*	−0.0358*	0.0103	−0.0042	0.0334*	0.1165*	0.0223	0.0407*	0.0632*	0.0354*	0.0147	−0.8527*	1.000		
15.	Gender_mixed	0.0338*	0.0190	0.0261*	−0.0063	0.0104	0.0347*	−0.0170	0.0040	−0.0292*	0.0383*	−0.0033	0.0118	−0.0820*	−0.1121*	1.000	
16.	Individual	−0.0455*	−0.0316*	−0.0319*	0.0042	−0.0042	−0.0387*	0.0319*	−0.0063	0.0396*	−0.0434*	−0.0015	−0.0162	0.0719*	0.1071*	−0.8304*	1.000
17.	Keyword_Family	0.0478*	0.0461*	0.0173	−0.0009	0.0143	0.0093	−0.0173	0.0142	−0.0143	0.1570*	0.0283*	0.0793*	−0.1311*	0.1262*	0.0203	−0.0161
18.	Keyword_Community	0.0170	0.0301*	−0.0164	−0.0179	−0.0125	0.0029	0.0164	0.0329*	0.0213	0.0966*	0.0541*	0.0230	−0.0278*	0.0267*	0.0046	−0.0048
19.	Immigration	0.0886*	0.1353*	−0.0755*	−0.0268*	−0.0102	−0.0665*	0.0755*	−0.0166	0.0802*	−0.1665*	0.0242	−0.0760*	−0.0632*	0.0701*	−0.0332*	0.0277*
20.	Keyword_Negative	0.0242	0.0248	0.0013	−0.0086	0.0069	0.0012	−0.0013	0.0151	−0.0068	0.1731*	−0.0177	0.0390*	−0.0454*	0.0497*	−0.0209	0.0144
21.	Principal per month	0.1321*	0.0427*	0.2830*	0.0853*	0.1276*	0.1801*	−0.2830*	0.1640*	−0.2708*	0.1706*	0.0046	0.0659*	0.0789*	−0.0901*	0.0535*	−0.0651*
22.	Keyword_Positive	0.0218	−0.0018	0.0323*	0.0007	0.0139	0.0272*	−0.0323*	−0.0101	−0.0327*	0.2327*	0.0470*	0.0253*	−0.0773*	0.0747*	0.0107	−0.0132
23.	Keyword_Purpose	0.0664*	0.0512*	0.0640*	0.0145	0.0240	0.0483*	−0.0640*	0.0444*	−0.0564*	0.1878*	0.0274*	0.0537*	−0.0723*	0.0741*	−0.0095	−0.0007
24.	Video	−0.0064	−0.0052	−0.0131	0.0035	−0.0006	−0.0181	0.0131	0.0232	0.0129	0.0001	−0.0034	−0.0085	0.0081	−0.0087	0.0031	0.0017
25.	Expiration	0.1365*	0.0374*	0.1838*	0.0781*	0.1196*	0.0690*	−0.1838*	−0.0221	−0.1808*	0.0613*	0.0037	0.0408*	0.0204	−0.0197	−0.0033	0.0044
26.	Funding Year 2011	0.0181	0.0242	0.0299*	0.0168	0.0261*	0.0042	−0.0267*	−0.0126	−0.0243	−0.0091	0.0027	−0.0070	0.0186	−0.0143	−0.0026	0.0033
27.	Funding Year 2012	0.0786*	0.0155	0.1558*	0.0704*	0.1202*	0.0473*	−0.1391*	−0.0420*	−0.1464*	0.0360*	−0.0215	0.0240	0.0343*	−0.0135	−0.0134	0.0173
28.	Funding Year 2013	0.1420*	0.0419*	0.2925*	0.0661*	0.1896*	0.1643*	−0.2607*	−0.0462*	−0.2731*	0.0901*	0.0067	0.0440*	0.0430*	−0.0151	0.0270*	−0.0383*
29.	Funding Year 2014	0.0795*	−0.0288*	0.2530*	0.0934*	0.1328*	0.1549*	−0.2142*	0.0062	−0.2608*	0.1240*	−0.0202	0.0710*	0.0585*	−0.0138	0.0165	−0.0093
30.	Funding Year 2015	−0.0368*	−0.0405*	0.0618*	0.0188	0.0297*	0.0703*	−0.0227	0.0245	−0.0520*	−0.0147	−0.0480*	0.0217	0.0682*	−0.0055	0.0062	−0.0125
31.	Funding Year 2016	−0.0973*	−0.0638*	−0.2144*	−0.0445*	−0.0840*	−0.1116*	0.2495*	−0.0335*	0.2090*	0.0055	0.0193	−0.0145	0.0426*	0.0391*	0.0039	0.0032
32.	Funding Year 2017	−0.0237	0.0884*	−0.2089*	−0.0478*	−0.1036*	−0.0805*	0.2483*	0.0432*	0.1986*	−0.1323*	0.0377*	−0.0809*	−0.0538*	0.1448*	−0.0144	0.0278*

		17.	18.	19.	20.	21.	22.	23.	24.	25.	26.	27.	28.	29.	30.	31.	32.
17.	Keyword_Family	1.000															
18.	Keyword_Community	0.0448*	1.000														
19.	Immigration	0.1306*	0.0139	1.000													
20.	Keyword_Negative	0.1073*	0.0420*	0.0438*	1.000												
21.	Principal per month	−0.0390*	−0.0167	−0.1089*	−0.0202	1.000											
22.	Keyword_Positive	0.0987*	0.0368*	−0.0771*	0.0208	0.0573*	1.000										
23.	Keyword_Purpose	0.0783*	0.0739*	0.0293*	0.0758*	−0.0033	0.0905*	1.000									
24.	Video	−0.0060	0.0126	−0.0209	0.0165	0.0075	−0.0284*	−0.0023	1.000								
25.	Expiration	−0.0086	−0.0028	−0.0181	0.0128	0.1638*	0.0213	0.0092	−0.0183	1.000							
26.	Funding Year 2011	0.0109	0.0050	0.0037	−0.0175	0.0379*	0.0157	0.0120	−0.0027	0.3281*	1.000						
27.	Funding Year 2012	−0.0157	−0.0338*	−0.0116	0.0236	0.1651*	0.0062	0.0074	−0.0142	0.5958*	−0.0032	1.000					
28.	Funding Year 2013	−0.0010	−0.0362*	−0.0066	−0.0123	0.0898*	0.0289*	0.0452*	−0.0267*	0.0819*	−0.0060	−0.0311*	1.000				
29.	Funding Year 2014	0.0125	−0.0465*	−0.0515*	0.0098	0.0136	0.0290*	0.0460*	−0.0386*	0.0309*	−0.0086	−0.0448*	−0.0839*	1.000			
30.	Funding Year 2015	−0.0568*	−0.0719*	−0.0947*	−0.0386*	0.1345*	0.0126	−0.0514*	−0.0517*	−0.0424*	−0.0115	−0.0598*	−0.1120*	−0.1615*	1.000		
31.	Funding Year 2016	−0.0625*	0.0319*	−0.0770*	−0.0168	−0.0208	−0.0080	−0.0279*	−0.0680*	−0.0893*	−0.0150	−0.0782*	−0.1465*	−0.2111*	−0.2817*	1.000	
32.	Funding Year 2017	0.1035*	0.0882*	0.1948*	0.0423*	−0.1946*	−0.0396*	0.0132	0.1526*	−0.1246*	−0.0164	−0.0856*	−0.1605*	−0.2313*	−0.3088*	−0.4037*	1.000

The variables Activity sector and US state are not explicitly reported due to the number of categories and minor explanatory power. Values labeled with the symbol * are significant at the 5% level. The variables are defined in Table 2.2.

2.5 Results

2.5.1 Results regarding the funding success

To commence, we focus on the empirical results of the estimated logistic models regarding the probability of funding success on Kiva. The respective logistic regression models are presented in Table 2.7. Model I is the basic model consisting of details that are obvious in the loan applications. It is extended by adding the different types of trustees in model II. Model III is the main model, including visible and less-visible details on credit risk indicators and social performance indicators of loan applications as determinants of funding success.

The dummy variable clarifying whether or not a loan application is endorsed by a trustee provides a clear picture as it is positive and significant at the 1% level (coeff.: 1.5398, st.err.: 0.0886[7]). Loans that are endorsed by a trustee are more likely to be funded than loans without a trustee endorsement. The result is further strengthened by the dummy variables depicting the type of trustee in model II. While loans without an endorsement are less likely to be funded compared with loans endorsed by an individual, loans promoted by a non-profit organization are even more likely to be funded.

Furthermore, the foundation of trust between investors and borrowers is expected to play a role. The length of the description text is used as a measurement for the borrower's willingness to share information. The coefficient of *# of words* is positive and significant (coeff.: 0.2465, st.err.: 0.0468). Therefore, the longer the text description, the higher the probability of successful funding. However, *# of words*^{2} has a significant and negative coefficient indicating an overall inverse u-shaped relation (coeff.: -0.0096, st.err.: 0.0032). Regarding the coefficients of *Keyword_Business* and *Keyword_Education*, we are unable to find any evidence as both of them are not significant.

Concerning H3 and H4, we observe the following. The dummy variable demonstrating female borrowers is positive and significant in all model specifications (coeff.: 0.5298, st.err.: 0.0698). Female borrowers are more likely to receive funding than their male counterparts. Regarding the question of individual vs. group loans, we can ascertain that individual applicants have more difficulties receiving funding than groups of borrowers (coeff.: -0.9528, st.err.: 0.5608). The variable *Keyword_Family* proves to be negatively related to funding success (coeff.: -0.1965, st.err.: 0.0915). The result is significant and contradictory to

[7] If not otherwise specified, the coeff. and st.err. in parentheses are from model III.

Table 2.7: Coefficients of logistic models on funding success

	All observations			with trustee		w/o trustee
	(I)	(II)	(III)	(IV)	(V)	(VI)
Trustee endorsement						
Trustee	1.5852***		1.5398***			
	(0.0881)		(0.0886)			
Type_Non_Profit		0.3114*		-0.0535	0.0455	
		(0.1828)		(0.1903)	(0.1920)	
Type_Others		-0.0463		-0.3274*	-0.2496	
		(0.1661)		(0.1783)	(0.1798)	
Type_No_End.		-1.5132***				
		(0.1510)				
Trustee's experience				0.0004***	0.0002	
				(0.0002)	(0.0002)	
Trustee's proximity				0.6564***	0.6614***	
				(0.2168)	(0.2150)	
Trust						
# of words			0.2465***	0.2647***	0.2645***	0.2683***
			(0.0468)	(0.0872)	(0.0873)	(0.0587)
# of words^2			-0.0096***	-0.0093	-0.0089	-0.0109***
			(0.0032)	(0.0063)	(0.0063)	(0.0041)
Keyword_Business			0.0294	0.3369	0.3248	-0.1019
			(0.2981)	(0.4354)	(0.4378)	(0.3940)
Keyword_Education			-0.0226	0.1292	0.1525	-0.0332
			(0.0696)	(0.1344)	(0.1336)	(0.0839)
Empowerment						
Individual	-1.0796*	-1.0952**	-0.9528*	-1.3781	-1.4042	-0.8295
	(0.5519)	(0.5507)	(0.5608)	(1.1968)	(1.1501)	(0.7150)
Gender_female	0.5735***	0.5769***	0.5298***	0.2388*	0.2280*	0.6389***
	(0.0683)	(0.0683)	(0.0698)	(0.1369)	(0.1365)	(0.0849)
Gender_mixed	-0.3578	-0.3598	-0.2615			-0.4520
	(0.6348)	(0.6339)	(0.6404)			(0.8156)
Keyword_Family			-0.1965**	-0.0707	-0.1169	-0.2381**
			(0.0915)	(0.1710)	(0.1706)	(0.1116)
Keyword_Community			0.0729	-0.1594	-0.2733	0.1845
			(0.1642)	(0.3191)	(0.3264)	(0.2048)
Vulnerability						
Immigration			0.5991***	-0.1029	-0.1323	0.7473***
			(0.0969)	(0.1803)	(0.1801)	(0.1095)
Keyword_Negative			-0.0162	-0.1147	-0.1093	0.0261
			(0.0715)	(0.1378)	(0.1375)	(0.0854)
Controls						
Keyword_Positive			-0.1092	-0.1465	-0.1414	-0.1024
			(0.0751)	(0.1461)	(0.1452)	(0.0909)
Keyword_Purpose			0.1307	0.1193	0.1520	0.1046
			(0.0838)	(0.1709)	(0.1692)	(0.1007)
Principal per month	0.0002	0.0002	-0.0001	-0.0003	-0.0001	0.0002
	(0.0004)	(0.0004)	(0.0004)	(0.0008)	(0.0008)	(0.0005)
Video	-0.1297	-0.1395	-0.1438	-0.6957	-0.8832*	0.2408
	(0.3123)	(0.3142)	(0.3107)	(0.4658)	(0.4799)	(0.3371)
Expiration	0.0342***	0.0341***	0.0335***	0.0213***	0.0262***	0.0372***
	(0.0035)	(0.0035)	(0.0035)	(0.0046)	(0.0059)	(0.0046)
Year Index	0.3287***	0.3309***	0.3196***		0.2132**	0.3169***
	(0.0407)	(0.0408)	(0.0419)		(0.0844)	(0.0526)
Activity sector	yes	yes	yes	yes	yes	yes
US state	yes	yes	yes	yes	yes	yes
_cons	-1.5408*	-0.0480	-2.7005***	1.3975	0.2327	-3.0215***
	(0.8953)	(0.9101)	(1.0089)	(1.4965)	(1.6047)	(1.0334)
N	6,121	6,121	6,121	2,550	2,550	3,533
Pseudo R^2	0.260	0.261	0.273	0.135	0.140	0.213

Models I - III include all observations. Models IV - VI consider the subsamples of loans with and without a trustee endorsement separately. Model I is extended by including the different types of trustees, resulting in Model II. Model III is the main model including several social performance indicators that have been extracted through keywords from the description text. Models IV - VI follow the main model. Eicker-Huber-White heteroskedastic-consistent errors are used. Values labeled with the symbols *, **, and *** are significant at the 10% level, the 5% level, and the 1% level. The variables are defined in Table 2.2.

our expectations. One possible reason could be that the borrower's responsibility for his or her family members appears to be obstructive in terms of entrepreneurship as opposed to being positively perceived in terms of empowerment. The second variable representing community empowerment is positive but not significant (coeff.: 0.0729, st.err.: 0.1642).

Furthermore, the coefficient of the immigration dummy variable is positive and significant at the 1% level (coeff.: 0.5991, st.err.: 0.0969), providing evidence that immigrants are more likely to be successfully funded through the crowd of socially-oriented investors. One reason behind this finding may be that the investors perceive immigrants as being needier and more vulnerable as they often suffer from exclusion in the United States. In contrast, the borrower's previous social dislocation stated by negative words does not appear to be a significant determinant.

The considered control variable for the time until the expiration of the loan application shows a positive and significant coefficient (coeff.: 0.0335, st.err.: 0.0035). Loans without a strict time limit for fundraising are more likely to be funded. It is interesting that *Year Index* is positive and significant (coeff.: 0.3196, st.err.: 0.0419), which could be considered as an indication for the investor's learning curve in terms of supporting more US direct loans over time. Taking into account that the volume of US direct loans on Kiva has increased significantly over the last years (see Table 2.5) as well as the positive development of funding success, it appears promising that the investors are becoming more confident when providing capital directly to US inhabitants in need. None of the other controls such as *Keyword_Positive*, *Principal per month*, and *Video* provide any further insights.

Additionally, we divide the data sample into two subsamples with and without a trustee endorsement and run subsample regressions. In the subsample of endorsed loans, 38 observations are lost as all loans requested by a mixed group of female and male individuals are successfully funded. The focus on the subsample of loans with a trustee endorsement in models IV and V allows us to include further variables that provide details about the trustees and the investors' responses to them. *Trustee's experience* is positive and significant in model IV, but not in model V, which also includes *Year Index*. Consequently, as the trustee's experience in days increases with the years, the result appears to be time-dependent and should not be overvalued. A noteworthy observation is the positive and significant coefficient of *Trustee's proximity* (Model V: coeff.: 0.6614, st.err.: 0.2150). The fact that the trustee and the borrower are located in the same US state is positively related to funding success. One reason behind this finding could be that the endorsement from a trustee who is geographically closer to the borrower is more recognized and valued by the investors.

Regarding the borrower's willingness to share information in the description texts, the results are similar to those in the total data set. The coefficient of *# of words* is positive and significant (Model V: coeff.: 0.2645, st.err.: 0.0873 / Model VI: coeff.: 0.2683, st.err.: 0.0587). The inverse u-shaped relation is only significant for the subsample of loans without a trustee endorsement in column VI (Model VI: coeff.: -0.0109, st.err.: 0.0041). *Keyword_Business* and *Keyword_Education* remain insignificant. Regarding empowerment, in contrast to the main models, the individual dummy is not significant for either of the subsamples, but female borrowers still appear to be targeted by the investors (Model V: coeff.: 0.2280, st.err.: 0.1365 / Model VI: coeff.: 0.6389, st.err.: 0.0849). *Keyword_Family* remains negative and significant in the subsample of non-endorsed loans (Model VI: coeff.: -0.2381, st.err.: 0.1116). This may signal the investors' doubt about the possibility of the explicitly mentioned care of family members being brought into line with successful entrepreneurship, especially without a trustee endorsement. *Keyword_Community* remains insignificant.

Interestingly, the vulnerability of borrowers emphasized by the immigration background does not appear to have any impact on the funding probability in the subsample of endorsed loans. In contrast, the immigration dummy is positive and significant in the subsample of loans without a trustee endorsement (Model VI: coeff.: 0.7473, st.err.: 0.1095). One possible explanation could be that direct loan investors are not a homogenous group. One group of investors that support non-endorsed loans could have a higher weighting factor α for the social return in their utility function. These more socially-oriented investors would still choose to support an immigrant without a trustee endorsement if the contribution of the social return to the personal utility is enough to compensate for a more negative financial return, indicated by the lack of a trustee endorsement. On the contrary, another group of less socially-oriented investors could focus more on the credit profile of the borrowers and pay less attention to the social impact of lending. However, an alternative explanation could be that all of direct loan investors simply apply different selection criteria for endorsed and non-endorsed loan applications. It is possible that the investors would emphasize more on the possibility of making a social contribution for non-endorsed applications as they are riskier. But if the loan application is endorsed by a trustee, the investors may worry less about associated credit risk and not ask for further evidence indicating possible social impact. The above two possible explanations could also explain why the coefficient of the female dummy is smaller and less significant in the subsample of endorsed loans (Model V: coeff.: 0.2280, st.err.: 0.1365) than in the subsample of non-endorsed loans (Model V: coeff.: 0.6389, st.err.: 0.0849). In addition, the coefficients of *Keyword_Negative* have contrary signs in the subsample analysis, though they are insignificant (Model V: coeff.: -0.1093, st.err.: 0.1375 / Model VI: coeff.: 0.0261, st.err.: 0.0854).

The results of all included control variables remain fairly unchanged compared with the models on the total data set.

2.5.2 Results regarding the funding time

In addition to funding success, we can also observe the funding time of loan applications. We use the reversed funding time as a dependent variable, thereby measuring the funding speed. The model set up is analogous to the logistic models. The results are displayed in Table 2.8. Models I to III include the entire 6,121 observations for funded and non-funded loans independently of whether the loan has a trustee endorsement.

All variables reveal a similar significance pattern as compared to the funding success analysis, implying that the same variables can explain funding success and speed. When inspecting the controls, the coefficient of *Keyword_Purpose* is positive and significant (coeff.: 0.1030, st.err.: 0.0545), which marks a first indication that the investors are attracted by the borrower's expectations related to receiving the loan. The coefficient of *Principal per month* is negatively related to the reversed funding time (coeff.: −0.0008, st.err.: 0.0003). This could be due to the investor's distrust in the borrower's ability to repay a proportionally high loan amount succeeding a short loan period. Loan applications including positive keywords appear to experience a slower funding process (coeff.: −0.0854, st.err.: 0.0473).

All the other control variables demonstrate the same significant relations as those in the funding success analysis.

2.5.3 Implication regarding the hypotheses

All in all, H1, which states that the existence of a trustee is positively related to funding success, is supported. Moreover, the borrower's willingness to share information is positively related to funding success and the reversed funding time as it appears to build trust and attracts the investors, which supports our expectation in H2. However, text descriptions that are too long tend to deter the investors. Signals of entrepreneurship and education in the text description do not appear to influence the investors' behavior.

Evidence in favor of H3 is observed in terms of empowering women as female borrowers are favored by the investors. Groups of borrowers are more likely to be funded and

Table 2.8: Coefficients of Tobit models on reversed funding time

	all observations			with trustee		w/o trustee
	(I)	(II)	(III)	(IV)	(V)	(VI)
Trustee endorsement						
Trustee	1.0305***		0.9765***			
	(0.0483)		(0.0484)			
Type_Non_Profit		0.1144		-0.0708	-0.0015	
		(0.0883)		(0.0626)	(0.0634)	
Type_Others		-0.0537		-0.1754***	-0.1236**	
		(0.0841)		(0.0599)	(0.0602)	
Type_No_End.		-1.0187***				
		(0.0798)				
Trustee's experience				0.0003***	0.0001***	
				(0.0000)	(0.0001)	
Trustee's proximity				0.2308***	0.2315***	
				(0.0790)	(0.0784)	
Trust						
# of words			0.1671***	0.0819**	0.0832**	0.2569***
			(0.0289)	(0.0351)	(0.0348)	(0.0476)
# of words^2			-0.0070***	-0.0030	-0.0028	-0.0106***
			(0.0019)	(0.0024)	(0.0023)	(0.0030)
Keyword_Business			-0.0240	0.0843	0.0684	-0.0396
			(0.1942)	(0.1717)	(0.1704)	(0.4057)
Keyword_Education			-0.0497	0.0022	0.0097	-0.0458
			(0.0443)	(0.0455)	(0.0452)	(0.0797)
Empowerment						
Individual	-0.6085**	-0.6186**	-0.5539**	-0.2963	-0.3212	-0.7161
	(0.2753)	(0.2752)	(0.2727)	(0.2497)	(0.2479)	(0.5373)
Gender_female	0.4591***	0.4606***	0.4209***	0.1370***	0.1291***	0.6938***
	(0.0432)	(0.0432)	(0.0434)	(0.0434)	(0.0431)	(0.0805)
Gender_mixed	-0.2460	-0.2450	-0.2003	-0.0741	-0.0932	-0.2567
	(0.3317)	(0.3314)	(0.3285)	(0.2985)	(0.2964)	(0.6573)
Keyword_Family			-0.0887	-0.0152	-0.0308	-0.2404**
			(0.0590)	(0.0591)	(0.0587)	(0.1081)
Keyword_Community			0.0653	0.0341	-0.0232	0.1597
			(0.1111)	(0.1046)	(0.1043)	(0.2169)
Vulnerability						
Immigration			0.4342***	-0.0694	-0.0749	0.7463***
			(0.0553)	(0.0581)	(0.0577)	(0.0988)
Keyword_Negative			0.0121	-0.0643	-0.0621	0.0642
			(0.0441)	(0.0445)	(0.0442)	(0.0801)
Controls						
Keyword_Positive			-0.0854*	-0.0717	-0.0667	-0.0945
			(0.0473)	(0.0481)	(0.0477)	(0.0855)
Keyword_Purpose			0.1030*	0.0250	0.0446	0.1095
			(0.0545)	(0.0578)	(0.0575)	(0.0957)
Principal per month	-0.0007***	-0.0007***	-0.0008***	-0.0004	-0.0004	-0.0008*
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0005)
Video	-0.1104	-0.1180	-0.0970	-0.4076*	-0.5013**	0.1679
	(0.1956)	(0.1956)	(0.1940)	(0.2143)	(0.2132)	(0.3242)
Expiration	0.0008***	0.0008***	0.0007***	0.0001	0.0003***	0.0156***
	(0.0002)	(0.0002)	(0.0002)	(0.0001)	(0.0001)	(0.0014)
Year Index	0.1702***	0.1724***	0.1660***		0.1032***	0.3287***
	(0.0189)	(0.0190)	(0.0191)		(0.0189)	(0.0476)
Activity sector	yes	yes	yes	yes	yes	yes
US state	yes	yes	yes	yes	yes	yes
_cons	2.0411***	3.0306***	1.3009**	3.1608***	2.8117***	-0.3850
	(0.4986)	(0.5013)	(0.5417)	(0.3698)	(0.3726)	(0.8794)
N	6,121	6,121	6,121	2,588	2,588	3,533
Pseudo R^2	0.072	0.072	0.078	0.040	0.044	0.087

Models I - III include all observations. Models IV - VI consider the subsamples of loans with and without a trustee endorsement separately. Model I is extended by including different types of trustees, resulting in Model II. Model III is the main model including several social performance indicators, which have been extracted through keywords from the description text. Models IV - VI follow the main model. Eicker-Huber-White heteroskedastic-consistent errors are used. Values labeled with the symbols *, **, and *** are significant at the 10% level, the 5% level, and the 1% level. The variables are defined in Table 2.2.

receive funding faster when considering the total sample, but this is not apparent in the subsample regressions. Empowerment beyond the borrowers themselves does not appear to attract the investors. In the subsample of loans without a trustee endorsement, the investors are even reluctant to provide capital to applicants who explicitly mention their responsibility towards family members.

H4 on the vulnerability of the borrowers is partly confirmed for the complete sample and the subsample of loans without a trustee endorsement. The financial needs of immigrants are recognized and the investors strive to support these applicants. But this is not the case for those borrowers with a trustee endorsement. This is preliminary evidence that direct loan investors do react to the vulnerability of the borrowers, but—dependently on an endorsement—to a different extent.

2.5.4 Robustness checks

To assess the robustness of our main findings, Cox proportional hazard models, which analyze the 'survival time' of the loan application, are carried out. There, both funding success and the funding time, are jointly considered as the time interval until the event of being successfully funded is estimated. For non-funded loans, the time until expiration is employed as the survival time. The survival time is multiplied by 100 and logarithmized to derive the variable *Cox survival time*[8]. The regression results are shown in Table 2.9.

Considering all observations in the columns I, II, and III, the majority of variables reveals itself to be consistent with our main results. A difference arises regarding the signals that build trust. The inverse u-shaped relation between the dependent variable and the text length is not confirmed anymore. However, the tendency remains unchanged (coeff.: -0.0024, st.err.: 0.0017). *Keyword_Education* turns out to be negative and slightly significant (coeff.: -0.0662, st.err.: 0.0353). Furthermore, the coefficient of *Keword_Family* becomes significant (coeff.: -0.1028, st.err.: 0.0470). In summary, the overall picture is robust as our hypotheses are supported by the main indicators.

The results of Cox models for the subsamples of loans both with and without a trustee endorsement are presented in columns IV - VI. Most of the results remain stable with the same values and slightly changed confidence levels. A considerable gain in insight can be derived from the fact that the borrower's vulnerability can attract the investors in the

[8]As 7 loan applications are funded within one day, the survival time is multiplied by 100 to avoid negative logarithmic values.

Table 2.9: Coefficients of Cox proportional hazard models

	Cox proportional hazard models					
	all observations			with trustee		w/o trustee
	(I)	(II)	(III)	(IV)	(V)	(VI)
Trustee endorsement						
Trustee	0.3727***		0.3525***			
	(0.0386)		(0.0385)			
Type_Non_Profit		0.0270		-0.2021***	-0.0680	
		(0.0655)		(0.0693)	(0.0697)	
Type_Others		-0.0695		-0.2387***	-0.1499**	
		(0.0611)		(0.0654)	(0.0644)	
Type_No_End.		-0.3966***				
		(0.0608)				
Trustee's experience				0.0006***	0.0002***	
				(0.0001)	(0.0001)	
Trustee's proximity				0.2757***	0.2432***	
				(0.0936)	(0.0908)	
Trust						
# of words			0.0739***	0.0706**	0.0800**	0.1609***
			(0.0247)	(0.0346)	(0.0370)	(0.0427)
# of words^2			-0.0024	-0.0030	-0.0032	-0.0071**
			(0.0017)	(0.0024)	(0.0027)	(0.0030)
Keyword_Business			0.0181	-0.0990	-0.0675	0.1496
			(0.1458)	(0.1589)	(0.1850)	(0.2435)
Keyword_Education			-0.0662*	-0.0266	-0.0260	-0.0521
			(0.0353)	(0.0484)	(0.0494)	(0.0517)
Empowerment						
Individual	-0.5542***	-0.5560***	-0.5333***	-0.3673	-0.4348*	-0.6644*
	(0.1683)	(0.1701)	(0.1737)	(0.2673)	(0.2423)	(0.3895)
Gender_female	0.2716***	0.2726***	0.2642***	0.1411***	0.1444***	0.4054***
	(0.0339)	(0.0339)	(0.0345)	(0.0463)	(0.0471)	(0.0533)
Gender_mixed	-0.3459*	-0.3372	-0.3364	-0.2237	-0.2741	-0.4695
	(0.2066)	(0.2080)	(0.2135)	(0.3050)	(0.2881)	(0.4505)
Keyword_Family			-0.1028**	-0.0105	-0.0359	-0.2186***
			(0.0470)	(0.0638)	(0.0655)	(0.0724)
Keyword_Community			0.0160	0.1690	0.0707	0.0542
			(0.0849)	(0.1056)	(0.1055)	(0.1400)
Vulnerability						
Immigration			0.1873***	-0.1366**	-0.1513**	0.4527***
			(0.0437)	(0.0603)	(0.0613)	(0.0589)
Keyword_Negative			-0.0012	-0.1017**	-0.0950**	0.0971*
			(0.0353)	(0.0465)	(0.0465)	(0.0523)
Controls						
Keyword_Positive			-0.0941**	-0.0881*	-0.0991*	-0.0309
			(0.0376)	(0.0505)	(0.0508)	(0.0554)
Keyword_Purpose			0.0599	-0.0106	0.0180	0.0577
			(0.0445)	(0.0620)	(0.0621)	(0.0637)
Principal per month	-0.0006***	-0.0006***	-0.0007***	-0.0002	0.0000	-0.0018***
	(0.0002)	(0.0002)	(0.0002)	(0.0003)	(0.0003)	(0.0003)
Video	-0.0519	-0.0564	-0.0440	-0.0692	-0.2775	0.1088
	(0.1258)	(0.1255)	(0.1240)	(0.1910)	(0.1948)	(0.1889)
Expiration	-0.0020	-0.0020	-0.0020	-0.0007**	-0.0002	-0.0385***
	(0.0017)	(0.0017)	(0.0017)	(0.0003)	(0.0002)	(0.0015)
Year Index	0.2647***	0.2651***	0.2687***		0.2365***	0.1755***
	(0.0201)	(0.0202)	(0.0206)		(0.0220)	(0.0306)
Activity sector	yes	yes	yes	yes	yes	yes
US state	yes	yes	yes	yes	yes	yes
N	6,121	6,121	6,121	2,588	2,588	3,533
Pseudo R^2	0.017	0.017	0.017	0.013	0.018	0.054

Robustness analysis through Cox proportional hazard models for the total data sample and exclusively for the subsamples of loans with a trustee endorsement as well as loans without a trustee endorsement. Eicker-Huber-White heteroskedastic-consistent errors are used. Values labeled with the symbols *, **, and *** are significant at the 10% level, the 5% level, and the 1% level. The variables are defined in Table 2.2.

subsample of non-endorsed loans but outfaces the investors in the subsample of endorsed loans. Both variables—*Immigration* (Model V: coeff.: −0.1513, st.err.: 0.0613 / Model VI: coeff.: 0.4527, st.err.: 0.0589) and *Keyword_Negative* (Model V: coeff.: −0.0950, st.err.: 0.0465 / Model VI: coeff.: 0.0971, st.err.: 0.0523)—demonstrate significant and opposite coefficients in the subsamples. The control variable *Keywords_Positive* is negative and significant in the subsample of endorsed loans, indicating that these investors do not appreciate positive emotions (Model V: coeff.: −0.0991, st.err.: 0.0508).

Table 2.10: Robustness analysis through further logistic and probit models on funding success

	all observations			with trustee		w/o trustee
	(I)	(II)	(III)	(IV)	(V)	(VI)
Trustee endorsement						
Trustee	1.6851***	1.5450***	0.9357***			
	(0.0933)	(0.0896)	(0.0516)			
Type_Non_Profit				-0.0116	0.0203	
				(0.1025)	(0.1038)	
Type_Others				-0.1623*	-0.1373	
				(0.0966)	(0.0973)	
Trustee's experience				0.0002**	0.0001	
				(0.0001)	(0.0001)	
Trustee's proximity				0.3744***	0.3769***	
				(0.1197)	(0.1197)	
Trust						
# of words	0.2555***	0.2351***	0.1521***	0.1531***	0.1528***	0.1690***
	(0.0479)	(0.0475)	(0.0272)	(0.0522)	(0.0525)	(0.0339)
# of words^2	-0.0101***	-0.0089***	-0.0063***	-0.0059	-0.0058	-0.0073***
	(0.0033)	(0.0033)	(0.0019)	(0.0038)	(0.0038)	(0.0023)
Keyword_Business	0.0660	0.1604	-0.0262	0.1390	0.1375	-0.0891
	(0.3021)	(0.2849)	(0.1725)	(0.2454)	(0.2474)	(0.2303)
Keyword_Education	-0.0118	-0.0431	-0.0028	0.0725	0.0788	-0.0051
	(0.0699)	(0.0710)	(0.0411)	(0.0735)	(0.0735)	(0.0503)
Empowerment						
Individual	-0.9245	-0.9519*	-0.5706*	-0.6569	-0.6827	-0.4851
	(0.5678)	(0.5609)	(0.3065)	(0.5611)	(0.5499)	(0.4119)
Gender_female	0.5145***	0.5297***	0.3051***	0.1331*	0.1315*	0.3754***
	(0.0704)	(0.0710)	(0.0406)	(0.0724)	(0.0724)	(0.0498)
Gender_mixed	-0.2172	-0.2475	-0.1154			-0.2196
	(0.6495)	(0.6398)	(0.3534)			(0.4718)
Keyword_Family	-0.2099**	-0.1521	-0.1032*	-0.0597	-0.0734	-0.1252*
	(0.0927)	(0.0925)	(0.0533)	(0.0936)	(0.0938)	(0.0663)
Keyword_Community	0.0683	0.0956	0.0574	-0.0516	-0.0823	0.0892
	(0.1650)	(0.1662)	(0.1029)	(0.1704)	(0.1738)	(0.1229)
Vulnerability						
Immigration	0.8397***	0.5919***	0.3340***	-0.0688	-0.0726	0.4254***
	(0.1033)	(0.0987)	(0.0566)	(0.0957)	(0.0957)	(0.0662)
Keyword_Negative	-0.0162	-0.0113	-0.0039	-0.0609	-0.0591	0.0209
	(0.0718)	(0.0728)	(0.0423)	(0.0741)	(0.0742)	(0.0521)
Interaction						
Trustee * Immigration	-1.0468***					
	(0.1982)					
Controls						
Keyword_Positive	-0.1057	-0.1257	-0.0683	-0.0693	-0.0688	-0.0629
	(0.0754)	(0.0768)	(0.0439)	(0.0788)	(0.0787)	(0.0535)
Keyword_Purpose	0.1171	0.1432*	0.0798	0.0785	0.0902	0.0620
	(0.0845)	(0.0849)	(0.0490)	(0.0924)	(0.0923)	(0.0591)
Principal per month	0.0000	0.0000	-0.0001	-0.0001	-0.0001	0.0001
	(0.0004)	(0.0004)	(0.0002)	(0.0004)	(0.0004)	(0.0003)
Video	-0.1572	-0.1880	-0.0848	-0.3969	-0.4572	0.1377
	(0.3164)	(0.3130)	(0.1803)	(0.2793)	(0.2829)	(0.2031)
Expiration	0.0338***	0.0332***	0.0129***	0.0090***	0.0101***	0.0140***
	(0.0035)	(0.0036)	(0.0018)	(0.0020)	(0.0024)	(0.0026)
Year Index	0.3151***	0.3271***	0.1492***		0.0668	0.1788***
	(0.0418)	(0.0433)	(0.0237)		(0.0411)	(0.0301)
Activity sector	yes	yes	yes	yes	yes	yes
US state	yes	yes	yes	yes	yes	yes
_cons	-2.8633***	-2.6913**	-0.9911*	0.8583	0.5497	-1.3529**
	(1.0358)	(1.0850)	(0.5492)	(0.7242)	(0.7653)	(0.6039)
N	6,121	5,927	6,121	2,550	2,550	3,533
Pseudo R^2	0.276	0.269	0.256	0.128	0.130	0.192

Logit Model I includes an additional interaction term of trustee endorsement and immigration background. Logit Model II is based on loan applications with a loan amount > 1,000 USD. Models III - VI are probit models analogous to the main Logit models for the total data sample and exclusively for the subsamples of loans with and without a trustee endorsement. Eicker-Huber-White heteroskedastic-consistent errors are used. Values labeled with the symbols *, **, and *** are significant at the 10% level, the 5% level, and the 1% level. The variables are defined in Table 2.2.

Furthermore, we run additional logistic regressions and probit regressions on funding success, which are shown in Table 2.10. First, we include an interaction term of *Trustee* and *Immigration* in the main logistic model to further investigate how various investors of endorsed and non-endorsed loans behave in regards to loan applicants with an immigration background. The interaction term is negative and significant at the 1% level (Model I: coeff.: -1.0468, st.err.: 0.1982). This implies that in the subsample of endorsed loans the immigration background indeed is not appreciated and lowers the probability of funding, while it increases the funding probability in the subsample of non-endorsed loans.

Second, all loan applications with an amount of less than 1,000 USD are excluded as these are less likely to properly support or enable entrepreneurship. The majority of variables does not change. The negative coefficient of *Keyword_Family* is not significant anymore. *Keyword_Purpose* turns out to be significant, indicating that the borrower's expectation increases in importance concerning higher volume loans (Model II: coeff.: 0.1432, st.err.: 0.0849).

Third, probit models analogous to the logistic models on all observations and the subsamples of endorsed and non-endorsed loans are run. The results are shown in columns III to VI. All variables employed to test the hypotheses on credit risk and social impact remain stable and are consistent with our main results.

2.6 Conclusion

In this paper, we study the funding determinants of interest-free P2P lending by utilizing a unique data set of direct loans requested by US inhabitants on the microfinancing platform Kiva during the observation interval from 2011 to 2017. The data set is unique as it represents social financing without interest compensation for credit risk to a borrower group from a developed country and utilizes textual information from original loan application texts.

The underlying Kiva model enables direct P2P lending between microentrepreneurs and investors. Although the investors bear the full credit risk, they are willing to grant interest-free loans to the borrowers, who are US inhabitants facing financial exclusion from the formal capital market.

Logistic regressions on funding success and Tobit regressions on the reversed funding time provide interesting insights into the investors' behavior regarding investment decision making. The existence of social underwriting through a trustee endorsement appears to have a highly positive impact on funding success and the reversed funding time. Furthermore, the description length as a measurement to share information and generate the investor's trust is highly related to the probability of funding success as well as the funding time. Empowerment representing the investment's social impact appears to be a crucial predictor. Female borrowers are clearly preferred by all investors. Furthermore, groups of borrowers are more likely to be both funded and funded faster in the total sample. However, we do not find evidence that the investors appreciate empowerment of other people beyond the borrowers. At first glance, the borrower's vulnerability measured by the immigration background is positively related to funding success and the reverse funding time in the total sample. Further subsample analysis indicates that the investors respond to the borrower's vulnerability to a varying extent.

In summary, our findings lead to the conclusion that the investment decisions of the involved interest-free P2P investors take into account the credit risk as well as the social impact of the respective investment. There are no indications that they only use 'play money' to generate some amusement for themselves, as they appear to invest very seriously and goal-oriented. Our research provides insights into the investors' financial and ethical considerations in the context of online P2P microfinancing in developed countries such as the United States. As a practical implication for potential borrowers, it can be stated that for them it is advantageous to be able to acquire a trustee endorsement. If this is not possible, then the applicants can be advised to at least write a comprehensive text in which they reveal their need for empowerment and/or their vulnerability. For the observation period, it is arguable that the borrows could not know exactly which features of their texts would boost the probability of being funded so that we regard the bias through purposeful dishonesty as negligible for this study. However, in the future it cannot be excluded that applicants use the findings revealed herein.

Last, this research also has some limitations: First, the data set does not contain information about the repayment of the granted loans, which would be essential to investigate the drivers of credit risk. Thus we cannot evaluate to which extent those variables that we used to test the credit risk hypothesis are really correlated with defaults. Second, as Kiva not only hosts the P2P lending platform subject to this research, but also the much larger intermediary-based model, the investors active there may be influenced by the latter model. Therefore, investors on a different interest-free P2P platform may behave differently. Due to this and the possibly different institutional features on other platforms, our results should only be generalized with caution. Third, the way we use keywords in the description texts as proxies for different financial and ethical aspects may still be connected with some inaccuracy.

This leaves room for further research. For instance, with our data set we cannot finally clarify why the investors respond differently to the signals of vulnerability when screening endorsed and non-endorsed loans. Moreover, more precise proxies such as deeper linguistic features or surveys among active investors are needed to better understand the investors' behavior in such a prosocial P2P context. Summarizing, further research on the innovative interest-free P2P model appears to have a promising potential.

Chapter 3

The pricing of green bonds: external reviews and the shades of green

This research paper is joint work with Gregor Dorfleitner and Sebastian Utz. The paper has been published as: Dorfleitner, G., Utz, S. & Zhang, R. The pricing of green bonds: external reviews and the shades of green. Rev Manag Sci (2021).

Abstract: We investigate the asset pricing implications of the greenness of bonds. To estimate a green-pricing effect, we determine the 'green bond premium' as the difference between the yields of matched conventional and green-labeled bonds. On a cross-sectional average, green bonds experience a statistically significant positive premium. This premium increases with external greenness evaluations, i.e., investors accept premiums of up to 5 basis points for bonds with a substantial environmental agenda. This external validation effect, which is strongest for bonds that are rated dark-green, may offset not incurring information costs, as this effect decreases with increasing age of bonds.

Keywords: Green bond premium, External review, Second-party opinion, Shade of green, Climate finance, Impact investing

MSC Classification: 91B76 91B16 62P20

3.1 Introduction

In recent finance literature, there has been a lively debate on the asset pricing implications of sustainable and particularly green investment opportunities (Bolton and Kacperczyk, 2021; Cheema-Fox et al., 2019). While existing studies focus mainly on equity, green bonds are also an important innovative financing tool for addressing environmental and climate challenges (Ehlers and Packer, 2017). In the last decade, green bonds have become increasingly appealing to investors (Krueger et al., 2020). Moreover, since the *European Investment Bank* issued the first green bond in 2007, the green bond market has experienced exponential development. According to the Climate Bond Initiative (2020), the worldwide annual issue volume has grown from less than 40 billion USD in 2014 to over 160 billion USD in 2018 and 257.7 billion USD in 2019 worldwide. This rapid growth which indicates an increasing amount of funds to finance climate change adaptation and mitigation, has also attracted the attention of academics. Existing studies explore whether issuers of such green securities enjoy lower costs of financing, and at the same time, whether investors request lower returns.

In this study, we exploit the green bond market as a laboratory for testing the asset pricing implications of investment vehicles dedicated to actions related to climate change. We measure the asset pricing implications of bond greenness in terms of the so-called 'green bond premium', i.e., the difference between the yields of matched green and conventional bonds. In particular, we systematically examine the existence of the green bond premium and analyze how it is influenced by external valuation for a bonds greenness. We extend the methodological frameworks of earlier related studies (Hachenberg and Schiereck, 2018; Nanayakkara and Colombage, 2019; Zerbib, 2019) by a stricter matching approach, a more precise measurement of the green bond premium, and a larger sample to analyze the green bond premium as the yield difference between green bonds and synthetic conventional bonds. The latter bonds are created by matching a pair of conventional bonds to each green bond and adjusting the maturity by interpolation. Our main finding is that investors reward green bonds that are approved by external reviews, documenting the bond's serious and genuine green purposes, with a premium in the sense of lower yields and higher bond prices.

The existence of such a green bond premium is in contrast to the modern portfolio theory, which makes the assumptions of rational investors, efficient markets, and expected returns as a function of risk. Nevertheless, asset pricing literature has shown that, additionally to these assumptions, several anomalies predict asset prices (Harvey et al., 2016). More specifically, behavioral finance literature generally assumes that investors are imperfect and subject to many emotional biases, i.e., behavioral finance differs from traditional finance in that it focuses on how investors actually behave, rather than theorizing how they should behave. In our particular case, green bonds cater for both the traditional financial and green objectives of bond investors. Therefore, green impact investors may achieve utility from the green investment outcome besides the utility gained from financial performance. Thus, following this utility paradigm, which goes one step beyond the irrationality-based behavioral finance perspective, green bonds could be priced higher than comparable conventional bonds, as the non-financial utility component may compensate for a lower financial return for green impact investors. To identify whether this pattern applies on financial markets and how the level of greenness impacts the green bond premium, is the research gap that needs to be filled. Some evidence on whether such a premium really prevails amongst investors exists (e.g., Baker et al., 2019; Bachelet et al., 2019; Zerbib, 2019), but no consensus has been reached, and the findings so far paint an unclear picture. While there is some evidence supporting a positive green bond premium and the appreciation of the greenness of these instruments (Baker et al., 2019), other studies elicit no green bond premium or even a negative one (Bachelet et al., 2019; Climate Bond Initiative, 2019b).

Our paper develops a theoretical framework for a 'greenness bias' in expectations of green bond investors and conducts empirical tests based on a sample of 250 matched bond triplets in the period from 2011 to 2020 containing more than 90,000 daily observations. To determine the matched bond triplets we applied a rigorous matching process. Moreover, we build a comprehensive dataset by consolidating various sources of information on green bonds and their comparable counterparts. To investigate the pricing mechanism of green bonds, we run hybrid regressions and focus on different types of external review reports and their evaluation results, to explain the variations in the distribution of green bond premiums. Furthermore, we control liquidity difference via a hybrid model, in order to extract the real green bond premium.

Our results show that, on average, green bonds enjoy an expected positive premium (approximately 1 BP) over comparable conventional bonds. Indeed, some green bonds do have an evidently higher premium than others. Reports from independent external reviewers are a main driver for investors to pay a significant green bond premium. However, the type of external review is crucial. While there is no evidence that external reviews such as a certification assigned by the Climate Bond Initiative (CBI) and a green rating from traditional credit rating agencies have a positive influence on the green bond premium, green bonds with a second-party opinion and a verification enjoy significantly lower yields, i.e., are traded at a positive premium (3 to 4 BP). Particularly second-party opinions asserting a 'dark green' or 'medium green' shade tend to be associated with a positive premium (5 or 4 BP, respectively). This finding adds to the discussion of whether investors are willing to pay more for certified green and sustainable investments (see e.g., Gutsche and Ziegler, 2019).

This finding also adds to the debate on whether green bonds could be used for financing regular non-green projects (Flammer, 2020) and as a tool for green-washing (Walker and Wan, 2012; Nyilasy et al., 2014). If this were the case, the green bond market would lose much of its credibility, and investors might start to ignore the green label. Therefore, preserving the integrity and credibility of green bonds is at the core of building a healthy green bond market. To mitigate the risk of greenwashing, the *International Capital Market Association* (ICMA) recommends issuers to appoint an independent external reviewer to confirm the alignment of their green bonds with the 'Green Bond Principles' (GBPs). Consequently, external reviews are the main approach to enhancing the integrity and credibility of the green bond market (Shishlov et al., 2016). With a growing market, the role of independent external reviewers is becoming even more prominent. However, to the best of our knowledge, very few studies present empirical indications that external reviews impact investor decisions and thereby the pricing of green bonds (Baker et al., 2019; Bachelet et al., 2019; Larcker and Watts, 2020). The question of whether different types of external reviews create value for investors has not yet been answered. Moreover, even though an increasing number of external reviewers explicitly evaluate detailed greenness issues instead of a general assessment, it is unanswered how investors react to the external greenness assessments.

Our empirical results provide evidence that investors rely on external reviews, especially second-party opinions and verifications, as a source of proven information on the greenness of green bonds. In particular, investors reward the integrity (expressed by second-party opinions) of green bond issuers with lower expected returns. We document that the effect of external validation on the green bond premium is strongest for bonds that are rated dark-green in second-party opinions, which affirms investors' positive perception of the shade of greenness of the project. This pattern may offset not incurring information costs, as the external validation effect decreases with increasing age of bonds. Thus, a second-party opinion, especially one with a clear evaluation conclusion in terms of a shade of green, can be one channel for investors to reduce information search costs aimed at confirming the greenness of a bond, reduces the uncertainty that the respective bond is reliably and consistently green, and thus motivates investors to buy the bond at higher prices, i.e., lower expected returns. This major finding is in accordance with the recent finding that retail investors, especially socially responsible investors, have significant preferences for socially responsible equity funds with certification and transparency logos (Gutsche and Zwergel, 2020).

With these findings, our study makes the following two contributions. First, our study considers almost all green bonds which provides adequate information for our analysis and covers most of their yield development between 2011 and 2020. In contrast to earlier studies that focus on a specific time frame or a relatively small sample (e.g., Ehlers and Packer, 2017; Nanayakkara and Colombage, 2019), our setting for analyzing the green bond premium is comprehensive and minimizes potential bias that could influence the statistical estimations. Furthermore, our stringent matching process ensures that the observed yield premium between green and corresponding conventional bonds can be regarded as the 'real' green bond premium. Second, this is the first study to systematically examine the impact of all four different kinds of external review reports on the pricing of green bonds.[1] To this end, we collect all available external review reports from major green bond databases or official issuer websites and classify them into different categories based on their formats and evaluation results. This dataset enables us to determine that serious climate action confirmed by 'dark green' and 'medium green' second-party opinions have a significant impact on the green bond premium. Nevertheless, the value of confirming climate-protection-related issues for investors declines with the age of the green bond. In terms of practical and policy implications, the reliability of external reviews in the green bond market is important to investors, and thus has implications for the cost of capital for financing climate-change adaption and mitigation.

The remainder of the paper is organized as follows. In Section 3.2, we discuss the importance of green credentials and the role of external review reports in the green bond market. We review the literature on the green bond premium and develop several hypotheses in Section 3.3. Section 3.4 presents our sample and Section 3.5 the methodological approach. Section 3.6 contains the empirical results and Section 3.7 concludes.

3.2 The green bond market: Institutional details

3.2.1 Green bond labels

The development of the green bond market in the past decade demonstrates the huge demand for climate adaptation and mitigation investments. Indeed, studies show that

[1] Several studies also touch upon this question (Baker et al., 2019; Bachelet et al., 2019; Larcker and Watts, 2020), but only to a minimal extent. For instance, Bachelet et al. (2019) adopt subsample analysis to examine the role of 'third-party verification', while Baker et al. (2019) include only a dummy variable to investigate the general influence of the CBI certification. None of these studies treats different types of external reviews separately, or examine the evaluated greenness (shade of green).

both institutional and retail investors with a focus on sustainable investment have a strong interest in investing in green bonds (Climate Bond Initiative, 2019a,b). Also, from an issuer perspective, green bonds can provide an ideal financing source for green projects. Besides fulfilling their commitment to the environment, green bond issuers may enjoy lower costs of capital in the primary market (Ehlers and Packer, 2017).

Originally, the proceeds from green bonds were intended to be used for green projects such as renewable energy or energy efficiency projects. As more and more issuers from various sectors entered the market, the concern arose that green bonds could be misused to finance greenwashing projects (Flammer, 2020). Shishlov et al. (2016) point out that one of the two major challenges for the green bond market is to ensure its environmental integrity so as to mitigate the green-washing criticism that could threaten its survival. Investors are also aware of the greenwashing risk. According to an investor survey conducted by CBI (Climate Bond Initiative, 2019a), green credentials and issuer transparency are the most important factors for green bond investors making investment decisions.

However, the green bond market is generally not subject to government regulation and there are only a few voluntary rules to prevent the possibility of greenwashing. Currently, the voluntary process guidelines proposed by ICMA, called 'Green Bond Principles' (GBPs), are regarded as the most widely accepted standards to promote the integrity of the green bond market. To ensure that green bonds make the expected contribution to the environment, issuers can disclose an overall green bond framework which has four core components, comprising 1) the use of proceeds, 2) process for project evaluation and selection, 3) managing of the proceeds, and 4) reporting, as defined by the GBPs. Yet, the fact that issuers can label their bonds as green and draft a green bond framework on their own, results in a need to seek independent and professional external reviewers to examine the alignment with the GBPs and the greenness of the bonds. Therefore, the GBPs encourage green bond issuers to seek external reviews, besides releasing statements on the four core components.

3.2.2 Different external green bond reviewers

According to the GBPs, there are generally four types of external review report, namely second-party opinion, verification, certification, and green rating. Each green bond can have just one or several types of external review. External reviewers are usually independent research institutions dedicated to environmental research such as the *Center for International Climate Research* (CICERO) and *ISS-Oekom*. They examine the alignment of green bonds with the GBPs, or evaluate greenness based on their specific criteria and methodologies. These external reviewers are intended to facilitate communication between investors and issuers, and thus contribute to a healthy and prosperous green bond market.

Second-party opinions (SPOs) are the most popular external reviews for green bonds. Each green bond can have an SPO issued by an independent research institution such as CICERO, *ISS-Oekom*, and *Sustainalytics*. SPOs are usually detailed and comprehensive, providing a thorough analysis of the four core components of the GBPs and other related issues. An SPO released by CICERO mainly contains a description and an assessment of the issuer's green bond framework, rules, and procedures for climate-related activities. The assessment part of the report comprises strengths, weaknesses, and pitfalls of the green bond framework. Moreover, some SPOs even provide a broad qualitative indication of the true greenness of green bonds. For instance, CICERO's SPOs are graded into several shades, namely 'dark green', 'medium green', 'light green', and 'brown', indicating the possible environmental impact of the green bond and the robustness of the issuer's governance structure that supports the framework. According to CICERO's criteria[2], 'dark green' is only awarded to green projects and solutions that represent the best way to realize the long-term vision of a climate-resilient future. For instance, the 2015 green bond framework of the German state-owned development bank KfW obtained such a 'dark green' shade from CICERO, because of its clear and exclusive focus on renewable energy and robust procedures for project screening. However, in 2019 the KfW green bond framework received the 'medium green' shade from CICERO. Even though the proceeds are allocated to provide favorable loans for renewable projects and the construction of energy efficiency buildings to push forward the usage of fossil-free sources, the 2019 green bond framework cannot fully guarantee the exclusion of fossil fuels (see CICERO, 2019) and thus regarded as somewhat less green as the 2015 one[3]. Moreover, 'light green'-shaded bonds finance mere quick-fix solutions that help initiate the transition towards the long-term vision, such as improvement of energy efficiency in fossil-based activities. The so-called 'brown' shade (which does not occur for any bond in our sample) indicates a bond's negative ecological impact. Besides CICERO, other SPO providers have a similar evaluation methodology in their SPO reports (see Table 3.1). For comparison purposes among different greenness evaluations, we convert the different schemes of greenness evaluation into one scale represented by the shades of 'dark green', 'medium green', 'brown', and 'no shade'. A green bond is awarded with the shade of dark green if it exhibits an above-average positive evaluation, while the shade of medium green indicates a level of greenness that SPO provider considers to be standard in the green bond market. Bonds

[2]The CICERO's shade of methodology: https://www.cicero.green
[3]See also CICERO (2015). We compare those two SPOs and come to this conclusion.

classified as brown shade show below-average or negative evaluation results. When no specific shade of green is explicitly expressed in an SPO, it is classified as having no shade in this study. In that case, green bond investors must be able to draw on their own overall judgment, based on positive and negative signals implicitly delivered by SPO providers.

Table 3.1: Different shade of green schemes

Shade	CICERO	Vigeo	ISS-Oekom	Sustainlytics
dark green	dark green	reasonable	excellent good	leader outperformer
medium green	medium green light green	moderate	medium	average performer
brown	brown	weak	poor	underperformer laggard
no shade	no clear shade	no clear shade	no clear shade	no clear shade

Verification reports are, compared with SPOs, generally less lengthy and detailed[4], and issued by auditing companies such as *KPMG* and *PwC*. In verification reports, reviewers accomplish predefined tasks such as examining whether the use of proceeds is aligned with the GBPs or other related national regulatory rules. Finally, they provide a statement on the question of whether the issuer has violated any requirements defined by the GBPs or by the issuer (on voluntary basis). Therefore, it can be stated that verification reviewers evaluate green bonds more objectively, while SPO reviewers deliver subjective and comprehensive opinions on green bonds, according to their own standards.

CBI certification is another type of external reviews. CBI as a well-known international organization dedicated to the development of the green bond market, offers a certification scheme which is based on scientific criteria ensuring consistency with the 2 degree Celsius warming of the Paris Agreement. CBI can award green bonds a certification through the approved verifiers.[5] When assigned a CBI certification, a green bond obtains the recognition of CBI regarding its greenness.

Green rating reports are issued by traditional credit rating agencies such as Moody's and S&P. For instance, Moody's assigns five grades of green ratings to green bonds, ranging from 'excellent' to 'poor'. At first glance, green ratings are similar to SPOs with a shade of green, since they both provide a greenness assessment. However, green rating reports from credit rating agencies are regarded as a different type of external review, as they are more quantitative and focus on issuers' environmental performance data. Moreover, they are far less frequent than SPO reports in the green bond market.

[4]A verification report normally has only 2-3 pages.

[5]A complete list of approved verifiers can be seen on the official website of CBI.

3.3 Literature review and hypothesis development

3.3.1 The green bond premium and its determinants

Several studies analyze the green bond premium and its determinants. Regarding the question of whether green bonds enjoy a significant premium, earlier studies show mixed empirical evidence. While some studies find evidence that green bonds enjoy a positive premium (e.g., Baker et al., 2019: 6 BP; Nanayakkara and Colombage, 2019: 63 BP; Zerbib, 2019: 2 BP), other studies cannot confirm its existence (Climate Bond Initiative, 2019b; Larcker and Watts, 2020; Flammer, 2020). Bachelet et al. (2019) even find that green bonds are slightly underpriced and thus have a negative premium (−2 BP). Differences in the identification strategy, sample selection, and observation period potentially cause diversity in the results (see the overview on the methodological spectrum of studies in Table 1 of Zerbib, 2019). Concerning the identification strategy, a comparison of the yield from green and conventional bonds could be conducted in the primary market (Ehlers and Packer, 2017; Climate Bond Initiative, 2019b) or on the secondary market by indirectly examining the impact of the green label by regressing the bond yield on a green label indicator (e.g., Baker et al., 2019; Nanayakkara and Colombage, 2019). Some recent studies extract the green bond premium by adopting a matching approach (Hachenberg and Schiereck, 2018; Bachelet et al., 2019; Zerbib, 2019), which enables researchers more precisely to estimate the premium.

Besides the inconsistent findings on the existence of a green bond premium, a few approaches analyze possible green bond determinants. Hachenberg and Schiereck (2018) and Zerbib (2019) show that basic bond features such as the credit rating and issuer type influence the green bond premium. Also, liquidity is confirmed as a major determinant of yield spreads of green bonds (Wulandari et al., 2018; Zerbib, 2019). Moreover, some preliminary findings show that green credentials are important for the cost of green bonds (Baker et al., 2019; Bachelet et al., 2019; Li et al., 2019). In particular, Baker et al. (2019) investigate the pricing of 2,083 U.S. municipal and 19 corporate green bonds and find that green bonds with a CBI certification have yields 26 BP lower than ordinary bonds with similar characteristic. Bachelet et al. (2019) focus on 89 matched green bonds and find that those green bonds issued by private firms with external reviews show a small premium (1 BP). Kapraun and Scheins (2019) analyze 641 green bonds and observe that certified green bonds have yields 2 BP lower than green bonds without a certification and green bonds traded on green exchanges show lower yields (7 BP) because they are required to meet some standards set by green exchanges. In contrast, Larcker and Watts

(2020) examine a matched sample of 640 municipal green bonds and find that the CBI certification make no significant difference in the pricing of municipal green bonds.

Nevertheless, these earlier studies have several drawbacks. For instance, some of them do not apply a strict matching process (see e.g., Baker et al., 2019; Kapraun and Scheins, 2019) to gain more observations and thus may be subjected to estimation biases. Some of them focus only on a sub-sector of the green bond market such as the U.S. municipal bonds (see e.g., Baker et al., 2019; Larcker and Watts, 2020). Most importantly, neither of these studies investigates the impact of the four different categories of external reviews that we have separately discussed above (earlier related studies mostly focus on the CBI certification, see e.g., Baker et al., 2019; Larcker and Watts, 2020). Moreover, these studies ignore the impact of specific evaluations of greenness levels in external reviews, i.e., the shade of green methodology in SPOs, on the green bond premium, despite their existence and increasing popularity in recent years.

3.3.2 Hypotheses development

A green bond premium, defined as the difference between the yield of a green bond and a comparable conventional bond may be due to the price impact of investor preferences regarding the climate-change exposure of assets (Painter, 2020). For instance, Gutsche et al. (2019) find that there is a significant positive correlation between socially responsible investments and the dummy variable for retail investors' environmental values (i.e., whether a respondent is a member of an environmental organization) based on a representative survey. As Fama and French (2007) put it, the demand for green assets is an investors' taste that adjusts equilibrium prices.

Our theoretical framework relies on the investors' taste argument, i.e., investors appreciate non-financial aspects of an investment.[6] In this regard, some investors are willing to sacrifice a certain proportion of the return in order to achieve a non-financial utility from the investment (Dorfleitner and Utz, 2014; Riedl and Smeets, 2017).[7] Moreover, Höchstädter and Scheck (2015) reveal that impact investors aiming at environmental (or social) impact accept some curtailment of the achievable financial return.

[6] Note that there is mixed evidence on whether sustainable and especially environmental-friendly (stock) investments yield a financial under- or out-performance (Orlitzky et al., 2003). However, this question is not relevant in our context, as the green bond premium corresponds by definition to a lower return, compared to a comparable conventional investment.

[7] One might see this phenomenon as the flip side of the well-known sin-stock effect, according to which stocks of especially unethical firms yield a higher return than otherwise comparable stocks because of investor preferences (Hong and Kacperczyk, 2009).

We measure the bond performance through the yield r_B of a bond B. In a setting with combined financial and non-financial investor preferences (see Dorfleitner and Utz, 2012), investors assess a bond B with the functional

$$r_B + \alpha g_B, \tag{3.1}$$

where the parameter α represents preferences for the (expected) greenness g_B of the bond with a yield r_B. Accordingly, we derive an equation describing an investor's preferences regarding a green bond (GB) and a comparable conventional bond (CB). If a green bond has a yield r_{GB} and an expected greenness g_{GB}, while the conventional bond has a yield r_{CB} and no greenness, then an α^* exists such that at a given point in time

$$r_{GB} + \alpha^* g_{GB} = r_{CB} \,. \tag{3.2}$$

Different investors may – depending on their non-financial preferences – yield different values for $\alpha \geq 0$ (Berry and Yeung, 2013). If specific investors have a higher appreciation of greenness than the market-related α^*, they will have a preference for the green bond. Therefore, they are willing to buy the bond at the current yield level. Aggregated over all investors, this effect yields a positive value of α^* and thus explains a positive green bond premium

$$r_{CB} - r_{GB} = \alpha^* g_{GB} \tag{3.3}$$

if there is a sufficiently large share of investors with positive values for α. An α^* different from zero in this framework indicates the existence of a price impact of bond greenness.

Based on these theoretical considerations, we deduce the following hypotheses on the link between investor preferences and the green bond premium. Since the greenness of a bond is a natural component of investor decision-making, investors have a need to objectify this type of non-financial information. On a company level, ESG issues are disclosed in the non-financial reporting, which signals a company's commitment to increasing transparency, and causes a reduction of information asymmetry (Dhaliwal et al., 2012). Existing evidence suggests a negative relationship between environmental performance and cost of capital (Heinkel et al., 2001; Ghoul et al., 2011; Chava, 2014).

On the green bond market, a majority of bond issuers report on the use of proceeds, typically by releasing their green bond framework or social impact reports based on the GBPs. Non-financial disclosure reduces information asymmetry regarding the implemented sustainability practices (Hahn and Kühnen, 2013). Given that non-financial disclosure on environmental activities increases the green bond transparency, thus reducing uncertainty and idiosyncratic risk, investors may accept lower risk compensation, leading to a reduction in the cost of debt of green bonds. Thus, if the label 'green bond' substantially increases the transparency of green bonds for a sufficiently large proportion of green investors who imply a positive greenness g_{GB} from the voluntary disclosure of the green bond issuer, then a positive α^* and thus also a positive premium $r_{CB} - r_{GB}$ can emerge.

H1: Green bonds are priced at a premium in the secondary market, compared to conventional bonds with similar characteristics.

We continue with disentangling the 'substantial increase' in green bond transparency from the 'sufficiently large' proportion of green investors by focusing on different levels of non-financial disclosure on green bonds. While we are capable of analyzing different levels of non-financial green bond disclosure as a measure of transparency, the 'sufficiently large' proportion of green investors is an implicit measure in our approach.

A majority of green bond issuers release non-financial disclosures to increase green bond transparency, and thereby reduce information asymmetry. Nevertheless, some information asymmetry remains regarding its validity. The validity of the released information is crucial in the following considerations, since it is difficult to obtain credible factual information on the use of green bond proceeds. Voluntary non-financial disclosure could even be misused for greenwashing purposes. It is well-known that credit ratings from external rating providers can overcome information asymmetry issues to some extent (Tang, 2009). Analogously, intermediaries such as external reviewers (e.g., SPO issuers and verification providers) and certification bodies can play an important role in mitigating information asymmetries regarding the non-financial aspects. From an investor perspective, external reviews thus can make green bond investments more reliable and instil more confidence (Climate Bond Initiative, 2019a).

Reverting to our model, we consider the investors' possibility to obtain a stronger differentiation between bonds. To keep things simple, let us assume that the greenness can be either $G > 0$ or zero. Due to information asymmetry, investors calculate the expected greenness according to

$$g_{GB} = p \cdot G + (1 - p) \cdot 0 = G \cdot p \tag{3.4}$$

with p denoting the probability of the greenness being G and reflecting their uncertainty in objectifying the real facts. Consider two bonds for clarification, the first (GB1) having an external review regarding its greenness, while the second (GB2) has no such confirmation. It is plausible that such an external review would in most cases increase the probability of p, namely in those cases in which the review comes to a positive conclusion and investors trust the assessment of the external reviewer more than their own assessment of the voluntary disclosure.[8] The expected greenness of the first bond is then higher ($g_{GB1} > g_{GB2}$). Thus, for an investor with a fixed value of α, a higher accepted premium αg_{GB} is implied.

This argumentation is in line with informational transaction cost theory. Without a review, investors who are potentially interested in a specific bond have to invest some information costs to verify whether or not the greenness is really there. If an external review is available, these research-related costs can be avoided. Such information costs are generally priced in bond markets (Fenn, 2000). In our case, investors will subtract the necessary transaction costs from the decision functional $r_{GB} + \alpha g_{GB}$ when considering their effective net value, and thus demand a higher yield for GB2 than for GB1. This directly implies a lower green bond premium for GB2. Either way of argumentation leads to the second hypothesis.

H2: Green bonds with a statement from an external reviewer on their true greenness enjoy a higher green bond premium in the secondary market.

The informational transaction cost argument implies another hypothesis. If a certain green bond without an external review has been traded on the market for some time, more and more potentially interested investors may already have spent their research-related transaction costs, which are inherently fixed costs. Therefore, following this reasoning, one can expect to observe an increasing green bond premium for those green bonds that do not have an external review, as these costs only accrue at each investor's first dealing with the specific bond. We subsume this consideration into another hypothesis.

H3: The premium of green bonds without a statement from an external reviewer on their true greenness is positively related to the duration for which the green bond has been traded on the secondary market.

Finally, we restrict our considerations to green bonds with external reviews, i.e., bonds having the same (high) value of p. However, different levels of greenness exist (see Table 3.1). Therefore, we substitute G in Equation (3.4) with one of the values G_1, G_2, and G_3 (with $G_1 > G_2 > G_3$). If one green bond (GB1) has a darker shade of green, say

[8] Indeed, in our sample employed in the empirical part, there are no bonds with external reviews claiming that the corresponding issuer is prone to greenwashing. However, theoretically, this is possible and would lead to a lower probability p.

G_1, than another (GB2), say G_2, then $g_{GB1} > g_{GB2}$ for equal values of p. Accordingly, a specific investor with a fixed α would, therefore, be willing to accept a higher premium for GB1 than for GB2. This discussion leads us directly to the fourth hypothesis.

H4: Green bonds with a higher level of greenness confirmed by external parties enjoy a higher premium in the secondary market.

3.4 Data description

3.4.1 The green bond dataset

Our main green bond database is *Environmental Finance* (EF), which lists self-labeled green bonds and contains information on bond issuance and related documents such as external review reports. We extract, from the EF database, a complete list of straight green bonds[9] issued since the inception of the green bond market in 2007 until April 2020. Moreover, we supplement the EF green bond dataset with those straight bonds marked as green bonds on Thomson Reuters Eikon. In particular, we collect any external review reports and ICMA green bond templates[10] from the EF database. However, even though EF provides a comprehensive record of documents regarding external review reports, some data on external review is still missing. Therefore, we also download green bond datasets maintained by ICMA and CBI, both of which contain valuable information on external review reports.[11] Moreover, since none of the existing data sources provides a complete record of all types of external review reports, we manually check each issuer's official website to further validate or supplement the existing information on external reviews. Furthermore, we augment the green bond database with basic bond features such as structure, seniority, and credit rating, from Thomson Reuters Eikon. In the end, we build a dataset of 1,248 straight green bonds with adequate data for further analysis.

[9] We do not consider green bonds with embedded options, since different types of options have a different impact on bond pricing, and thus disable the comparison that is necessary in the matching process.

[10] To promote the transparency of the green bond market, ICMA designs a template on which issuers can publish information about their issuance of green bonds and the corresponding external reviews. Some issuers may voluntarily upload the template on ICMA's official website.

[11] The green bond dataset from ICMA: https://www.icmagroup.org/green-social-and-sustainability-bonds/green-social-and-sustainability-bonds-database. The green bond dataset of CBI: https://www.climatebonds.net/cbi/pub/data/bonds

3.4.2 Matching conventional and green bonds

To isolate the impact of the green label on the bond yield, i.e., the green bond premium, the ideal setting would comprise one bond that exists in both treatments, i.e., as a green bond and a conventional bond at the same time. Since this situation could not be observed from market data, we match treated (i.e., green) bonds to otherwise comparable conventional bonds (see e.g., Bachelet et al., 2019; Zerbib, 2019). Therefore, we identify, for each green bond, a list of conventional bonds with similar bond characteristics. Conventional counterfactual bonds resemble green bonds in all matching criteria, and therefore, we expect them to develop similarly to green bonds.

In general, a perfect match between green bonds and conventional bonds is unlikely, since only few parties issue such a bond pair. Therefore, we capture the remaining differences by further controls. Although a perfect matching is impossible, a rigorous matching-pair approach can derive a more reliable estimation of the green bond premium and would strip out any significant differences between green bonds and conventional bonds, other than that of the green label itself (Zerbib, 2019).

Our matching approach proceeds as follows (see Figure 3.2 in the appendix). We make use of the Eikon security screener, and extract, for each green bond, all straight conventional bonds of the respective issuer. For instance, we select the matching partner of a green bond issued by *Berlin Hyp AG* from a pool of more than 1,000 conventional bonds also issued by *Berlin Hyp AG*. We consider active and inactive, i.e., expired, plain vanilla bonds. Further, we apply the matching criteria of Zerbib (2019) for comparison reasons. For potential matching partners, we require conventional bonds to have the same currency denomination, coupon type, seniority and collateral status, and credit rating[12] as the green bond. Moreover, we select the issue amount of conventional bond candidates to be less than 4 times and higher than 1/4 of the issue amount of the green bond, so as to account for volume difference. Additionally, we exclude those conventional bonds with an issue date six years earlier or later than that of the green bond.[13] By considering every possible conventional bond in the matching approach, our matching can identify the globally optimal matching result.

One major issue in the matching of green and conventional bonds is the difference that may result from different maturities of the matching partners. Since not many issuers issue green bonds and similar conventional bonds with the same maturity date, we first

[12] Note that different credit rating regimes have been integrated into the same scale as that of S&P on Eikon.

[13] All the above mentioned matching criteria are exactly the same as those in Zerbib (2019).

choose conventional bonds with a close maturity date, i.e., a maturity date that differs less than two years from that of the green bond (see Zerbib, 2019). Moreover, we split the set of eligible conventional bond candidates ($C = I \cup J$, with $I \cap J = \emptyset$) into two groups: one including conventional bonds with an earlier maturity date (group cb_i with $i \in I$) and the other of conventional bonds with a maturity date later than that of the green bond (group cb_j with $j \in J$). A potential bond triplet consists of the considered green bond (gb), and one bond from each of the two groups (cb_i, cb_j). Therefore, the number of potential bond triplets for a green bond is the product of the numbers of conventional bonds in these two groups ($|I| \times |J|$).

In the next step, we consider the following two aspects in order to identify the best matched triplet from all potential bond triplets. First, we take the quality and availability of yield data into account. To this end, we download daily bid and ask yield data for bonds in each potential bond triplet from Bloomberg from respective issue dates to June 2020. We drop observations with the same bid and ask yield as the previous or next trading day, or with a bid-ask spread larger than 50 BP, because constant yields or large spreads indicate bond illiquidity and the bid and ask yield data may not reflect the market conditions in these cases. Thereafter, we merge the yield data of three bonds, based on the timestamp, and exclude bond triplets that provide less than 50 joint daily observations.

Second, we account for the remaining difference in maturity by applying the following correction for each potential bond triplet (gb, cb_i, cb_j). We construct all possible synthetic bonds (cb_{ij}) by linear interpolation. Each synthetic bond has exactly the same maturity as the green bond, and we choose this respective synthetic bond as the final counterfactual of the green bond. Unlike some studies allowing both interpolation and extrapolation (Bachelet et al., 2019; Zerbib, 2019), for consistency reasons, we implement only interpolation.[14] Accordingly, the yield of the synthetic conventional bond can be calculated through

$$r_{cb_{ij}} = r_{cb_i} + \frac{r_{cb_j} - r_{cb_i}}{D_{cb_j} - D_{cb_i}} \cdot (D_{gb} - D_{cb_i}) \tag{3.5}$$

where $D_{cb_i} < D_{gb} < D_{cb_j}$ and D represents maturity. Thus, the yield difference between the green bond and its comparable synthetic bond can be calculated.

$$\Delta r = r_{cb_{ij}} - r_{gb} \tag{3.6}$$

We remove all observations with an absolute yield difference $|\Delta r|$ larger than 100 BP as a

[14] Linear interpolation and extrapolation tend to have different impacts on the yield estimation and additional noises might be introduced if both of them are allowed. This is actually a stricter requirement, as it makes the search for comparable conventional bonds more difficult.

signal for data irregularities.[15] To minimize the error resulting from linear interpolation, we solve the problem

$$\min_{i,j} |D_{cb_i} - D_{gb}| + |D_{cb_j} - D_{gb}| \tag{3.7}$$
$$\text{s.t. } i \in I \text{ and } j \in J$$

to determine the triplet with the smallest sum of absolute maturity differences as the final matched triplet (gb, cb_{i^*}, cb_{j^*}).

3.4.3 Liquidity adjustment

One important determinant of the bond pricing is liquidity (Amihud and Mendelson, 1986; Chen et al., 2007). Therefore, we apply the following approach to capture a possible liquidity difference in the yield difference Δr for each potential bond triplet and to provide an accurate estimation of the green bond premium. We choose the daily bid-ask spread as the measure of liquidity in bond markets (see e.g., Schestag et al., 2016). For a single bond, we calculate the bid-ask spread L as the difference between the bid and the ask yield:

$$L = r_{bid} - r_{ask} \, . \tag{3.8}$$

For the synthetic bonds, we interpolate the liquidity measure based on the liquidity of the two comparable conventional bonds:

$$L_{cb} = L_{cb_{i^*}} + \frac{L_{cb_{j^*}} - L_{cb_{i^*}}}{D_{cb_{j^*}} - D_{cb_{i^*}}} \cdot (D_{gb} - D_{cb_{i^*}}) \, . \tag{3.9}$$

Thereafter, the corresponding liquidity difference ΔL between green bonds and synthetic conventional bonds is

$$\Delta L = L_{cb} - L_{gb} \, . \tag{3.10}$$

We use the liquidity difference ΔL to capture the influence of distinct liquidity on the yield difference between green and conventional bonds in the following.

[15] This data cleaning procedure leads to a reduction of only 112 daily observations. We also remove this procedure or change the 100 BP yield difference requirement to 150 BP to see whether it may lead to biases. These additional checks show similar empirical results as the main results reported in this paper.

3.4.4 Sample and descriptive statistics

After the matching process, we identify 250 best matched bond triplets (250 green bonds matched with 500 conventional bonds).[16] We document the reduction in sample size from 1,248 to 250 during the whole matching process when adding matching criteria step by step in Table 3.11.

In total, our sample comprises 92,774 daily observations for the period from 2011 to 2020 and for various variables defined in Table 3.12. On average, the yield and liquidity difference between green bonds and comparable conventional bonds, are both close to zero (see Table 3.2). The maturity of green bonds has an average value of 4.20 years and ranges from less than one month to more than 28 years. Green bonds have a maximum yield of 23% and a minimum of −0.97%, with a mean of 1.62%. The average issue volume of green bonds is 0.43 billion USD, which is lower than that of comparable conventional bonds (0.67 billion USD).

Table 3.2: Descriptive statistics for metric variables

Variable	Obs.	Mean	Std.	Min	Median	Max
Panel: time-variant						
Δr (%)	92,774	−0.0012	0.1208	−0.9969	0.0012	0.9834
ΔL (%)	92,774	−0.0019	0.0460	−0.4477	0.0001	0.3898
gb_yield (%)	92,774	1.6214	2.1388	−0.9720	0.9430	23.0020
maturity (in years)	92,774	4.2459	3.0428	0.0548	3.5863	28.8110
Panel: time-invariant						
maturity[1] (in years)	250	6.2407	3.2921	1.9973	5.0055	30.0192
gb_volume (bn USD)	250	0.4267	0.4243	0.0018	0.3727	3.3456
cb_volume (bn USD)	250	0.6695	0.8690	0.0015	0.3235	5.5760

This table reports summary statistics on time-variant and time-invariant green bond characteristics. The entire data sample contains 92,774 daily observations from 250 bond triplets (250 green bonds matched with 500 conventional bonds). The variables are defined in Table 3.12.
[1] Maturity of the green bond at issuance.

Around half of the green bonds in the final sample are denominated in USD or EUR, while those denominated in currencies such as HKD, MXN, and SGD have a share of less than 1% (see Table 3.3). Regarding issuer type, the largest share (26.80%) of green bonds are from supranational institutions such as the *World Bank* and the *International Finance Corporation*, and financial institutions such as banks. Furthermore, green bonds with an AAA credit rating comprise almost a third of the sample while those with a credit rating lower than A+ have a share of 10%.

Besides basic bond features, we observe information related to external review reports. SPOs are the most popular type of external reviews. 196 out of 250 green bonds are

[16] Our green bonds are at least representative for plain vanilla green bonds for which a GBP can be identified. Tables 3.14 and 3.13 contain the respective summary statistics on the sample of all 1,248 plain vanilla green bonds.

Table 3.3: Descriptive statistics for categorical variables

Variable	Obs.	Relative	Variable	Obs.	Relative
SPO			CNY	10	4.00
Yes	196	78.40	EUR	72	28.80
No	54	21.60	GBP	3	1.20
			HKD	1	0.40
shade			INR	4	1.60
dark green	49	19.60	JPY	5	2.00
medium green	52	20.80	MXN	1	0.40
no shade	95	38.00	NOK	6	2.40
no SPO	54	21.60	SEK	48	19.20
			SGD	1	0.40
verification			TRY	3	1.20
Yes	51	20.40	USD	52	20.80
No	199	79.60	ZAR	6	2.40
CBI_certification			*issuer_type*		
Yes	17	6.80	agency	46	18.40
No	233	93.20	corporate	47	18.80
			financial	67	26.80
green_rating			municipal	21	8.40
Yes	10	4.00	sovereign	2	0.80
No	240	96.00	supranational	67	26.80
seniority[1]			*credit_rating*		
MTG	8	3.20	AAA	79	31.60
SEC	2	0.80	AA+	11	4.40
SR	203	81.20	AA	13	5.20
SRBN	4	1.60	AA−	12	4.80
SRP	15	6.00	A+	15	6.00
SRSEC	4	1.60	A	5	2.00
UN	14	5.60	A−	4	1.60
			BBB+	8	3.20
currency			BBB	4	1.60
AUD	23	9.20	BBB−	4	1.60
CAD	7	2.80	NR[2]	95	38.00
CHF	8	3.20			

This table contains summary statistics on the green bond sample of this study. The entire data sample contains 92,774 daily observations from 250 bond triplets (250 green bonds matched with 500 conventional bonds). The variables are defined in Table 3.12.

[1] *Seniority* indicates the combined information on bond seniority and collateral status on Eikon. MTG: senior secured and mortgage backed; SEC: secured; SR: senior unsecured; SRBN: senior non-preferred; SRP: senior preferred; SRSEC: senior secured; UN: unsecured.

[2] NR means that the green bond does not have a S&P equivalent crediting rating on Eikon.

assigned to an SPO. Among the green bonds with an SPO, 49 are categorized as dark green, and 52 as medium green. However, the other 95 have no specific shade of green, despite the existence of an SPO. Moreover, no green bond is classified as brown by SPO providers in our sample.[17] Verification reports and the CBI certification appear to be less popular than SPOs in the green bond market. In our final sample, 20.40% green bonds have a verification report and only 6.80% have a certification from CBI. Only ten green bonds have a green rating from traditional credit rating agencies. Eight green bonds reveal an 'excellent' green rating from Moody and two green bonds a 'Green 1' green rating from *Japan Credit Rating Agency* (JCR).

3.5 Empirical methodology

3.5.1 Estimating the green bond premium

To eliminate the impact of the liquidity difference on the green bond premium, we regress the yield difference on the liquidity difference in a hybrid model (see e.g., Mundlak, 1978; Bell and Jones, 2015):

$$\Delta r_{it} = \beta_0 + \beta_1(\Delta L_{it} - \overline{\Delta L_i}) + \beta_2\overline{\Delta L_i} + (u_i + e_{it}) \tag{3.11}$$

where $\overline{\Delta L_i}$ is the mean of the liquidity difference within a specific bond i, u_i represents the individual error term, and e_{it} is the overall error term. In the hybrid model, the variable ΔL_{it} is decomposed into a within-effects component $\Delta L_{it} - \overline{\Delta L_i}$ and a between-effects component $\overline{\Delta L_i}$. The estimate of the within-effects β_1 is unbiased, regardless whether u_i is correlated with ΔL_{it} (Schunck, 2013; Bell and Jones, 2015). Moreover, it is also possible to estimate the between-effects β_2 in the hybrid model. Given that the bonds in this study are collected from various countries and traded on various platforms, there could be between-effects in the bond pricing dynamics.

We further subtract the influence of the liquidity difference from the yield difference and estimate the green bond premium as follows:

$$\hat{p_{it}} = \Delta r_{it} - \hat{\beta}_1(\Delta L_{it} - \overline{\Delta L_i}) - \hat{\beta}_2\overline{\Delta L_i} \tag{3.12}$$

[17]This does not mean that our sample is not representative. SPO providers seldom release a negative shade. For instance, CICERO's SPOs are graded as dark green or medium green in most instances, if there is a clear evaluation result.

where $\hat{\beta}_1$ and $\hat{\beta}_2$ are estimated coefficients from the hybrid model in Equation (3.11). In this way, the estimated green bond premium $\hat{p_{it}}$ varies across different bonds and over time.

3.5.2 Determinants of the green bond premium

We investigate the determinants of the green bond premium in another hybrid regression model. Besides its advantages mentioned in the previous subsection, the hybrid model enables incorporating time-invariant variables (Bell and Jones, 2015). Since some time-invariant variables related to external reviews such as *SPO* and *shade* are of particular interest and important for testing our hypotheses, we adopt the hybrid model to investigate the determinants. We run the model with the time-variant green bond premium $\hat{p_{it}}$ extracted from the initial hybrid regression in Equation (3.11) as the dependent variable:

$$\hat{p_{it}} = \gamma_0 + \gamma_1(TV_{it} - \overline{TV_i}) + \gamma_2\overline{TV_i} + \gamma_3 TI_i + (u_i + e_{it})\,. \tag{3.13}$$

TV_{it} represents time-variant control variables, i.e., *maturity* and *gb_yield*. Accordingly, each time-variant variable is transformed into two variables (one in the within-effects vector $TV_{it} - \overline{TV_i}$ and the other in the between-effects vector $\overline{TV_i}$) in the hybrid regression. TI_i comprises time-invariant variables of interest, namely dummy or categorical variables regarding the existence of a specific type of external review or related greenness evaluation results. Moreover, TI_i includes other time-invariant control variables related to basic bond features such as *currency*, *issuer_type*, and *credit_rating* that have been extensively investigated in earlier studies (see e.g., Zerbib, 2019).

3.6 Results

3.6.1 The green bond premium

This section tests our first hypothesis of whether investors trade green bonds at a premium in the secondary market in general. Accordingly, we apply the hybrid model in Equation (3.11). Thereby, we estimate the green bond premium for each green bond on each trading day following Equation (3.12). The variation of liquidity difference at the bond level explains part of the variation of yield difference as the coefficient of $\Delta L_{it} - \overline{\Delta L_i}$ which is significant at the 1% level (see Table 3.4). Therefore, it is important to control for the liquidity difference when estimating the green bond premium. The significance of the coefficient of the second term $\overline{\Delta L_i}$ at the 5% level shows the existence of between-effects among different bonds. Moreover, the constant term (0.94 BP) in Table (3.4) is significant at the 5% level. This constant term is the estimate for the expected value of the overall green bond premium. Considering Equations (3.11) and (3.12), this constant term is the estimate for the expected value of β_0 of Equation 3.11. Thus, the expected overall green bond premium in our model is the average over the premiums of each green bond. Based on the significances presented in Table (3.4), we find statistical evidence that supports H1 stating that investors trade green bonds, on average, at a premium over comparable conventional bonds.

Table 3.4: Hybrid model to extract the green bond premium

	Coef.	Robust S.E.
$\Delta L_{it} - \overline{\Delta L_i}$	0.2882***	0.0953
$\overline{\Delta L_i}$	0.9210**	0.3754
_cons	0.0094**	0.0039
N	92,774	
Wald chi^2	43.7700	
Prob > chi^2	0.0000	
Rho	0.4968	

This table contains the results of the hybrid model explaining the difference in the yields of green and matched conventional bonds by the variation of liquidity. $\Delta L_{it} - \overline{\Delta L_i}$ measures the within-variability in liquidity, i.e., at the bond level. $\overline{\Delta L_i}$ represents the between-variability to capture cross-sectional effects. *_cons* represents the estimate for the average overall green bond premium in our sample. The full sample includes 92,774 daily observations for 250 bond triplets. Standard errors are cluster-robust at the issuer level. $^*p < .1$, $^{**}p < .05$, $^{***}\ p < .01$

To illustrate the time-variant green bond premium, we calculate the cross-sectional average of $\hat{p_{it}}$ on a daily basis to show the general development of the estimated green bond premium over time (see Figure 3.1).[18] The green bond premium was rather volatile in earlier years and became stable in recent years.[19] It appears that overall the green bond premium was more likely to be negative before 2015, and increased in the following years.

3.6.2 Shades of green and time-variant green bond premium

We continue with the test of hypotheses H2 to H4 regarding whether an external review and the greenness of green bonds impact on the premium in the secondary market. Therefore, we run hybrid model regressions defined in Equation (3.13) with robust standard errors clustered at the issuer level.

To test Hypothesis 2, we include four dummy variables (*SPO*, *green_rating*, *verification*,

[18] The spikes and dips are reasonable, as for some trading days, there are fewer daily observations.

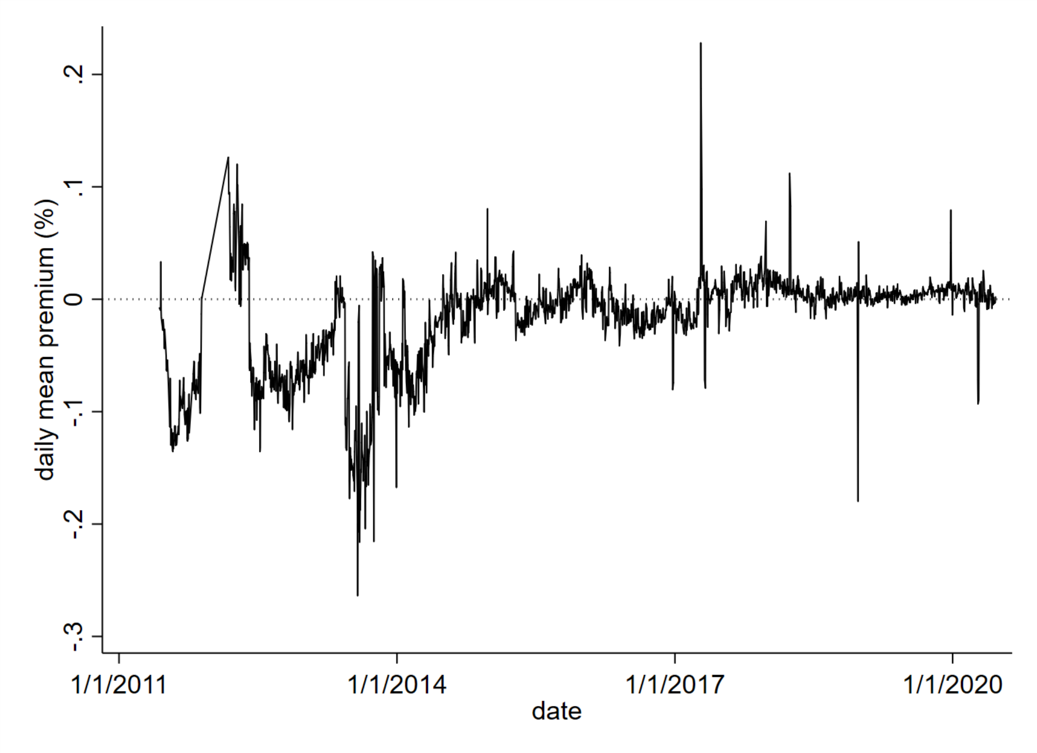
[19] It should be noted that the panel dataset is unbalanced and there are fewer green bonds in the first few years.

Table 3.5: Determinants of the green bond premium: main hybrid models

		Hybrid1	Hybrid2	Hybrid3	Hybrid4
H2	SPO	0.0355***	0.0425*		
		(0.0126)	(0.0231)		
	verification	0.0246**	0.0248**	0.0255***	
		(0.0099)	(0.0100)	(0.0095)	
	CBI_certification	−0.0257	−0.0260	−0.0244	
		(0.0261)	(0.0258)	(0.0264)	
	green_rating	0.0179	0.0187	0.0210*	
		(0.0120)	(0.0116)	(0.0112)	
H3	$d\_maturity * \overline{SPO}$		−0.0070		
			(0.0062)		
	$m\_maturity * \overline{SPO}$		0.0018		
			(0.0037)		
H4	dark_green			0.0536***	0.0227**
				(0.0157)	(0.0094)
	medium_green			0.0376***	0.0076
				(0.0145)	(0.0093)
	no_shade			0.0330***	
				(0.0126)	
Controls	d_maturity	0.0076	0.0103*	0.0076	0.0131**
		(0.0052)	(0.0055)	(0.0052)	(0.0061)
	d_gb_yield	−0.0257**	−0.0262**	−0.0257**	−0.0366**
		(0.0131)	(0.0131)	(0.0131)	(0.0156)
	m_maturity	−0.0026**	−0.0029**	−0.0026**	−0.0024*
		(0.0012)	(0.0014)	(0.0012)	(0.0013)
	m_gb_yield	0.0112	0.0113	0.0100	0.0151
		(0.0071)	(0.0072)	(0.0076)	(0.0096)
	gb_volume	−0.0173	−0.0168	−0.0152	−0.0245*
		(0.0139)	(0.0142)	(0.0131)	(0.0142)
issuer_type	agency	−0.0175*	−0.0186*	−0.0161	0.0013
		(0.0098)	(0.0100)	(0.0106)	(0.0122)
	financial	0.0327**	0.0326**	0.0343**	0.0340**
		(0.0163)	(0.0163)	(0.0161)	(0.0171)
	municipal	−0.0067	−0.0075	−0.0067	0.0020
		(0.0109)	(0.0112)	(0.0109)	(0.0104)
	sovereign	0.0185	0.0175	0.0297	0.0527
		(0.0376)	(0.0373)	(0.0372)	(0.0381)
	supranational	−0.0041	−0.0048	0.0032	0.0004
		(0.0100)	(0.0103)	(0.0110)	(0.0127)
credit_rating	AAA	0.0745***	0.0748***	0.0676**	0.0376
		(0.0269)	(0.0270)	(0.0270)	(0.0284)
	AA+	0.0573*	0.0574*	0.0551*	0.0262
		(0.0295)	(0.0296)	(0.0299)	(0.0278)
	AA	0.0023	0.0027	−0.0029	−0.0001
		(0.0221)	(0.0223)	(0.0238)	(0.0240)
	AA−	0.0484	0.0482	0.0486	0.0417
		(0.0296)	(0.0297)	(0.0302)	(0.0294)
	A+	0.0106	0.0114	0.0091	0.0248
		(0.0357)	(0.0358)	(0.0360)	(0.0398)
	A	−0.0055	−0.0054	−0.0141	−0.0104
		(0.0380)	(0.0379)	(0.0381)	(0.0414)
	A−	0.0976***	0.0988***	0.0955***	0.0786***
		(0.0303)	(0.0313)	(0.0286)	(0.0294)
	BBB+	0.0328	0.0334	0.0321	0.0546**
		(0.0215)	(0.0214)	(0.0218)	(0.0228)
	BBB	0.0137	0.0140	0.0070	−0.0055
		(0.0209)	(0.0208)	(0.0204)	(0.0244)
	NR	0.0381*	0.0384*	0.0338	0.0184
		(0.0225)	(0.0225)	(0.0229)	(0.0246)
	seniority	Yes	Yes	Yes	Yes
	currency	Yes	Yes	Yes	Yes
	_cons	−0.0442	−0.0498	−0.0412	0.0206
		(0.0404)	(0.0440)	(0.0429)	(0.0443)
	N	92,774	92,774	92,774	68,215
	Rho	0.4971	0.4993	0.4971	0.4871

This table reports the results of the hybrid model regressions with the green bond premium $\hat{p_{it}}$ as the dependent variable. Standard errors are cluster-robust at the issuer level. The full sample includes 92,774 daily observations for 250 matched bond triplets. The subsample in Model *Hybrid4* only includes 68,215 daily observations for 196 green bonds with an SPO. $^{*}p < .1$, $^{**}p < .05$, $^{***}\ p < .01$.

Figure 3.1: Premium development over time



This figure shows the daily, cross-sectional average green bond premium over time.

and *CBI_certification*) indicating whether a specific type of external review is available, besides control variables in the hybrid model (Model *Hybrid1*). The coefficients of *SPO* and *verification* are both significantly positive, at the 1% and 5% level, respectively. Thus, green bonds with an SPO and a verification face higher premiums than green bonds without such external reviews. In our theoretical framework, this finding supports the hypothesis that non-financial disclosure from external reviewers increases transparency substantially and there is a sufficiently large group of investors with an $\alpha > 0$ that influence equilibrium prices of green bond investments. However, we cannot find evidence that a CBI certification or a green rating makes an additional marginal contribution to a higher green bond premium. Therefore, even though there are four types of external reviews available in the green bond market, we find that the more popular type of external reviews, i.e., SPOs and verifications, are really valued by green bond investors.

Furthermore, we include an interaction term ($maturity * \overline{SPO}$) of *maturity* and $\overline{SPO}$ ($= 1 - SPO$) in Model *Hybrid2* to test whether the influence of an SPO is time-dependent (H3). In a hybrid model, the time-variant interaction term is transformed into two terms, namely the within-effects term ($d\_maturity * \overline{SPO}$), denoted by the prefix 'd', and the between-effects term ($m\_maturity * \overline{SPO}$), denoted by the prefix 'm'. Nevertheless, the coefficient of the within-effects term is not significant in the entire sample and thus does not support H3 stating that the premium of green bonds without an SPO will increase as investors become more familiar with these bonds.

Model *Hybrid3* takes advantages of the classification of SPOs into different categories, namely *dark_green*, *medium_green*, *no_shade*, and *no_SPO*, according to external reviewers' evaluation results. The coefficients of the different shades of green in Model *Hybrid3* are all significantly positive at the 1% level compared to the *no_SPO* category (reference category), with that of *dark_green* being the highest (5.36 BP) and *no_shade* the lowest (3.30 BP). In line with H2, green bonds reviewed by an SPO provider show significantly higher premiums compared to green bonds without an SPO for all shades of green. More specifically, green bonds with a "better" shade of green tend to have a higher green bond premium, which is in accordance with H4. Thus, investors also integrate the greenness of green bonds, as suggested by an SPO provider, into the pricing. Investors are willing to pay a higher premium if the green bond has proved to contribute seriously to climate adaptation and mitigation.

To investigate the significance of the differences in the impact on the premium among different shades of green, we analyze the impact of the level of greenness on the green bond premium in the subsample of green bonds with an SPO. Accordingly, we run the estimation of Model *Hybrid4*, Table 3.5, on the subsample of green bonds with an SPO. When *no_shade* is taken as the reference category, the coefficient of *dark_green* is significantly positive at the 5% level, while that of *medium_green* is not significant. Thus, investors trade a 'dark green'-shaded green bond at a significantly higher premium than ones with no shade. This finding confirms our theoretical expectation that investors appreciate a higher level of greenness, and supports H4 to some extent. The pricing effect of the greenness level on the green bond premium prevails only for the *dark_green* vs. *no_shade* comparison, but is insignificant for the *dark_green* vs. *medium_green* comparison.[20]

Besides the above findings from variables that are of special interest, it is noteworthy that the coefficient of *d_maturity* is significantly positive at the 5% level in Model *Hybrid4* (the subsample analysis). This pattern indicates that the premium of green bonds with an SPO is positively related to their maturity. In other words, as green bonds with an SPO have been traded on the market for a longer time (the maturity decreases), the green bond premium decreases. This fact provides some weak supporting evidence in the context of H3 to the extent that the premiums of green bonds with an SPO diminish, when the green bonds approach maturity. In terms of informational transaction cost theory, the documented premium difference pattern can be explained by searching costs for information,

[20]This is analyzed by making medium green the reference category and redoing the regression. Given the absence of a significant result, the corresponding table is omitted.

which is already provided by SPOs. The more mature a green bond becomes, the more information on the respective greenness is available. This reduces information costs and therefore, the requirement of higher yields to compensate for idiosyncratic greenness risk and to cover search costs.

Regarding the other control variables, we observe the following. The coefficient of *d_gb_yield* indicates that the green bond yield is negatively related to the green bond premium. As regards issuer type, green bonds issued by agencies have a lower premium (in Model *Hybrid1* - *Hybrid2*), while those issued by financial institutions enjoy a significantly higher premium compared to those issued by corporates (in all model specifications). Moreover, green bonds with a credit rating of AAA or AA+, which constitute a considerable percentage of the sample, evidently enjoy a higher premium in the full sample regressions (Model *Hybrid1* - *Hybrid3*), which may suggest that green bond investors prefer those with the highest credit ratings. Investors are also interested in green bonds with a A− rating, as all the coefficients are significantly positive in different model setups.

3.6.3 Additional analyses

To gauge the robustness of our results from the main models, we adjust the selection filters of yield data and redo the matching process. We limit the maximum bid-ask spread to 30 BP (instead of 50 BP in the main matching process), so as to exclude daily observations of less liquid bonds. Moreover, we increase the required minimum number of daily observations for each bond triplet to 100 (instead of 50 in the main matching process). These changes in filters result in a reduction of around 3,500 daily observations and a total of 34 bond triplets in the sample.

Similarly, we run the hybrid model in Equation (3.11) and extract the green bond premium from Equation (3.12) (see Table 3.6 and Table 3.7). In Table 3.6, we observe a relatively small intercept term which may indicate an overall small positive premium of 0.64 BP (significant at the 10% level). We rerun the hybrid model in Equation (3.13) to investigate the determinants, and present the regression results for the subsample in Table 3.8. The main results for Models *Hybrid1a* - *Hybrid4a* are similar to those of Models *Hybrid1* - *Hybrid4* in Table 3.5, despite some deviations in the significance levels. Both the coefficients of *SPO* and *verification* are significantly positive across different models. Again, all shades of green lead to a higher premium in Model *Hybrid3a*. Moreover, the coefficient of *dark_green* remains significantly positive when the hybrid model is run in a subsample of green bonds with an SPO, and thus supports H4. Nevertheless, we still do not find a statistically significant difference when comparing a dark green shade with a medium green one, or a medium green with no clear shade. Lastly, the coefficient of the control variable *d_maturity* in Model *Hybrid4a* is also significantly positive, which indicates that the premiums of SPO and non-SPO green bonds converge. Regarding control variables *issuer_type* and *credit_rating*, we cannot confirm those findings in the main models with the restricted sample.

Table 3.6: Hybrid model to extract the green bond premium - robustness check

	Coef.	Robust S.E.
$\Delta L_{it} - \overline{\Delta L_i}$	0.4603***	0.1309
$\overline{\Delta L_i}$	0.8898***	0.3364
_cons	0.0065*	0.0037
N	89,285	
Wald chi^2	25.9900	
Prob > chi^2	0.0000	
Rho	0.4917	

This table contains the results of the hybrid model explaining the difference in the yields of green and matched conventional bonds by the variation of liquidity for the restricted sample. $\Delta L_{it} - \overline{\Delta L_i}$ measures the within-variability in the liquidity, i.e., at the bond level. $\overline{\Delta L_i}$ represents the between-variability to capture cross-sectional effects. The restricted sample includes only 89,285 daily observations for 216 bond triplets, due to stricter data filters. Standard errors are cluster-robust at the issuer level. $^*p < .1$, $^{**}p < .05$, *** $p < .01$

Table 3.7: Descriptive statistics for the green bond premium - robustness check

	Obs.	Mean	Std.	p-value	Min	Median	Max
Panel: smaller sample							
$\hat{p_{it}}$	89,285	−0.0002	0.1058	0.4881	−0.9698	0.0024	0.9509
Panel: Zerbib's approach							
$\hat{p_i}$	250	0.0095	0.0820	0.0678	−0.4799	0.0031	0.6327

The restricted sample includes only 89,285 daily observations for 216 bond triplets, due to stricter data filters. The green bond premium $\hat{p_{it}}$ is estimated by Equation 3.12. The green bond premium $\hat{p_i}$ is extracted from the fixed-effects model in Equation 3.14. p-value is from a t-test identifying whether $\hat{p_{it}}$ or $\hat{p_i}$ is significantly different from zero.

Additionally, we rerun the hybrid models in the full sample with robust standard errors clustered at the bond level (instead of at the issuer level in the main models in Table 3.5) as another robustness check (see Table 3.9). The overall significance pattern for Models *Hybrid1b* - *Hybrid4b* remains relatively stable and robust, compared with the main results.

For a further robustness check, we follow the empirical approach of most comprehensive existing analysis of the green bond premium by Zerbib (2019) to estimate the green bond premium in a fixed-effects model as follows:

$$\Delta r_{it} = p_i + \beta \Delta L_{it} + \epsilon_{it} \tag{3.14}$$

where the time-invariant individual effects p_i is treated as the green bond premium. The estimated green bond premium $\hat{p_i}$ has a mean value of 0.95 BP. It is significantly different from zero at the 10% level. This value is very close to the intercept terms in the hybrid

Table 3.8: The green bond premium determinants in the restricted sample

		Hybrid1a	Hybrid2a	Hybrid3a	Hybrid4a
H2	SPO	0.0213**	0.0263*		
		(0.0103)	(0.0160)		
	verification	0.0150*	0.0151*	0.0183**	
		(0.0080)	(0.0080)	(0.0078)	
	CBI_certification	−0.0250	−0.0251	−0.0218	
		(0.0265)	(0.0263)	(0.0270)	
	green_rating	0.0059	0.0060	0.0127	
		(0.0106)	(0.0106)	(0.0107)	
H3	$d\_maturity * \overline{SPO}$		−0.0076		
			(0.0063)		
	$m\_maturity * \overline{SPO}$		0.0011		
			(0.0029)		
H4	dark_green			0.0516***	0.0383***
				(0.0157)	(0.0127)
	medium_green			0.0251**	0.0112
				(0.0116)	(0.0091)
	no_shade			0.0178*	
				(0.0093)	
Controls	d_maturity	0.0095*	0.0124**	0.0095*	0.0147**
		(0.0054)	(0.0062)	(0.0054)	(0.0070)
	d_gb_yield	−0.0236**	−0.0241**	−0.0236**	−0.0325**
		(0.0116)	(0.0119)	(0.0116)	(0.0158)
	m_maturity	−0.0008	−0.0009	−0.0006	0.0001
		(0.0011)	(0.0011)	(0.0011)	(0.0012)
	m_gb_yield	−0.0241*	−0.0241*	−0.0283**	−0.0400**
		(0.0126)	(0.0127)	(0.0128)	(0.0161)
	gb_volume	−0.0230**	−0.0227*	−0.0206**	−0.0296**
		(0.0117)	(0.0116)	(0.0103)	(0.0133)
issuer_type	agency	−0.0200	−0.0204	−0.0183	−0.0134
		(0.0126)	(0.0127)	(0.0144)	(0.0166)
	financial	0.0241	0.0242	0.0262	0.0257
		(0.0167)	(0.0168)	(0.0164)	(0.0174)
	municipal	−0.0052	−0.0056	−0.0065	−0.0045
		(0.0116)	(0.0119)	(0.0118)	(0.0121)
	sovereign	0.0129	0.0128	0.0299	0.0478
		(0.0416)	(0.0416)	(0.0413)	(0.0416)
	supranational	−0.0084	−0.0085	0.0038	0.0010
		(0.0133)	(0.0134)	(0.0130)	(0.0125)
credit_rating	AAA	−0.0098	−0.0099	−0.0242	−0.0577*
		(0.0366)	(0.0365)	(0.0315)	(0.0341)
	AA+	−0.0240	−0.0238	−0.0286	−0.0597*
		(0.0370)	(0.0369)	(0.0325)	(0.0326)
	AA	−0.0646*	−0.0646*	−0.0754**	−0.0877**
		(0.0359)	(0.0359)	(0.0330)	(0.0354)
	AA−	−0.0202	−0.0204	−0.0207	−0.0375
		(0.0418)	(0.0418)	(0.0374)	(0.0368)
	A+	−0.0509	−0.0504	−0.0547	−0.0472
		(0.0491)	(0.0490)	(0.0442)	(0.0459)
	A	−0.0536	−0.0536	−0.0702	−0.0833*
		(0.0470)	(0.0470)	(0.0468)	(0.0498)
	A−	0.0302	0.0308	0.0261	0.0066
		(0.0468)	(0.0470)	(0.0389)	(0.0409)
	BBB+	0.0007	0.0009	−0.0032	0.0000
		(0.0369)	(0.0368)	(0.0315)	(0.0335)
	BBB	−0.0327	−0.0325	−0.0438	−0.0548*
		(0.0343)	(0.0343)	(0.0282)	(0.0297)
	NR	−0.0414	−0.0413	−0.0495	−0.0724**
		(0.0360)	(0.0360)	(0.0309)	(0.0340)
	seniority	Yes	Yes	Yes	Yes
	currency	Yes	Yes	Yes	Yes
	_cons	0.0985	0.0942	0.1112	0.1771*
		(0.0755)	(0.0735)	(0.0739)	(0.0913)
	N	89,285	89,285	89,285	65,962
	Rho	0.4810	0.4841	0.4763	0.4764

This table reports the results of the hybrid model regressions with the green bond premium $\hat{p_{it}}$ as the dependent variable. Standard errors are cluster-robust at the issuer level. Due to stricter data filters, this smaller sample includes 89,285 daily observations for 216 bond triplets. For Model *Hybrid4a*, the subsample includes 65,962 daily observations for 170 green bonds with an SPO. $^{*}p < .1$, $^{**}p < .05$, $^{***}\ p < .01$

Table 3.9: Green bond premium determinants for standard errors clustered at the bond level

		Hybrid1b	Hybrid2b	Hybrid3b	Hybrid4b
H2	SPO	0.0355**	0.0425*		
		(0.0163)	(0.0239)		
	verification	0.0246**	0.0248**	0.0255**	
		(0.0112)	(0.0112)	(0.0111)	
	CBI_certification	−0.0257	−0.0260	−0.0244	
		(0.0281)	(0.0278)	(0.0280)	
	green_rating	0.0179	0.0187	0.0210*	
		(0.0129)	(0.0129)	(0.0128)	
H3	$d\_maturity * \overline{SPO}$		−0.0070		
			(0.0067)		
	$m\_maturity * \overline{SPO}$		0.0018		
			(0.0034)		
H4	dark_green			0.0536***	0.0227*
				(0.0194)	(0.0126)
	medium_green			0.0376**	0.0076
				(0.0175)	(0.0114)
	no_shade			0.0330**	
				(0.0165)	
Controls	d_maturity	0.0076*	0.0103**	0.0076*	0.0131**
		(0.0042)	(0.0050)	(0.0042)	(0.0054)
	d_gb_yield	−0.0257**	−0.0262**	−0.0257**	−0.0366**
		(0.0111)	(0.0112)	(0.0111)	(0.0146)
	m_maturity	−0.0026**	−0.0029**	−0.0026**	−0.0024*
		(0.0013)	(0.0013)	(0.0013)	(0.0015)
	m_gb_yield	0.0112	0.0113	0.0100	0.0151
		(0.0131)	(0.0130)	(0.0135)	(0.0156)
	gb_volume	−0.0173	−0.0168	−0.0152	−0.0245**
		(0.0118)	(0.0119)	(0.0109)	(0.0111)
issuer_type	agency	−0.0175	−0.0186	−0.0161	0.0013
		(0.0131)	(0.0133)	(0.0138)	(0.0171)
	financial	0.0327*	0.0326*	0.0343*	0.0340*
		(0.0176)	(0.0176)	(0.0177)	(0.0187)
	municipal	−0.0067	−0.0075	−0.0067	0.0020
		(0.0128)	(0.0132)	(0.0127)	(0.0122)
	sovereign	0.0185	0.0175	0.0297	0.0527
		(0.0400)	(0.0399)	(0.0406)	(0.0398)
	supranational	−0.0041	−0.0048	0.0032	0.0004
		(0.0185)	(0.0187)	(0.0184)	(0.0211)
credit_rating	AAA	0.0745***	0.0748***	0.0676**	0.0376
		(0.0274)	(0.0274)	(0.0275)	(0.0297)
	AA+	0.0573**	0.0574**	0.0551*	0.0262
		(0.0290)	(0.0292)	(0.0297)	(0.0279)
	AA	0.0023	0.0027	−0.0029	−0.0001
		(0.0225)	(0.0225)	(0.0241)	(0.0251)
	AA−	0.0484	0.0482	0.0486	0.0417
		(0.0308)	(0.0309)	(0.0315)	(0.0304)
	A+	0.0106	0.0114	0.0091	0.0248
		(0.0358)	(0.0358)	(0.0363)	(0.0401)
	A	−0.0055	−0.0054	−0.0141	−0.0104
		(0.0395)	(0.0394)	(0.0402)	(0.0431)
	A−	0.0976***	0.0988***	0.0955***	0.0786***
		(0.0309)	(0.0314)	(0.0294)	(0.0303)
	BBB+	0.0328	0.0334	0.0321	0.0546**
		(0.0214)	(0.0213)	(0.0221)	(0.0240)
	BBB	0.0137	0.0140	0.0070	−0.0055
		(0.0208)	(0.0208)	(0.0204)	(0.0247)
	NR	0.0381	0.0384	0.0338	0.0184
		(0.0244)	(0.0244)	(0.0249)	(0.0264)
	seniority	Yes	Yes	Yes	Yes
	currency	Yes	Yes	Yes	Yes
	_cons	−0.0442	−0.0498	−0.0412	0.0206
		(0.0605)	(0.0607)	(0.0629)	(0.0645)
	N	92,774	92,774	92,774	68,215
	Rho	0.4971	0.4993	0.4971	0.4871

This table reports the results of the hybrid model regressions with the green bond premium $\hat{p_{it}}$ as the dependent variable. Standard errors are cluster-robust at the bond level. The full sample includes 92,774 daily observations for 250 matched bond triplets. For Model *Hybrid4b*, the subsample only includes 68,215 daily observations for 196 green bonds with an SPO. $^{*}p < .1$, $^{**}p < .05$, *** $p < .01$.

models (see Table 3.4 and Table 3.6). Thus, overall we find some evidence that there is in general a small green bond premium.

Lastly, we run additional cross-sectional OLS regressions with the estimated time-invariant individual effects $\hat{p_i}$ as the dependent variable to investigate the determinants:

$$\hat{p_i} = \beta_0 + \beta_1 B_i + \beta_2 G_i + \epsilon_i \tag{3.15}$$

where B_i represents a vector of variables covering basic bond features, and G_i is a vector of variables related to information from external review reports. Note that two control variables, namely *maturity* and *gb_yield*, can no longer be included in cross-sectional OLS regressions. For this reason, we cannot test H3 regarding whether the impact of external reviews is time-dependent.

Model *OLS_1* of Table 3.10 shows that the coefficient of *SPO* is significantly positive at the 1% level. This provides evidence supporting H2 stating that external review reports have a positive influence on the premium. Regarding *CBI_certification*, *green_rating*, and *verification*, we do not find strong evidence for H2 except that *green_rating* is significant at the 10% level in Model *OLS_2*. When SPOs with different shades of green are treated separately in Model *OLS_2*, the coefficient of *dark_green*, *medium_green*, and *no_shade* yield a similar pattern as in the main models and thus support H4. Lastly, the coefficient of *dark_green* shows significantly higher premium in Model *OLS_3*. In summary, the OLS regression results support most of our main findings from the hybrid models regarding H2 and H4.

3.7 Conclusion

In this paper, we revisit the existence of the green bond premium in a comprehensive dataset and examine systematically the impact of all four different types of external reviews and their greenness evaluation on the bond yields. To estimate the green bond premium, we adopt a strict matching between green and conventional bonds. After the matching process, the final sample contains 250 green bonds matched with 500 conventional ones, and more than 92,774 daily observations from 2011 to 2020. On this sample, we perform a two-step regression procedure based on a hybrid model to elicit the green bond premium and its determinants. The first main finding is that, on average, the expected green bond premium is positive and statistically significant.

Table 3.10: Green bond premium determinants in an OLS regression approach

		OLS1	OLS2	OLS3
H2	SPO	0.0360***		
		(0.0131)		
	verification	0.0112	0.0119	
		(0.0114)	(0.0112)	
	CBI_certification	−0.0227	−0.0213	
		(0.0290)	(0.0294)	
	green_rating	0.0171	0.0192*	
		(0.0107)	(0.0107)	
H4	dark_green		0.0521***	0.0232**
			(0.0170)	(0.0105)
	medium_green		0.0389***	0.0102
			(0.0143)	(0.0082)
	no_shade		0.0335**	
			(0.0132)	
Controls	gb_volume	−0.0212*	−0.0189	−0.0236**
		(0.0124)	(0.0118)	(0.0111)
issuer_type	agency	−0.0169	−0.0155	−0.0039
		(0.0119)	(0.0127)	(0.0140)
	financial	0.0260*	0.0273*	0.0251
		(0.0149)	(0.0149)	(0.0158)
	municipal	−0.0091	−0.0086	0.0005
		(0.0118)	(0.0116)	(0.0099)
	sovereign	−0.0163	−0.0051	0.0242
		(0.0446)	(0.0444)	(0.0452)
	supranational	−0.0085	−0.0015	−0.0080
		(0.0133)	(0.0143)	(0.0138)
credit_rating	AAA	0.0223	0.0163	−0.0154
		(0.0354)	(0.0327)	(0.0315)
	AA+	−0.0050	−0.0071	−0.0335
		(0.0350)	(0.0320)	(0.0294)
	AA	−0.0471	−0.0518	−0.0602*
		(0.0344)	(0.0322)	(0.0313)
	AA−	−0.0040	−0.0036	−0.0166
		(0.0402)	(0.0371)	(0.0360)
	A+	−0.0267	−0.0281	−0.0166
		(0.0477)	(0.0450)	(0.0472)
	A	−0.0376	−0.0452	−0.0540
		(0.0494)	(0.0485)	(0.0509)
	A−	0.0336	0.0317	0.0195
		(0.0379)	(0.0329)	(0.0329)
	BBB+	0.0094	0.0085	0.0144
		(0.0353)	(0.0316)	(0.0320)
	BBB	−0.0261	−0.0327	−0.0449
		(0.0321)	(0.0283)	(0.0297)
	NR	−0.0195	−0.0231	−0.0333
		(0.0324)	(0.0290)	(0.0292)
	seniority	Yes	Yes	Yes
	currency	Yes	Yes	Yes
	_cons	0.0182	0.0180	0.0741*
		(0.0399)	(0.0368)	(0.0442)
	N	250	250	196
	R^2	0.23	0.24	0.34
	Adjusted R^2	0.08	0.08	0.19

This table reports the results of the OLS regressions with the green bond premium $\hat{p_{it}}$ as the dependent variable. The dependent variable is the estimated individual effects $\hat{p}_i$ derived from the fixed-effects regression. Standard errors are cluster-robust at the issuer level. The full sample includes 250 matched bond triplets. For Model *OLS3*, the subsample only includes 196 green bonds with an SPO. $^*p < .1$, $^{**}p < .05$, $^{***}\ p < .01$

However, some green bonds are priced evidently higher than their counterparts. In particular, green bonds with an SPO or a verification c.p. enjoy a higher green bond premium. This relationship indicates that credible and assured non-financial disclosure seems valuable for investors. In particular, investors trade green bonds with SPOs at prices that increase with the level of greenness evaluation of the green bond, i.e., a darker shade of green is more likely to have a higher premium. This pattern implies that the shade-of-green methodology adopted by external reviewers has the potential to function as a tool for assessing the greenness of green bonds in a pricing-relevant manner, analogous to credit ratings. Issuers of green bonds can thereby lower the financing costs, at least for such green bonds that finance deeply green projects related to mitigating climate change.

Our results also have significant policy and research implications. Independent external reviews appear to be one of the most important pillars of a healthy green bond market, through reducing information asymmetry between issuers and investors. The importance of external reviews and shade of green methodology in green bond pricing reveals that investors are sensitive to information asymmetry on the green asset market. If more public information regarding the greenness of green bonds is available, the investor base of green assets may be extended as investors have more confidence in green assets and are subject to a lower risk of greenwashing. Thus, a reduction of information asymmetry is indeed crucial to the development of climate finance. For instance, there could be more deliberately designed mandatory rules that foster transparency in the industry besides current voluntary-based industry guidelines such as the GBPs. Easier access to third-party reports and evaluations should be promoted to facilitate communication among market participants. Governmental policies supporting issuers of green bonds to achieve standardized, affordable, and independent greenness assessments may contribute to a prosperous climate finance market. This observations on financial markets also highlight the need for more theoretical and empirical research on green finance aspects. Clearly, our findings are not in line with traditional finance theory. Thus, from a behavioral finance perspective they provide some evidence for a greenness bias in the prices of green bonds. However, a contemporary view on such phenomena is rather to rationalize them, i.e., to view them as rational and not as irrational effects. The theoretical reasoning pursued in this study adheres to such an approach. However, more future research on the rationale – and even the calculus – of impact investors appears to be in urgent need. Moreover, future research may analyze further green pricing anomalies and, if applicable, develop a new asset pricing model for bonds. Furthermore, these results are not limited to the bond market but may be applied to other asset classes.

3.8 Appendix

Table 3.11: How the sample size is reduced during the matching process

	Criterium Description	Sample Size
	Initial sample	1,248
1	Same bond structure (i.e. straight conventional bonds)[1]	−292
2	Same currency type	−74
3	Same coupon type	−1
4	Same seniority and collateral status	−67
5	Same credit rating	−25
6	Issue amount: 0.25 to 4 times	−66
7	Issue date: -6 to 6 years	−38
8	Duration difference: -2 to 2 years	−241
9	50 joint daily yield observations	−194
	Final sample	250

This table shows how the sample size is reduced during the matching process step by step. The initial sample size of green bonds is 1,248. We extract a complete list of conventional bonds for each green bond issuer and start the matching process from step 1 to step 9.

[1]This requirement means that for 292 green bonds we do not find straight conventional bonds which can be matched with green bonds.

Table 3.12: Definition of variables

Variable	Description
H1	
Δr	Yield difference between green bonds and comparable synthetic conventional bonds.
$\hat{p_{it}}$	Green bond premium extracted from the hybrid model in Equation 3.11.
$\hat{p_i}$	Individual effects extracted from the fixed-effects model in Equation 3.14.
H2	
SPO	Binary variable with a value of one if a second-party opinion is assigned to the green bond, zero otherwise.
verification	Binary variable with a value of one if a verification is assigned to the green bond, zero otherwise.
CBI_certification	Binary variable with a value of one if a CBI certification is assigned to the green bond, zero otherwise.
green_rating	Binary variable with a value of one if the green bond has a green rating from a traditional credit rating agency, zero otherwise.
H3	
$\overline{SPO}$	Binary variable with a value of one if a second-party opinion is not available, zero otherwise. $\overline{SPO} = 1 - SPO$.
H4	
shade	Categorical variable indicating the shade of green. Green bonds are classified into four categories, namely dark green, medium green, no shade and no SPO. The default reference category is no SPO.
Controls	
ΔL	Liquidity difference between a green bond and its comparable synthetic conventional bond.
maturity	Maturity of the green bond.
gb_yield	Daily bid yield of the green bond.
gb_volume	Issue volume of the green bond.
cb_volume	Issue volume of the synthetic bond. The issue volume of the synthetic bond is calculated as the mean of the issue volumes of the two conventional bonds ($cb1$ and $cb2$).
seniority	Categorical variable indicating the seniority and the collateral status of the green bond on Eikon. The reference category is 'unsecured'.
currency	Categorical variable indicating which currency the green bond is denominated in. The reference category is USD.
issuer_type	Green bond issuers are classified into six categories, such as agency, corporate and financial institution. The reference category is corporate.
credit_rating	Credit rating of the green bond. Credit ratings from different rating agencies have been transformed into the same scale. The reference category is BBB−.

Figure 3.2: The matching process

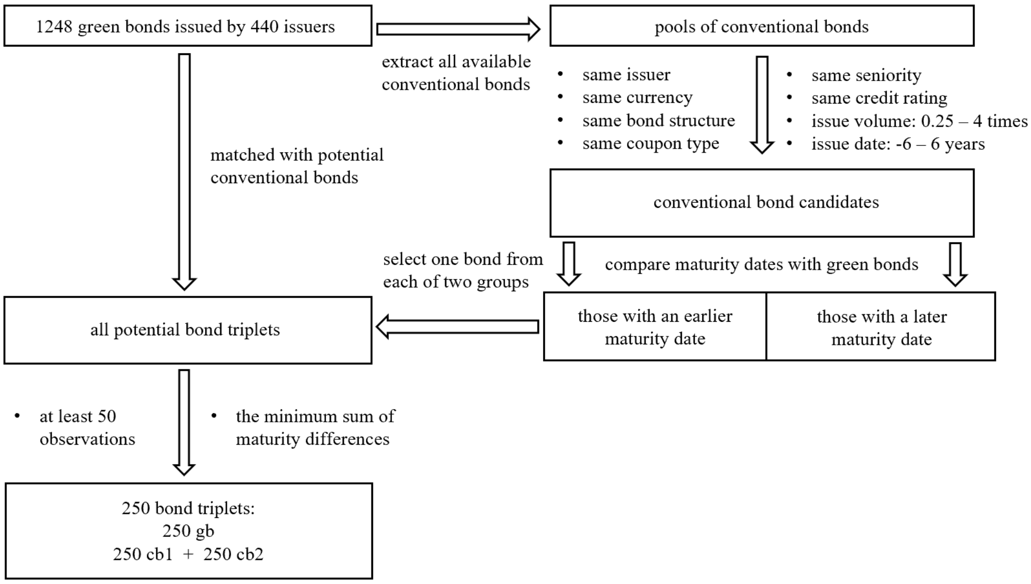


Table 3.13: Descriptive statistics for the green bond sample before matching - metric variables

Variable	Obs.	Mean	Std.	Min	Median	Max
maturity[1] (in years)	1,248	7.4109	6.2719	0.9945	5.0055	100.0658
gb_volume (bn USD)	1,248	0.3000	0.5010	0.0000	0.1029	6.6912

This table reports summary statistics on characteristics of the green bond sample before the matching process. The sample includes 1,248 green bonds. The variables are defined in Table 3.12.
[1] Maturity of the green bond at issuance.

Table 3.14: Descriptive statistics for the green bond sample before matching - categorical variables

Variable	Obs.	Relative	Variable	Obs.	Relative
seniority[1]			SEK	148	11.86
MTG	12	0.96	SGD	4	0.32
SEC	6	0.48	THB	5	0.40
SR	904	72.44	TRY	12	0.96
SRBN	11	0.88	TWD	23	1.84
SRP	20	1.60	USD	264	21.15
SRSEC	38	3.04	VND	2	0.16
UN	257	20.59	ZAR	17	1.36
currency			*issuer_type*		
AUD	58	4.65	agency	183	14.66
BRL	15	1.20	corporate	345	27.64
CAD	27	2.16	financial	389	31.17
CHF	18	1.44	municipal	67	5.37
CNY	119	9.54	sovereign	11	0.88
COP	1	0.08	supranational	253	20.27
CZK	2	0.16			
DKK	2	0.16	*credit_rating*		
EUR	245	19.63	AAA	299	23.96
GBP	10	0.80	AA+	57	4.57
HKD	20	1.60	AA	82	6.57
HUF	3	0.24	AA−	78	6.25
IDR	8	0.64	A+	105	8.41
INR	19	1.52	A	23	1.84
JPY	105	8.41	A−	50	4.01
KRW	2	0.16	BBB+	37	2.96
MXN	12	0.96	BBB	26	2.08
MYR	54	4.33	BBB−	14	1.12
NGN	1	0.08	BB+	1	0.08
NOK	21	1.68	BB	2	0.16
NZD	20	1.60	B+	1	0.08
PEN	2	0.16	B	2	0.16
PHP	2	0.16	B−	1	0.08
PLN	3	0.24	NR[2]	470	37.66
RUB	4	0.32			

This table reports summary statistics on characteristics of the green bond sample before the matching process. The sample includes 1,248 green bonds. The variables are defined in Table 3.12.

[1] *Seniority* indicates the combined information on bond seniority and collateral status on Eikon. MTG: senior secured and mortgage backed; SEC: secured; SR: senior unsecured; SRBN: senior non-preferred; SRP: senior preferred; SRSEC: senior secured; UN: unsecured.

[2] NR means that the green bond does not have a S&P equivalent crediting rating on Eikon.

Chapter 4

The Pricing of ESG News: A Comprehensive Investigation via BERT

This research paper is joint work with Gregor Dorfleitner. It will be soon submitted to a renowned journal for peer review.

Abstract: In this paper, we examine the pricing implication of ESG news on financial markets in a systematic manner. Instead of acquiring proprietary ESG news datasets directly from specific ESG data providers, we extract ESG news from massive raw news articles on Thomson Reuters Eikon in an innovative and relatively accurate way. We showcase how the newest development in NLP (i.e., the BERT model) can be applied to build a comprehensive and unique ESG news dataset, and how news sentiment efficiently recognized by machine could be applied to examine soft factors on financial markets. Specifically, we adopt this methodology to investigate the pricing mechanism of ESG news on major stock markets for more than 13,000 listed companies from all over the world. We find that the market reacts to ESG news based on news sentiment. On the event day, positive ESG news have an average abnormal return of 0.31% while negative ESG news lead to a mean value of -0.75%. More interestingly, we find that the impact of ESG news may depend on the company's historical ESG record. The negative impact of negative ESG news have less severe consequence for companies with an overall better ESG record, while the positive impact of positive ESG news may be more pronounced for companies with a worse ESG record.

Keywords: ESG, , NLP, BERT, Sentiment analysis

JEL Classification: G12 Q51 Q56

With the increasing awareness of ethical issues such as environment protection and social care, the conception of ESG has become more and more prominent and urgent, not only in our everyday lives but also on the financial markets. As Van Duuren et al. (2016) and Amel-Zadeh and Serafeim (2018) suggest, ESG is already regarded as one of the important considerations for fund managers. Since the outbreak of the COVID-19 pandemic, the social pillar of ESG gains much more attention than ever. On January 13 2021, as the file hosting service provider Dropbox Inc announced the laid-off of 11% of its workforce due to the need to shift business resources to response to the challenge of the pandemic, its stock price dropped nearly 6%. The stock market reaction shows vividly that besides financially material news, *instant ESG news* can also be an important factor and price driver on financial markets.

In the past decade, ESG has also become one of the hottest topics in finance literature. However, the research of ESG issues is still at its initial stage. Most ESG studies (e.g. Bennani et al., 2018; Hartzmark and Sussman, 2019) rely heavily on ESG data such as different ESG ratings provided by specific ESG data providers based on their in-house developed methodologies (Fiaschi et al., 2020). As Dorfleitner et al. (2015) suggest, there is an evident lack in the convergence of ESG measurement concepts and the different ratings neither coincide in distribution nor in risk. Therefore, empirical studies focusing on proprietary ESG performance proxies may be subjected to the problem of proxy biases. Also, the low-frequency of those ESG ratings and various rating methodologies make it almost impossible to understand how the market reacts to ESG issues in real time. Most recently, several studies (see e.g. Krüger, 2015; Capelle-Blancard and Petit, 2019; Taleb et al., 2020; Naumer and Yurtoglu, 2020) start to focus more on ESG news, which appear much more frequently than ESG ratings. Krüger (2015) finds some evidence that investors may react to ESG events and reveal their possible pricing implications. However, due to the difficulty to process unstructured raw text data, these studies have to acquire ESG news data from ESG data providers.[1] The reliance on proprietary ESG news dataset may raise the concern that empirical results regarding the pricing implications of ESG news on the financial markets could be sensitive to how ESG news data providers collect (e.g., different ESG news coverage) and process ESG news (e.g., different implementation of sentiment analysis). Therefore, despite some efforts being made, whether ESG news are priced and how they influence financial markets are far from being fully understood.

In this study, we show how a comprehensive ESG news dataset is built upon massive raw ESG news and how news sentiment is extracted in a transparent way before empirical investigations are conducted. Compared with related studies which often adopt

[1] To the best of our knowledge, all related studies acquire ESG news from ESG data providers. For instance, Krüger (2015) acquires ESG events data directly from MSCI KLD.

ready-for-use ESG news data from data providers, this study builds a ESG news dataset based on raw ESG news published by more than 10,000 news sources on Thomson Reuters Eikon. We introduce the recent development in Nature Language Processing (NLP), i.e. the BERT model, to construct a comprehensive and fresh ESG news dataset from raw ESG news. Moreover, we extract sentiment signals from the unstructured textual data by applying the fined-tuned BERT sentiment classifier, which is considered as more accurate than classical sentiment analysis methods such as lexicon-based sentiment analysis (Kotelnikova et al., 2021; Alaparthi and Mishra, 2021).

With such a comprehensive dataset including almost all listed stocks with ESG news coverage for the past two years, we conduct for the first time a complete empirical investigation of the pricing effects of ESG news on major stock markets. It sheds some light on the market reactions to instant ESG information. We find that the market responses to ESG news parallel to the news sentiment. The market reacts positively to positive ESG news while negatively to negative ESG news. Yet these reactions appear to be asymmetric. The market reaction to negative ESG news are stronger as compared with positive ESG news. These patterns exist not only on American stock markets, but also on European stock markets. At last, we discover an interesting point regarding the relationship between ESG news shock and historical ESG records. When investors are confronted with ESG news, they also take the overall ESG performance of the target company into consideration. Companies with better ESG record suffer less from market value loss due to negative ESG news, while those with worse ESG performance enjoy more market value gain when facing positive ESG news.

These findings add to the discussion of integrating ESG factors in asset pricing (see e.g. Pedersen et al., 2020). Since ESG issues are found to be perceived seriously by investors, they should be considered and included as important factors in related research. Moreover, the pricing implications of ESG news questions the efficiency of financial markets, as systematic arbitrage by closely monitoring ESG news could be viable. Our study also adds to the debate whether companies tend to exaggerate their ESG performance (see e.g. Kim and Lyon, 2015). Our data shows that positive ESG news prevail on the market, which might suggest the existence of performance exaggeration regarding ESG issues. Meanwhile, the fact that the overwhelming positive ESG news are still perceived positively suggests that investors might not be able to completely detect false claim of good ESG performance. Our study shows that companies could possibly game the system by releasing more ESG information to their advantage.

The contribution of this study is twofold. First, we show how to apply the BERT model to build our own unique and massive ESG news dataset and judge news sentiment. Especially, we believe that the newest breakthrough in NLP can also contribute to the advancement in financial studies focusing particularly on soft factors and provide a new and better approach in the toolbox of financial researchers to gain deeper insight into their role on the financial markets. Second, to the best of our knowledge, we examine the pricing implications of ESG news in a comprehensive and complete framework for the first time. We do not rely on ESG news dataset from specific data providers and thus avoid the possible biases and errors associated with such tailored datasets. The way we build the ESG news dataset enables us to come to more credible conclusions regarding the pricing implication of ESG news. Even though some earlier studies find that only negative ESG news matters (Krüger, 2015; Capelle-Blancard and Petit, 2019), we find evidence that investors may also value positive ones, albeit to a smaller extent. This finding has a policy implication for companies that it really matters to improve their ESG profile, but not just to avoid negative ESG news. Moreover, this study gives some clues regarding how investors deal with the relationship between newest and past ESG performance, which is rarely touched upon in ESG studies (see e.g. Serafeim and Yoon, 2021). Our study suggests that a good long-term ESG profile might serve as a buffer to moderate the impact of short-term ESG news.

The remainder of the paper is organized as follows. In Section 4.1, we discuss basic background information regarding different types of ESG information, especially ESG news. We review the literature on the pricing implications of ESG news and propose hypotheses in Section 4.3. Section 4.4 describes how we build ESG news dataset step by step. In Section 4.5, we discuss necessary empirical methodological approaches. Section 4.6 presents the empirical results and Section 4.7 concludes.

4.1 ESG information processing

As the interest and demand of stakeholders in ESG issues grows, companies are subject to an increasing amount of ESG reporting guidance or requirements (KPMG, 2019). EU requires that large companies with more than 500 employees should report policies they implement regarding issues such as environmental protection, social responsibility and respect for human rights (EU Non-Financial Reporting Directive) (Grewal et al., 2019). More and more companies response to disclose ESG information, either voluntarily or mandatorily, by releasing stand-alone or integrated ESG reports. According to the survey conducted by KPMG (2020), the percentage of the biggest companies which report on sustainability has increased from 53% in 2008 to 80% in 2020. Nevertheless, ESG disclosure as a source of ESG information has several obvious drawbacks for stakeholders. In general, most ESG reports are released only once a year and may not reflect recent company ESG concerns. Also, the practice of ESG disclosure is relatively flexible, vary in different companies and countries (Baldini et al., 2018) due to different regulations and management (McBrayer, 2018; Li et al., 2018), and thus often lacks credibility, timeliness and relevance (Maniora, 2017).

Due to the difficulty to process ESG disclosure directly, stakeholders often rely on a third-party assessment, especially ESG ratings from ESG rating agencies (Berg et al., 2019). ESG rating agencies are the major ESG information distributors. They usually apply a qualitative and quantitative methodology to assess corporate ESG performance by constructing ESG rating metrics based on information collected from different sources such as ESG disclosure, ESG news and questionaries (Escrig-Olmedo et al., 2019; Del Giudice and Rigamonti, 2020). ESG ratings from these agencies are widely adopted by market participants (Escrig-Olmedo et al., 2019) and thus play a central role in reducing information asymmetry between companies and stakeholders to contribute positively to sustainable development (Del Giudice and Rigamonti, 2020; Lopez et al., 2020). However, a few studies raise the concern that whether ESG ratings are good proxies of corporate ESG performance (Dorfleitner et al., 2015; Drempetic et al., 2019). This question is of great importance, as market participants may be misled by ESG rating if they fail to measure true ESG performance (Christensen et al., 2019). Despite the growth and development in ESG ratings, it is argued that ESG ratings as a ESG performance measurement have many problems (Escrig-Olmedo et al., 2014, 2019; Drempetic et al., 2019; Lopez et al., 2020). First, ESG rating agencies may fail to measure ESG performance with their rating schemes. Escrig-Olmedo et al. (2019) conduct an analysis of ESG rating agencies and find that ESG rating agencies do not fully integrate sustainability principles into their assessment process. Drempetic et al. (2019) document the influence of firm size on ESG ratings and raise the question whether ESG rating agencies really measure what market participants expect. Another issue is that ESG ratings from different agencies can reveal significant differences (Dorfleitner et al., 2015; Berg et al., 2019; Lopez et al., 2020). Berg et al. (2019) investigate the divergence of ESG ratings and find three main sources, i.e., scope divergence, measurement divergence and weights divergence. Those divergences lead to difficulties in choosing the right ESG rating scheme as decision-markers try to evaluate ESG performance (Berg et al., 2019; Escrig-Olmedo et al., 2019). At last, the fact that most ESG ratings are updated on a yearly basis poses a challenge for tracking corporate ESG performance in time. Event though ESG rating agencies consider various sources of ESG information including high-frequency data such as ESG news (Escrig-Olmedo et al., 2019), they are often embedded in rating scores from time to time and cannot reflect the recent development of ESG performance.

Besides official ESG disclosure and ESG ratings from ESG agencies, *instant ESG news* can be another important source of information for investors. The media nowadays plays a central role in diffusing information on financial markets and contribute to the efficiency of the stock market by improving the dissemination of information (Peress, 2014). On financial terminals such as Thomson Reuters and Bloomberg or main stream websites, news stories related to specific companies, including company ESG news, are updated at a lightning speed. If investors care about ESG issues just like traditional financial fundamentals, they could possibly be influenced by reading these ESG news articles. However, unlike ESG ratings as numeric values, ESG news from different news sources are unstructured text data which is difficult to quantify. While ESG rating values can be homogeneously interpreted as the overall ESG performance, ESG news cannot be easily standardized and transformed into a common index which is easy to comprehend. Although *instant ESG news* may be consumed by individual or institutional investors and thus integrated into their investment decision-making process, it is unclear how and to what extent they may react to these instant non-financial information. To answer this question, a comprehensive stream of *instant ESG news* should be available and processed in a plausible way. Nevertheless, a ready-for-use ESG news dataset is usually not for free and should always be purchased from specific ESG data providers. Earlier related studies often adopt such a ESG news dataset from several popular ESG data providers such as MSCI KLD, Ravenpack and Truvalue Labs (Capelle-Blancard and Petit, 2019; Krüger, 2015; Taleb et al., 2020). The key problem of this approach is that these proprietary ESG data providers may have different news coverage and textual processing methodologies, which are in most cases not transparent to researchers.

4.2 Advancement in Nature Language Processing: the BERT model

As the need to understand the role of soft factors extracted from unstructured text data on financial markets grows, classical textual analysis has been more commonly adopted in financial studies in recent years (see e.g. Dorfleitner et al., 2016). Despite some preliminary progress, it appears that the research with classical textual analysis has reached the stage of stagnation as its benefits appear to have been fully exploited.

The progresses in NLP in the past few years, however, give new hope for a further quantification of unstructured text data. Devlin et al. (2018) propose an exciting language presentation model, which is called Bidirectional Encoder Representations from Transformers (BERT). The BERT model is designed to pre-train deep bidirectional textual representation from unlabelled text data. Since its introduction, it has been recognized widely as the state-of-the-art language model in various language tasks. The power of the BERT model originates from several parts. First, the massive size of the BERT model is unprecedented: the base BERT model contains 110 million parameters. Second, the deliberately designed neural networks can grasp the complex relationship among words and sentences. The neutral network architecture of the BERT model is based on several encoder layers of the popular *Transformer* model proposed by Vaswani et al. (2017), of which the most important part is the so-called *self-attention* mechanism. Third, the BERT model is pre-trained with unprecedentedly massive text datasets including the BookCorpus and English Wikipedia (Devlin et al., 2018) over two different pre-training tasks.[2] With such large training inputs, the BERT model can be pre-trained to the extent that meaningful word or sentence representations can arise.

The BERT model is a transfer learning framework and its usage is often separated into two stages: pre-training and fine-tuning. Various pre-trained BERT models have been pre-trained on different unlabelled text datasets with different training settings and can be accessed by researchers who seek to quantify textual information for their purposes. They can be applied directly to a wide range of down stream tasks such as text classification, named entity recognization and question answering, and has obtained the best results for many language tasks (Devlin et al., 2018). For a specific language task such as sentiment classification, researchers can continue training a pre-trained BERT model with their own labelled datasets.

After the introduction of the original BERT model (Devlin et al., 2018), some more refined and robust BERT-like models, such as RoBERTa (Liu et al., 2019) and ALBERT (Lan et al., 2019), are proposed based on the basic architecture of the BERT model and achieve better performance by slightly modifying some parts of the model design or the pre-training hyper-parameters. These models are also available to scholars and can be further fine-tuned for different language tasks.[3]

[2]The discussion of the BERT model is not the key part of this study. Please refer to Devlin et al. (2018) and Vaswani et al. (2017).

[3]For example, the *Hugging Face* team maintains a list of pre-trained BERT-like models: https://huggingface.co

4.3 Literature review and hypothesis development

While numerous studies report a positive relationship between ESG performance and corporate financial performance (Friede et al., 2015), there is less consensus about how ESG performance may influence the financial markets. Although the investment community consider ESG information during investment decision-making process (Amel-Zadeh and Serafeim, 2018; Van Duuren et al., 2016), the role of ESG issues on financial markets is not well understood (Bennani et al., 2018). Pedersen et al. (2020) theoretically propose a ESG-adjusted CAPM and predict that a security with a higher ESG score has a higher demand from ESG investors, which is also supported by the empirical evidence that ESG performance proxies correlate positively with institutional holdings. Hartzmark and Sussman (2019) examine the relationship between the sustainability rating rankings of the US mutual funds and fund flows and present evidence that investors do value sustainability. Regarding the market performance related to ESG investment, Mănescu (2011) find that only some ESG attributes such as community relations, has an impact on stock returns by analyzing a long panel dataset of US firms. Bennani et al. (2018) document that the impact of ESG screening on stock performance is highly time-dependent: they find no evidence of a consistent reward for ESG integration during the 2010-2013 period while a significant excess return for the 2014-2017 period.

Despite their different perspectives and results, these earlier studies usually adopt some kind of ESG performance proxies provided by ESG data providers such as ESG rating. Nevertheless, very few studies focus on the pricing implication of high-frequency news in the field of ESG studies (see e.g. Krüger, 2015; Capelle-Blancard and Petit, 2019), despite the existence of a stream of studies investigating ESG events such as announcements or disclosure (Flammer, 2013; Naughton et al., 2019; Grewal et al., 2020)[4]. However, there are a significant number of studies analyzing the relationship between financial news and stock markets (Alanyali et al., 2013; Boudoukh et al., 2019). For instance, Alanyali et al. (2013) find that financial news are found to be closely linked to trading movements. Boudoukh et al. (2019) find evidence that there is a close relationship between identified relevant firm-level financial news and stock prices. In particular, the tone of news can be of great importance to investors. Many studies apply semantic analysis to extract sentiment signals in financial news articles and investigate their possible influence. Tetlock (2007) uses a word count program to analyze texts, to investigate the interaction between financial news and the stock market and observes that the extracted media sentiment predicts stock prices and trading volume. In recent years, the development of machine learning techniques has enabled researchers to investigate the role of news tonality on

[4]They do not really touch upon high-frequency ESG news in our context.

financial markets in deeper detail. Heston and Sinha (2017) measure news sentiment with proprietary neural network and find that daily financial news can predict stock returns for one to two days. Ke et al. (2019) introduce a supervised learning framework that can extract sentiment information from financial news articles and find that those extracted sentiment signals can predict stock returns to a large extent.

Similarly, *instant ESG news* as an important source of ESG information for (ESG) investors could possibly influence their investment decisions. Positive (negative) ESG news indicate the marginal improvement (deterioration) of company ESG performance and could be considered by investors in two ways. On the one hand, an improvement (deterioration) of ESG performance may lead to an improvement (deterioration) in corporate financial performance (Friede et al., 2015) and thus have an impact on the stock performance via the incorporation of this positive cash flow news into prices. On the other hand, an improvement (deterioration) of ESG performance may attract (repel) ethical investors who have the incentive to promote ESG development (Pedersen et al., 2020). Therefore, we expect that the market reaction to ESG news is closely related to the news sentiment.

H1: Positive (negative) ESG news is associated with stock over-performance (under-performance).

However, the market reaction to positive and negative ESG news could be different in terms of scale. Capelle-Blancard and Petit (2019) investigate about 33,000 ESG news for 100 listed companies from 2002 to 2010 provided by the ESG data provider Covalence and find that companies facing negative ESG news experience a drop of 0.10% in the market value, whereas gain nothing on average from positive ones. This could be explained by investors' concern that companies have the incentive to exaggerate their ESG performance (Yu et al., 2020). With the increasing attention paid to ESG from various stakeholders, some companies find it beneficial to overstate their commitment to ESG topics (Bazillier and Vauday, 2009). For instance, greenwahsing, which describes the intention of companies to label non-green products or practices as green, has been a hot topic in the past two decades (Flammer, 2020). Nevertheless, a pretending of unsubstantiated ethical engagement can cause public mistrust (Jahdi and Acikdilli, 2009). If companies disclose more frequently ESG information or exaggerate their ESG performance, the probability that companies do good to the society decreases or the overall contribution is less valued. Therefore, investors may react less actively to overwhelming positive ESG news. Another explanation can be the so-called "negativity bias", in which the market reacts significantly to negative news while remains relatively calm when good news arrive. In psychology literature, negativity bias refers to the phenomenon that humans give greater weight to negative events, which is manifested in different ways such as negative potency, steeper negative gradients, negativity dominance and negative differentiation as described by Rozin and Royzman (2001). Several studies examine this negativity bias on the financial markets. Edmans et al. (2007) observe a strong negative stock market reaction to losses of national sport teams while no evidence of a corresponding reaction to wins. Akhtar et al. (2011) investigate the market responses to consumer sentiment announcements and document the existence of negativity bias on the Australian stock market.

Likewise, it can be expected that the market reactions related to negative and positive ESG news are asymmetric. More precisely, negative ESG news may be perceived more seriously by the market and lead to stronger reactions as compared to positive ESG news. We summarize the hypothesis as follows.

H2: The market reaction related to negative ESG news is stronger than to positive ESG news.

Last, we discuss the possible linkage between the historical ESG record and the reaction to *instant ESG news*. As mentioned above, the ESG score and *instant ESG news* are two different types of ESG information. The former can be seen as a mid- or long-term ESG record of the company, in which all of the past ESG news are aggregated. As opposed to that, the latter reflects short-term changes of the ESG performance. Previous studies indicate that low-frequency ESG performance proxies such as ESG ratings are important to investors (see e.g. Amel-Zadeh and Serafeim, 2018; Bennani et al., 2018).

To model the impact of *instant ESG news* in the light of an existing long-term ESG rating, we propose a simple adaptive model to depict how investors adapt their perception of company ESG performance to the arrival of *instant ESG news*. Considering the fact that ESG agencies often update their ESG ratings based on the aggregated ESG information since the last evaluation period (e.g. Escrig-Olmedo et al., 2019), we propose a steady adaption to the arrival of ESG news. Let, $ESG_{i,t-1}$ denote the present ESG performance figure, based on past ESG information, while $esg_{i,t}$ measures the additional ESG contribution inherent in the instant news under consideration. We regard $esg_{i,t}$ as exogenous, while its expected value can depend on the company's past ESG profile to some extent. This is because past ESG ratings may have already embedded some part of future ESG activities, and positive (negative) news are more anticipated for companies with a good (bad) ESG record (Serafeim and Yoon, 2021). The new ESG performance $ESG_{i,t}$ then results as the sum of past ESG performance $ESG_{i,t-1}$ and the ESG performance change $esg_{i,t}$ due to the news, i.e.:

$$ESG_{i,t} = ESG_{i,t-1} + esg_{i,t}\,. \tag{4.1}$$

Note that the sign of $esg_{i,t}$ is positive (negative) in case of positive (negative) ESG news, while $ESG_{i,t-1}$ can without loss of generality be assumed to lie between 0 and 100, where 100 (0) describes a perfectly sustainable (unsustainable) company. Furthermore, usually $ESG_{i,t}$ is not immediately published by the ESG score provider, however, it can be seen as the theoretical new value for an investor who considers both the old ESG score and the content value of the new instant news.

As for a company with a high ESG score it is less easy to increase its ESG score as compared to a company with a low ESG score, we consider the relative ESG performance change

$$\Delta ESG_{i,t} = \frac{esg_{i,t}}{ESG_{i,t-1}}\,. \tag{4.2}$$

Given the same value of $esg_{i,t}$, it is obvious that $\Delta ESG_{i,t}$ is higher (lower) for companies with lower $ESG_{i,t-1}$ when they encounter positive (negative) ESG news. Consequently, the market may behave differently to the same kind of instant news for companies with different past ESG ratings. If ESG performance enhances value, as claimed by H1, then the relative value can increase much more for a company with a low ESG score, while for a company with an already high ESG score positive and negative instant news with the same absolute value $|esg_{i,t}|$ will yield a lower value change. This view is suported by Glück et al. (2021), who argue that companies with a good ESG profile may face diminishing marginal benefits of ESG performance improvement, which is consistent with the over-investment view proposed by Goss and Roberts (2011). Combining the expectation argumentation that companies with a bad ESG record may enjoy even more ESG performance increase from good ESG news as such news are less anticipated and more surprising to the market, we can expect stronger market reactions for these companies. However, it is less clear regarding how differently the market may react to bad ESG news for companies with different ESG records. On the one hand, the expectation argumentation indicates that bad ESG news are less anticipated for companies with a good ESG record and thus $|esg_{i,t}|$ may be higher. On the other hand, it should be noted that companies with a good ESG profile are still perceived as doing relatively good despite the slight downgrade of ESG performance (Glück et al., 2021) due to negative ESG news. Several studies (Lins et al., 2017; Shiu and Yang, 2017; Bartov et al., 2021) show that an overall good ESG reputation can alleviate the negative impact of negative ESG events. If the latter aspect outweighs the former, we can expect that the market reacts less strongly to ESG news of companies with a good ESG record. To sum up these considerations, we state our third hypothesis as follows.

H3: The market reacts more favorably to positive ESG news of companies with a bad ESG record while less severely to negative ESG news of companies with a good ESG record.

4.4 Data description

4.4.1 The drawbacks of proprietary ESG news dataset

In earlier related studies (Krüger, 2015; Capelle-Blancard and Petit, 2019) focusing specifically on ESG news, scholars usually adopt ESG news datasets from specific ESG data providers. However, this common approach has several obvious drawbacks for studying the pricing implication of ESG news. First, this kind of specific ESG news datasets acquired directly from data providers are often characterized by a relatively small sample size, which may lead to biased empirical result. For instance, Krüger (2015) adopt a sample of 2,116 ESG events for 745 listed companies, which obviously only represent a small part of the whole stock market. Second, the way ESG data providers process text data is usually opaque. When it comes to sentiment classification of ESG news, it remains obscure to the researchers how the data providers evaluate the sentiment of ESG news. Some data providers may still rely on personal judgement to rate ESG news (Krüger, 2015). However, personal judgement could be unreliable and inconsistent over time and across involved persons. Moreover, these datasets are built by ESG data providers based on limited news resources they have, and could be less representative for the whole sample universe. For example, positive ESG news are clearly under-represented in the sample of Krüger (2015). In contrast, positive ESG news prevail in our final ESG news sample from the general news vendor Thomson Reuters Eikon.

4.4.2 Building a comprehensive ESG news dataset

In this study, we propose an alternative and general way to obtain a representative ESG news dataset which is less likely to be subject to the above problems. The original raw ESG news dataset is directly extracted from the general data provider Thomson Reuters

Eikon, which covers more than 10,000 news sources and serve as one of the most important news vendors in the world. With such a wide news coverage, it is more likely that we consider the majority of *instant ESG news*. We first build a complete list of stocks (more than 58,000 primarily quoted stocks on Eikon) traded from all over the world, and query their raw English ESG news on Eikon one by one in the period from May 2019 to March 2021. In total, we obtain a full original sample of 245,723 raw news entries tagged as ESG news by Thomson Reuters.

Before we conduct empirical study, we clean the ESG news dataset in the following steps. First, we remove those ESG news records without a complete title or article text, and exact duplicate news, identified as those with an exact title or article text as earlier news for the same company. Accordingly, 59,519 ESG news records are dropped from the sample. Second, we further remove ESG news for which we do not have enough data for conducting event study (i.e., those without stock or index price data). This cleaning procedure leads to a further reduction of 27,846 ESG news.

The way we construct the ESG news sample makes sure that it is less likely to be subject to serious selection biases. Nevertheless, while we may enjoy the benefit of a wide coverage of *instant ESG news*, another challenge arises at the same time. There are still many fuzzy duplicate news in the sample as more than one source may publish similar ESG news on different dates or at different times on the same date, which constitutes an obstacle to further empirical investigations.

4.4.3 Identifying and eliminating fuzzy duplicate ESG news

To tackle the problem of fuzzy duplicate or stale ESG news, we leverage the power of BERT-like language models. We apply the pre-trained *Sentence-BERT* model (Reimers and Gurevych, 2019) to derive sentence embeddings of ESG news titles. The *Sentence-BERT* model has already been pre-trained on Natural Language Inference (NLI) datasets SNLI and MultiNLI[5] and can produce meaningful vectors for sentences. Those sentence embeddings derived from the pre-trained model are numeric representations of ESG news titles. Therefore, news titles with similar semantic meanings should be close to each other in such a high-dimensional space.

We take the following steps to figure out fuzzy duplicate or stale news entries. First of all, we sort ESG news for the same company according to their release timestamp

[5]Data: https://nlp.stanford.edu/projects/snli/

in an ascending order. As news titles generally represent main ideas of news articles, we identify similar ESG news as those with a relatively high cosine similarity between sentence embeddings of ESG news titles. For each ESG news, we calculate the cosine similarity of its title sentence embedding and that of ESG news with an earlier timestamp and the same stock symbol. If we find any earlier ESG news which has a value of cosine similarity higher than 0.8 with the investigated ESG news, we identify the investigated ESG news as fuzzy duplicate or stale ESG news. We repeat this routine until all fuzzy duplicate ESG news are identified. Table 4.1 shows some examples to demonstrate how stale ESG news are recognized. In the end, 73,523 fuzzy duplicate news are dropped and the final sample consists of 84,835 unique and fresh ESG news.

Table 4.1: Examples of identifying fuzzy duplicate or stale ESG news

Date	Company	ESG news title	cosine similarity
2019-05-13 14:41:11	Apple	CBOE Holdings Inc. - US Supreme Court Has Ruled Against Apple In App Store Antitrust Dispute	base[1]
2019-05-13 23:12:02	Apple	iPhone owners can sue Apple over its apps, US Supreme Court decides Customers argue that company's control over the App Store is unfair	0.8517[2]
2019-11-21 13:58:04	Microsoft	Vattenfall and Microsoft pilot world's first hourly matching of renewable energy	base
2019-11-22 17:19:33	Microsoft	Sweden : Vattenfall and Microsoft pilot world's first hourly matching (24/7) of renewable energy	0.8535
2020-09-04 08:00:00	Daimler	daimler ag joins forces with terre des hommes and the responsible mica initiative to improve mica supply chains and eliminate child labour	base
2020-09-09 13:04:16	Daimler	daimler collaborates with terre des hommes and responsible mica initiative to improve mica supply chains and eliminate child labour	0.9723

[1] The earlier ESG news for the same company.
[2] The cosine similarity between the title of the base (earlier) ESG news and that of the investigated ESG news. Since the similarity is over 0.8, we remove the investigated ESG news.

4.4.4 Sentiment classification with fine-tuned BERT model

Sentiment analysis identifies the overall emotion within the text, to inspect whether the author holds a positive, neutral or negative opinion towards the event mentioned in the news article in general. In this study, our ESG news samples from Eikon are classified into three categories: positive, neutral and negative ESG news based on the classification results of a sentiment classifier. ESG news are classified as positive ESG news when the overall positive emotion or attitude such as praise and recognization is identified while classified as negative ESG news when they show negative emotion or attitude such as disappointment and criticism. Otherwise, ESG news without clear indication or direction of sentiment are classified as neutral ESG news.

Sentiment analysis has long been applied in financial studies (see e.g. Kearney and Liu, 2014; Li et al., 2014). However, most studies adopt classical dictionary-based sentiment analysis (Kearney and Liu, 2014), which is often considered as inefficient to understand texts written by humans. As far as we know, few studies apply the BERT-like language model to do semantic analysis for finance research (e.g. Araci, 2019). We are the first study to introduce the recent ground-breaking development of NLP in the field of ESG studies. For fine-tuning a ESG news sentiment classifier, we need an extra training dataset of ESG news tagged with sentiment labels. To this end, we first extract raw news records from an open-source news database called *The GDELT Project*[6]. *The GDELT Project* monitors and collects news articles from nearly every country on the planet and claims to be the largest and most comprehensive open database of human society ever created. We choose *The Global Entity Graph* (GEG), a sub database of *The GDELT Project* as our training dataset for the sentiment classifier, because of its comprehensiveness and richness.[7] Most importantly, this news database has an overall sentiment score for each news article. These news articles have already been processed by the Google nature language API and assigned with document-level sentiment scores.[8] With these sentiment scores available, we can tag news with sentiment labels.

Moreover, since our target is to classify ESG news sentiment, we explicitly focus on company ESG news in the GEG. We adopt a two-step approach to pick up company ESG news from the GEG, in which we first extract company news from the whole news universe and then extract company ESG news from the identified company news. Accordingly, we train two other BERT-like classifiers (BERT model I and II) which can tell whether news are company news and whether company news are ESG related. For fine-tuning the first classifier, we collect 20,000 company news directly on Eikon and 20,000 non-company news from another sub database of *The GDELT Project*, i.e., the GDELT Event Database (GED).[9] The GED provides news entries in which type of event and major event participants have been identified. We remove those news with participant types identified as BUS and MNCs[10] and take the rest as non-company news. For fine-tuning the second classifier, we focus on company news exclusively extracted from Eikon. We collect 20,000 ESG news and 20,000 non-ESG news by changing the query criterium on

[6]See: https://www.gdeltproject.org

[7]On average, there are more than 100,000 news collected by the GEG for a single day.

[8]For more information about the sentiment score, see: https://cloud.google.com/natural-language/docs/basics

[9]See: https://www.gdeltproject.org/data.html

[10]BUS: businessmen, companies, and enterprises, not including MNCs. MNC: multi-national corporations.

Eikon. These two classifiers show the ability (with an accuracy of 99% on the evaluation datasets) to identify whether general news are company news, and whether company news are ESG related (see BERT model I and II in Table 4.2). With these two additional classifiers, we are able to extract explicitly company ESG news from the massive news sample of the GEG database. We scan over 38 million news[11] of the GEG published in 2020 and identify 0.66 million company news[12] with the first classifier, from which we identify 50,332 company ESG news using the second classifier.

At last, we tag each company ESG news extracted from the GEG according to their overall sentiment scores. For ESG news with an overall sentiment score not lower than 0.2, we label them as positive and those with an overall sentiment score not higher than -0.2 as negative. The rest of news in the sample are labelled as neural ESG news. We summarize all the additional news datasets and how we derive a labelled ESG news dataset as described above in Figure 4.1.

Table 4.2: BERT models for classifying news types and news sentiment

	BERT model I	BERT model II	BERT model III
Description	distinguish between company news and non-company news in general news sample	distinguish between company ESG news and non-ESG news	identify ESG news sentiment (negative, neutral, positive)
Data	20,000 non-company news from the GDELT Event Database and 20,000 company news from the Thomson Reuters Eikon	20,000 company ESG news and 20,000 company non-ESG news from Thomson Reuters Eikon	50,332 ESG news extracted from the GDELT GEG database with the help of BERT model I and II, in which 5,667 are labelled as negative ESG news, 29,862 as neutral ESG news and the rest 14,803 as positive ESG news
Accuracy rate[1]	99%	99%	81%
Max. length	128 word pieces	512 word pieces	512 word pieces
Model variant	RoBERTa	RoBERTa	RoBERTa
Key parameters	training ratio 0.8[2], training epoch 3, learning rate 2e-5, batch size 8		

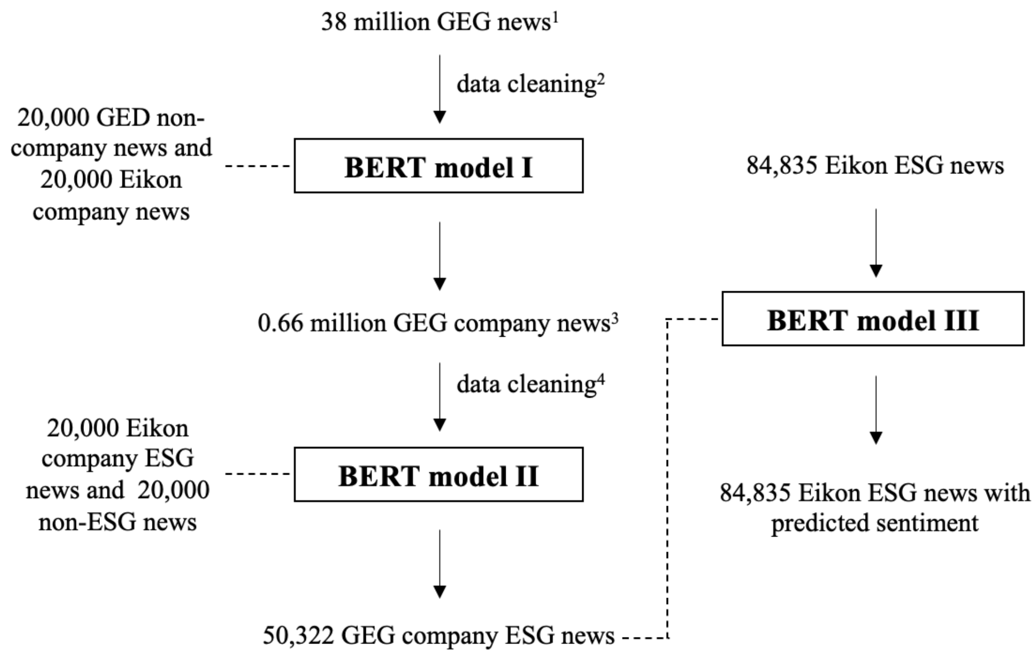
[1] Accuracy rate on the evaluation dataset.
[2] News dataset is split into a training dataset (80%) and an evaluation dataset (20%).

Given the sentiment labelled ESG news, we can finally fine-tune a BERT-like model (BERT model III) to identify ESG news sentiment. We choose a maximum possible text length of 512 word pieces, which means that news articles more than 512 word pieces will be truncated. For more detail regarding the model, please refer to Table 4.2 and Figure 4.2. With an accuracy rate of 81% on the evaluation set, our fine-tuned BERT

[11] The GEG does not provides original text information. We extract news titles from news URLs and use them as inputs for the first classifier. Therefore, we drop those news entries from which we fail to extract news titles.

[12] We try to scrape down original article texts for these company news and use them as inputs for the second classifier. If it fails, we drop these news.

Figure 4.1: How a sentiment labelled ESG news dataset is derived

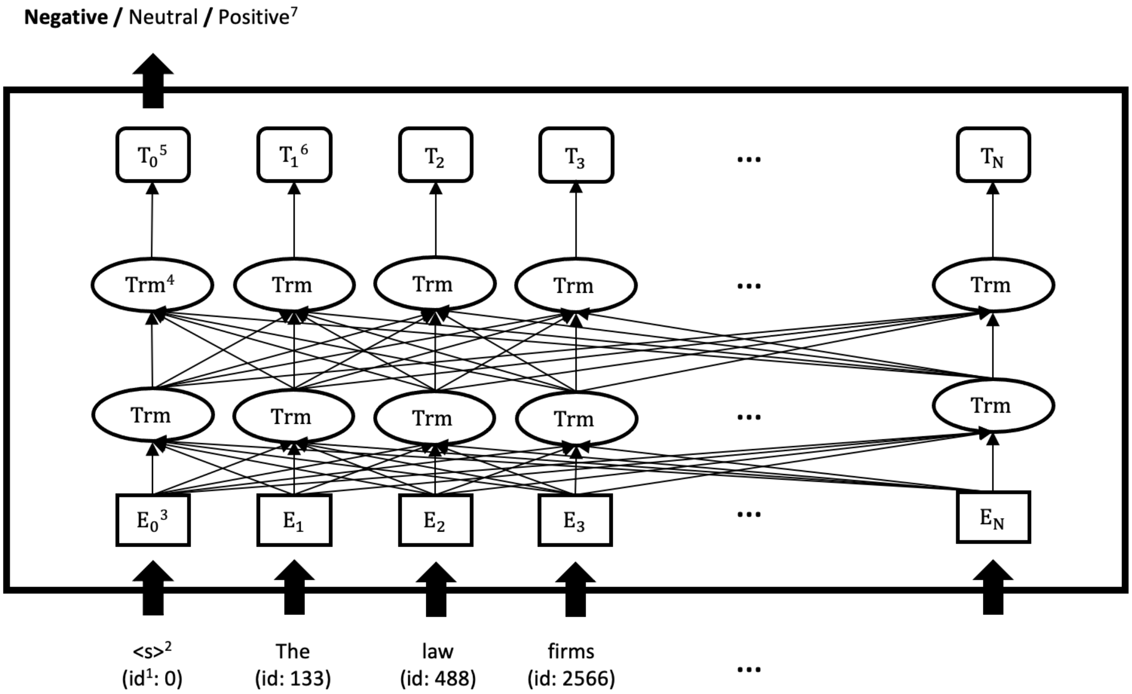


The BERT model I and II are designed to pick up company ESG news from the GEG news universe. These company ESG news with sentiment labels are used to fine-tune a ESG news sentiment classifier, i.e. BERT model III, which is supposed to help us identify Eikon ESG news sentiment. Arrows indicate the inference process and dash lines refer to corresponding training and evaluation datasets.
[1] News titles are extracted from the given URLs and used as model input.
[2] News without a title or exact duplicate news are removed.
[3] Main article texts are scraped down from the corresponding servers and used as model input.
[4] News without a main article text or exact duplicate news are removed.

model III, for most of the time, is able to determine the overall sentiment direction of company ESG news. In fact, this accuracy rate is quite satisfying, especially given the fact that the text input is relatively long (i.e., 512 word pieces) and there exists three sentiment labels instead of two with only negative or positive sentiment[13]. Da Silva et al. (2014) review many studies applying classical machine learning models which aim to classify tweets (relatively short texts) into positive or negative ones (only two labels) and document that most of the time the accuracy rates of these models are lower than 80%. Therefore, we are confident that this sentiment classifier can provide satisfying classification results and differentiate ESG news with different sentiment.[14] We feed all unique Eikon ESG news into BERT model III and classify them into positive, neutral and negative ESG news. See some representative ESG news entries with different sentiment in Table 4.3.

[13]Models with three labels are more difficult to achieve better results than those with only two labels, as they need to extract more useful information from data to distinguish among three classes.
[14]Even when humans are asked to do the classification task, it is possible that they may be inconsistent in judging news sentiment and have different opinions.

Figure 4.2: BERT sentiment classifier: How ESG news is scored



This figure shows the main architecture of the BERT model and how ESG news is classified by the fine-tuned BERT sentiment classifier. The example ESG news is the last ESG news for BMW in Table 4.3: "The law firms of Waddell Phillips Professional Corporation and Podrebarac Barristers Professional Corporation announced today that ...". The maximum text input length is 512 word pieces. News with longer text length are truncated.
[1] id is the word piece ID in the adopted vocabulary.
[2] <s> is a special symbol indicating the beginning of text inputs.
[3] The embedding (a 768-dimensional vector) of the input token <s>. For detailed implementation of the embedding procedure, please refer to Devlin et al. (2018).
[4] The intermediate representations of the token in the *Transformer* encoder layers (see Vaswani et al., 2017).
[5] The output of the special symbol (i.e., <s>) after the *Transformer* encoder layers.
[6] The output of the corresponding input of a word piece after the *Transformer* encoder layers.
[7] The final sentiment label is identified by converting the previous output (corresponding to T_0 in the figure) into probabilities through a Softmax output layer and choosing the highest one.

4.4.5 Basic descriptive statistics

In total, the final ESG news sample contains 84,835 ESG news from 13,327 listed companies from all over the world. In Table 4.4, we show where ESG news originate from. More than half of the ESG news final sample comes from America, while around 27% from Europe. Asia and Oceania also have a share of 16% and 3%, respectively. Moreover, we show the number of ESG news for each top 20 countries (regions) in the full sample in Table 4.13. USA has the biggest share of 42% of the overall sample, followed by Canada with 11% and UK with 9%. Our sample covers almost every corner of the world and should be representative to study the pricing implication of ESG news. As regards sector distribution, the top five sectors are Industrials (18%), Information Technology (12%), Materials (12%), Financials (11%) and Consumer Discretionary (10%), according to the Global Industry Classification Standard (GICS).

Table 4.3: Examples of ESG news sentiment classification

Date	Company	ESG news text	sentiment
2019-06-05 07:35:29	BMW	As it develops its plans for the mobility of the future, the BMW Group is increasingly focusing on co-operations to help make next-level electrification technology more widely available to customers by the start of the coming decade...	positive
2019-09-14 20:19:16	BMW	BERLIN, Sept 14 (Reuters) - BMW's engine development and purchasing expert, Markus Duesmann, is set to become the CEO of Volkswagen's Audi premium brand....	neutral
2020-04-07 16:17:54	BMW	The law firms of Waddell Phillips Professional Corporation and Podrebarac Barristers Professional Corporation announced today that the Ontario Superior Court of Justice has certified a national class action against luxury automaker Bayerische Motoren Werke AG....	negative
2020-03-24 15:26:23	Dow, Inc	Mar 24, 2020. Dow Inc. introduced two innovations that simplify the formulation of water-based, high temperature-resistant industrial coatings...	positive
2020-06-17 12:35:54	Dow, Inc	Jun 17, 2020. Dow Inc. inked a joint development deal with Shell to speed up the development of technology that can electrify ethylene steam crackers....	neutral
2020-05-21 03:06:51	Dow, Inc	Catastrophic flooding triggered by dam failures in Michigan could potentially release toxic pollution from a site contaminated by the industrial giant Dow Chemical.Dow's facility in Midland, Michigan, where the company is headquartered along the Tittabawassee River, manufactured chlorine-based products beginning in the early 1900s...	negative

In Table 4.5, we provide more basic descriptive statistics for company level features. Note that these company basic features are from the previous year end for each ESG news. Overall, ESG news have an average company asset of 69.7 billion USD. Only 50,722 out of 84,835 ESG news are paired with an Eikon ESG score. On average, ESG news have an ESG score of 59.

Table 4.4: Descriptive statistics for categorical variables

	obs.	%	stocks	%		obs.	%	stocks	%
Continent									
America	45,688	53.86	5,085	38.16	Consumer Discretionary	8,371	9.87	1,406	10.55
Europe	22,926	27.02	3,387	25.41	Consumer Staples	4,129	4.87	666	5.00
Asia	13,269	15.64	4,033	30.26	Health Care	7,137	8.41	1,486	11.15
Oceania	2,275	2.68	563	4.22	Financials	9,499	11.20	1,527	11.46
Africa	677	0.80	259	1.94	Information Technology	10,077	11.88	1,457	10.93
					Communication Services	3,521	4.15	606	4.55
Sector					Utilities	7,038	8.30	406	3.05
Energy	6,394	7.54	742	5.57	Real Estate	2,561	3.02	746	5.62
Materials	9,853	11.61	1,701	12.76	None[1]	1,092	1.29	442	3.32
Industrials	15,161	17.87	2,142	16.07					

[1] None means there is no GICS sector classification.

Moreover, we show the sentiment distribution of ESG news in Table 4.6. Overall, 44% ESG news are classified as positive news by our sentiment classifier, while only 5% as negative news. The remaining half of ESG news is classified as neutral ESG news, which means that there is no clear positive or negative sentiment direction revealed in texts in our context. The sentiment distributions of ESG news for America, Europe and Asia are

Table 4.5: Descriptive statistics for metric variables

	Obs.	Mean	Std.	Min	Median	Max
asset[1] (in million USD)	82,571	69,700	248,000	0.01	6,020	4,320,000
esg[1]	50,722	59.08	21.51	0.92	63.15	94.47
num_news[2]	84,835	32.00	41.53	1	16	302

[1] *asset* and *esg* are firm-level data at the end of previous year of the event date.
[2] *num_news* indicates how many pieces of ESG news are released for the same company during the sample period.

similar. Negative ESG news contribute to only 4% to 6% of the corresponding continent subsamples, except for Oceania (16%).

Table 4.6: Sentiment distribution of ESG news

Continent	Obs.	negative	%	neutral	%	positive	%
America	45,688	2,147	4.70	23,171	50.72	20,370	44.59
Europe	22,926	854	3.73	11,576	50.49	10,496	45.78
Asia	13,269	828	6.24	6,711	50.58	5,730	43.18
Oceania	2,275	355	15.60	1,265	55.60	655	28.79
Africa	677	37	5.47	479	70.75	161	23.78
Total	84,835	4,221	4.98	43,202	50.92	37,412	44.10

This table presents the ESG news sentiment distribution across different continents. In toal, there are 84,835 unique and fresh ESG news for 13,327 stocks.

Unlike other ESG news samples adopted in related studies (e.g. Krüger, 2015; Capelle-Blancard and Petit, 2019), our sample is constructed based on massive raw ESG news from comprehensive sources from all over the world. Therefore, it is much more presentative and should reflect how company ESG issues are reported as a whole. It can be said that in general news media prefer to report positive ESG issues rather than negative ones. Given the sentiment classification result of the three groups of ESG news, we investigate whether there is stock performance differences among them and what are the possible determinants.

4.5 Empirical methodology

4.5.1 Event study

In order to examine the pricing implication of ESG news, we conduct event study for each ESG news. We define the day when ESG news is released as the event day T_0, and choose an event window which covers a period several days before and after the event day, i.e., from $T_0 - \tau$ to $T_0 + \tau$. For each day u in the event window, we calculate daily log-returns for ESG news i as

$$r_{i,u} = \ln p_{i,u} - \ln p_{i,u-1} \,. \tag{4.3}$$

where u is the event window days relative to the event day T_0. Next, we calculate abnormal returns for ESG news i by estimating the market model as

$$r_{i,t} = \alpha + \beta R_{i,t} \tag{4.4}$$

where $R_{i,t}$ is the daily return of the corresponding stock index[15]. We adopt an estimation period of 200 trading days which has a distance of 50 trading days to the event date. Accordingly, daily abnormal return for each ESG news event can be calculated as

$$ar_{i,u} = r_{i,u} - \hat{\alpha} - \hat{\beta} R_{i,u} \tag{4.5}$$

where $\hat{\alpha}$ and $\hat{\beta}$ are estimated coefficients from the market model in Equation (4.4). Moreover, cumulative abnormal returns are defined by

$$CAR_{i,\tau} = \sum_{u=T_0-\tau}^{T_0+\tau} ar_{i,u} \tag{4.6}$$

where $2\tau + 1$ is the length of the ESG news event window.

The test of the statistic significance of stock performance is often based on the following t-statistic:

$$t_\tau = \frac{\overline{CAR_\tau}}{\sqrt{var[\overline{CAR_\tau}]}} \tag{4.7}$$

where $\overline{CAR_\tau}$ is the average of the cumulative abnormal returns across the same type of events. However, $var[\overline{CAR_\tau}]$ should be estimated with caution. Kolari and Pynnönen (2010) find that cross-sectional correlations among abnormal returns in the case of event-date clustering with the same event window may lead to biased standard tests and therefore should be considered when designing the t-statistic. In our case, we have ESG news events across many stocks and over a more than 1.5 year timeframe. Some ESG news concern the same company and event windows may partly overlap with each other. Therefore, the corresponding cumulative abnormal returns may be subject to correlation. To address this concern, we adopt the cross-sectional and time serial correlation robust $var[\overline{CAR_\tau}]$ proposed by Kolari et al. (2018). Kolari et al. (2018) consider both cross-sectional and time serial correlation when estimating $var[\overline{CAR_\tau}]$ by grouping abnormal returns in both cross-sectional and time dimensions:

$$var[\overline{CAR_\tau}] = \frac{1}{n^2}\left(\sum_{i=1}^{n} var[CAR_{i\tau}] + \sum_{t=1}^{T} var[AR_t] - \sum_{i=1}^{n}\sum_{u=\tau_1}^{\tau_2} var[ar_{iu}]\right) \tag{4.8}$$

[15]We choose the major stock index for each country.

where n is the number of events, T is the number of calendar days covered by any ESG news event for the whole sample, and AR_t is the aggregated abnormal returns on the calendar day t. The first term $\frac{1}{n^2}\sum_{i=1}^{n} var[CAR_{i\tau}]$ itself equals to $var[\overline{CAR_\tau}]$ under the assumption that events are independent, and can be consistently estimated by

$$v\hat{a}r_a[\overline{CAR_\tau}] = \frac{1}{n^2}\sum_{i=1}^{n}(CAR_{i\tau} - \overline{CAR_\tau})^2 \,. \tag{4.9}$$

The second term $\frac{1}{n^2}\sum_{t=1}^{T} var[AR_t]$ itself also equals to $var[\overline{CAR_\tau}]$ under the assumption of serial independence, and can be consistently estimated by

$$v\hat{a}r_b[\overline{CAR_\tau}] = \frac{1}{n^2}\sum_{t=1}^{T}(AR_t - \overline{AR})^2 \tag{4.10}$$

where $AR_t = \sum_{ar_{iu}\in D_t} ar_{iu}$ (D_t denotes the set of all ar_{iu} on the same calendar day t). The sum of the first and second term embeds both serial correlation and cross-sectional correlation terms and thus is serial and cross-section correlation robust. However, it double counts the individual variances $var[ar_{iu}]$. Therefore, we subtract the third term from the sum of the first and second term to get the robust $var[\overline{CAR_\tau}]$. The third term $\frac{1}{n^2}\sum_{i=1}^{n}\sum_{u=\tau_1}^{\tau_2} var[ar_{iu}]$ is estimated by

$$v\hat{a}r_{ar}[\overline{CAR_\tau}] = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{u=\tau_1}^{\tau_2}(ar_{iu} - \overline{ar})^2 \,. \tag{4.11}$$

Moreover, besides the significance test for the same group of ESG news, we also test whether the mean difference of stock performance between the positive group and the negative group is statistically significant. Accordingly, we adopt the following t statistic

$$t_d = \frac{\overline{CAR_{pos}} - \overline{CAR_{neg}}}{\sqrt{var[\overline{CAR_{pos}}] + var[\overline{CAR_{neg}}]}} \tag{4.12}$$

where both $var[\overline{CAR_{pos}}]$ and $var[\overline{CAR_{neg}}]$ are estimated as described in Equation 4.8.

One concern of event studies is that the empirical results may be driven by confounding events. In our context this means that synchronuous non-ESG news could have an impact on the financial markets and thus could blur the real influence of ESG news. However, we regard this as very unlikely for the following reason. As mentioned in Section 4.4, our dataset is very comprehensive and covers most ESG news in the observation period for more than 10,000 listed companies from all over the world. If the empirical results were driven by confounding events, the non-ESG news would need to be aligned with the ESG news in a systematic way. More precisely, positive (negative) ESG news would need to be systematically accompanied by positive (negative) non-ESG news from the same company—published close to the announcement date. However, while such a news disclosure behavior is unlikely but possible for any arbitrary company, there is no reason to assume that even thousands of companies behave in the same manner. For this reason, we hold that non-ESG news within the event window are diverse in nature and sentiment and therefore their influence on the results cancels out within our large samples. With other words, confunding events can be a problem for event studies with a relatively small number of events and a small number of different stocks. None of both is the case here.

4.5.2 Regressions

Apart from event study, we regress stock performance, measured by abnormal returns on the event day or cumulative abnormal returns of different event windows, on several independent variables to investigate whether news sentiment is a key determinant of stock performance. Moreover, we are interested in whether the past ESG ratings as assigned by ESG raters such as Thomson Reuters may have an impact on stock performance when *instant ESG news* is released. The regression setup is as follows:

$$R_i = \beta_0 sentiment_i + \beta_1 esg_i + \beta_2 sentiment_i \cdot esg_i + \beta_3 controls_i + e_i \tag{4.13}$$

where R_i represents stock performance, measured by abnormal returns ar_0 on the event day T_0, or by cumulative abnormal returns CAR_1 and CAR_2. The variable $sentiment_i$ represents the overall ESG news sentiment, i.e., positive, neutral and negative sentiment, as predicted by our fine-tuned BERT model III. The variable esg_i is the Eikon ESG score for the company under investigation. We include interaction terms between $sentiment_i$ and esg_i to further test whether their impact on stock performance are intertwined, as predicted by H3. As regards control variables $controls_i$, we have the following setups. To control for the possible size effect, we include the variable $asset_i$ in regressions. We also add num_news_i, which indicates the number of ESG news for the same company in the sample period to control difference in media exposure. We further add $sector_i$ and $continent_i$ to control for sector and geographic differences. For detailed explanation of variable definitions, please refer to Table 4.14.

4.6 Results

4.6.1 Event study results from the overall sample

We show descriptive statistics of stock performance as abnormal return on the event day T_0 and cumulative abnormal returns during different sizes of event windows for each group of ESG news in Table 4.7. Note that we adopt robust t-statistic to test whether stock performance is significantly different from zero as described in Section 4.5. On average, the positive group shows a significant 0.31% average abnormal return while the negative group has a significant -0.75% average abnormal return on the event day. The neutral group has a relatively smaller scale of average abnormal return of 0.20% on the even day T_0. The univariate analysis on the event day provides evidence that positive ESG news is associated with outperformance while negative ones may lead to underperformance, especially on the event day. Moreover, we observe that the market reactions to positive and negative may be asymmetric. This provides first evidence supporting H1 and H2. When stock performance is evaluated by CAR_1, positive ESG news lead to an average cumulative abnormal return of 1.17% while the negative group suffer from a significant loss of -1.28%. Again, neutral ESG news show a smaller average cumulative abnormal return. When we further expand the window size, i.e., change CAR_1 to CAR_2 and CAR_5, we get similar result patterns but do not see more obvious performance difference. For CAR_{10}, only the positive group has a significant mean cumulative abnormal return of 1.24%.

Next, we show the average abnormal returns across all ESG news for the whole event window in Figure 4.3. The difference between the negative group and the other two groups is obvious. The stock performance of the negative group is most significantly negative on the event day and one day before. For the positive group, we observe notably positive abnormal return only on the event day. In contrast, the neutral group shows more mild performance throughout the whole event window. In Figure 4.4, we show cumulative abnormal returns. The performance difference between the negative group and the other two groups is evident. The difference between the positive group and the neutral group only becomes more clear on the event day and thereafter.

Table 4.7: (Cumulative) abnormal returns

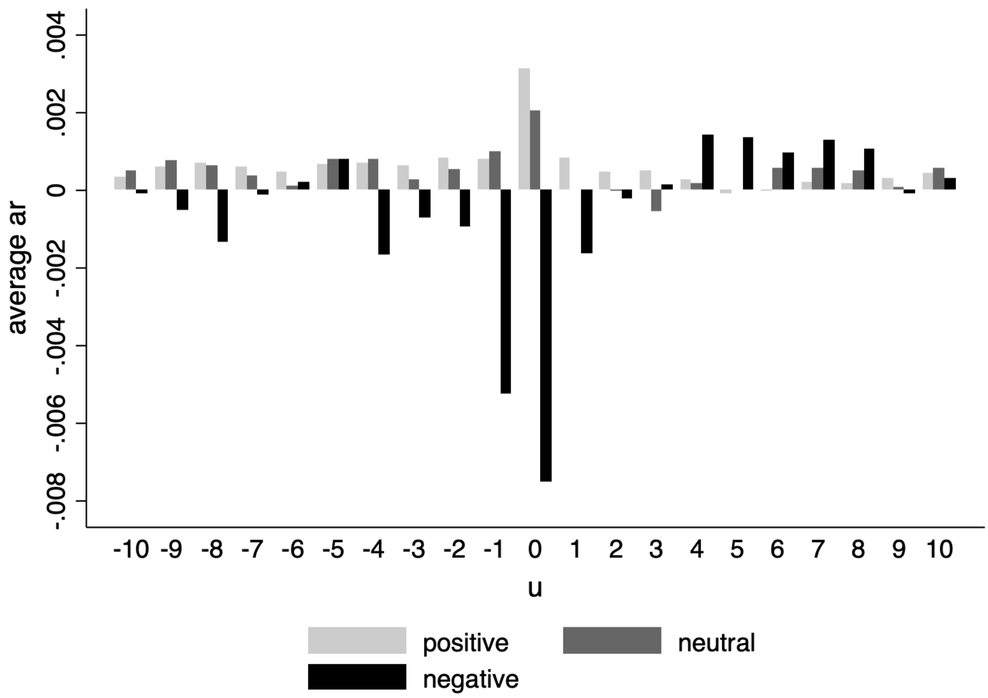
Group	Obs.	Mean (%)	S.D. (%)	t_τ	Min (%)	Med (%)	Max (%)	Diff.[1]
ar_0								
positive	37,412	0.31***	4.30	8.36	−62.65	0.05	165.16	
neutral	43,202	0.20***	6.96	3.02	−152.95	0.04	262.61	***
negative	4,221	−0.75***	7.03	−4.25	−119.23	−0.19	53.32	
CAR_1								
positive	37,412	1.17***	14.99	11.44	−173.88	0.41	225.16	
neutral	43,202	0.87***	20.14	5.24	−375.77	0.27	237.69	***
negative	4,221	−1.28***	21.89	−3.24	−341.51	−0.63	154.03	
CAR_2								
positive	37,412	1.24***	15.64	9.36	−196.96	0.46	242.88	
neutral	43,202	0.97***	20.87	4.24	−365.50	0.28	215.70	***
negative	4,221	−1.26***	22.66	−2.72	−358.13	−0.54	135.72	
CAR_5								
positive	37,412	0.87***	11.34	3.73	−212.76	0.26	188.84	
neutral	43,202	0.51	16.28	1.28	−361.99	0.10	231.45	***
negative	4,221	−1.42**	19.80	−2.26	−414.32	−0.53	155.64	
CAR_{10}								
positive	37,412	1.24***	15.64	2.89	−196.96	0.46	242.88	
neutral	43,202	0.97	20.87	1.53	−365.50	0.28	215.70	***
negative	4,211	−1.26	22.66	−1.51	−358.13	−0.54	135.72	

This table presents descriptive statistics of abnormal returns on the event day or cumulative abnormal returns during the event window. In total, there are 84,835 unique and fresh ESG news for 13,327 stocks. $^*p < .1$, $^{**}p < .05$, *** $p < .01$.
[1] Test results (t-statistic: t_d as described in Section 4.5) indicating whether the mean of the positive group is statistically different from that of the negative group

4.6.2 Event study results from the America subsample

The America subsample contributes 54% of the overall sample and thus is our main focus in this study. Overall, the America subsample shows a similar or even more clear picture. Table 4.8 presents the average one-day abnormal return on the event day and cumulative abnormal returns for the whole event windows. In the America subsample, the difference among different ESG news groups is more evident. The positive group may enjoy an average abnormal return of 0.37% on the event day, while the negative group is associated with a stronger and negative abnormal return of -1.01%. When we look at the cumulative abnormal returns for different sizes of event windows (CAR_1, CAR_2 and CAR_5, we see significant difference between the positive and negative groups. Specifically, the negative group suffers an average cumulative abnormal return of -2.10% and the positive group enjoys 1.38% over a three-day event window. Again, we find evidence supporting H1 and H2, which state that stock performance of ESG news is related to the news sentiment and stock performance is asymmetric for positive and negative ESG news. Figure 4.5 and Figure 4.6 show daily and cumulative abnormal returns during the event window for the America subsample.

Figure 4.3: Abnormal returns during ($T_0 - 10$, $T_0 + 10$)



4.6.3 Event study results from the Europe subsample

Besides the America subsample, we take a closer look at the Europe subsample. With a share of 27%, it is the second largest subsample. We examine the Europe subsample with special care, also because English is popular or often official language for European countries. As always, we investigate stock performance of the three groups of ESG news in terms of abnormal return on the event day and cumulative abnormal returns for the whole event window (see Table 4.9). The positive group enjoys a significant average abnormal return of 0.34%. In contrast, the negative group is associated with a significant negative average abnormal return of -0.78%. When stock performance is measured over a small event window (CAR_1), the positive group enjoys 1.16% average cumulative abnormal returns while the negative group suffers from a mean loss of -1.82%.

Following the same examination procedure, we present daily and cumulative abnormal returns for the whole event window in Figure 4.7 and Figure 4.8. The results for the Europe subsample show a very similar pattern as the America subsample. Apart from the distinct difference between the negative and the other two groups, the difference between the positive and neutral groups is more observable.

Overall, we find evidence in favor of H1 and H2 not only for the overall sample, but also

Table 4.8: (Cumulative) abnormal returns: the America subsample

Group	Obs.	Mean (%)	S.D. (%)	t_τ	Min (%)	Med (%)	Max (%)	Diff.[1]
ar_0								
positive	20,370	0.37***	4.84	6.45	−62.65	0.08	165.16	
neutral	23,171	0.21*	8.05	1.92	−152.95	0.04	262.61	***
negative	2,147	−1.01***	8.00	−3.70	−119.23	−0.33	53.32	
CAR_1								
positive	20,370	1.38***	16.84	8.81	−149.62	0.46	216.51	
neutral	23,171	0.88***	22.64	3.28	−199.69	0.29	219.05	***
negative	2,147	−2.10***	22.71	−3.60	−188.76	−1.11	154.03	
CAR_2								
positive	20,370	1.46***	17.48	7.19	−162.49	0.49	242.88	
neutral	23,171	1.05***	23.34	2.72	−190.68	0.33	215.70	***
negative	2,147	−2.07***	23.80	−2.92	−188.09	−0.98	135.72	
CAR_5								
positive	20,370	1.00***	12.71	2.80	−122.81	0.30	188.84	
neutral	23,171	0.40	18.19	0.59	−199.82	0.11	231.45	***
negative	2,147	−2.14**	20.06	−2.22	−199.82	−0.85	155.64	
CAR_{10}								
positive	20,370	1.46**	17.48	2.20	−162.49	0.49	242.88	
neutral	23,171	1.05	23.34	1.01	−190.68	0.33	215.70	**
negative	2,147	−2.07	23.80	−1.49	−188.09	−0.98	135.72	

This table presents descriptive statistics of abnormal returns on the event day or cumulative abnormal returns during the event window for the America subsample. In the America subsample, there are 45,688 unique and fresh ESG news for 5,085 stocks. $^*p < .1$, $^{**}p < .05$, $^{***}\ p < .01$.

[1] Test results (t-statistic: t_d as described in Section 4.5) indicating whether the mean of the positive group is statistically different from that of the negative group

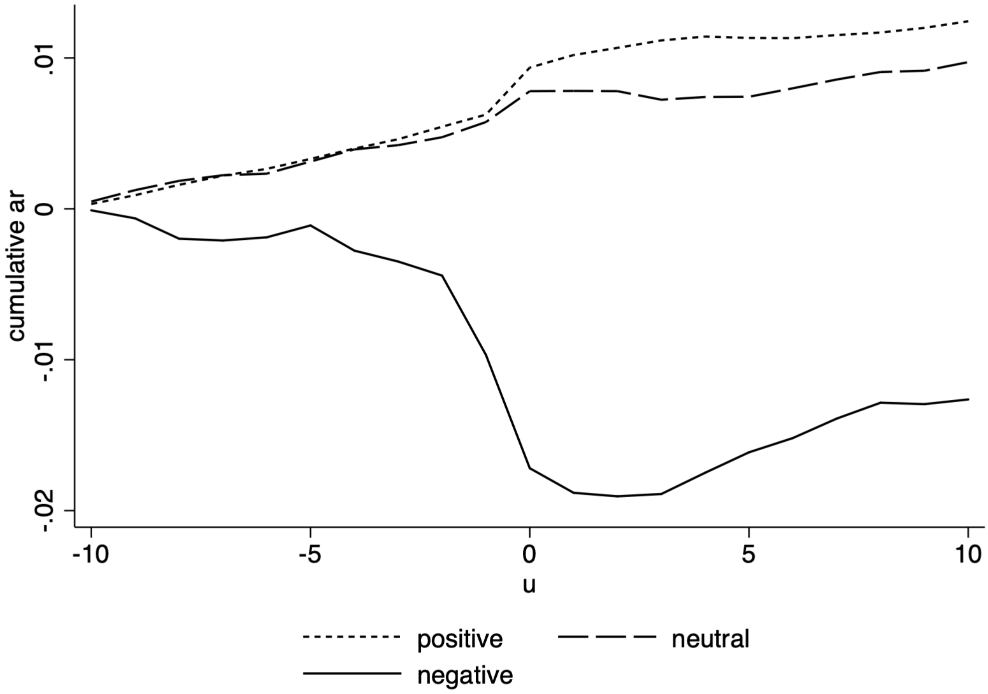
Table 4.9: (Cumulative) abnormal returns: the Europe subsample

Group	Obs.	Mean (%)	S.D. (%)	t_τ	Min (%)	Med (%)	Max (%)	Diff.[1]
ar_0								
positive	10,496	0.34***	3.91	6.10	−60.78	0.06	142.26	
neutral	11,576	0.19**	5.79	2.23	−143.91	0.05	90.93	***
negative	854	−0.78%**	6.33	−2.46	−61.96	−0.16	21.55	
CAR_1								
positive	10,496	1.16***	12.40	8.72	−128.03	0.62	225.16	
neutral	11,576	0.45**	17.40	2.36	−375.77	0.21	237.69	***
negative	854	−1.82**	24.98	−1.96	−341.51	−0.23	83.19	
CAR_2								
positive	10,496	1.25***	12.96	8.19	−140.59	0.59	195.43	
neutral	11,576	0.44*	18.25	1.91	−365.50	0.18	210.15	***
negative	854	−1.54	25.87	−1.51	−358.13	0.10	87.04	
CAR_5								
positive	10,496	0.83***	9.43	4.44	−105.73	0.33	141.58	
neutral	11,576	0.40	14.54	1.16	−361.99	0.10	176.54	**
negative	854	−2.13*	24.90	−1.75	−414.32	−0.14	70.40	
CAR_{10}								
positive	10,496	1.25***	12.96	3.73	−140.59	0.59	195.43	
neutral	11,576	0.44	18.25	0.73	−365.50	0.18	210.15	**
negative	854	−1.54	25.87	−1.09	−358.13	0.10	87.04	

This table presents descriptive statistics of abnormal returns on the event day or cumulative abnormal returns during the event window for the Europe subsample. In the Europe subsample, there are 22,926 unique and fresh ESG news for 3,387 stocks. $^*p < .1$, $^{**}p < .05$, $^{***}\ p < .01$.

[1] Test results (t-statistic: t_d as described in Section 4.5) indicating whether the mean of the positive group is statistically different from that of the negative group

Figure 4.4: Cumulative abnormal returns during $(T_0 - 10, T_0 + 10)$



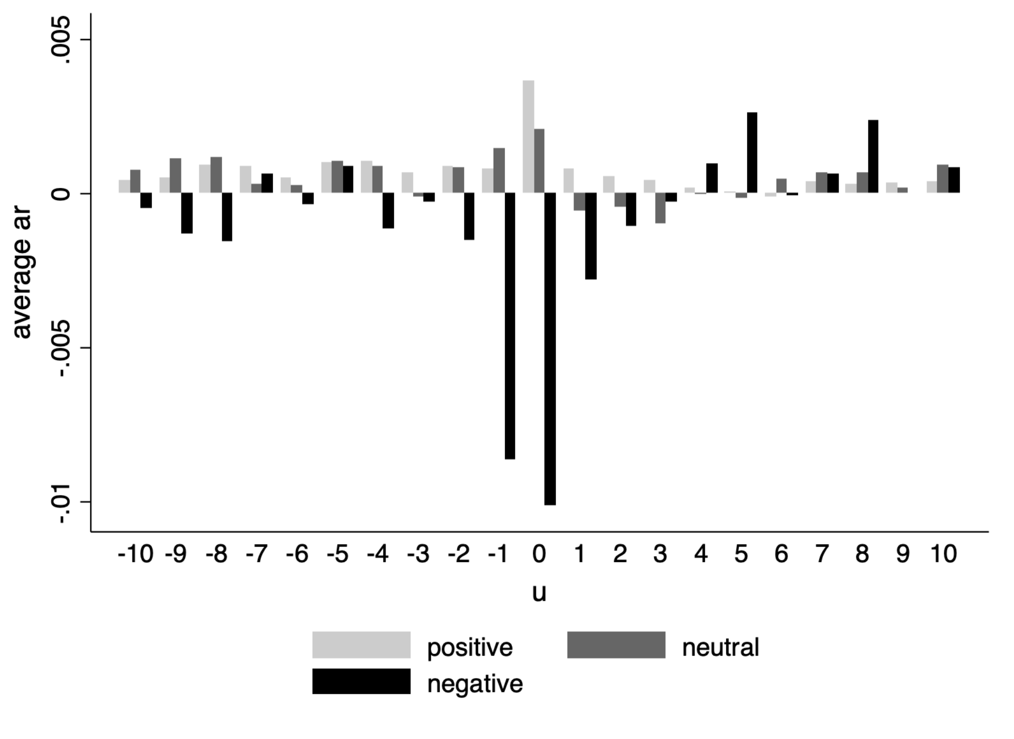
for the America and Europe subsamples.

4.6.4 Regression results

Besides event study, we run multiple linear regressions to investigate possible determinants of stock performance related to ESG news. We regress stock performance on ESG news sentiment, ESG score, interaction terms between sentiment and ESG score, and some other control variables. Note that we choose cluster-robust standard errors at the company level in all regressions. Table 4.10 shows the regression results for the overall sample. We regress stock performance, i.e., ar_0, CAR_1, and CAR_2 on possible determinants and controls. In model I, III and V, we include only the categorical sentiment variable *sentiment* and controls. In model II, IV and VI we add interaction terms between *sentiment* and *esg* to check whether the influence of ESG news sentiment depends on the past ESG reputation of the target company.

Overall, we find evidence that whether ESG news sentiment is negative or positive has a significant effect on stock performance, which is in favor of H1. First of all, we find that the coefficient of *negative* is significantly negative across different model setups, which means that the release of negative ESG news has a noticeable and negative impact on stock

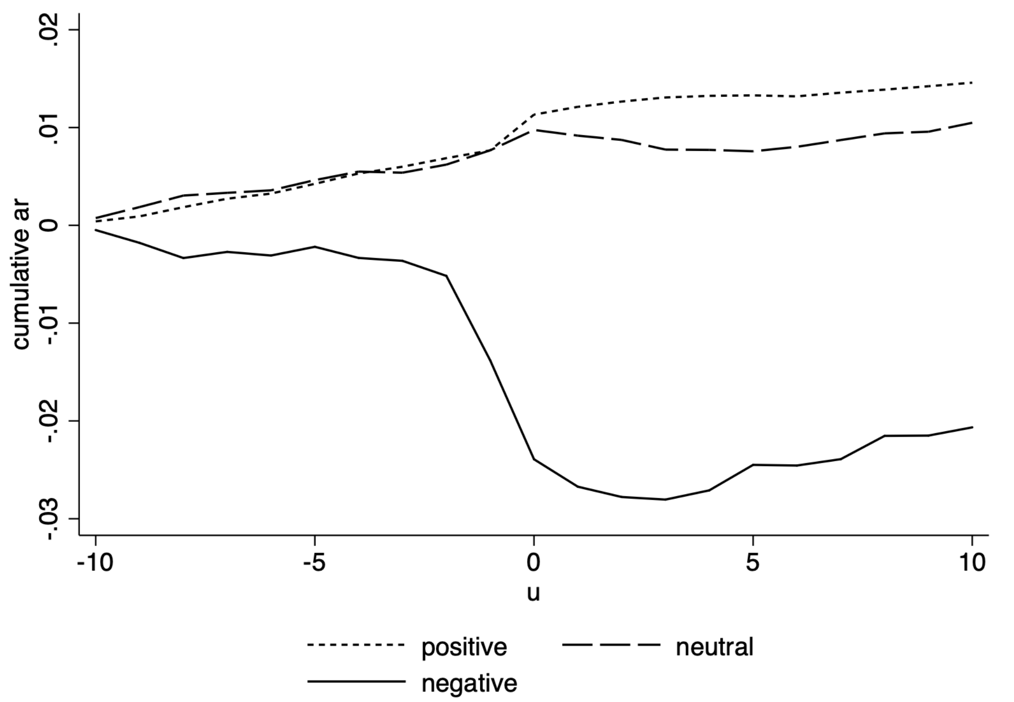
Figure 4.5: Abnormal returns during ($T_0 - 10$, $T_0 + 10$): the America subsample



performance, as compared to neutral ESG news. This indicates that negative ESG news are perceived seriously and priced by investors on stock markets. Regarding positive ESG news, we can observe significant and positive coefficients of *positive* in all models. This is evidence that positive ESG news are digested in a positive way on financial markets. Despite the fact that positive ESG news prevails, they are still positively perceived by investors. Nevertheless, when compared with *negative*, *positive* has obviously smaller coefficients across different models and thus the impact of positive ESG news may be lower than that of negative ESG news. This provides some support for H2.

When interaction terms between *sentiment* and *esg* are added, we gain more insights into market reactions to ESG news under different conditions. Interestingly, the coefficients of the interaction term *negative*esg* are significantly positive in model II, IV and VI. One possible explanation is that the past ESG record of the company may play a role in the impact of negative ESG news on stock performance. If a company has a good ESG record, the negative impact of negative ESG news could be dampened. Therefore, even though a company may suffer from bad stock performance when bad ESG news is released, a good historical ESG image may help relieve the problem. We also observe the significantly negative coefficient of *positive*esg* in model II and IV when the one-day performance ar_0 and three-day performance CAR_1 are taken as the dependent variable. It could be possible that when positive ESG news is released for a company with a bad ESG record, investors react more favorably since the company performs marginally better

Figure 4.6: Cumulative abnormal returns during (T_0-10, T_0+10): the America subsample



in ESG issues. Overall, our regression results suggest that stock performance related to ESG news depends not only on the news sentiment, but also on the historical ESG record. Therefore, H3 is also supported by our empirical results.

Similarly, we run the same regression routine for the America subsample. The regression results are reported in Table 4.11. Just like in the overall sample, we find that negative ESG news tend to have significantly negative influence on stock performance, regardless of different model setups. We also find that the coefficients of *positive* are positive and significant, which indicates that investors react positively to positive ESG news. Again, negative ESG news appear to be taken more seriously than positive ESG news as the scale of the coefficients is larger. Moreover, we observe that the interaction term *negative*esg* is positive and significant in different models. The coefficient of *positive*esg* is also significant in model I and IV. These are indications that the historical ESG record may have influence on investors' perception of ESG news.

We also conduct similar regressions for the Europe subsample and report results in Table 4.12. Even though the Europe sample presents a less clear picture, the overall patterns still hold. *negative* is significantly negative in most models except for model III and V. *positive* is significantly positive in model III and V. Moreover, *negative*esg* is significantly positive in all models at the 10% level. This provides further support to our previous findings in the overall sample and the America subsample. No matter in America or Europe,

Table 4.10: Regressions: the overall sample

	ar_0		CAR_1		CAR_2	
	(I)	(II)	(III)	(IV)	(V)	(VI)
H1/H2						
sentiment						
negative	−0.0081***	−0.0383**	−0.0187***	−0.1484**	−0.0187***	−0.1572***
	(0.0017)	(0.0152)	(0.0055)	(0.0576)	(0.0053)	(0.0546)
positive	0.0013***	0.0159***	0.0056***	0.0364**	0.0061***	0.0346*
	(0.0004)	(0.0047)	(0.0017)	(0.0183)	(0.0018)	(0.0189)
H3						
negative*esg		0.0076**		0.0327**		0.0349***
		(0.0037)		(0.0139)		(0.0132)
positive*esg		−0.0037***		−0.0077*		−0.0071
		(0.0011)		(0.0044)		(0.0045)
Controls						
esg	0.0014*	0.0028**	−0.0012	0.0007	−0.0012	0.0004
	(0.0007)	(0.0011)	(0.0032)	(0.0044)	(0.0033)	(0.0046)
asset	−0.0003	−0.0003	−0.0011	−0.0012	−0.0013	−0.0014
	(0.0002)	(0.0002)	(0.0008)	(0.0009)	(0.0009)	(0.0009)
num_news	−0.0003	−0.0002	−0.0011	−0.0010	−0.0008	−0.0009
	(0.0006)	(0.0006)	(0.0024)	(0.0025)	(0.0026)	(0.0028)
continent						
africa	−0.0008	−0.0010	0.0100	0.0096	0.0108	0.0103
	(0.0025)	(0.0024)	(0.0120)	(0.0120)	(0.0136)	(0.0136)
asia	0.0003	0.0004	0.0013	0.0014	0.0014	0.0015
	(0.0007)	(0.0007)	(0.0039)	(0.0039)	(0.0039)	(0.0040)
europe	0.0006	0.0007	0.0022	0.0025	0.0022	0.0025
	(0.0005)	(0.0005)	(0.0020)	(0.0020)	(0.0021)	(0.0021)
oceania	0.0049***	0.0044***	0.0070	0.0052	0.0067	0.0048
	(0.0013)	(0.0013)	(0.0056)	(0.0056)	(0.0057)	(0.0057)
sector						
communication_services	0.0016	0.0017	−0.0036	−0.0032	−0.0039	−0.0034
	(0.0016)	(0.0016)	(0.0050)	(0.0050)	(0.0053)	(0.0052)
consumer_discretionary	−0.0003	−0.0004	0.0061	0.0061	0.0073	0.0073
	(0.0008)	(0.0008)	(0.0054)	(0.0055)	(0.0054)	(0.0054)
consumer_staples	0.0001	0.0001	−0.0028	−0.0027	−0.0031	−0.0030
	(0.0008)	(0.0008)	(0.0031)	(0.0031)	(0.0033)	(0.0033)
energy	0.0012	0.0014	0.0030	0.0033	0.0032	0.0034
	(0.0011)	(0.0011)	(0.0049)	(0.0049)	(0.0051)	(0.0051)
financials	−0.0005	−0.0005	0.0000	−0.0002	−0.0001	−0.0003
	(0.0009)	(0.0009)	(0.0050)	(0.0050)	(0.0051)	(0.0051)
health_care	−0.0041***	−0.0041***	−0.0070*	−0.0068	−0.0083*	−0.0081*
	(0.0011)	(0.0011)	(0.0042)	(0.0044)	(0.0044)	(0.0045)
information_technology	−0.0023**	−0.0024***	−0.0148***	−0.0151***	−0.0157***	−0.0161***
	(0.0009)	(0.0009)	(0.0032)	(0.0032)	(0.0034)	(0.0034)
materials	−0.0012*	−0.0011*	0.0037	0.0038	0.0028	0.0029
	(0.0007)	(0.0007)	(0.0032)	(0.0032)	(0.0035)	(0.0035)
real_estate	0.0001	0.0003	−0.0074*	−0.0069	−0.0086*	−0.0082*
	(0.0010)	(0.0010)	(0.0043)	(0.0043)	(0.0045)	(0.0045)
utilities	−0.0012*	−0.0012*	−0.0121***	−0.0120***	−0.0142***	−0.0141***
	(0.0007)	(0.0007)	(0.0032)	(0.0032)	(0.0035)	(0.0035)
_cons	0.0005	−0.0041	0.0333*	0.0277	0.0394**	0.0353*
	(0.0041)	(0.0052)	(0.0170)	(0.0201)	(0.0179)	(0.0213)
N	50,532	50,532	50,532	50,532	50,532	50,532
F_Statistic	4.81***	4.47***	6.09***	5.81***	6.33***	5.95***
R_Squared	0.0040	0.0051	0.0046	0.0056	0.0049	0.0059
Adj_R_Squared	0.0036	0.0046	0.0042	0.0052	0.0045	0.0055

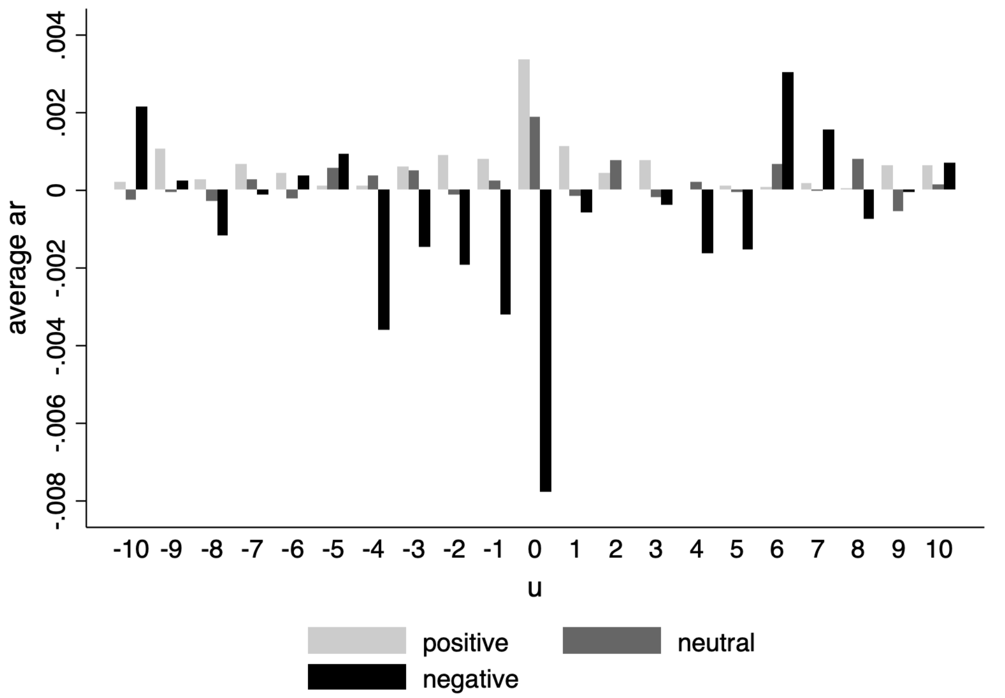
This table shows regression results for the overall sample. Standard errors are cluster-robust at the company level. All variables are defined in Table 4.14. $^*p < .1$, $^{**}p < .05$, *** $p < .01$.

Table 4.11: Regressions: the America sample

	ar_0		CAR_1		CAR_2	
	(I)	(II)	(III)	(IV)	(V)	(VI)
H1/H2						
sentiment						
negative	−0.0092***	−0.0367**	−0.0263***	−0.1650**	−0.0264***	−0.1755***
	(0.0022)	(0.0176)	(0.0069)	(0.0683)	(0.0066)	(0.0638)
positive	0.0023***	0.0183***	0.0066**	0.0436*	0.0069**	0.0433*
	(0.0007)	(0.0062)	(0.0027)	(0.0239)	(0.0028)	(0.0245)
H3						
negative*esg		0.0072*		0.0363**		0.0390**
		(0.0043)		(0.0165)		(0.0156)
positive*esg		−0.0041***		−0.0095*		−0.0094
		(0.0015)		(0.0057)		(0.0059)
Controls						
esg	0.0016	0.0032**	−0.0028	0.0001	−0.0032	−0.0005
	(0.0010)	(0.0015)	(0.0046)	(0.0063)	(0.0047)	(0.0065)
asset	−0.0005*	−0.0005*	−0.0015	−0.0015	−0.0019*	−0.0019*
	(0.0003)	(0.0003)	(0.0010)	(0.0010)	(0.0011)	(0.0011)
num_news	0.0011*	0.0011*	0.0027	0.0024	0.0030	0.0028
	(0.0006)	(0.0006)	(0.0035)	(0.0036)	(0.0038)	(0.0039)
sector						
communication_services	0.0036	0.0036	−0.0007	−0.0004	−0.0001	0.0003
	(0.0026)	(0.0026)	(0.0070)	(0.0070)	(0.0077)	(0.0076)
consumer_discretionary	−0.0014	−0.0014	0.0098	0.0096	0.0106	0.0105
	(0.0012)	(0.0012)	(0.0092)	(0.0094)	(0.0091)	(0.0093)
consumer_staples	0.0005	0.0006	−0.0028	−0.0025	−0.0033	−0.0031
	(0.0012)	(0.0012)	(0.0047)	(0.0047)	(0.0051)	(0.0051)
energy	0.0029	0.0030	0.0028	0.0031	0.0037	0.0039
	(0.0019)	(0.0019)	(0.0083)	(0.0084)	(0.0087)	(0.0088)
financials	0.0005	0.0005	−0.0053	−0.0054	−0.0048	−0.0049
	(0.0011)	(0.0011)	(0.0048)	(0.0048)	(0.0050)	(0.0050)
health_care	−0.0043***	−0.0041***	−0.0035	−0.0029	−0.0043	−0.0037
	(0.0016)	(0.0016)	(0.0056)	(0.0056)	(0.0058)	(0.0058)
information_technology	−0.0023*	−0.0024**	−0.0193***	−0.0196***	−0.0208***	−0.0211***
	(0.0012)	(0.0012)	(0.0046)	(0.0046)	(0.0047)	(0.0047)
materials	−0.0007	−0.0006	0.0014	0.0016	−0.0004	−0.0002
	(0.0012)	(0.0012)	(0.0055)	(0.0055)	(0.0061)	(0.0061)
real_estate	0.0003	0.0005	−0.0128**	−0.0120**	−0.0154**	−0.0145**
	(0.0014)	(0.0014)	(0.0060)	(0.0060)	(0.0063)	(0.0063)
utilities	−0.0014	−0.0014	−0.0181***	−0.0180***	−0.0210***	−0.0209***
	(0.0012)	(0.0012)	(0.0051)	(0.0051)	(0.0055)	(0.0055)
_cons	0.0043	−0.0015	0.0486**	0.0394	0.0596***	0.0512*
	(0.0055)	(0.0071)	(0.0204)	(0.0253)	(0.0218)	(0.0270)
N	27,613	27,613	27,613	27,613	27,613	27,613
F_Statistic	3.89***	3.65***	6.72***	6.40***	6.88***	6.50***
R_Squared	0.0046	0.0057	0.0064	0.0076	0.0069	0.0082
Adj_R_Squared	0.0041	0.0050	0.0058	0.0070	0.0064	0.0076

This table shows regression results for the America sample. Standard errors are cluster-robust at the company level. All variables are defined in Table 4.14. $*p < .1$, $**p < .05$, $***\ p < .01$.

Figure 4.7: Abnormal returns during $(T_0 - 10, T_0 + 10)$: the Europe subsample



good historical reputation could be an asset when a company has some bad ESG news coverage, while a liability when it encounters good ones.

4.7 Conclusion

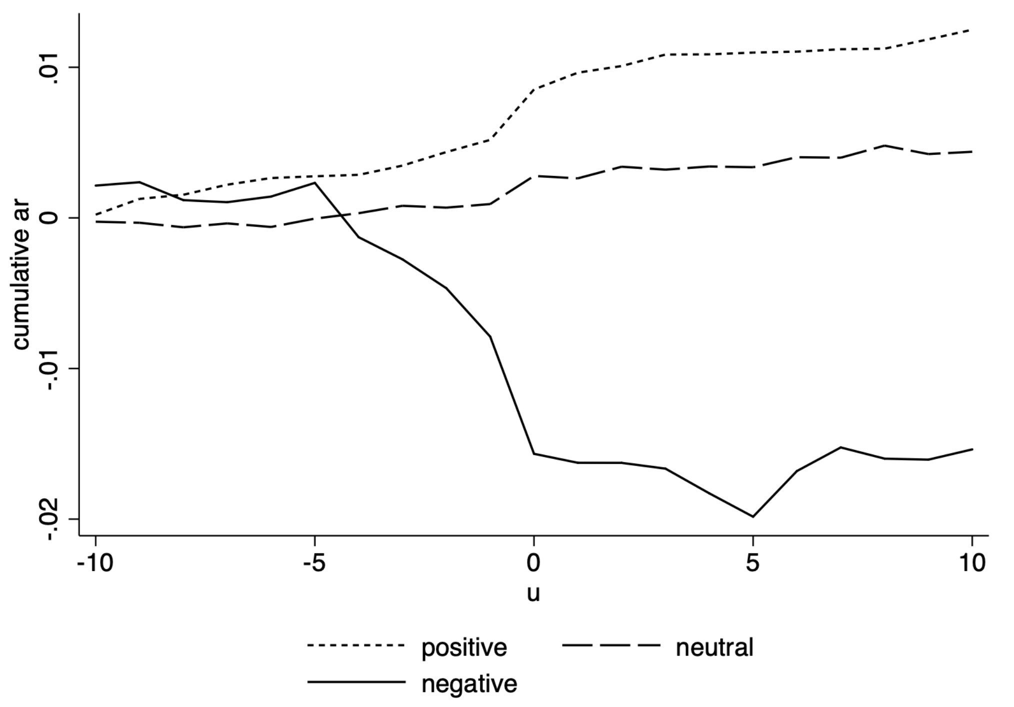
In this study, we examine the pricing mechanism of ESG news on the major stock markets. We show how the newest development in NLP can be applied in understanding the market reactions to *instant ESG news*. Instead of directly adopting a proprietary ESG news dataset from ESG data providers, we construct our sample by extracting raw ESG news from Thomson Reuters Eikon and clean the news data in a consistent way. Based on a pre-trained sentence-BERT model, we are able to remove fuzzy duplicate or stale news and retain only fresh and unique ESG news to a large extent. This procedure makes sure that we have a unique and fresh news dataset while enjoying the wide coverage of ESG news from all over the world. Moreover, we fine-tune a ESG news sentiment classifier based on the BERT-like language model and achieve relatively good predictive performance. We apply it to judge the sentiment of ESG news instead of using classical lexicon-based sentiment analysis methods.

Table 4.12: Regressions: the Europe sample

	ar_0		CAR_1		CAR_2	
	(I)	(II)	(III)	(IV)	(V)	(VI)
H1/H2						
sentiment						
negative	−0.0081**	−0.0708*	−0.0175	−0.2840*	−0.0157	−0.3068*
	(0.0031)	(0.0401)	(0.0123)	(0.1651)	(0.0127)	(0.1717)
positive	−0.0003	0.0144	0.0055***	0.0149	0.0063***	0.0120
	(0.0005)	(0.0091)	(0.0020)	(0.0237)	(0.0021)	(0.0269)
H3						
negative*esg		0.0152*		0.0645*		0.0704*
		(0.0092)		(0.0378)		(0.0393)
positive*esg		−0.0035		−0.0022		−0.0013
		(0.0021)		(0.0056)		(0.0063)
Controls						
esg	0.0028**	0.0038*	0.0031	0.0013	0.0048	0.0023
	(0.0014)	(0.0020)	(0.0036)	(0.0047)	(0.0039)	(0.0054)
asset	−0.0004	−0.0005*	−0.0006	−0.0008	−0.0005	−0.0007
	(0.0003)	(0.0003)	(0.0016)	(0.0015)	(0.0016)	(0.0016)
num_news	−0.0011*	−0.0009	−0.0059	−0.0055	−0.0060	−0.0057
	(0.0006)	(0.0006)	(0.0042)	(0.0043)	(0.0046)	(0.0046)
sector						
communication_services	−0.0031	−0.0030	−0.0123	−0.0119	−0.0145	−0.0141
	(0.0030)	(0.0029)	(0.0118)	(0.0117)	(0.0119)	(0.0117)
consumer_discretionary	0.0004	0.0007	−0.0045	−0.0033	−0.0024	−0.0010
	(0.0016)	(0.0015)	(0.0060)	(0.0055)	(0.0064)	(0.0058)
consumer_staples	0.0005	0.0006	−0.0045	−0.0045	−0.0054	−0.0054
	(0.0012)	(0.0012)	(0.0048)	(0.0048)	(0.0049)	(0.0049)
energy	0.0003	0.0004	0.0014	0.0015	−0.0001	0.0000
	(0.0012)	(0.0012)	(0.0062)	(0.0062)	(0.0067)	(0.0067)
financials	−0.0002	0.0000	0.0021	0.0027	0.0011	0.0017
	(0.0015)	(0.0014)	(0.0077)	(0.0074)	(0.0080)	(0.0077)
health_care	−0.0046***	−0.0050***	−0.0134*	−0.0149*	−0.0146**	−0.0162**
	(0.0017)	(0.0018)	(0.0071)	(0.0078)	(0.0073)	(0.0080)
information_technology	−0.0006	−0.0008	−0.0056	−0.0061	−0.0048	−0.0053
	(0.0012)	(0.0012)	(0.0064)	(0.0063)	(0.0070)	(0.0068)
materials	−0.0008	−0.0008	0.0074*	0.0075*	0.0069	0.0069
	(0.0008)	(0.0008)	(0.0038)	(0.0038)	(0.0042)	(0.0042)
real_estate	0.0003	0.0003	0.0025	0.0022	0.0050	0.0046
	(0.0019)	(0.0020)	(0.0089)	(0.0090)	(0.0097)	(0.0098)
utilities	−0.0015**	−0.0015*	−0.0066	−0.0064	−0.0073	−0.0070
	(0.0008)	(0.0008)	(0.0043)	(0.0043)	(0.0046)	(0.0046)
_cons	−0.0004	−0.0025	0.0093	0.0209	−0.0002	0.0142
	(0.0071)	(0.0084)	(0.0343)	(0.0309)	(0.0366)	(0.0339)
N	14,457	14,457	14,457	14,457	14,457	14,457
F_Statistic	2.95***	2.73***	2.40***	2.20***	2.45***	2.29***
R_Squared	0.0050	0.0073	0.0056	0.0082	0.0053	0.0081
Adj_R_Squared	0.0039	0.0062	0.0046	0.0070	0.0043	0.0069

This table shows regression results for the Europe sample. Standard errors are cluster-robust at the company level. All variables are defined in Table 4.14. $^{*}p < .1$, $^{**}p < .05$, $^{***}\ p < .01$.

Figure 4.8: Cumulative abnormal returns during (T_0-10, T_0+10): the Europe subsample



We find that the impact of ESG news is closely related to the ESG news sentiment. However, the market reactions to positive and negative ESG news are asymmetric. Positive ESG news have positive influence on the stock price while negative ESG news have stronger and negative influence on stock performance. This indicates that positive ESG news may add some value to the firm while negative ESG news do harm to the firm value to a considerable extent. Moreover, the historical ESG image of a company may influence the impact of ESG news on stock markets. More specifically, the market reaction to negative ESG news is related to the ESG record of the company. If the company has a good ESG record in the past, the negative influence of negative ESG news could be dampened and less severe. In contrast, if the company has a bad ESG record, the market reacts more favorably to marginal improvement of ESG performance.

This study has clear research implications for other financial studies. We show how the recent development in NLP could possibly facilitate and advance the research on ESG topics in different ways. The proposed text processing methodologies can also be applied in related studies, especially those investigating the role of non-financial factors. We focus specifically on the possible pricing effect of *instant ESG news* and provide new insights on how the market reacts to *instant ESG news* on the major stock markets. The empirical findings suggest the importance of ESG issues on the financial markets. Investors may incorporate daily ESG information into their investment decisions, instead of merely depending on company ESG disclosure and ESG ratings from agencies. Therefore, more attention should be given to more frequent ESG information such as *instant ESG news* in order to better understand the role of ESG issues.

The practical implications of this study are obvious and straightforward. Firstly, we show the importance of timely tracking of *instant ESG news* for investors. Investors can monitor real-time ESG news and incorporate this information in a timely manner in investment practice. Secondly, companies should not only avoid negative ESG news, but also work on improving their ESG performance since positive ESG news are also valued by investors. Moreover, companies should build up their own media monitoring system as part of their investor-relationship management, so as to build a better ESG image and avoid any misunderstandings with investors and the general public. At last, for related policy makers, our study indicates the possibility of ESG performance fraudulence or exaggeration in ESG news. Regulations or policies that can detect or increase the cost of such behavior should be considered and implemented. One possible countermeasure is the establishment of a third-party reviewing system in which independent external reviewers validate and evaluate these ESG news on a regular basis. Moreover, the advancement in NLP could also be applied to alleviate the problem. By constructing a ESG news dataset with a label indicating the authenticity of the news, a classifier can be be trained to detect fraudulence or exaggeration.

We are also aware of the limitations of this study and thus provide some future research directions. First, despite the relatively good sentiment classification result, our sentiment classifier is trained on a labelled dataset pre-processed by a third party and thus its validity is restrained by the given training dataset. A better (but more costly and complicated) solution would be constructing a sentiment labelled dataset by designing an experiment in which participants are asked to read company ESG news and evaluate the news sentiment. Second, we do not differentiate various types of ESG news in this study and may not know whether investors may perceive them differently. For example, whether sustainability issues are financially material has a significant impact on the firm value (Khan et al., 2016). It would be interesting to integrate the financial materiality aspect into the pricing implication analysis of ESG news.

4.8 Appendix

Table 4.13: ESG news (in English) volume: Top 20 countries/regions

no.	Country/Region	ESG news	no.	Country/Region	ESG news
1.	USA	35,284	11.	Switzerland	1,243
2.	Canada	9,710	12.	Italy	1,166
3.	UK	7,319	13.	Finland	1,095
4.	Japan	3,372	14.	Korea	1,061
5.	France	3,307	15.	Spain	1,057
6.	India	3,090	16.	China	985
7.	Australia	2,087	17.	Netherlands	888
8.	Germany	1,883	18.	Norway	821
9.	Sweden	1,436	19.	Thailand	815
10.	Hongkong	1,325	20.	Russia	642

Table 4.14: Definition of variables

Variable	Description
H1	
ar_0	Abnormal return on the event date based on the estiamted market model.
CAR_1	Cumulated return during the ($T_0 - 1$, $T_0 + 1$) event window.
CAR_2	Cumulated abnormal return during the ($T_0 - 2$, $T_0 + 2$) event window.
CAR_5	Cumulated return during the ($T_0 - 5$, $T_0 + 5$) event window.
CAR_{10}	Cumulated abnormal return during the ($T_0 - 5$, $T_0 + 5$) event window.
H2	
sentiment	Categorical variable indicating ESG news sentiment judged by the BERT model, indicating whether the news sentiment is "positive", "neutral", and "negative". The reference category is "neutral".
H3	
$negative * esg$	Interaction term between *negative* and *esg*.
$positive * esg$	Interaction term between *positive* and *esg*.
Controls	
esg	ESG score for the corresponding company in the previous year of the event day, logarithmized in regressions
asset	Total asset of the company, logarithmized in regressions.
num_news	Number of ESG news for the company during the sample period, divided by 100 in regressions.
continent	Continent where the company is located. The reference category is "America".
sector	Sector to which the company belong. The Global Industry Classifiction Standards (GICS) is adopted. The reference category is "industrials".

Chapter 5

Conclusion

This dissertation focuses on non-financial factors such as altruistic intentions and sustainability preferences in alternative and sustainable finance. With the rise of alternative and sustainable finance in recent years, the role of non-financial factors becomes more prominent in understanding the funding and pricing mechanism of related financing instruments. However, many of these factors are implicitly embedded in unstructured data such as descriptive texts. The difficulty in extracting non-financial factors from non-numeric data poses a great challenge to the investigation of their role and influence on financial markets.

In this dissertation, several linguistic analysis techniques are applied to measure non-financial factors in the setting of alternative and sustainable finance. In the first paper, keyword analysis of microloan applications provides several proxies of soft factors including investors' social concerns. The second paper consolidates various kinds of external reviews and integrates different greenness evaluation schemes into a unified scale by hand, and thus enables the measurement of the authenticity and greenness of green bonds. In the last paper, the BERT language model, which is widely recognized as a substantial breakthrough in NLP, is applied in processing raw ESG news sample in several different aspects such as removing fuzzy duplicate news and identifying news sentiment. By evaluating the fluctuation of ESG performance through sentiment signals extracted from instant ESG news, it also avoids the reliance on ESG scores in the research of sustainable finance.

Given these quantitative indicators of non-financial factors, further empirical investigations can be conducted to inspect their role in the funding and pricing mechanism in alternative and sustainable finance. The first paper examines the funding determinants of interest-free P2P lending in the US, paying special attention to philanthropic needs of investors on the Kiva platform. Logistic regression, Tobit regression and Cox regression are run to investigate the funding determinants. We find that investors prefer to grant loans to women and groups of borrowers besides favoring a social endorsement and creditworthiness signals in texts. Moreover, borrowers' vulnerability appears to be only of interest to investors of endorsed loans, indicating the presence of investors' heterogeneous preference for altruism. In the second paper, we apply comparison analysis, i.e. a rigorous matching procedure, to estimate true green bond premiums. By matching green bonds with conventional bonds with similar credit characteristics, we can contribute the yield difference after excluding the impact of liquidity difference to be the market value of greenness. While overall green bonds enjoy a small-scale premium compared to ideally matched conventional bonds, investors can differentiate the authenticity and greenness level of green bonds by referring to signals extracted from four types of external reviews and shade of green evaluation results. Based on a large sample of fresh ESG news with a sentiment label, the last research conducts event study and regressions to give a glance at how ESG news are priced with different sentiment and given different historical ESG profiles. There is clear evidence that the stock markets react to ESG news parallel to the news sentiment. However, negative ESG news tend to have stronger influence than positive ESG news. We also find that the impact of ESG news depends on the past ESG profile.

This dissertation contributes to the research of non-financial factors on different aspects. First of all, the way we adopt linguistic analysis techniques in different settings can be a first step towards better measurement of these soft factors. Moreover, our research has clear theoretical implications. Overall, this dissertation reveals several interesting and important aspects of the role of non-financial factors in both alternative and sustainable finance. The empirical findings, which may be considered as abnormalities in classical finance theory, suggest that non-financial factors are perceived and priced by investors. Therefore, these factors should be considered and integrated into future research. Lastly, this dissertation has practical implications for related market participants. For instance, philanthropic crowdfunding platforms can design their loan application in a way that applicants' characteristics that may interest investors the most are underscored. Climate bond issuers have a better view of how investors perceive the effectiveness of different external review reports and thus can choose the ideal ones. Sustainable investors can build a real-time and more sensitive ESG performance evaluation mechanism to guide their investment decisions.

However, the three papers are also subject to several limitations, which may indicate further research directions on this topic. Our endeavors to quantify non-financial factors are imperfect and need further improvement. Keyword analysis of microloan applications may be insufficient to extract targeted information. Manual process and integration of external review reports are not scalable and subject to human errors when the size of sample increases. The last paper shows potential for a systematic way of quantifying non-financial factors. With the recent breakthrough in NLP, we are confident that more innovative applications of language models could take the research of non-financial factors to a completely different level. Moreover, our research is conducted in distinct settings in alternative and sustainable finance and thus does not provide a unified analysis framework of non-financial factors. While the first paper focuses exclusively on social care, the other two papers both take the environmental protection aspect into consideration. Future research on non-financial factors can propose a standardized taxonomy of non-financial considerations and conduct analysis in an integrated framework. At last, our research does not consider rigorously possible interactions of financial and non-financial factors. In other words, it could be possible that investors' taste for non-financial factors is partly or completely driven by financial motives. Thus, a promising research direction could be the integration of financial materiality into the research of non-financial factors.

Bibliography

Akerlof, G. A., 1970. The market for 'lemons': Quality uncertainty and the market mechanism. Quarterly Journal of Economics, 84, 488–500.

Akhtar, S., Faff, R., Oliver, B., Subrahmanyam, A., 2011. The power of bad: The negativity bias in australian consumer sentiment announcements on stock returns. Journal of Banking & Finance, 35 (5), 1239–1249.

Alanyali, M., Moat, H. S., Preis, T., 2013. Quantifying the relationship between financial news and the stock market. Scientific Reports, 3 (1), 1–6.

Alaparthi, S., Mishra, M., 2021. Bert: A sentiment analysis odyssey. Journal of Marketing Analytics, 9 (2), 118–126.

Aldén, L., Hammarstedt, M., 2016. Discrimination in the credit market? Access to financial capital among self-employed immigrants. Kyklos, 69 (1), 3–31.

Allet, M., 2014. Why do microfinance institutions go green? An exploratory study. Journal of Business Ethics, 122 (3), 405–424.

Allet, M., et al., 2011. Measuring the environmental performance of microfinance. CEB Working Paper.

Allison, T. H., Davis, B. C., Short, J. C., Webb, J. W., 2015. Crowdfunding in a prosocial microlending environment: Examining the role of intrinsic versus extrinsic cues. Entrepreneurship Theory and Practice, 39 (1), 53–73.

Allison, T. H., McKenny, A. F., Short, J. C., 2013. The effect of entrepreneurial rhetoric on microlending investment: An examination of the warm-glow effect. Journal of Business Venturing, 28 (6), 690–707.

Alsos, G. A., Ljunggren, E., 2017. The role of gender in entrepreneur–investor relationships: A signaling theory approach. Entrepreneurship Theory and Practice, 41 (4), 567–590.

Amel-Zadeh, A., Serafeim, G., 2018. Why and how investors use ESG information: Evidence from a global survey. Financial Analysts Journal, 74 (3), 87–103.

Amihud, Y., Mendelson, H., 1986. Asset pricing and the bid-ask spread. Journal of Financial Economics, 17 (2), 223–249.

Araci, D., 2019. Finbert: Financial sentiment analysis with pre-trained language models. arXiv preprint arXiv:1908.10063.

Armendáriz, B., Morduch, J., 2010. The economics of microfinance. MIT press.

Bachelet, M. J., Becchetti, L., Manfredonia, S., 2019. The green bonds premium puzzle: The role of issuer characteristics and third-party verification. Sustainability, 11 (4), 1098.

Baker, M., Bergstresser, D., Serafeim, G., Wurgler, J., 2019. Financing the response to climate change: The pricing and ownership of U.S. green bonds. NBER working paper.

Baldini, M., Dal Maso, L., Liberatore, G., Mazzi, F., Terzani, S., 2018. Role of country- and firm-level determinants in environmental, social, and governance disclosure. Journal of Business Ethics, 150 (1), 79–98.

Barasinska, N., Schäfer, D., 2014. Is crowdfunding different? Evidence on the relation between gender and funding success from a German peer-to-peer lending platform. German Economic Review, 15 (4), 436–452.

Barinaga, E., 2014. Microfinance in a developed welfare state: A hybrid technology for the government of the outcast. Geoforum, 51, 27–36.

Bartov, E., Marra, A., Momente, F., 2021. Corporate social responsibility and the market reaction to negative events: Evidence from inadvertent and fraudulent restatement announcements. The Accounting Review, 96 (2), 81–106.

Battilana, J., Dorado, S., 2010. Building sustainable hybrid organizations: The case of commercial microfinance organizations. Academy of management Journal, 53 (6), 1419–1440.

Bazillier, R., Vauday, J., 2009. The greenwashing machine: Is CSR more than communication. HAL, hal-00448861.

Beatriz, A., Marc, L., 2011. The handbook of microfinance. World Scientific.

Bell, A., Jones, K., 2015. Explaining fixed effects: Random effects modeling of time-series cross-sectional and panel data. Political Science Research and Methods, 3 (1), 133–153.

Bendig, M., Unterberg, M., Sarpong, B., 2012. Overview of the microcredit sector in the European Union 2010-2011. European Microfinance Network.

Bendig, M., Unterberg, M., Sarpong, B., 2014. Overview of the microcredit sector in the European Union 2012-2013. European Microfinance Network.

Bennani, L., Le Guenedal, T., Lepetit, F., Ly, L., Mortier, V., Roncalli, T., Sekine, T., 2018. How ESG Investing has impacted the asset pricing in the equity market. Available at SSRN.

Berg, F., Koelbel, J. F., Rigobon, R., 2019. Aggregate confusion: The divergence of ESG ratings. MIT Sloan School of Management.

Berger, S. C., Gleisner, F., 2009. Emergence of financial intermediaries in electronic markets: The case of online P2P lending. BuR-Business Research, 2, 39–65.

Berns, J. P., Figueroa-Armijos, M., da Motta Veiga, S. P., Dunne, T. C., 2018. Dynamics of lending-based prosocial crowdfunding: Using a social responsibility lens. Journal of Business Ethics.

Berry, R., Yeung, F., 2013. Are investors willing to sacrifice cash for morality? Journal of Business Ethics, 117, 477–492.

Bialkowski, J., Starks, L. T., 2016. SRI funds: Investor demand, exogenous shocks and ESG profiles. Working paper. University of Canterbury.

Bolton, P., Kacperczyk, M., 2021. Do investors care about carbon risk? Journal of Financial Economics.

Boudoukh, J., Feldman, R., Kogan, S., Richardson, M., 2019. Information, trading, and volatility: Evidence from firm-specific news. The Review of Financial Studies, 32 (3), 992–1033.

Bourlès, R., Cozarenco, A., 2018. Entrepreneurial motivation and business performance: Evidence from a french microfinance institution. Small Business Economics, 51, 943–963.

Bruhn-Leon, B., Eriksson, P.-E., Kraemer-Eis, H., 2012. Progress for microfinance in Europe. EIF Working Paper.

Bruton, G., Khavul, S., Siegel, D., Wright, M., 2015. New financial alternatives in seeding entrepreneurship: Microfinance, crowdfunding, and peer–to–peer innovations. Entrepreneurship Theory and Practice, 39, 9–26.

Bruton, G. D., Khavul, S., Chavez, H., 2011. Microlending in emerging economies: Building a new line of inquiry from the ground up. Journal of International Business Studies, 42 (5), 718–739.

Burtch, G., Ghose, A., Wattal, S., 2014. Cultural differences and geography as determinants of online pro-social lending. MIS Quarterly, 38 (3), 773–794.

Calic, G., Mosakowski, E., 2016. Kicking off social entrepreneurship: How a sustainability orientation influences crowdfunding success. Journal of Management Studies, 53 (5), 738–767.

Camilleri, M. A., et al., 2017. Corporate sustainability, social responsibility and environmental management. Cham, Switzerland: Springer International Publishing.

Capelle-Blancard, G., Petit, A., 2019. Every little helps? ESG news and stock market reaction. Journal of Business Ethics, 157 (2), 543–565.

Carboni, B. J., Calderón, M. L., Garrido, S. R., Dayson, K., Kickul, J., 2010. Handbook of Microcredit in Europe. Edward Elgar Publishing.

Chava, S., 2014. Environmental externalities and cost of capital. Management Science, 60 (9), 2111–2380.

Cheema-Fox, A., LaPerla, B. R., Serafeim, G., Turkington, D., Wang, H., 2019. Decarbonization Factors. SSRN Working Paper.

Chen, L., Lesmond, D. A., Wei, J., 2007. Corporate yield spreads and bond liquidity. The Journal of Finance, 62 (1), 119–149.

Cheston, S., Kuhn, L., 2002. Empowering women through microfinance. Publication sponsored by UNIFEM.

Christensen, D., Serafeim, G., Sikochi, A., 2019. Why is corporate virtue in the eye of the beholder? The case of ESG ratings. Working paper.

CICERO, 2015. Kfw green bond second opinion 2015. Tech. rep., CICERO.

CICERO, 2019. Kfw green bond second opinion 2019. Tech. rep., CICERO.

Climate Bond Initiative, 2019a. Green bond european investor survey 2019. Tech. rep., Climate Change Initiative.

Climate Bond Initiative, 2019b. Green bonds pricing in the primary market: January - june 2019. Tech. rep., Climate Change Initiative.

Climate Bond Initiative, 2020. 2019 green bond market summary. Tech. rep., Climate Change Initiative.

Collier, B. C., Hampshire, R., 2010. Sending mixed signals: Multilevel reputation effects in peer-to-peer lending markets. In: Proceedings of the 2010 ACM conference on Computer supported cooperative work. ACM, pp. 197–206.

Connelly, B. L., Certo, S. T., Ireland, R. D., Reutzel, C. R., 2011. Signaling theory: A review and assessment. Journal of Management, 37 (1), 39–67.

Cozarenco, A., Szafarz, A., 2018. Gender biases in bank lending: Lessons from microcredit in France. Journal of Business Ethics, 147 (3), 631–650.

Cozarenco, A., Szafarz, A., et al., 2014. Microcredit in developed countries: Unexpected consequences of loan ceilings. CEB Working Paper.

Da Silva, N. F., Hruschka, E. R., Hruschka Jr, E. R., 2014. Tweet sentiment analysis with classifier ensembles. Decision Support Systems, 66, 170–179.

Del Giudice, A., Rigamonti, S., 2020. Does audit improve the quality of ESG scores? Evidence from corporate misconduct. Sustainability, 12 (14), 5670.

Demirguc-Kunt, A., Klapper, L., Singer, D., Ansar, S., 2018. The Global Findex Database 2017: Measuring financial inclusion and the fintech revolution. World Bank Publications.

Devlin, J., Chang, M.-W., Lee, K., Toutanova, K., 2018. BERT: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.

Dhaliwal, D., Radhakrishnan, S., Tsang, A., Yang, Y. G., 2012. Nonfinancial disclosure and analyst forecast accuracy: International evidence on corporate social responsibility disclosure. The Accounting Review, 87 (3), 723–759.

Dhar, V., 2013. Data science and prediction. Communications of the ACM, 56 (12), 64–73.

Dichter, T. W., Harper, M., 2007. What's wrong with microfinance? Practical Action Publishing Rugby.

Diriker, D., Landoni, P., Benaglio, N., et al., 2018. Microfinance in Europe: Survery Report 2016-2017. European Microfinance Network.

Doms, M., Lewis, E., Robb, A., 2010. Local labor force education, new business characteristics, and firm performance. Journal of Urban Economics, 67 (1), 61–77.

Dorfleitner, G., Halbritter, G., Nguyen, M., 2015. Measuring the level and risk of corporate responsibility–an empirical comparison of different ESG rating approaches. Journal of Asset Management 16 (7), 450–466.

Dorfleitner, G., Oswald, E.-M., 2016. Repayment behavior in peer-to-peer microfinancing: Empirical evidence from kiva. Review of Financial Economics, 30, 45–59.

Dorfleitner, G., Oswald, E.-M., Röhe, M., 2020. The access of microfinance institutions to financing via the worldwide crowd. The Quarterly Review of Economics and Finance, 75, 133–146.

Dorfleitner, G., Priberny, C., Schuster, S., Stoiber, J., Weber, M., de Castro, I., Kammler, J., 2016. Description-text related soft information in peer-to-peer lending–evidence from two leading european platforms. Journal of Banking & Finance, 64, 169–187.

Dorfleitner, G., Utz, S., 2012. Safety first portfolio choice based on financial and sustainability returns. European Journal of Operational Research, 221 (1), 155–164.

Dorfleitner, G., Utz, S., 2014. Profiling german-speaking socially responsible investors. Qualitative Research in Financial Markets, 6 (2), 118–156.

Drempetic, S., Klein, C., Zwergel, B., 2019. The influence of firm size on the ESG score: Corporate sustainability ratings under review. Journal of Business Ethics, 1–28.

Duarte, J., Siegel, S., Young, L., 2012. Trust and credit: The role of appearance in peer-to-peer lending. The Review of Financial Studies, 25 (8), 2455–2484.

Edmans, A., Garcia, D., Norli, Ø., 2007. Sports sentiment and stock returns. The Journal of Finance, 62 (4), 1967–1998.

Ehlers, T., Packer, F., 2017. Green bond finance and certification. BIS Quarterly Review, September.

Escrig-Olmedo, E., Fernández-Izquierdo, M. Á., Ferrero-Ferrero, I., Rivera-Lirio, J. M., Muñoz-Torres, M. J., 2019. Rating the raters: Evaluating how ESG rating agencies integrate sustainability principles. Sustainability, 11 (3), 915.

Escrig-Olmedo, E., Muñoz-Torres, M. J., Fernández-Izquierdo, M. Á., Rivera-Lirio, J. M., 2014. Lights and shadows on sustainability rating scoring. Review of Managerial Science, 8 (4), 559–574.

Fama, E. F., French, K. R., 2007. Disagreement, tastes, and asset prices. Journal of Financial Economics, 83 (3), 667–689.

Fenn, G. W., 2000. Speed of issuance and the adequacy of disclosure in the 144a high-yield debt market. Journal of Financial Economics, 56 (3), 383 – 405.

Fiaschi, D., Giuliani, E., Nieri, F., Salvati, N., 2020. How bad is your company? measuring corporate wrongdoing beyond the magic of esg metrics. Business Horizons, 63 (3), 287–299.

Flammer, C., 2013. Corporate social responsibility and shareholder reaction: The environmental awareness of investors. Academy of Management Journal, 56 (3), 758–781.

Flammer, C., 2020. Corporate green bonds. Journal of Financial Economics, Forthcoming.

Forcella, D., Hudon, M., 2016. Green microfinance in Europe. Journal of Business Ethics, 135 (3), 445–459.

Freedman, S., Jin, G., 2017. The information value of online social networks: Lessons from peer-to-peer lending. International Journal of Industrial Organization, 51, 185–222.

Freedman, S., Jin, G. Z., 2008. Do social networks solve information problems for peer-to-peer lending? Evidence from prosper. com. NET Institute Working Paper. Bloomington. Indiana University.

Friede, G., Busch, T., Bassen, A., 2015. ESG and financial performance: Aggregated evidence from more than 2000 empirical studies. Journal of Sustainable Finance & Investment, 5 (4), 210–233.

Gaiha, R., Thapa, G., 2006. A methodology for assessment of the impact of microfinance on empowerment and vulnerability. Working Paper. International Fund for Agricultural Development.

Galema, R., Lensink, R., Spierdijk, L., 2011. International diversification and microfinance. Journal of International Money and Finance, 30 (3), 507–515.

Ghosh, S., Van Tassel, E., 2013. Funding microfinance under asymmetric information. Journal of Development Economics, 101, 8–15.

Ghoul, S. E., Guedhami, O., Kwok, C. C., Mishra, D. R., 2011. Does corporate social responsibility affect the cost of capital? Journal of Banking & Finance, 35 (9), 2388–2406.

Glück, M., Hübel, B., Scholz, H., 2021. ESG rating events and stock market reactions. Available at SSRN.

Goss, A., Roberts, G. S., 2011. The impact of corporate social responsibility on the cost of bank loans. Journal of Banking & Finance, 35 (7), 1794–1810.

Grewal, J., Hauptmann, C., Serafeim, G., 2020. Material sustainability information and stock price informativeness. Journal of Business Ethics, 1–32.

Grewal, J., Riedl, E. J., Serafeim, G., 2019. Market reaction to mandatory nonfinancial disclosure. Management Science, 65 (7), 3061–3084.

Guo, L., Shi, F., Tu, J., 2016. Textual analysis and machine leaning: Crack unstructured data in finance and accounting. The Journal of Finance and Data Science, 2 (3), 153–170.

Gutsche, G., Köbrich León, A., Ziegler, A., 2019. On the relevance of contextual factors for socially responsible investments: An econometric analysis. Oxford Economic Papers, 71 (3), 756–776.

Gutsche, G., Ziegler, A., 2019. Which private investors are willing to pay for sustainable investments? Empirical evidence from stated choice experiments. Journal of Banking & Finance, 102, 193–214.

Gutsche, G., Zwergel, B., 2020. Investment barriers and labeling schemes for socially responsible investments. Schmalenbach Business Review, 72, 1–47.

Hachenberg, B., Schiereck, D., 2018. Are green bonds priced differently from conventional bonds? Journal of Asset Management, 19 (6), 371–383.

Hahn, R., Kühnen, M., 2013. Determinants of sustainability reporting: A review of results, trends, theory, and opportunities in an expanding field of research. Journal of Cleaner Production, 59, 5–21.

Hammill, A., Matthew, R., McCarter, E., 2008. Microfinance and climate change adaptation. International Institute for Sustainable Development.

Hartzmark, S. M., Sussman, A. B., 2019. Do investors value sustainability? A natural experiment examining ranking and fund flows. The Journal of Finance, 74 (6), 2789–2837.

Harvey, C. R., Liu, Y., Zhu, H., 1 2016. … and the Cross-Section of Expected Returns. Review of Financial Studies, 29 (1), 5–68.

Heinkel, R., Kraus, A., Zechner, J., 2001. The effect of green investment on corporate behavior. Journal of Financial and Quantitative Analysis, 36 (4), 431–449.

Heller, L. R., Badding, K. D., 2012. For compassion or money? The factors influencing the funding of micro loans. The Journal of Socio-Economics, 41 (6), 831–835.

Herzenstein, M., Sonenshein, S., Dholakia, U. M., 2011. Tell me a good story and I may lend you money: The role of narratives in peer-to-peer lending decisions. Journal of Marketing Research, 48, 138–149.

Heston, S. L., Sinha, N. R., 2017. News vs. sentiment: Predicting stock returns from news stories. Financial Analysts Journal, 73 (3), 67–83.

Höchstädter, A. K., Scheck, B., 2015. What's in a name: An analysis of impact investing understandings by academics and practitioners. Journal of Business Ethics, 132 (2), 449–475.

Hong, H., Kacperczyk, M., 2009. The price of sin: The effects of social norms on markets. Journal of Financial Economics, 93 (1), 15–36.

Hudon, M., Ashta, A., 2013. Fairness and microcredit interest rates: From Rawlsian principles of justice to the distribution of the bargaining range. Business Ethics: A European Review, 22, 277–291.

Hudon, M., Traca, D., 2006. Subsidies and sustainability in microfinance. Centre Emile Bernheim, WP-CEB 06 20.

Hudon, M., Traca, D., 2011. On the efficiency effect of subsidies in microfinance: An empirical inquiry. World Development, 39 (6), 966–973.

Imai, K. S., Arun, T., Annim, S. K., 2010. Microfinance and household poverty reduction: New evidence from India. World Development, 38 (12), 1760–1774.

Jahdi, K. S., Acikdilli, G., 2009. Marketing communications and corporate social responsibility (CSR): Marriage of convenience or shotgun wedding? Journal of Business Ethics, 88 (1), 103–113.

Jancenelle, V. E., Javalgi, R. R. G., Cavusgil, E., 2018. The role of economic and normative signals in international prosocial crowdfunding: An illustration using market orientation and psychological capital. International Business Review, 27 (1), 208–217.

Jayo, B., González, A., Conzett, C., 2010. Overview of the microcredit sector in the European Union 2008-2009. European Microfinance Network.

Jenq, C., Pan, J., Theseira, W., 2015. Beauty, weight, and skin color in charitable giving. Journal of Economic Behavior & Organization, 119, 234–253.

Jiang, C., Wang, Z., Wang, R., Ding, Y., 2018. Loan default prediction by combining soft information extracted from descriptive text in online peer-to-peer lending. Annals of Operations Research, 266 (1-2), 511–529.

Johnson, S., Ashta, A., Assadi, D., 2010. Online or offline: The rise of 'peer-to-peer' lending in microfinance. Journal of Electronic Commerce in Organizations, 8 (3), 26–37.

Kabeer, N., 2001. Conflicts over credit: Re-evaluating the empowerment potential of loans to women in rural bangladesh. World Development, 29, 63–84.

Kabeer, N., 2005. Is microfinance a 'magic bullet' for women's empowerment? Analysis of findings from South Asia. Economic and Political weekly, 40, 4709–4718.

Kapraun, J., Scheins, C., 2019. (In)-Credibly Green: Which Bonds Trade at a Green Bond Premium? SSRN working paper.

Ke, Z. T., Kelly, B. T., Xiu, D., 2019. Predicting returns with text data. Tech. rep., National Bureau of Economic Research.

Kearney, C., Liu, S., 2014. Textual sentiment in finance: A survey of methods and models. International Review of Financial Analysis, 33, 171–185.

Kennedy, P., 2008. A guide to econometrics, 6th Edition. Blackwell, Malden, MA.

Khan, M., Serafeim, G., Yoon, A., 2016. Corporate sustainability: First evidence on materiality. The Accounting Review, 91 (6), 1697–1724.

Khandker, S. R., 2005. Microfinance and poverty: Evidence using panel data from Bangladesh. The World Bank Economic Review, 19 (2), 263–286.

Khavul, S., 2010. Microfinance: Creating opportunities for the poor? Academy of Management Perspectives, 24 (3), 58–72.

Kim, E.-H., Lyon, T. P., 2015. Greenwash vs. brownwash: Exaggeration and undue modesty in corporate sustainability disclosure. Organization Science, 26 (3), 705–723.

Kiva, 2018a. Webpage – information on Kiva statistics. Https://www.kiva.org/about, last verified 2018-07-08.

Kiva, 2018b. Webpage – requirements for Kiva direct loans. Https://www.kiva.org/about/due-diligence/direct-loans, last verified 2018-06-09.

Kiva, 2019a. Webpage – information on Kiva trustees. Https://www.kiva.org/trustees, last verified 2019-06-10.

Kiva, 2019b. Webpage – information on the risk of lending. Https://www.kiva.org/about/due-diligence/risk, last verified 2019-06-10.

Kiva, 2019c. Webpage – information on the social performance. Https://www.kiva.org/about/impact/socialperformance, last verified 2019-06-10.

Kolari, J. W., Pape, B., Pynnonen, S., 2018. Event study testing with cross-sectional correlation due to partially overlapping event windows. Mays Business School Research Paper.

Kolari, J. W., Pynnönen, S., 2010. Event study testing with cross-sectional correlation of abnormal returns. The Review of Financial Studies, 23 (11), 3996–4025.

Kotelnikova, A., Paschenko, D., Razova, E., 2021. Lexicon-based methods and BERT model for sentiment analysis of russian text corpora. CEUR Workshop Proceedings.

KPMG, 2019. Impact of ESG disclosures. Tech. rep., KPMG.

KPMG, 2020. The kpmg survey of sustainability reporting 2020. Tech. rep., KPMG.

Kraemer-Eis, H., Conforti, A., 2009. Microfinance in Europe: A market overview. EIF Working Paper.

Krauss, N., Walter, I., 2009. Can microfinance reduce portfolio volatility? Economic Development and Cultural Change, 58 (1), 85–110.

Krueger, P., Sautner, Z., Starks, L. T., 2020. The Importance of Climate Risks for Institutional Investors. Review of Financial Studies, 33 (3), 1067–1111.

Krüger, P., 2015. Corporate goodness and shareholder wealth. Journal of Financial Economics, 115 (2), 304–329.

Lan, Z., Chen, M., Goodman, S., Gimpel, K., Sharma, P., Soricut, R., 2019. ALBERT: A lite BERT for self-supervised learning of language representations. arXiv preprint arXiv:1909.11942.

Larcker, D. F., Watts, E. M., 2020. Where's the greenium? Journal of Accounting and Economics, 69, 101312.

Larrimore, L., Jiang, L., Larrimore, J., Markowitz, D., Gorski, S., 2011. Peer to peer lending: The relationship between language features, trustworthiness, and persuasion success. Journal of Applied Communication Research, 39 (1), 19–37.

Ledgerwood, J., Earne, J., Nelson, C., 2013. The new microfinance handbook: A financial market system perspective. The World Bank.

Lee, E., Lee, B., 2012. Herding behavior in online p2p lending: An empirical investigation. Electronic Commerce Research and Applications, 11 (5), 495–503.

Lester, R. H., Certo, S. T., Dalton, C. M., Dalton, D. R., Cannella Jr, A. A., 2006. Initial public offering investor valuations: An examination of top management team prestige and environmental uncertainty. Journal of Small Business Management, 44 (1), 1–26.

Li, X., Xie, H., Chen, L., Wang, J., Deng, X., 2014. News impact on stock price return via sentiment analysis. Knowledge-Based Systems, 69, 14–23.

Li, Y., Gong, M., Zhang, X.-Y., Koh, L., 2018. The impact of environmental, social, and governance disclosure on firm value: The role of CEO power. The British Accounting Review, 50 (1), 60–75.

Li, Z., Tang, Y., Wu, J., Zhang, J., Lv, Q., 2019. The interest costs of green bonds: Credit ratings, corporate social responsibility, and certification. Emerging Markets Finance and Trade, forthcoming.

Lin, M., Prabhala, N. R., Viswanathan, S., 2013. Judging borrowers by the company they keep: Friendship networks and information asymmetry in online peer-to-peer lending. Management Science, 59 (1), 17–35.

Lins, K. V., Servaes, H., Tamayo, A., 2017. Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. The Journal of Finance, 72 (4), 1785–1824.

Liu, D., Brass, D., Lu, Y., Chen, D., 2015. Friendships in online peer-to-peer lending: Pipes, prisms, and relational herding. MIS Quarterly, 39 (3), 729–742.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., Stoyanov, V., 2019. RoBERTa: A robustly optimized BERT pretraining approach. arXiv preprint arXiv:1907.11692.

Lopez, C., Contreras, O., Bendix, J., 2020. Disagreement among ESG rating agencies: Shall we be worried? MPRA paper.

Ly, P., Mason, G., 2012a. Competition between microfinance NGOs: Evidence from Kiva. World Development, 40 (3), 643–655.

Ly, P., Mason, G., 2012b. Individual preferences over development projects: Evidence from microlending on Kiva. Voluntas: International Journal of Voluntary and Nonprofit Organizations, 23 (4), 1036–1055.

Maniora, J., 2017. Is integrated reporting really the superior mechanism for the integration of ethics into the core business model? An empirical analysis. Journal of Business Ethics, 140 (4), 755–786.

McBrayer, G. A., 2018. Does persistence explain ESG disclosure decisions? Corporate Social Responsibility and Environmental Management, 25 (6), 1074–1086.

Michels, J., 2012. Do unverifiable disclosures matter? Evidence from peer-to-peer lending. The Accounting Review, 87 (4), 1385–1413.

Morduch, J., 1999. The microfinance promise. Journal of Economic Literature, 37 (4), 1569–1614.

Morduch, J., 2000. The microfinance schism. World development, 28 (4), 617–629.

Moss, T. W., Neubaum, D. O., Meyskens, M., 2015. The effect of virtuous and entrepreneurial orientations on microfinance lending and repayment: A signaling theory perspective. Entrepreneurship Theory and Practice, 39 (1), 27–52.

Moss, T. W., Renko, M., Block, E., Meyskens, M., 2017. Funding the story of hybrid ventures: Crowdfunder lending preferences and linguistic hybridity. Journal of Business Venturing, 33 (5), 643–659.

Mundlak, Y., 1978. On the pooling of time series and cross section data. Econometrica, 46 (1), 69–85.

Mănescu, C., 2011. Stock returns in relation to environmental, social and governance performance: Mispricing or compensation for risk? Sustainable Development, 19 (2), 95–118.

Nanayakkara, M., Colombage, S., 2019. Do investors in green bond market pay a premium? Global evidence. Applied Economics, 51 (40), 4425–4437.

Naughton, J. P., Wang, C., Yeung, I., 2019. Investor sentiment for corporate social performance. The Accounting Review, 94 (4), 401–420.

Naumer, H.-J., Yurtoglu, B., 2020. It is not only what you say, but how you say it: ESG, corporate news, and the impact on CDS spreads. Global Finance Journal.

Nofsiger, J., Sulaeman, J., Varma, A., 2016. Institutional investors? Socially responsible investments: It just makes (economic) sense. Working paper. University of Alaska Anchorage, National University of Singapore, and Illinois State University.

Nyilasy, G., Gangadharbatla, H., Paladino, A., 2014. Perceived greenwashing: The interactive effects of green advertising and corporate environmental performance on consumer reactions. Journal of Business Ethics, 125 (4), 693–707.

Obaidullah, M., Shirazi, N. S., 2014. Integrating philanthropy with microfinance: Models of community empowerment. In: Social Impact Finance. Springer, pp. 75–96.

Orlitzky, M., Schmidt, F. L., Rynes, S. L., 2003. Corporate social and financial performance: A meta-analysis. Organization Studies, 24 (3), 403–441.

Painter, M., 2020. An inconvenient cost: The effects of climate change on municipal bonds. Journal of Financial Economics, 135 (2), 468–482.

Parhankangas, A., Renko, M., 2017. Linguistic style and crowdfunding success among social and commercial entrepreneurs. Journal of Business Venturing, 32 (2), 215–236.

Pedersen, L. H., Fitzgibbons, S., Pomorski, L., 2020. Responsible investing: The ESG-efficient frontier. Journal of Financial Economics.

Pedrini, M., Bramanti, V., Minciullo, M., Ferri, L. M., 2016. Rethinking microfinance for developed countries. Journal of International Development, 28 (2), 281–302.

Peress, J., 2014. The media and the diffusion of information in financial markets: Evidence from newspaper strikes. The Journal of Finance, 69 (5), 2007–2043.

Pietraszkiewicz, A., Soppe, B., Formanowicz, M., 2017. Go pro bono: Prosocial language as a success factor in crowdfunding. Social Psychology, 48 (5), 265–278.

Pope, D., Sydnor, J., 2011. What's in a picture? Evidence of discrimination from prosper.com. Journal of Human Resources, 46, 53–92.

Reimers, N., Gurevych, I., 2019. Sentence-BERT: Sentence embeddings using siamese BERT-networks. arXiv preprint arXiv:1908.10084.

Riedl, A., Smeets, P., 2017. Why do investors hold socially responsible mutual funds? The Journal of Finance, 72 (6), 2505–2550.

Robinson, M. S., 2001. The microfinance revolution. Sustainable finance for the poor. Washington, DC. The World Bank.

Robinson, P. B., Sexton, E. A., 1994. The effect of education and experience on self-employment success. Journal of Business Venturing, 9 (2), 141–156.

Rozin, P., Royzman, E. B., 2001. Negativity bias, negativity dominance, and contagion. Personality and Social Psychology Review, 5 (4), 296–320.

Schestag, R., Schuster, P., Uhrig-Homburg, M., 2016. Measuring liquidity in bond markets. Review of Financial Studies, 29 (5), 1170–1219.

Schicks, J., 2014. Over-indebtedness in microfinance – an empirical analysis of related factors on the borrower level. World Development, 54, 301 – 324.

Schröder, M., 2004. The performance of socially responsible investments: Investment funds and indices. Financial Markets and Portfolio Management, 18 (2), 122.

Schulz, A. J., Israel, B. A., Zimmerman, M. A., Checkoway, B. N., 1995. Empowerment as a multi-level construct: Perceived control at the individual, organizational and community levels. Health Education Research, 10 (3), 309–327.

Schunck, R., 2013. Within and between estimates in random-effects models: Advantages and drawbacks of correlated random effects and hybrid models. The Stata Journal, 13 (1), 65–76.

Serafeim, G., Yoon, A., 2021. Stock price reactions to esg news: The role of ESG ratings and disagreement. Harvard Business School Accounting & Management Unit Working Paper (21-079).

Shishlov, I., Morel, R., Cochran, I., 2016. Beyond transparency: Unlocking the full potential of green bonds. Institute for Climate Economics Report.

Shiu, Y.-M., Yang, S.-L., 2017. Does engagement in corporate social responsibility provide strategic insurance-like effects? Strategic Management Journal, 38 (2), 455–470.

Sparkes, R., Cowton, C. J., 2004. The maturing of socially responsible investment: A review of the developing link with corporate social responsibility. Journal of Business Ethics, 52 (1), 45–57.

Spence, M., 1973. Job market signaling. Quarterly Journal of Economics, 87 (3), 355–374.

Spence, M., 2002. Signaling in retrospect and the informational structure of markets. American Economic Review, 92 (3), 434–459.

Stewart, F., 2005. Groups and capabilities. Journal of Human Development, 6 (2), 185–204.

Swain, R. B., Floro, M., 2012. Assessing the effect of microfinance on vulnerability and poverty among low income households. Journal of Development Studies, 48 (5), 605–618.

Swain, R. B., Wallentin, F. Y., 2009. Does microfinance empower women? Evidence from self-help groups in india. International Review of Applied Economics, 23 (5), 541–556.

Taleb, W., Le Guenedal, T., Lepetit, F., Mortier, V., Sekine, T., Stagnol, L., 2020. Corporate ESG news and the stock market. Available at SSRN.

Tang, T. T., 2009. Information asymmetry and firms' credit market access: Evidence from Moody's credit rating format refinement. Journal of Financial Economics, 93 (2), 325 – 351.

Tchouassi, G., 2011. Microfinance, inequality and vulnerability: Empirical analysis from central african countries. Journal of Development and Agricultural Economics, 3 (3), 150–156.

Tetlock, P. C., 2007. Giving content to investor sentiment: The role of media in the stock market. The Journal of Finance, 62 (3), 1139–1168.

Thorp, R., Stewart, F., Heyer, A., 2005. When and how far is group formation a route out of chronic poverty? World Development, 33 (6), 907–920.

Underwood, T., 2006. Overview of the microcredit sector in Europe 2004-2005. European Microfinance Network.

Unger, J. M., Rauch, A., Frese, M., Rosenbusch, N., 2011. Human capital and entrepreneurial success: A meta-analytical review. Journal of Business Venturing, 26 (3), 341–358.

UNIDO, 2018. Webpage – gender equality and the empowerment of women. Https://www.unido.org/our-focus/cross-cutting-services/gender-equality-and-empowerment-women, last verified 2018-09-03.

Van Duuren, E., Plantinga, A., Scholtens, B., 2016. ESG integration and the investment management process: Fundamental investing reinvented. Journal of Business Ethics, 138 (3), 525–533.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., Polosukhin, I., 2017. Attention is all you need. In: Advances in neural information processing systems. pp. 5998–6008.

Walker, K., Wan, F., 2012. The harm of symbolic actions and green-washing: Corporate actions and communications on environmental performance and their financial implications. Journal of Business Ethics, 109 (2), 227–242.

Wulandari, F., Schäfer, D., Stephan, A., Suna, C., 2018. The impact of liquidity risk on the yield spread of green bonds. Finance Research Letters, 27, 53–59.

Yu, E. P.-y., Van Luu, B., Chen, C. H., 2020. Greenwashing in environmental, social and governance disclosures. Research in International Business and Finance, 52, 101192.

Yum, H., Lee, B., Chae, M., 2012. From the wisdom of crowds to my own judgement in microfinance through online peer-to-peer lending platforms. Electronic Commerce Research and Applications, 11 (5), 469–483.

Zaman, H., 1999. Assessing the poverty and vulnerability impact of micro-credit in Bangladesh: A case study of BRAC. The World Bank.

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

Ziegler, T., Reedy, E., Le, A., Zhang, B., Kroszner, R. S., Garvey, K., 2017. The Americas alternative finance industry report 2017. Cambridge Center for Alternative Finance, the Cambridge Judge Business School.