



ESG Disclosure and Short-term Return

After IPO

*A study of the ESG indicators of Norwegian new listing companies in 2020
and their returns*

Mengting Zhao and Xifei Pan

Supervisor: Tore Leite and Xunhua Su

Master thesis, Economics and Business Administration

Major: Business Analytics and Finance

NORWEGIAN SCHOOL OF ECONOMICS

This thesis was written as a part of the Master of Science in Economics and Business Administration at NHH. Please note that neither the institution nor the examiners are responsible – through the approval of this thesis – for the theories and methods used, or results and conclusions drawn in this work.

Acknowledgements

We feel deeply honored to have had you as our supervisor, as well as our career path guide and humorous friend. Professor Su Xunhua, thank you for your generosity in providing us with your knowledge and assistance. We will always miss you. R.I.P.

At such a fraught moment, we appreciate Professor Tore Leite for accepting to guide us halfway. Thank you for your advice and encouragement to let us be able to finish our work. We also want to express our gratitude to our families, for no matter what time it is, you always support us. Finally, we would like to thank our school, NHH, for its free resources, which allow us to find any literature we need for this article at any time.

Norwegian School of Economics

Bergen, December 2021

Mengting Zhao

Xifei Pan

Abstract

With the concept of sustainability gaining traction in the business world, ESG indicators are beginning to be used by investors as one of their investment metrics. While global finances were hit hard by the pandemic in 2020, there was an extraordinary surge of new listing companies in Norway. In this paper, we try to investigate whether the ESG disclosure in the prospectus or information document (admission document) filed by Norwegian companies during their listing procedure in 2020 has an impact on their short-term share price returns. We divide the textual analysis into two dimensions, ESG term frequency and sentiment analysis. We use the ESG basis vector to construct our own ESG dictionaries and use them as the benchmark for our textual analysis. We mainly use multiple linear regression to explore the relationship between variables but also attempt to fit several machine learning models. From the results of the regression analysis, underpricing is the most sensitive to ESG indicators, as environmental term frequency, environmental sentiment and governance sentiment are all significantly correlated with it. In contrast, 3-day return is not significantly related to any ESG indicator; 1-week return and 1-month return are significantly related to environmental term frequency and governance sentiment, respectively.

Keywords: *ESG, IPO, 2020, short-term return, textual analysis, Norwegian stock market*

Contents

1	Introduction	1
1.1	Motivation	4
1.2	Purpose and research questions	5
1.3	Research structure	6
2	Concepts and Theoretical Foundations	9
2.1	Introduction of ESG	9
2.2	Introduction of IPO	10
2.2.1	Pros and cons of going public	11
2.2.2	IPO and underpricing	12
2.3	Introduction of private placement	13
2.3.1	Why choose private placement?	14
2.3.2	Listing requirements comparison	14
2.4	Related theoretical foundations	15
2.4.1	Information asymmetry theory	15
2.4.2	Behavioral finance theory	18
2.4.3	The ESG criteria in IPO	19
3	Literature Review	20
3.1	Background of ESG	20
3.1.1	The relationship between ESG and CSR	20
3.1.2	European and Norwegian regulations	21
3.1.3	ESG ratings	22
3.2	Previous research on the impact of ESG	23
3.2.1	ESG and financial performance of companies	23
3.2.2	ESG and investment performance of stocks	25
3.3	Previous research on ESG and IPO	26
4	Data and Methodology	30
4.1	Independent variable	30
4.1.1	Material	30
4.1.2	Vector basis for ESG dictionary	31
4.1.3	ESG dictionary construction	32
4.1.3.1	Word2vec algorithm	32
4.1.4	ESG indicators	35
4.1.4.1	ESG-term frequency	35
4.1.4.2	ESG sentiment score	35
4.2	Independent variable	36
4.3	Control variables	38
4.3.1	Listing market (document type)	38
4.3.2	Covid-19 measurement	38
4.3.3	Green company	40
4.4	Variables summary	41
4.5	Multiple regression	41
4.6	Attempts in machine learning	43

5	Analysis and Results	44
5.1	Descriptive statistics	44
5.2	Multiple regression	46
5.2.1	ESG-term frequency	47
5.2.2	Sentiment of ESG content	51
5.2.3	Summary of multiple regression results	55
5.3	Machine learning	55
6	Discussion	58
6.1	Discussion of textual analysis	58
6.2	Discussion of regressions	59
6.3	Implications	61
6.4	Limitations	62
7	Conclusion	65
7.1	Summary of findings	65
7.2	Further research	66
	Appendix	73
A1	Output of ESG initial word list by using Word2vec algorithm	73
A2	ESG dictionary	74
A3	Part of Python-code for textual analysis	74

List of Figures

1.1	Global growth of sustainable investment 2016-2020	2
1.2	New listings on Euronext Growth Oslo 2016-2020	3
2.1	The range of ESG	9
2.2	Underpricing percentages in selected European countries and USA	13
3.1	The Action Plan on Sustainable Finance	21
4.1	Skip-gram model	33
4.2	CBOW model	33
4.3	Social extended words	34
4.4	Textual analysis steps	36
4.5	Changes in four indicators of 48 companies	37
4.6	New Covid19 cases in Norway (smoothed)	39
5.1	Correlation Matrix	46
A1.1	Social extended words	73
A1.2	Environmental extended words	73
A1.3	Governance extended words	73

List of Tables

2.1	Listing requirements on two Norwegian stock marketplaces	15
3.1	Summary of the previous literature on the impact of ESG	26
4.1	Two different document types	31
4.2	ESG vector basis	32
4.3	Word <i>clean</i> similarity	34
4.4	Variables summary	41
5.1	Descriptive statistics	44
5.2	Regression models description	47
5.3	Regression of underpricing on ESG freq	48
5.4	Regression of 3-day return	49
5.5	Regression of 1-week return	50
5.6	Regression of 1-month return on ESG freq	51
5.7	Regression of underpricing on ESG senti	52
5.8	Regression of 3 day and 1 week return on ESG senti	53
5.9	Regression of 1-month return on ESG senti	54
5.10	Multiple regression summary	55
5.11	Machine Learning Results	57
A2.1	Final ESG Dictionary	74

1 Introduction

In the last century, the economy developed to a certain level and humans gradually realized the environmental and resource pressures brought about by economic growth, urbanization, and population growth. Due to the emergency, world leaders convened at the United Nations in New York in 2015 to discuss the need for more sustainable global development. They devised Agenda 2030, a set of seventeen sustainable development objectives including economic development, environmentally responsible production and consumption for the year 2030. A few months later, a second conference was held in Paris to address climate change and greenhouse gas emissions specifically: 195 nations signed the Paris Agreement. The convention's objective is to limit the average global temperature rise below 2 degrees Celsius over pre-industrial levels.

Europe alone would need at least €180 billion in climate expenditures over the next decade to fulfil the Paris Agreement's goals (Eurosif, 2018). Additionally, it is projected that achieving the UN Sustainable Development Goals would need yearly global expenditures ranging from \$5 to \$7 trillion. Thus, the financial system is critical for channeling significant money toward the Paris agenda's objectives.

With the gradual improvement of investors' awareness of environmental protection and social responsibility, the concept of responsible investment has entered the eyes of investors. Responsible investment means that while investors pay attention to corporate financial performance, they must comprehensively consider the company's impact on society, the environment, and the overall economy to make investment decisions. The content covered by the responsibility investment is constantly evolving with the changes of the times, developing and forming three pillars of value, namely the E (Environment), S (Social Responsibility) and G (Corporate Governance) are the concepts of ESG investment.

The scale of ESG investment has maintained rapid growth for many years. The GSIR (2020) report pointed out that as of the beginning of 2020, the total sustainable investment assets in the world's five largest markets (Europe, United States, Canada, Australia/New Zealand, and Japan) were US\$35.3 trillion, a 15% increase in 2018-2020 and 55% in 2016-2020. Among them, the scale of ESG investment has increased significantly,

with a four-year growth of 143%, reaching US\$25.1 trillion, of which Europe and the United States accounted for 80% of the global ESG investment scale. Japan, Canada, Australia/New Zealand are regions that have experienced rapid growth in recent years.

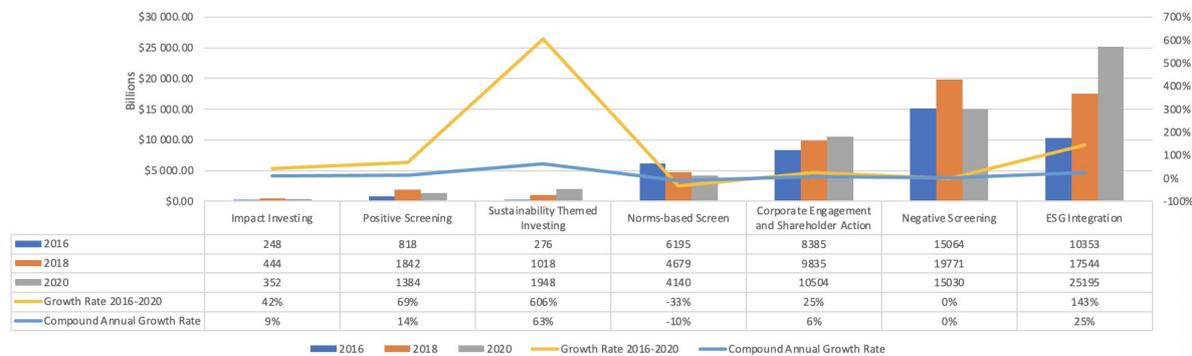


Figure 1.1: Global growth of sustainable investment 2016-2020

ESG integration is seen as a solution to many global issues, including climate change, human rights, and legal compliance. Socially aware investors allegedly utilize ESG criteria to screen possible investments, either due to the intrinsic value of their sustainability investment or to avoid investing in firms that do not adequately manage their environmental, social, and governance activities. However, the proliferation of ESG practices unavoidably impacts companies' operations, financial performance, and culture. Thus, it is questioned whether market players really consider ESG factors when making choices.

Recognizing and evaluating current trends, developing sustainability plans include incorporating sustainability into a business's strategy. This integration occurs on many levels and impacts numerous elements of a company, including mergers and acquisitions, credit ratings, and IPOs.

Integrating sustainability into a company's vision is a critical component of the long-term vision required for a successful Initial Public Offering (IPO). Sustainability issues cannot be addressed in isolation from an IPO process since they influence shareholder perceptions of the business and, therefore, the IPO itself.

Compared to the rest of Europe, the Nordic nations have a relatively modest population and GDP but have recently accounted for a sizable portion of Europe's IPOs. The Nordic

countries' stability and transparent financial markets remain advantageous qualities in market uncertainty. Historically, Sweden has been the primary driver of Nordic IPOs. However, there was a significant surge in Norwegian IPOs in 2020. With 59 new listings, including private placements, IPOs, and all other types of listings, the nation is considerably over the 10-year average of 13 every year despite Covid-19 and decreasing oil prices. It is worth noting that there was a considerable increase in IPOs of green companies in industries such as renewable energy, carbon capture and storage, and waste management. Merkur Market, a growth-oriented Norwegian marketplace renamed a "Euronext Growth Market," has experienced a significant increase (see Figure 2). Merkur has grown in popularity as a listing destination, particularly in this modern virtual world, because of its simple, cost-effective, and rapid listing procedure. Another factor contributing to this is the involvement of well-known cornerstone investors and certain high-profile listings, which provide validity to the marketplace (Wiersholm, 2021). Folketrygdfondet, KLP, Delphi, Pareto, and JP Morgan are among the investors, while high-profile corporations include Kahoot, the Aker enterprises, and Meltwater.

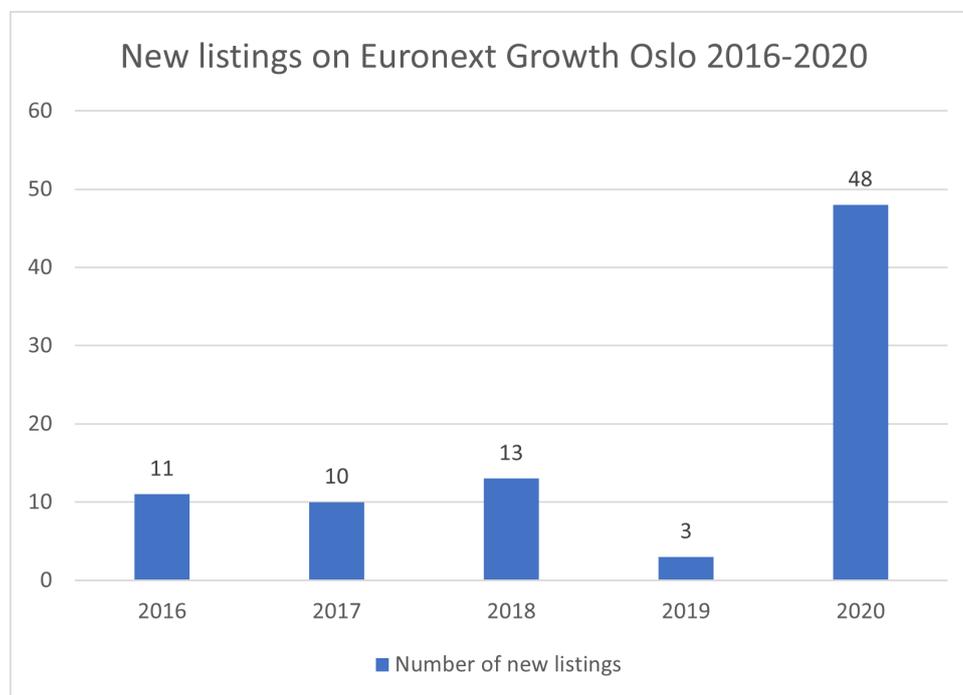


Figure 1.2: New listings on Euronext Growth Oslo 2016-2020

This thesis will examine the impact of Norwegian companies' ESG practices on their

listings and underpricing. However, although private placements were the majority of new listings in Norway, we did not find any study describing the performance or underpricing of private placements. Thus, we suggest that the extensive literature explaining IPO may be applied to private placements and the effects and incentives are essentially the same for both types of listings. Here we clarify that IPO activities in this article can refer to both public offerings and private placements and compare them in the following sections. We chose to study the Norwegian capital market in 2020, which includes the Oslo Stock Exchange (Oslo Børs) and Euronext Growth (formerly Merkur Market), with 48 firms (4 IPOs and 44 private placements). As such, it serves as a valid example for this thesis. Norwegian and international firms across various sectors opt to list on a Norway exchange due to the market's liquidity, transparency and diverse investor base. Such a market is an appealing subject for our thesis study.

Another reason Norway is an excellent target for our research is that no previous study has been undertaken in this area. Indeed, there is a substantial body of research on IPOs and underpricing at the moment, but none of it considers firms' ESG policies as a critical factor in this phenomenon. With the continuous improvement of the domestic ESG system, ESG investment has received more and more attention, and its research value will continue to increase. Therefore, it is an opportunity for us to add to the current literature with a new, fresh perspective and stimulate the Norwegian market's interest in sustainable finance.

1.1 Motivation

We took classes at NHH as Master students in Economics and Business Administration that introduced us to sustainable financing and reporting. This subject piqued our interest, and we decided to delve more into the sustainable aspects of the financial industry. The original intention was to research ESG and investment management, specifically how the performance of Norwegian ESG funds differs from that of conventional funds. However, we quickly found that relevant studies had already been undertaken on this subject in recent years. As a result, we decided to look for another topic. After consulting with our professor Su, we recognized the existence of an area in the field of IPO and underpricing. Thus, we outlined our final research question to determine the impact

of ESG factors on listings, particularly concerning the phenomenon of underpricing.

There is a wealth of literature on IPOs and underpricing since they are major market events. On the other side, we observed a research gap: just a few studies have been conducted on the influence of ESG issues on the undervaluation of IPOs. ESG is a prominent issue in today's world, and many businesses are attempting to catch up, but what is the actual effect of sustainability on markets? We feel that our study can contribute to this issue by providing a vital jigsaw piece. Analyzing the influence of ESG variables on underpricing in Norwegian IPO activities may provide us with reliable data on whether or not the market cares about sustainability in the Norwegian capital market. We picked the Norwegian market because it is one of the most transparent, and we saw a surge in Norway's 2020 IPO market. Data is a critical component of our study. We discovered that data on Norwegian listings and ESG indicators could be available on the internet and in the Oslo Stock Exchange's database. At this stage, we determined that the study is feasible and can provide value to the IPO and Private Placement literature in terms of sustainability.

Furthermore, most people think that the Covid-19 could be an important factor when investigating different subjects. No wonder that the pandemic not only hindered international trade and investment activities but also impacted the global industrial chain and supply chain, causing the world economy to face recession, and it also caused shocks to the stock market. However, considering that the pandemic outbreak is a sudden exogenous event, whether the virus impacts the underpricing of listings also triggers our interest.

1.2 Purpose and research questions

The article seeks to address the following question: **Does the ESG disclosure in the company's listing in the Norwegian market in 2020 affect its short-term return?**

The purpose of this thesis is to study the market's assessment of ESG elements in the context of IPOs and private placements by determining the extent to which they are

underpriced. In other words, our research seeks to determine if the market seriously considers environmental, social, and governance factors in connection with different listings.

First, this article hopes to use textual analysis to see whether ESG in the IPO prospectus or private placement admission documents in the Norwegian market has effective investment guidance, meaning whether sustainability factors can result in underpricing. One method to gauge the true value of ESG in practice is to examine the market response to ESG companies' listing by calculating the underpricing. We may refer to how market players perceive a firm's ESG features favorably or negatively based on the underpricing situation.

Second, from the perspective of investor sentiment, explore the impact of ESG on the underpricing of listed companies, provide a new perspective for interpreting the effect, and supplement research in related fields. This research aims to demonstrate the impact of ESG practices on the market via the lens of IPOs and private placements. We want to ascertain the market's view of ESG in the context by examining the connection between underpricing and the ESG performance of businesses going public.

Third, this thesis may offer empirical data on market perceptions of environmental, social, and governance issues. It would be able to determine if the market is really concerned with ESG issues in the context of listings. The research contributes to incentivizing businesses to include ESG factors into their operations. It helps review Norway's ESG investment and the current development status of the ESG rating system, analyze and summarize the recent problems at this stage, and give corresponding suggestions.

1.3 Research structure

The present study is quantitative since we gathered numerical data regarding firms' listings and ESG characteristics from official websites rather than via interviews or personal information. As a result, we compose this thesis from an objectivist perspective. The procedure of obtaining financial data using web tools does not need our subjective interpretation and judgment. Additionally, our study applies to a cross-sectional design. According to Bryman (2012), this study's data collection includes "more than one

instance" of IPO transactions over a specific time period to see whether there is an ESG trend among IPO businesses and how the market's impression of ESG develops over time. However, we suggest that our study's breadth is adequate to ensure its reliability and validity.

To be more specific, the thesis is separated into chapters in which we have incorporated all pertinent issues that we believe are important to complete the research properly. The article uses six parts for argumentation and study:

First of all, the introduction part of the article fully elaborates and analyses the relevant background, research motivation, research purpose and research structure of the argument of this article.

The second part introduces the related concepts and theories to lay a good foundation for understanding the essential subjects - ESG rating system, IPO and Private Placement, the phenomenon of underpricing, information asymmetry theory and efficient market theory.

The third part analyses previous literature in this area, showing their research on the reasons for underpricing, relations between ESG and financial performance, and, most importantly, findings of ESG and underpricing, although there are only a few studies.

The fourth part introduces the data and methodology we employed, which is also the core focus of this article. The description is based on four aspects: sample selection, data source, variable description, and specific methodology. We selected the listed companies in the Norwegian market from 2020 as the research objects. Based on data availability, choose available companies' entering materials as samples and conduct screening and subsequent processing procedures, including the measurement basis, ESG dictionary construction and variable construction models, etc.

The fifth part is empirical analysis and results, including descriptive statistics, correlation testing, multiple regression analysis, and attempts in machine learning. The previous hypotheses were verified by analyzing the results of the actual proofs, and the analysis results were performed using the tool-variable method at the same time to ensure the rationality of the research model and the reliability of the research conclusions.

The sixth part is research conclusions and prospects. Based on the previous proof

results and theoretical analysis, the findings of this article are summarized, and relevant management suggestions are proposed based on the conclusions obtained. Finally, we point out the shortcomings of this article and provide a reference for further research in the future.

2 Concepts and Theoretical Foundations

2.1 Introduction of ESG

In 2004, the concept of ESG was put forward for the first time by the United Nations, pointing out that embedding ESG factors in the capital market has good commercial significance and is conducive to the sustainable development of society and the economy. Nowadays, environmental, social, and governance (ESG) issues have become inextricably linked to the domestic economy and almost every company (Henisz et al., 2019). Due to the lack of agreement on the exact definition of ESG elements, each company has its own presentation for ESG as well as an ESG framework or policy for implementation (see Figure below).

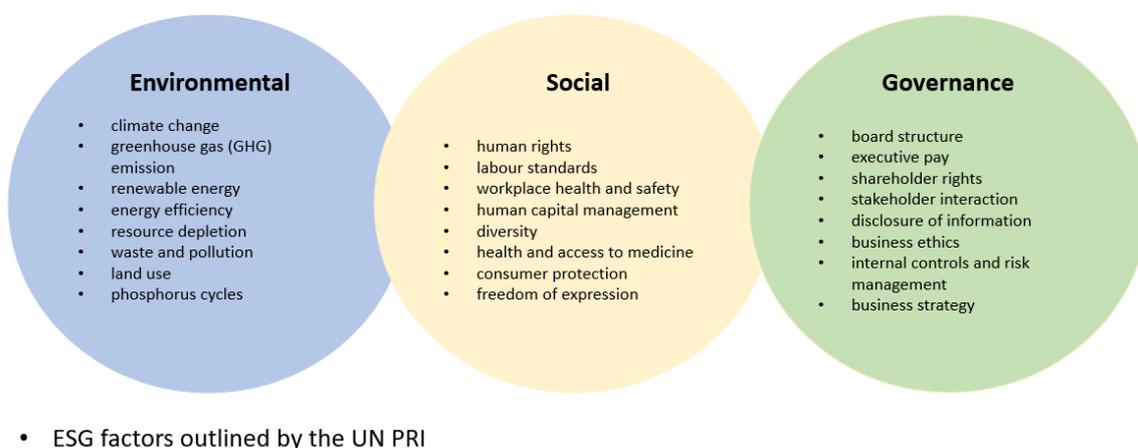


Figure 2.1: The range of ESG

Many regulatory agencies and non-governmental organizations also promote the development of ESG by emphasizing ESG information disclosure. Regulatory agencies in some countries require listed companies to disclose ESG-related content in their reports. Currently, most of the top 50 economies globally have issued ESG-related disclosure policies, and most of them are mandatory disclosure requirements. In terms of investors, investors in many national sovereign funds, pension funds and other financial institutions have voluntarily applied ESG investment concepts in their investment activities. In 2006,

the United Nations established the Principles for Responsible Investment (UN-PIR). As one of the most important international organizations in socially responsible investment, UN-PRI has always been committed to promoting major investment institutions to consider ESG factors in their investments and helping signatories improve their responsible investment capabilities. Many financial institutions worldwide have joined UN-PRI one after another, signed PRI partnerships, and consciously abide by PRI investment principles and incorporate ESG concepts into the investment decisions of their fund products. As of the end of 2018, UN-PRI had 1714 contracted institutions, and the total assets covered by it exceeded US\$80 trillion. UN-PRI also conducted a survey of more than 900 signatories in 48 countries. The results showed that 94% of asset owners and 93% of investment managers had formulated formal responsible investment policies.

ESG is a metric used to assess a company's social responsibility (CSR). It refers to activities done by businesses that go beyond their legal responsibilities to society and the environment. We shall use the term ESG to refer to both ESG and CSR activities in this article, and the definitions of associated words will be addressed in detail later.

It is debatable if increasing ESG standards require businesses to forgo a part of the profits. Economic efficiency has an impact on how market players see ESG integration. Given that the real-world data presented earlier raised the question of whether market participants really care about incorporating ESG into their decision-making, this issue should be discussed again in academic research.

2.2 Introduction of IPO

Initial Public Offering (IPO) is how private companies become publicly traded by first selling stock to the general public. The primary reason for coming public with capital expenditure and new investment plans is to provide funding for the firm's recent operations and growth (Eckbo et al., 2008). Numerous studies have shown that going public may increase a firm's value by boosting the liquidity of its shares. It makes the firm's actions visible to all stakeholders, and investors can monitor the firm's operations constantly, increasing its worth.

The underlying motivation for companies to go public has been studied extensively for

a long time. Consequently, many theories have explained this choice from the firm's viewpoint. The main objective is financial, followed by non-financial goals such as advancing the public interest (Ritter & Welch, 2002). As Zingales(1995) asserts, becoming public makes it simpler to attract acquirers and more likely to result in a higher value than an outright sale. Market-Timing Theories predict that companies would delay their IPOs if their stock offering is undervalued and IPO volume will rise during economic cycle expansions. Another reason companies go public is to optimize their ownership structure as the next stage of their life cycle approaches, a concept known as life-cycle theory. Black & Gilson(1998) and Chemmanur & Fulghieri(1999) proved that venture investors' exit is essential for the company to go to the next stage of its life.

2.2.1 Pros and cons of going public

Why does a business go public? The primary objective of going public is to generate new money to fund future operations (Geddes, 2003). However, it has been shown that companies do not go public for the goal of post-IPO investment and development, but rather to use this new resource to rebalance their financial accounts after a necessary life stage (Rydqvist & Högholm, 1995). According to Ritter and Welch(2002), the principal objectives remain financial. Therefore, non-financial objectives such as raising the public interest play a small role.

Firms must weigh the benefits and drawbacks of going public. The first and most obvious effect is that going public impacts the firm's value. Investors value listed companies more than private enterprises because listed firms are required to publish their IPO prospectus and annual report, thus decreasing information asymmetry and uncertainty. Additionally, investors are prepared to pay a premium for the liquidity of public company shares (Amihud et al., 1988). According to Geddes(2003), a company would earn more than 30% on its stock price if it went public or remained private. Moreover, the IPO process requires businesses to have a clear business plan under the oversight of public ownership in return for future access to additional financial resources. When public investors from the stock market watch the operations of public companies, their value may grow as well (Holmström & Tirole, 1993).

Apart from the economic benefits, Geddes (2003) and Ritter & Welch(2002) highlight the non-economic benefit of enhancing the firm's reputation. This advantage is often referred to as the firm's public image or public interest. Additionally, consumers with more information are more likely to purchase goods and pay attention to businesses' actions.

On the other hand, going public is a double-edged sword that businesses must wield prudently - becoming public reduces a company's degree of privacy as the amount of information disclosed grows, such as corporate operations and activities, management expenses, contracts, and consumers. The company may be under pressure to live up to shareholder expectations, and it must explain every detail if necessary (KPMG, 2015). Furthermore, the business must bear the significant costs associated with going public. On average, underwriters charge 4%-7% of the total proceeds. (PwC, 2017). Additionally, the business must pay a legal cost to prepare IPO papers, which includes the writing and approval of all contracts.

2.2.2 IPO and underpricing

Numerous academics have noticed and generated hypotheses regarding the features of IPOs. One of them is the underpricing issue. Underpricing is measured as the percentage difference between the offer and market prices at the close of the first trading day of a stock (Dolvin, 2012). Thus, if a firm offers shares to the public at \$10 and the offering closes at \$12, the underpricing is 20%. Underpricing is often in the range of 15-20% (or more), and it is usually believed that underpricing is the primary cost of issuance. On average, the IPO price is lower than the first day's closing price, and this first-day abnormal return is statistically significant (Chang, 2011). Ritter(2003) has also researched other countries to analyze how IPO results differ by nation. His 2003 research placed Norwegian IPOs in the low end of the distribution while observing high examples of IPO in China, Malaysia, and Brazil, with averages of 256.9%, 104.1% and 78.6%. The graphic below illustrates the average IPO return across a sample of European nations with parallels to the Norwegian stock market, as well as notable outliers to demonstrate how European countries underprice their IPOs.

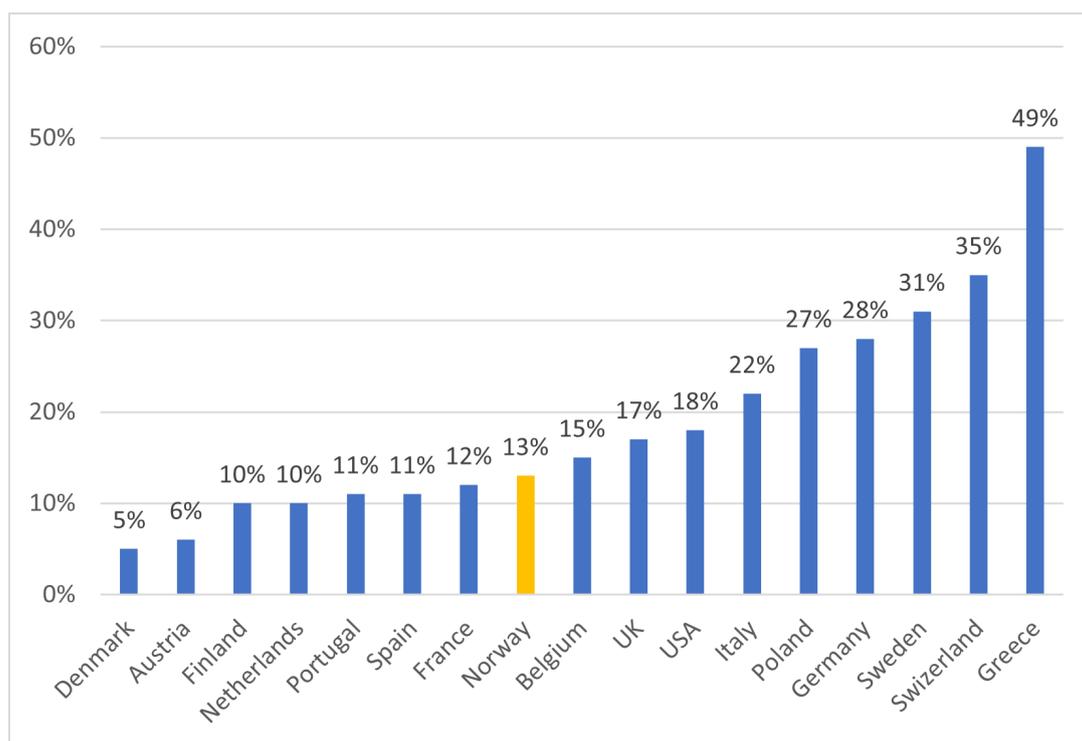


Figure 2.2: Underpricing percentages in selected European countries and USA

To be more specific, the underpricing of Norwegian IPOs is based on Emilsen et al's.(1997) research. The authors examined the Norwegian market from 1984 to 1996 and found an average underpricing of 12.5%. Fjesme(2011) estimates the initial return for 1993–2007 at 8%. Ritter(2021) collated data from several studies in different periods and determined that the average underpricing in the Norwegian market from 1984 to 2018 was 6.7%.

2.3 Introduction of private placement

Private placements vary from IPOs in that they are a less complicated and time-consuming method of raising cash and listing on a marketplace. According to Investopedia, securities issued for sale solely to approved buyers, such as investment banks, pension funds, or mutual funds, are known as private placement offers. Some high-net-worth people may also use these options to acquire shares. Private placements are used by businesses to raise a smaller amount of money from a few investors. These securities are excluded from many of the financial reporting requirements of public offerings, saving the issuing business time and money.

Companies must submit an admission document explaining their business before being listed as private placements. Among other criteria, the information document must include specific liability disclosures and a description of the issuers' business. According to Euronext regulations, private placements must have occurred within a year of the anticipated date of initial admission. Unless an exception is granted, the private placement must include at least three entities, not including significant insiders. Private placements may be made by issuing new shares or selling existing shares by insiders or major shareholders owning more than 20% of the company.

2.3.1 Why choose private placement?

Numerous articles have provided reasons for some corporations' private placement decisions. According to the studies of Wruck(1989) and Hertz et al.(1993), private placements are utilized to attract investors, such as monitoring or certification investors. These investors verify that businesses are operated optimally or certify corporate values. Zingales(1995) implies that private placements are employed when buyers premium private control advantages. Before listing on the OSE via private placements, several firms conduct follow-on public and employee offers.

There is little research analyzing the merits of a private vs. a public offering for firms seeking to list on a public marketplace. However, the less restrictive guidelines are likely to explain why private placements are more prevalent than IPOs. Companies listed on Euronext Growth may meet criteria via a private offering, avoiding the significantly more involved and expensive process of conducting an IPO. It is worth noting that many firms that list on Euronext Growth do so as a stepping stone to listing on Euronext Expand or Oslo Børs's central marketplace, and are open about their goals.

2.3.2 Listing requirements comparison

Euronext Growth is a multilateral trading facility, which is essentially exempt from the Norwegian Securities Trading Act (Abrahamsen & Sveen, 2021). The listing procedure on Euronext Growth is designed to make it simple for companies to seek a listing on the

exchange. Its listing procedure differs from the main list due to less stringent standards. Table 1 summarizes the significant variations in the criteria for listing on the Oslo Stock Exchange's two markets (Oslo Børs, 2020).

Table 2.1: Listing requirements on two Norwegian stock marketplaces

	Oslo Børs	Euronext Growth (formerly Merkur Market)
Marketplace Status	Stock exchange listing under EU requirements and Norwegian Securities Trading Act.	Multilateral trading facility.
Admission Process Duration	4-8 weeks	1-2 weeks
Spread of Share Ownership	25%	15%
Minimum Price Per Share	NOK 10	NOK 1
Number of Shareholders	500, with at least NOK 10,000 holding value at the time of admission.	30, with at least NOK 5000 holding value at the time of admission.
Market Capitalisation	NOK 300 million	No requirement
Financial History	At least three years of financial history.	At least one financial report, either interim or annual.
Prospectus	EEA compliant prospectus, supervised by the Norwegian Financial Supervisory Authority.	Information document (formerly admission document), which is less extensive than an EEA compliant prospectus.
Liquidity	Must demonstrate sufficient liquidity for 12 months of operations.	Not necessary to demonstrate sufficient liquidity for 12 months of operations.

2.4 Related theoretical foundations

2.4.1 Information asymmetry theory

In the actual business activities of enterprises, it is difficult to obtain sufficient comprehensive and effective information, and this process often costs. Usually, they can only get relevant information based on the report disclosed by the company to

society, and this information needs to be evaluated. Thus, companies' public disclosure of information related to the environment, social responsibility, and corporate governance can substantially reduce the current market situation of information asymmetry. The disclosure of ESG information by enterprises increases the amount of information that investors can obtain, and helps investors to make investment decisions after comprehensive analysis. The more serious the information asymmetry, the more likely investors will expect a higher risk compensation. However, at the same time, because the current ESG information disclosure standards are not yet complete, companies may selectively release information that is beneficial to them while stealing unfavorable information. This in turn exacerbates the information asymmetry to some extent.

In recent years, scholars have paid more and more attention to the signaling effect of corporate social responsibility in the capital market. Research has found that investors' understanding of this signal will affect their requirements and evaluation of enterprises. If investors regard the company's actions to improve its ESG performance as measures taken to achieve strategic goals, and the strategic goals can meet the expectations of stakeholders, in that case, these behaviors of the company can convey to investors that the company's business is in good condition. To a certain extent, this can avoid the problem of adverse selection caused by insufficient information collection due to poor information or prejudice.

Regarding IPO, here we include three main theories under the asymmetry information model:

First, Rock(1986) established the Winner curse's model, stating that underpricing is due to knowledge asymmetry between the business (the issuer) and the investors. Investors are classified as informed or uninformed. Informed investors may determine if an IPO is underpriced with better knowledge: the offer price is less than the market price, and they will only look into underpriced IPOs. Meanwhile, the uninformed lack the methods necessary to correctly evaluate an IPO and then subscribe to both offerings. As a result of both kinds of investors participating, the underpriced will be oversubscribed, and the uninformed will get a tiny percentage of underpriced IPOs but will get all overpriced IPOs. Due to the possibility of a low profit, demand for new share issuances among uninformed investors would likely decline, depending on the performance of the IPOs. Companies and underwriters think they must actively make the stocks underpriced to

generate demand and provide a fair return for uninformed investors. Issuers are inclined to reward uninformed investors by accepting the risk of underpricing for participating in an IPO when this investor has adequate knowledge.

Given that the Winner's curse hypothesis has been established over a long time and is often utilized to explain underpricing, numerous scholars have undertaken many implications and studies to examine this theory. One significant conclusion of the IPO research is that "the more the ex-ante uncertainty, the bigger the underpricing" (Beatty & Ritter, 1986). Since the value before IPO is unclear, uninformed investors will perceive an increased likelihood of getting the winner's curse. Historically, many proxies for ex-ante uncertainty have been broadly used in the literature. These proxies may be divided into four categories: company characteristics, offering features, prospectus disclosure and aftermarket information (Ljungqvist, 2008).

Second, the theories of information relevance take a different view of Rock's. Ljungqvist(2008) claimed that informed investors have an incentive to falsify positive information to convince the underwriter to give a lower price. Book-building is a frequent technique used to coerce investors into disclosing accurate information. Following the price investors suggest, they disclose beneficial information and get a greater share of stock distribution if they bid a higher price. However, underwriters continue underpricing the stocks to maintain an incentive for investors, or underpricing is utilized to compensate investors for practical information (Benveniste & Spindt, 1989).

Last, the signaling model is based on Rock's model's reversal premise that businesses have superior knowledge about themselves from outsiders. Allen and Faulhaber(1989) examine the underpricing phenomena for IPOs and argue why underpricing is a strategy to attract investors about their "real value." As with the previous asymmetric information model, the concept indicates that the higher level of ex-ante uncertainty about the firm's worth, the greater the degree of underpricing. It highlighted why, in a competitive market, businesses must underprice in order to differentiate themselves from rivals.

Regarding private placements, Hertz and Smith(1993) demonstrate that private placements may alleviate some of the information asymmetry associated with company quality and value. Recent research on issuance techniques indicates that high-quality enterprises with a higher level of information asymmetry prefer private placements

over public offers to decrease the costs of adverse selection and information creation (Maksimovic & Pichler, 2006). According to Sharpe and Woo(2005), the chance of employing private placements grows in direct proportion to the degree of information asymmetry and the reputation of privately owned businesses. These studies demonstrate that private placement is an effective issuance technique for mitigating information asymmetry about the issuing firm's quality.

2.4.2 Behavioral finance theory

Behavioral finance theory proposes that the intrinsic value of securities cannot determine the market price of securities, and the subject of investor behavior also affects the market price of securities. In addition to some uncertain factors, people's decision-making behavior is also affected by their internal motivations, resulting in systematic cognitive biases. It can be seen that the market price of stocks is not necessarily determined by the intrinsic value of stocks, the actual returns and expected returns of stocks are not usually the same, and they will be affected to a certain extent due to changes in investor behavior.

Welch(1992)'s cascades concept explains underpricing in behavioral finance. He believes that investors, despite their own knowledge, would depend on past investors' choices. This indicates that the underpricing may come from the first investors' good bids, which encouraged subsequent investors to bid in the IPO market. The first investors want more discounting to initiate positive cascades, thus increasing demand for IPOs. Ljungqvist et al.(2006) address the next behavioral explanation, saying that irrational investors create a "bubble" in the market, inflating IPO price over its actual worth. The issuers prefer to work with institutional investors to keep stocks and sell them later to sentiment investors. IPO underpricing compensates these cooperating investors for the chance that the "hot" period would expire sooner than anticipated. This process also accounts for the long-term underperformance of the stock after the IPO.

2.4.3 The ESG criteria in IPO

Numerous theories concerning underpricing in the IPO and ESG setting separately have been examined in the preceding sections. On the basis of these ideas, it is easy to see the critical nature of including ESG considerations while considering underpricing in the IPO scenario. According to asymmetric information theories, the degree of underpricing is determined by ex-ante uncertainty, often known as company risk. This means that if the business had additional knowledge, the market-determined amount of underpricing would be changed, and hence the IPO process would be affected as well. According to this, businesses would be more transparent about their ESG operations if they had an effect on underpricing and investor perception. Companies may disclose information regarding ESG problems in their prospectus or during the pre-IPO period. According to Park and Patel(2015), less transparency in the prospectus results in a greater likelihood of IPO underpricing. Adding precise information on environmental, social, and governance measures, this statement states, would reduce underpricing. To substantiate this argument, it is required to do empirical research on the problem of ESG underpricing.

On the other hand, the findings may suggest that ESG enterprises provide more financial advantages to clients. According to the signaling hypothesis (Allen Faulhaber, 1989), excellent businesses may want to communicate their true worth to potential investors by underpricing their IPOs. The answer to the issue of whether this theory applies to ESG enterprises is one worth investigating. Additionally, investors seeking value are becoming increasingly conscious of the ways in which environmental, social, and governance aspects may enhance or detract from a firm's value. These factors are important, particularly when it comes to investing. If investors are really concerned about ESG, they may initiate a cascade effect, in which subsequent investors act and respond to the pioneers' actions (Welch, 1992). Thus, if investors are more concerned about ESG, their demand for equities will increase, driving the price of the stocks up on the first day. To summarize, it is critical to include ESG practices into underpricing investigations since they have the ability to influence both underpricing and IPO success.

3 Literature Review

3.1 Background of ESG

3.1.1 The relationship between ESG and CSR

Since the concept of ESG investment appeared late, and its definitions and connotations overlap with Corporate Social Responsibility (CSR), many early scholars' studies on the environment, social responsibility and corporate governance are based on the CSR perspective. Thus, when summing up predecessors' related research in the ESG field, we need to combine them with the CSR field literature for comprehensive study. However, at the same time, it is also necessary to distinguish between the concepts of CSR and ESG.

The European Commission(2001) defines corporate social responsibility as "*A company incorporates social, environmental, ethical, human rights and consumer concerns into its business operations and core strategies*". In terms of definition, compared with CSR, the definition of ESG is relatively narrow. As social responsibility activities gradually become an essential business practice for enterprises, more and more companies choose to disclose social responsibility information in the form of CSR, corporate responsibility reports, corporate sustainability reports, corporate environmental reports, integrated corporate reports and more. Later, with the rise of ESG investment concepts, corporate ESG reports developed rapidly. From the perspective of the two reports, the main difference between the CSR report and the emerging ESG report is that the target audience of the CSR report is various stakeholders, including government regulatory agencies, employees, partners, communities, NGOs. The target audience group of the ESG report is mainly capital market participants, especially institutional investors. Companies have greater freedom in preparing CSR reports, while ESG reports generally have more detailed guidelines.

3.1.2 European and Norwegian regulations

Even while ESG criteria are not legally binding on the market, the increasing trend of incorporating ESG into business and investment decision-making necessitates the establishment of regulatory frameworks by governments and organizations. In 2018, the European Commission, the EU's executive branch, released the Action Plan on Sustainable Finance. The EU taxonomy was developed as a categorization system to achieve the EU's 2030 climate and energy goals and the European green deal's aims, providing a list of ecologically friendly economic activities.

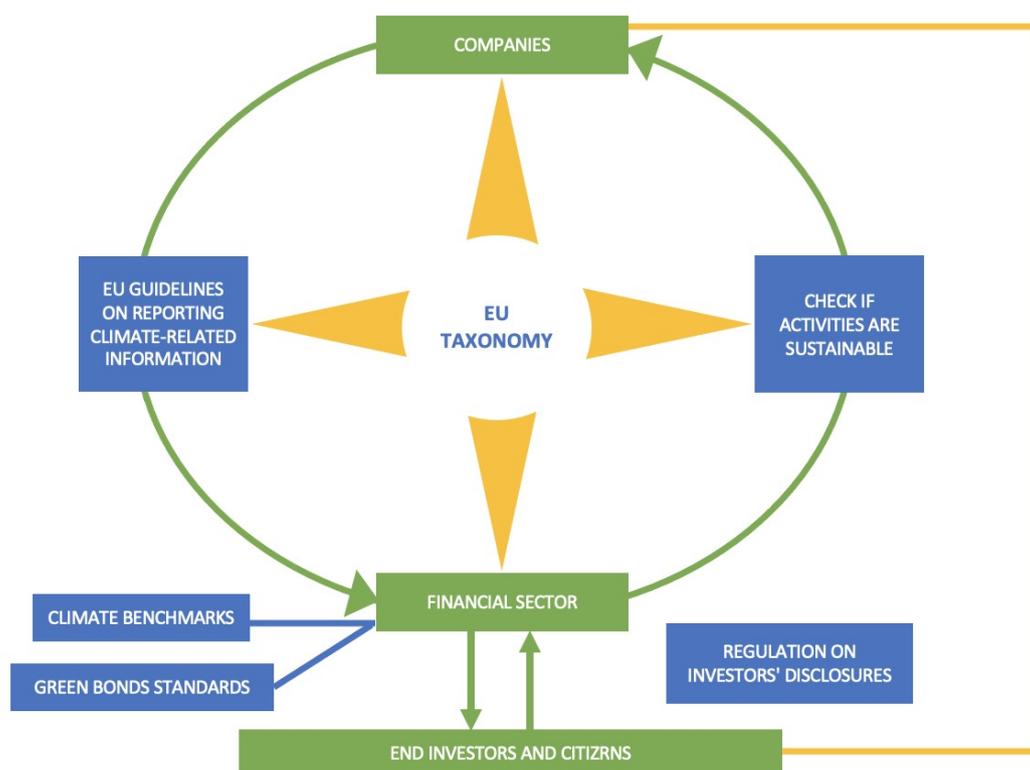


Figure 3.1: The Action Plan on Sustainable Finance

Corporate social responsibility reports—also known as CSR reports, impact reports, or sustainability reports—have grown in popularity over the last decade, a trend that is expected to continue. Norway has established itself as a leader in CSR policy, demonstrating that industry and government share responsibility for sustainability. Transparency and sustainability disclosure, the Norwegian government believes, are critical in establishing and developing trust between society and businesses. In Norway,

the primary drivers of CSR and sustainability reporting are respect for human rights and an awareness of the need to address environmental problems and combat corruption.

Companies listed on Oslo Børs are subject to corporate social responsibility reporting (Oslo Børs, 2018). The requirements for Norwegian firms are principally outlined by Section 3-3 a-d of the Norwegian Accounting Act, which includes a need to provide a corporate governance report. Corporate responsibility is also addressed in the Norwegian Code of Corporate Governance.

3.1.3 ESG ratings

The ESG evaluation system was first proposed in the "United Nations Principles for Responsible Investment" in 2006. It aims to help investors understand the impact of environmental, social, and corporate governance factors on investment value. Subsequently, various international organizations have successively issued several ESG information disclosure standards. Rating agencies use these ESG information disclosure standards as the logical basis for ESG assessment of companies. Therefore, corporate ESG performance can be reflected in the capital market.

The five mainstream ESG rating agencies globally are MSCI, Dow Jones, Thomson Reuters, FTSE, and Morningstar. These institutions have promoted the unification of ESG information disclosure standards to a certain extent through their market influence and coverage. Their systems usually have massive data, and the evaluation indicators are biased to quantify. They divide the rating indicators into three areas: E, S, and G, and involve the corresponding areas. From the point of view of the calculation method, each system is based on a weighted average. The rating company assigns a certain weight to each indicator and adjusts the weight according to industry conditions.

Even though several empirical studies have adopted third-party ESG ratings, there is still debate over the reliability of such ratings and which assessment is more reliable. Doyle(2018) outlines how the ESG methodology used by rating agencies has shown shortcomings and a lack of standardization. The ratings have three biases: firm size bias, geography bias, and industry weight and business alignment bias. Furthermore, it is critical to emphasize that rating agencies usually rate only publicly traded companies.

The scarcity of ESG data from private enterprises makes acquiring an ESG evaluation for these firms difficult. This constraint results in a knowledge gap regarding the learning of ESG factors in private enterprises. The primary impediment, particularly in our research, is establishing a mechanism for assessing ESG in private enterprises.

3.2 Previous research on the impact of ESG

3.2.1 ESG and financial performance of companies

Scholars have done much research on the relationship between ESG performance and corporate financial performance. Giese and Lee(2019) think there is a positive correlation between ESG performance and corporate financial performance. Scholars explain it from heterogeneous transmission (Gregory et al., 2013) and systematic risk transmission. First of all, from the perspective of heterogeneous transmission mechanisms, companies with higher ESG ratings are more competitive than their peers in terms of internal management, supply chain management, resource utilization, and human capital development, and thus have above-average risk control capabilities. Generally, companies with higher ESG ratings are better at formulating business plans and incentive plans for executives and using their competitive advantages to bring higher profits to themselves, which brings higher dividends. Secondly, from the perspective of the systemic risk transmission mechanism, companies with better ESG performance are less susceptible to systemic market shocks, that is, the beta coefficient is small.

Neilling and Webb(2009) find that the outcomes indicate not a very significant correlation between ESG and financial success. They recognize that the stock return influences the firm's ESG performance. After they apply panel fixed effects and Tobit models and experiment with various statistical techniques, it concludes that there is no significant connection between ESG and corporate financial performance.

Nollet et al.(2015) examine the connection between ESG and corporate financial performance using ROA, return on capital (ROC) and a market-based metric called excess stock return. They see a U-shaped connection between the ESG score and both the ROA and ROC. This finding indicates a minimum level of investment that a business must

make to improve its financial performance. If companies' efforts fall short of the threshold, ESG investments may be ineffective, negatively impacting earnings. Additionally, the authors advise that businesses should include ESG in their long-term strategy, although investments in ESG may not always pay off immediately. ESG initiatives need shareholder patience and significant money to achieve success. Furthermore, they explain how ESG's governance requirements are the most important variables affecting financial performance.

Friede et al.(2015) count nearly 2000 ESG-related studies from 1970 to 2014. Around 90% of the studies find a non-negative relationship between ESG-CFP (financial performance), and 52% of stock-related studies show ESG and company financial performance has a significant positive correlation. In general, the explanations for the positive correlation are mainly as follows: (1) companies with excellent ESG performance can significantly reduce company-specific risks; (2) companies with excellent ESG performance have a positive impact on reputation; (3) Excellent ESG performance can enhance customer stickiness, improve company operations, and enhance company financial performance. Nonetheless, it may be argued that, in general, ESG does not harm financial performance. Thus, the authors recommend that rational investors focus on long-term responsible investment and connect their interests more closely with ESG issues. It indicates that investors should put more weight on ESG in their long-term investing decisions.

Other scholars believe that ESG performance and corporate financial performance are irrelevant or negatively related. The non-financial obligations of companies in terms of environmental, social responsibility, and governance, or related activities will not improve the financial performance and may even harm the experience of the company. The neoclassical theoretical scholars represented by Friedman(1970) believe that social responsibility activities will increase corporate costs, weakening financial performance and lowering corporate valuations. A similar study is conducted by Vance(1975), who proposes that high-level performance of corporate social responsibility will increase costs and put the company at a financial disadvantage, limiting the company's strategic choices. Kim and Lyon(2014) find that companies that participated in environmental protection activities or received green rewards had negative abnormal returns and believed that environmental protection activities negatively impacted corporate financial performance. Studies have even shown that investing in stocks that perform poorly in environmental protection and social responsibility, such as tobacco, alcohol, and gambling games, can

also bring additional financial returns (Hong & Kacperczyk, 2009).

3.2.2 ESG and investment performance of stocks

In the research on environment responsibility(E), Shane and Spicer(1983) use the event study method to analyze the market's response to corporate social responsibility information and concluded that companies that invest more in environmental pollution treatment obtain higher excess returns. Fernandez-Kranz and Santalo(2010) find that investors in developed capital markets can realize that negative environmental responsibility information will negatively affect companies and incorporate it into investment decision-making. In terms of corporate social responsibility (S), McPeak and Tooley(2008) study on 56 American companies with good social responsibility performance records and find that their financial performance was outstanding, with stock gains greatly exceeding benchmark returns. Van Dijken(2007) systematically studies the relationship between corporate social responsibility and stock prices using 90 Dow Jones sustainable companies in the United States as a sample. The results show that the stocks of "value-driven corporate social responsibility" companies have been better than the market and peers for a long period of time. However, there is also an argument that excessive investment in social responsibility will cause a decline in corporate profitability, so social responsibility is negatively correlated with stock returns (Becchetti & Ciciretti, 2012). From the perspective of corporate governance responsibility (G), if listed companies are well-managed, their stock returns are significantly higher than those with relatively poor governance (Drobetz & Schillhofer, 2003).

Kumar et al.(2016) examine the connection between ESG and annualized volatility, the annual return, and risk-adjusted returns. Using a database of ESG firms in the Dow Jones Sustainability Index and around 800 companies not included there, they discover that ESG firms' stocks have less volatility than peers in the same area. This effect of ESG varies by industry: energy is the most volatile and insurance is the least. Additionally, the authors discover that ESG stocks provide better annual returns in most industries. These findings suggest that ESG companies' risk-adjusted returns will be greater than those of other enterprises.

Durán-Santomil et al.(2019) research the mutual funds from the Morningstar database to determine the effect of ESG ratings on their performance. The findings suggest that the more sustainable investments a fund makes, the lower risk it will suffer. Moreover, the authors discover that a high degree of ESG reduces the mutual fund’s value at risk (VaR), implying that ESG funds are better protected against severe loss.

To summarize the last part, although the research on ESG and corporate financial performance is fragmented, the overall trend indicates a non-negative connection. If market players are not motivated only by financial gain, integrating ESG should be promoted as a decision-making criterion. We can find that an important feature of ESG companies is their low-risk nature, shown in many studies. This feature may be used to explain the connection between ESG and underpricing. It may serve as an incentive for our research to uncover empirical proof that a firm’s ESG characteristics affect the listing underpricing.

Table 3.1: Summary of the previous literature on the impact of ESG

Perspective	Key Findings	Related Studies
Firm	There is a non-negative correlation between ESG and corporate financial performance.	Neilling & Webb (2009)
		Friede et al. (2015)
	There is a minimum level of investment that a business must make to improve its financial performance	Nollet et al. (2015)
		Giese & Lee (2019)
	Non-financial obligations of companies will not improve the financial performance.	Friedman (1970)
		Vance (1975)
	Kim & Lyon (2014)	
Stock/Investor	ESG firms’ stocks or funds are less risky than peers in the same area.	Ma (2019)
		Kumar et al. (2016)
		Durán-Santomil et al. (2019)
	ESG is negatively correlated with the following year’s performance.	Hong & Kacperczyk (2009)

3.3 Previous research on ESG and IPO

There are few studies regarding the impact of sustainability, CSR, ESG, SRI factors in the IPO setting. Bollazzi and Risalvato(2017)’s first article examines the impact of ESG

on the underpricing of several companies on the Italian stock market. Also, the authors analyze the effect of ESG on ROE in this research. They utilize a variety of factors to determine a firm's ESG level since ESG ratings are not accessible for companies at the time of IPO. This is because rating agencies can gather adequate data for ESG assessments only after companies have gone public. As a result, the authors are required to develop their own ESG evaluation factors. The dummy variable sustainable report assumes that ESG companies would confidently produce the sustainable report. However, only seven companies in this study's sample released the report since sustainability reports are not a legal obligation. The percentage of companies with sustainability reports is insufficient to conclude that the variables adequately reflect the firm's ESG level. Additionally, the authors construct certain independent variables by calculating all ESG criteria. However, they did not reveal how they get rating information on the companies. At the moment, analyzing the impact of being ESG on underpricing is difficult. The limitations of ESG evaluations must be carefully addressed since the variables reflect the extent to which ESG information is disclosed, which serves as a proxy for a firm's ESG level.

Following the ESG evaluation stages, the authors perform a univariate analysis, putting companies into two categories based on their characteristics: environmental and non-environmental. The t-test thus allows for two averages of a single dependent variable to determine if ESG variables affect ROE before the listing and the degree of underpricing. Consequently, they discover that companies with more sustainable financial reporting have a greater rate of underpricing. As a result, the authors argue that the market rewards companies that are comfortable releasing ESG data. The authors then conduct multivariate analysis on each ESG component, and the findings indicate that none of the factors is significant. This implies that the market does not benefit ESG companies throughout the IPO process. One may argue that the univariate approach is insufficient for establishing ESG's impact on underpricing. The authors simply find that a greater underpricing indicates that the market favors ESG companies. One might think that if ESG companies are valued more than non-ESG businesses, issuers and underwriters will learn and set a higher offer price for future IPOs. The findings need additional discussion in light of underpricing theories.

Francesco et al.(2017) study the influence of CSR on the performance of newly listed companies on the Italian Stock Exchange and the first day of listing performance. The

empirical analysis examines 84 firms (76% are SMEs) that were listed between 2009 and 2015, and it reveals that only 19 of the 84 are involved in all areas of CSR. According to univariate statistical research, firms that publish a sustainability report seem to have a higher underpricing, indicating that the market recognizes and rewards ESG corporate responsibility. According to univariate analysis, organizations who submit a sustainability report appear to have a substantially different underpricing (10% confidence level) than newly listed companies that do not prepare a sustainability report. There are consequently many engaged in the listing process that want to boost their ESG corporate responsibility initiatives by offering the IPO shares at a higher offer price, but the market seems to notice and reward responsible firms on the first day of trade.

Another article on the subject is from Huang et al.(2019), who analyze the connection between ESG performance and post-IPO performance. The textual information indexes for this research are constructed utilizing CSR information from IPO prospectuses to quantify company social performance and environmental performance, which are the two critical aspects of the ESG frame. The article discusses the framework for creating text indexes and the formulae for creating expenditure indexes. Holding period return and cumulative abnormal return are dependent variables to examine the impact of newly public companies' CSR information. The institutional investor ownership ratio variable is selected because financial institutions have more information than individual investors. If corporate environmental performance and company social performance are more valuable to a business, institutional investors will own a greater proportion of these shares.

The regression analysis shows that improved social disclosure of a company is positively correlated with its holding period returns. Furthermore, institutional investors are more concerned about corporate environmental performance disclosure. While the article does not examine IPO underpricing explicitly, it examines the extent to which investors consider ESG factors when investing in freshly issued IPO stocks. Additionally, it is instructive to understand the methodology used to extract ESG data from IPO prospectuses since ESG ratings are not accessible.

A recent study by Bui and Alessandra(2020) examines whether a firm's ESG variables influence its IPO underpricing in the US stock market. The relationship between a firm's ESG qualities and its degree of underpricing is investigated using quantitative

analysis. Based on the multiple regression analysis results, they find that the frequency of environmental and governance phrases in IPO prospectuses, the negative tone, and the overall attitude of the environment disclosure all contribute considerably to IPO underpricing. The market does not recognize a firm's social elements in the context of an IPO and is worried about the extent to which governance actions are disclosed and if these activities are adequately described in the prospectus. Meanwhile, the market takes a firm's environmental actions seriously by evaluating the quality. Additionally, the market is more receptive to negative information concerning environmental material than favorable information in an IPO. The paper indicates that some businesses' ESG aspects affect IPO underpricing, indicating that the market is concerned about ESG problems.

Alessandro and Carlo Raimondo(2021) develop a text-based measure of ESG disclosure in IPOs based on a sample of 783 US IPOs. The findings indicate that (a) the number of ESG disclosures in S-1s is negatively correlated with IPO underpricing and Tobin's Q; and (b) this impact is predominantly driven by ESG-related underpricing, as well as price revision. Rather than that, the results demonstrate that when evaluating Tobin's Q, investors place a premium on the governance component of sustainability. Thus, they claim that publishing more ESG information in the S-1 prospectus reduces the information asymmetry between the firm and investors, enhancing the company's financial performance, in this case via reduced underpricing and evaluation.

The articles argue that caution is required while doing ESG assessments because the information should be adequate and trustworthy, for example, an IPO prospectus. Additionally, the limitations of the two pieces of research motivate us to continue our research on ESG companies in the context of different listings. Examining ESG in the IPO underpricing setting makes it possible to ascertain the issuers', underwriters', and investors' preferences for ESG features. For instance, investors' perceptions of ESG companies may be contrasted against those that do not adhere to sustainability standards.

4 Data and Methodology

In this chapter, we introduce the data used in our study and detailed methodology. To be specific, data sources and details how we processed them from their initial appearance to their use in the article, covering the origin of the initial idea, the specific processing algorithm, and the final use, in addition to showing the model for linear regression analysis and attempt in machine learning method.

4.1 Independent variable

4.1.1 Material

We used the content of the prospectus or information(admission) document that is filed before the company goes public as the object we want to study. More specifically, we tried to quantify the disclosure and sentiment of ESG content in a company's prospectus information(admission) to see any connection with the company's short-term stock price return.

Historically conducted empirical studies on the ESG topic have frequently used ESG scores provided by professional institutions such as Bloomberg, etc. We tended to find a document that companies would disclose before going public and use as a basis for measuring their ESG performance, as there are no mandatory requirements or standards for disclosure reports to shareholders before a company goes public. In the Norwegian market, the prospectus or information (admission) document meets this need and is one of the documents that must be submitted before going public. The primary purpose of both the prospectus and the information (admission) document is to provide investors with information about the company's profile, risks, development, economic situation, etc., that will be of interest to them. In Norway, because of the different listing markets, Euronext Growth faces less stringent requirements than Oslo Bors in terms of filing documents, one of which is the possibility of replacing the more lengthy prospectus with the information (admission) document. Although there are no rules and requirements for ESG content disclosure in the prospectus or information(admission) document, as ESG is becoming one of the critical factors for investors to measure, companies going public will never ignore

this critical disclosure, even though they are not explicit or prominent presentations, thus, how to quantify the ESG disclosure and sentiment in this type of document is the next major issue we need to address in our methodology.

In the past 2020, there have been a total of 59 new listings in Norway. In Euronext’s all IPOs announcement, our initial filter was: listed companies with a listing in location Oslo from 01/01/2020 to 01/01/2021. 59 companies met this condition. However, only 48 of these 59 companies met our research criteria, which is that the company is an IPO/private placement and has a submission of an English version of the prospectus or admission document. With this selection requirement, we finally collected these 48 company documents and used them as the focus of our research from the beginning to the end. The challenge here is how to quantify the ESG profile in a non-ESG disclosure report.

Table 4.1: Two different document types

Number of firms	Document type	Listing market
4	Prospectus	Oslo Børs
44	Information(Admission) Document	MerkurMarket/Euronext Growth

4.1.2 Vector basis for ESG dictionary

There has been no consensus on how to define ESG metrics so far, especially in textual analysis. If we want to know what content in the text is ESG relevant, we must first have a vector benchmark, but which word vectors are ESG relevant? Baier et al.(2018) constructed such an ESG lexicon from 10K files of US-listed companies, but Baier et al.’s lexicon is too large for our study since the 10K file reports are not of the same order of magnitude as the prospectus files we use in this paper, and the content of annual reports of US-listed companies is also different from Norwegian prospectus. Therefore, after reviewing a large amount of relevant literature, we constructed a set of word vectors for each of the three ESG indicators, each containing ten words that we believe are the most relevant. We will use these three sets of words as the basis to create our own ESG dictionary.

Table 4.2: ESG vector basis

ENVIRONMENTAL	clean, environmental, climate, renewable, carbon, waste, emission, water, pollution, air
SOCIAL	health, human, labor, society, education, employment, access, public, train, hire
GOVERNANCE	governance, approve, transparency, audit, control, stakeholder, engagement, compliance, responsible, structure

4.1.3 ESG dictionary construction

By using the ESG base vector and the word2vec algorithm, we can expand our original vector into a larger ESG lexicon vector. In this step, we expanded each ESG keyword constructed from the previous step into the ten most relevant words by the word2vec algorithm, and this vector list is the final word frequency calculation lexicon that we would use in our subsequent textual analysis.

4.1.3.1 Word2vec algorithm

In NLP (Natural language processing), the most granular thing is words, which make up sentences, and then paragraphs, chapters, and documents. Therefore, to deal with NLP, we must first start with words. As a simple example, to determine the lexical nature of a word, whether it is a verb or a noun. Using the machine learning idea, we have a series of samples (x,y) , where x is the word, and y is their lexical nature, and we want to construct a mapping $f(x) \rightarrow y$. However, the mathematical model $f(x)$ (e.g., neural network, SVM) only accepts numerical input here, while the words in NLP, which are abstract summaries of humans, are in symbolic form (e.g., Chinese, English, Latin, etc.), so they need to be converted into numerical form, or rather to say: embedded in a mathematical space, this embedding method is called word embedding, and Word2vec, is a kind of word embedding.

We use a simple example to explain how to find similar words using Word2vec. For a sentence: *talents will get grade A*, if x here is *talents*, then y could be *get, grade, A*, etc. Now for another sentence: *we will get grade A*, if x here is *we*, then it is not hard to find that the output y here is the same as the previous sentence, thus $f(\text{talents})=f(\text{we})=y$, so

the model will tell us we=talents.

There are two main training models for word2vec: 1. CBOW (Continuous Bag-of-Words Model) and 2. Skip-gram (Continuous Skip-gram Model). Using the simplest definition to distinguish the difference between the two models, namely:

- (a) If a word is used as input to predict the surrounding context, then the model is called Skip-gram.
- (b) If the context of a word is used as an input to predict the word itself, then the model is CBOW.

Skip-gram



Figure 4.1: Skip-gram model

CBOW



Figure 4.2: CBOW model

According to the definition, obviously, in our article, we used Skip-gram model. Based on this model basis, we can identify those ESG most related 10 words from a large amount of text by using the rank of similarity. An example of the output in finding similar words using Word2vec shown below; in this case, we find similar words for *clean*:

Table 4.3: Word *clean* similarity

	clean	Similarity
0	environmentally	0.581879
1	pure	0.529368
2	mix	0.528612
3	chemical	0.512952
4	hydroponic	0.504759
5	advanced	0.502718
6	waste	0.499168
7	climate	0.494509
8	energy	0.489401
9	optimize	0.487519

Our output includes three vector matrices, which are for each of E, S, and G. For readability, we only present the Social matrix here:

	0	1	2	3	4	5	6	7	8	9
health	healthcare	hospital	population	welfare	payer	patient	age	connected	ehealth	medical
human	animal	routine	nutrition	food	computer	perceive	error	telecommunication	safety	technical
labor	wage	skilled	workforce	injury	gender	equality	litigation	death	certainty	skill
train	skilled	personnel	recruit	replacement	attract	experienced	shortage	competence	reliant	skill
society	transition	energy	europa	renewable	minimize	emission	climate	sector	utility	stakeholder
education	entertainment	elearning	chess	learning	career	nearly	child	learn	workforce	leadership
hire	personnel	recruit	skilled	continued	retain	skill	train	frame	key	klovning
employment	termination	severance	benefit	employee	terminate	salary	sverre	equality	options	gender
access	enable	help	opportunity	platform	sensitive	cloud	digital	easily	deliver	solution
public	council	repeal	parliament	incrimination	state	disqualify	review	offer	inquiry	expression

Figure 4.3: Social extended words

We can see from the display that we acquired pretty good results with the word2vec algorithm, but at the same time, there were a couple of problems:

- 1) Duplication. For example, [carbon], which is the base word in environmental, also appears in the extended word of governance, and the solution to this problem is to put it into the correct group and delete it from the other group by subjective judgment. In addition, the same word may appear in the same group several times, so we did de-duplication for this problem.
- 2) Words with no clear meaning. We did a manual screening and removed this type of word vector.

4.1.4 ESG indicators

4.1.4.1 ESG-term frequency

After doing many preliminaries, we obtained the first set of independent variables, the ESG-term frequency: we did count statistics for each group of word vectors of E, S, G and derived the percentage of that count to the total length of the full text. This counting and ratio calculation is a simple and intuitive method of textual analysis, as such ratio directly reflects the length of discussion of each indicator in the text, which reflects the extent to which the company emphasizes ESG in its listing disclosure. After this step, we got three independent variables: environment term frequency, social term frequency, and governance term frequency.

4.1.4.2 ESG sentiment score

The second textual analysis method for which we obtained independent variables is sentiment analysis, where the objects of sentiment analysis are those textual contents that contain ESG-related word vectors. We believe that the most important thing in measuring sentiment direction is the sentiment lexicon we are based on, because after reviewing a large amount of literature, we found that the vast majority of studies nowadays are based on a sentiment lexicon and then measure the frequency of negative and positive sentiment words in the text to be analyzed, so with this in mind, we chose LM-dictionary, which is a sentiment lexicon that Loughran & McDonald(2016) developed specifically for financial texts. A good example of why they constructed this dictionary is that the word "liability," a frequently but negative word, is not necessarily the case in financial texts. Therefore, we counted the number of negative words and positive words in ESG-related texts based on the LM-dictionary and did a straightforward subtraction to obtain our final sentiment score, i.e., $\text{Sentiment score (of E, S, G)} = \text{positive words count (in E, S, G)} - \text{negative words count (in E, S, G)}$. Therefore, in these variables, larger values imply more positive attitudes, while smaller values imply more negative sentiment. So after completing this step, we got another three more independent variables: environment sentiment score, social sentiment score, and governance sentiment score.

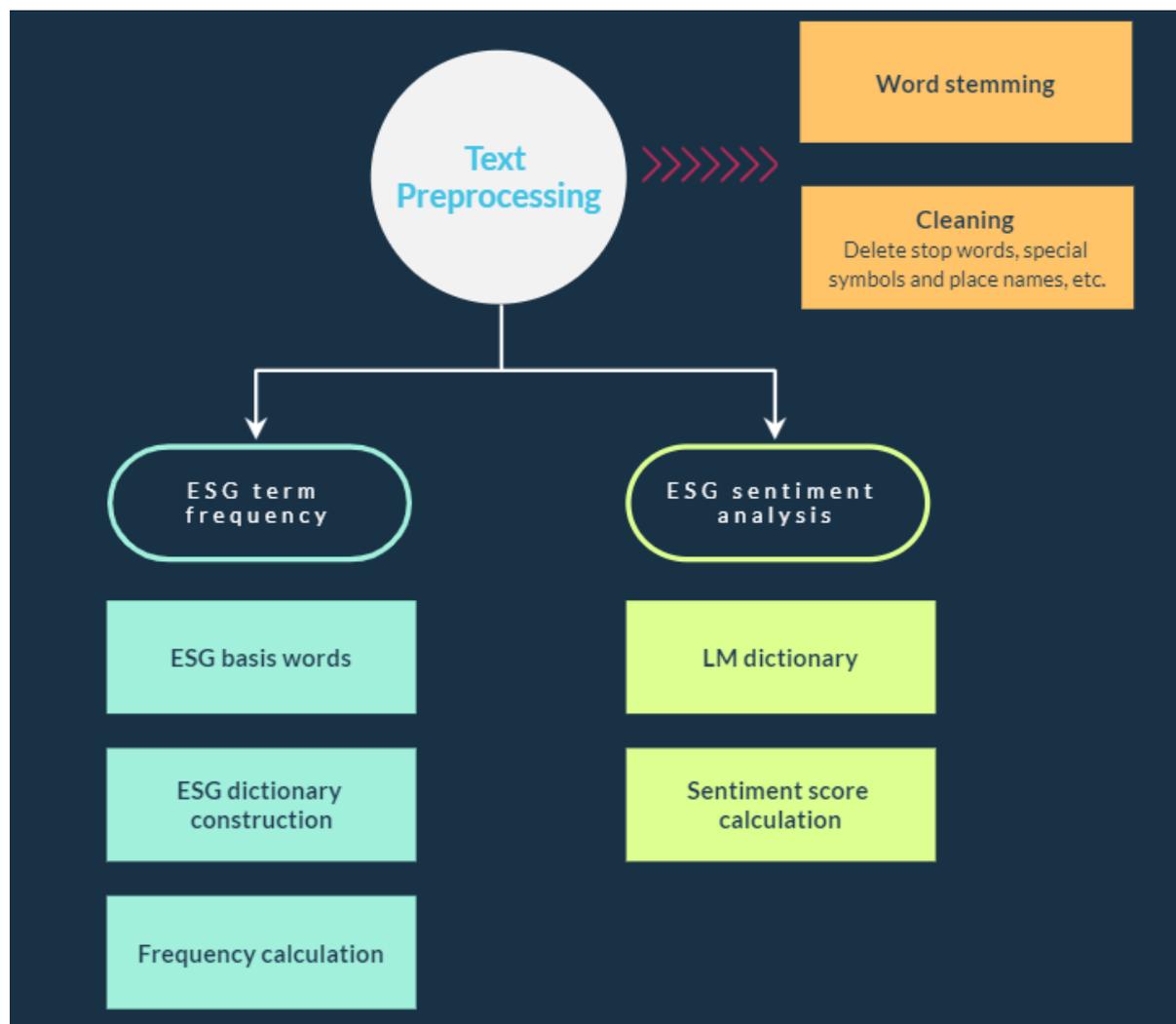


Figure 4.4: Textual analysis steps

4.2 Independent variable

Short-term stock price returns are the dependent variable in our study. In this paper, we separate the short term into four-time points: one day, three days, one week, and one month after the IPO. Typically, after the first day of trading, a stock's initial return has a technical term called underpricing, therefore, in stricter language, we divide the short-term return into four components: underpricing, three-day return, one-week return, and one-month return.

$$\text{Underpricing} = \frac{\text{Closing price on the first day} - \text{Offer price}}{\text{Offer price}} \quad (4.1)$$

$$3\text{-Day return} = \frac{\text{Closing price on the third day} - \text{Closing price on the first day}}{\text{Closing price on the first day}} \quad (4.2)$$

$$1\text{-week return} = \frac{\text{Closing price on the first week} - \text{Closing price on the first day}}{\text{Closing price on the first day}} \quad (4.3)$$

$$1\text{-month return} = \frac{\text{Closing price on the first month} - \text{Closing price on the first day}}{\text{Closing price on the first day}} \quad (4.4)$$

A common issue in the share price collection process is the non-trading day. For example, if a company goes public on Thursday, March 5, 2020, its third-day stock price collection should be the closing price on March 8, but March 8 (Sunday) is a non-trading day, in which case we would collect the closing price of the opening day closest to the target date, in the above example, if March 9 is a trading day, the third-day closing price of this stock would be the closing price on March 9. Similarly, we use this treatment for the third-day, first-week, and first-month stock price collection.

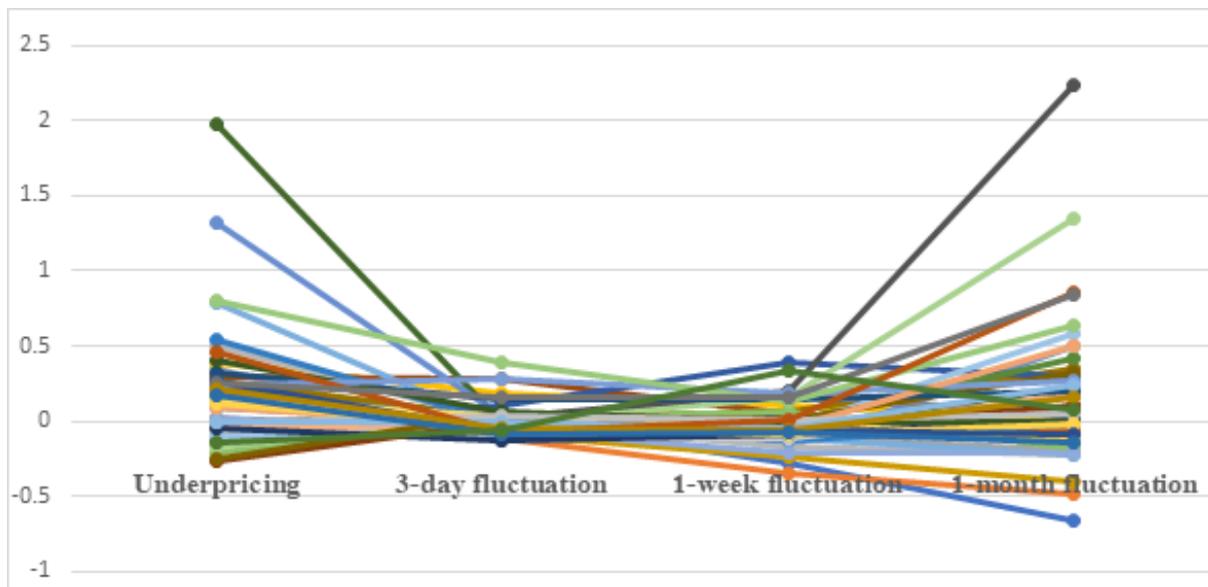


Figure 4.5: Changes in four indicators of 48 companies

Stock price volatility and returns are a complex result of multiple causes. The longer the time on the market, the less weight the impact of the information disclosed by the company at the time of listing, which is reasonable because outdated information does

not have much value, and conversely, as time goes by, more information that is relevant to the company's development will emerge that will have a different impact on the stock price. Although this is well-known common sense, we still want to include the share price return after one month in this study to explore the information it may contain.

4.3 Control variables

4.3.1 Listing market (document type)

There are three major stock markets in Norway, Oslo Børs, Oslo Axess¹ and Euronext Growth. In our screened study sample, 4 companies are listed on Oslo Børs and 44 on Merkur Market. As we mentioned earlier, the documents chosen for the textual analysis differ by market: for those listed in Oslo Børs, we used prospectus, while for those listed in Euronext Growth, we used the admission document. This is mainly due to the difference in the filing requirements between the two markets and the fact that the two markets have different qualifications in many aspects.

Since Merkur Market is a multilateral trading facility (MTF), both the entry standards and the subsequent reporting requirements are less stringent and thorough than those for Oslo Børs. Although the information(admission) document is an alternative to the prospectus, the structure and content preferences of the two documents are different, which is the main reason for including this indicator as a control variable.

4.3.2 Covid-19 measurement

The sudden arrival of the Covid-19 has brought an unprecedented blow to everything in 2020, and one of the most significant blows is to the economies. However, in 2020, an anomaly happened in the Norwegian stock market: there were three times more new listings in the Norwegian market in 2020 than in 2019, which is mainly attributed to Euronext Growth. Our sample companies are all listed in 2020 and have been affected by the outbreak to a very large extent, but to what extent? We looked back at the

¹We would not discuss this market since there is no new listing in Oslo Axess in 2020.

Norwegian outbreak, starting from the first case in Norway and ending on February 1, 2021, found that Norway experienced three outbreaks in that period:

- 1) 02/03/2020 - 28/03/2020;
- 2) 16/10/2020 - 18/11/2020;
- 3) 06/12/ 2020 - 09/01/2021.

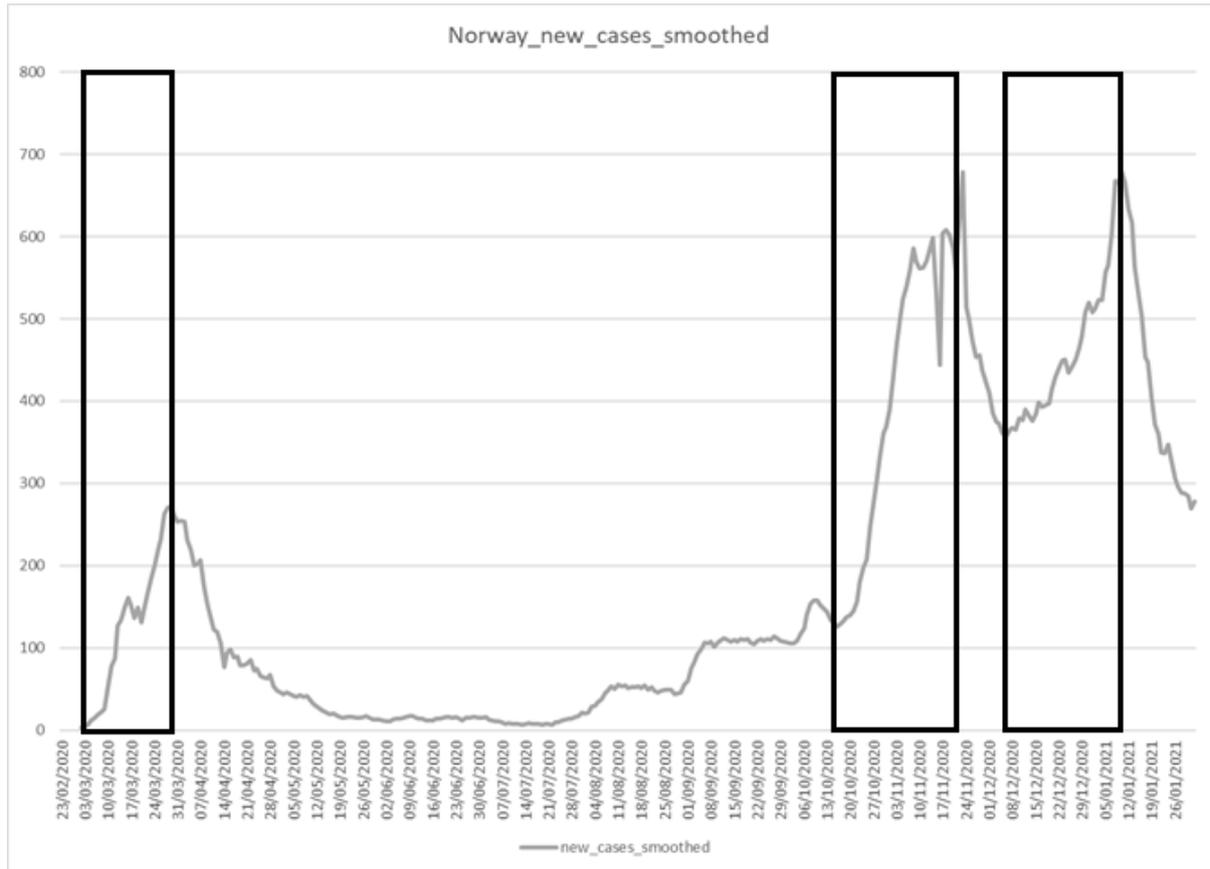


Figure 4.6: New Covid19 cases in Norway (smoothed)

As mentioned in the previous section introducing the dependent variable, we divided the short-term stock price collection into four-time points: the first day of the market, the third day, the first week, and the first month, so the quantification we use here to measure the impact of the pandemic is:

$$Covid - 19 \text{ indicator of company } k = \sum_{i \in I} \sum_{j \in J_k} w_{i,j} \quad \forall k \in K \quad (4.5)$$

Where:

$i \in I$: set of Covid – 19 outbreak period in Norway

$j \in J_k$: set of four – time points of the new listing company k

$k \in K$: set of new listing companies

$w_{i,j}$: equal to 1 if the time point j of the new listing company k is during the outbreak period i in Norway; otherwise, equal to 0.

This formula can assign a Covid19 index to each company to indicate the extent of the outbreak's impact: if the four points in time during the outbreak are collected, the company will have the highest score of 4. Conversely, if none of the four points in time are during the outbreak, the company will be assigned the lowest impact score of 0.

4.3.3 Green company

With Norway's ratification of the Paris Agreement on climate change, Norway has pledged to achieve at least a 40% reduction in greenhouse gas emissions by 2030 compared to 1990 levels. Norway's initiative to address climate change is also evident in their newly listed companies, with a growing share of companies in environmentally-friendly industries. Although the definition of a "green company" is not uniform, the increasing number of listed companies in areas such as renewable energy, waste treatment, etc., is an indication that environmental and climate issues are becoming more and more critical to the business sector as to investors. In our study, green companies may suggest that the company may have more ESG disclosure content at the time of IPO than companies in traditional industries.

4.4 Variables summary

Table 4.4: Variables summary

Variable	Type	Description
Environmental_frequency	Independent variable	Environmental related term frequency.
Social_frequency	Independent variable	Social related term frequency.
Governance_frequency	Independent variable	Governance related term frequency.
Environmental_sentiment_score	Independent variable	Sentiment score of Environmental related content.
Social_sentiment_score	Independent variable	Sentiment score of Social related content.
Governance_sentiment_score	Independent variable	Sentiment score of Governance related content.
Underpricing	Dependent variable	Initial return after the first day of IPO.
3-day return	Dependent variable	3-day return after the first day of IPO.
1-week return	Dependent variable	1-week return after the first day of IPO.
1-month return	Dependent variable	1-month return after the first day of IPO.
Market	Control variable	Dummy variable. Equal to 0 if listing on Euronext Growth, equal to 1 if listing on Oslo Børs.
Covid-19 measurement	Control variable	A measure of how much the company was affected by the covid outbreak.
Green company	Control variable	Dummy variable. Equal to 1 if the company is green, equal to 0 if not.

4.5 Multiple regression

We firstly fitted our variables with multiple linear regression. The purpose of the regression is that we wanted to see if the ESG indicators have explanatory power on the short-term returns after IPO listing, in this paper, i.e., the returns at four time points, our regression model is as follows:

$$\begin{aligned}
 \text{Short-term return indicator}_i &= \beta_0 + \beta_1 \cdot \text{ESG indicator}_i + \beta_2 \cdot \text{Market}_i \\
 &+ \beta_3 \cdot \text{Covid19 measurement}_i + \beta_4 \cdot \text{Green}_i + \epsilon
 \end{aligned}
 \tag{4.6}$$

Since the regression models are based on ordinary least squared (OLS), a few assumptions need to be considered:

1) Linearity of parameters

The assumption of Linearity in Parameters implies that the explanatory and dependent variables must have a linear relationship. According to the literature we have reviewed, numerous studies have used ESG indicator variables to demonstrate the linear relationship; therefore, we believe this assumption holds for the variables we use in this article.

2) No perfect collinearity

We do not want to include perfectly correlated variables because two variables can represent two distinct effects that may be relevant in explaining the dependent variables; thus, there is no excluded variable in the model. This article uses the Variation Inflation Factor (VIF) for each regression to measure multicollinearity

3) No heteroskedasticity

This assumption is to see if the error term will have the same variance in terms of the explanatory variables; otherwise, the model will be heteroskedastic. In our case, we conducted the Breusch-Pagan test.

4) No autocorrelation

Autocorrelation means that there is a correlation between the expectation values of the random error terms and that there is autocorrelation or serial correlation between the random error terms. OLS estimation has minimum variance only under the condition of homoskedasticity and non-autocorrelation. When the model is autocorrelated, the OLS estimation is still unbiased but no longer valid. In this article, we conducted the Durbin-Watson test.

5) Normality

This assumption means that the error terms of the model are normally distributed.

We will use Q-Q plot to visualize this assumption.

4.6 Attempts in machine learning

When dealing with giant data sets or the massive number of variables known as the "curse of dimensionality," algorithmic modeling is incomparably superior to classical modeling in the face of many competing models without assuming any distributional context for the data. Machine models can solve problems that are essentially powerless in classical statistics, and they are not constrained by the assumptions made in classical linear regression models as shown above, which we can also compare with cross-validation to check the effectiveness of the models. In our case, Although the massive data and dimensions were not a hindrance, linear regression models did not provide us with satisfactory results because most of them have a low R-square value, which means our regression model may barely have explanatory power for the variables. This was the reason we wanted to try modern algorithmic regression method to see if it can give us with more helpful information than the traditional method since it is a common sense that machine learning fit better models than traditional methods do . More specifically, we will try two machine learning models in our paper: 1) Random forest and 2) Support vector regression(SVR), and compare them with linear model.

5 Analysis and Results

5.1 Descriptive statistics

The role of statistical description of variables is to observe the range, mean, and fluctuations of the variables. In our case, we did descriptive statistics for eight independent variables, four dependent variables, and three control variables in 2 dimensions to get the maximum, minimum, mean, and standard deviation of each variable. The output of the software is shown in the following table:

Table 5.1: Descriptive statistics

Statistic	N	Mean	St. Dev.	Min	Max
Underpricing	48	0.181	0.400	-0.267	1.971
3-day return	48	0.014	0.105	-0.133	0.389
1-week return	48	-0.016	0.143	-0.351	0.385
1-month return	48	0.147	0.471	-0.668	2.242
Covid19	48	1.729	1.349	0	4
Green	48	0.271	0.449	0	1
Market	48	0.083	0.279	0	1
E_freq	48	0.007	0.009	0.0001	0.033
S_freq	48	0.016	0.007	0.0001	0.037
G_freq	48	0.028	0.007	0.001	0.041
E_senti	48	-167.708	167.401	-763	97
S_senti	48	-110.896	152.561	-579	269
G_senti	48	-215.208	195.504	-912	82

Among the dependent variables, Underpricing has the highest average return of 0.181. In contrast, the 1-week return has the lowest average return of -0.016, the only negative value among the four indicator averages, indicating that companies listed in 2020 have the worst return after one week of listing when compared across the four times points. In terms of maximum and minimum values, the maximum and minimum values of 1-month return are the most extreme of the four indicators, which is understandable because the longer the time on the market, the longer the stock price volatility accumulates, resulting in a more significant gap between the initial listing and the stock price. Underpricing and 1-month return both have relatively large standard deviations, indicating that there is a

large variability in individual stocks for these two indicators.

When it comes to the control variables, Green and market are dummy variables, and the mean value tells us that about 27% of the companies in our sample are green companies, and about 8.3% are listed on Oslo Børs. The mean value of the Covid-19 measurement is about 1.73, indicating that by using our measure of a covid pandemic outbreak, companies listed in 2020 are affected by the outbreak by an average of 1.73 units.

The independent variables are divided into two dimensions: 1) ESG-term frequency and 2) ESG content sentiment. As we mentioned earlier in the explanation of variables, these two dimensions correspond to the degree of disclosure of ESG and the company's attitude towards ESG. In terms of the degree of disclosure, among the three indicators of ESG, governance has the highest percentage of disclosure, followed by social and finally, environmental, which is the same as the vast majority of papers that study the degree of ESG disclosure in company documents, for example, in the ESG lexicon constructed by Baier et al. (2018), environmental and social keywords are only distributed to 2.3% and 8.8%. However, the remaining 88.9% are governance keywords, which indicates that the probability of governance-related words appearing in financial texts is much higher than environmental and social. Although our ESG lexicon is constructed based on the same number of word vectors, this still does not change that the percentage of governance disclosure is the highest among the three indicators. In terms of sentiment, all three indicators have negative means, i.e., more negative than positive sentiment, although such negative results do not affect our subsequent study, as they are still comparable across samples. The most negative sentiment was related to the governance content, compared to the more positive sentiment of related social content.

In addition, we also observed the correlations between the variables, and for readability, we visualized the correlation matrix. It is worth noting that the three indicators obtained in the sentiment analysis have a darker color in the graph, implying a higher correlation coefficient between them, which may be due to the fact that their sentiment analyses are based on the same sentiment lexicon, i.e. LM dictionary.

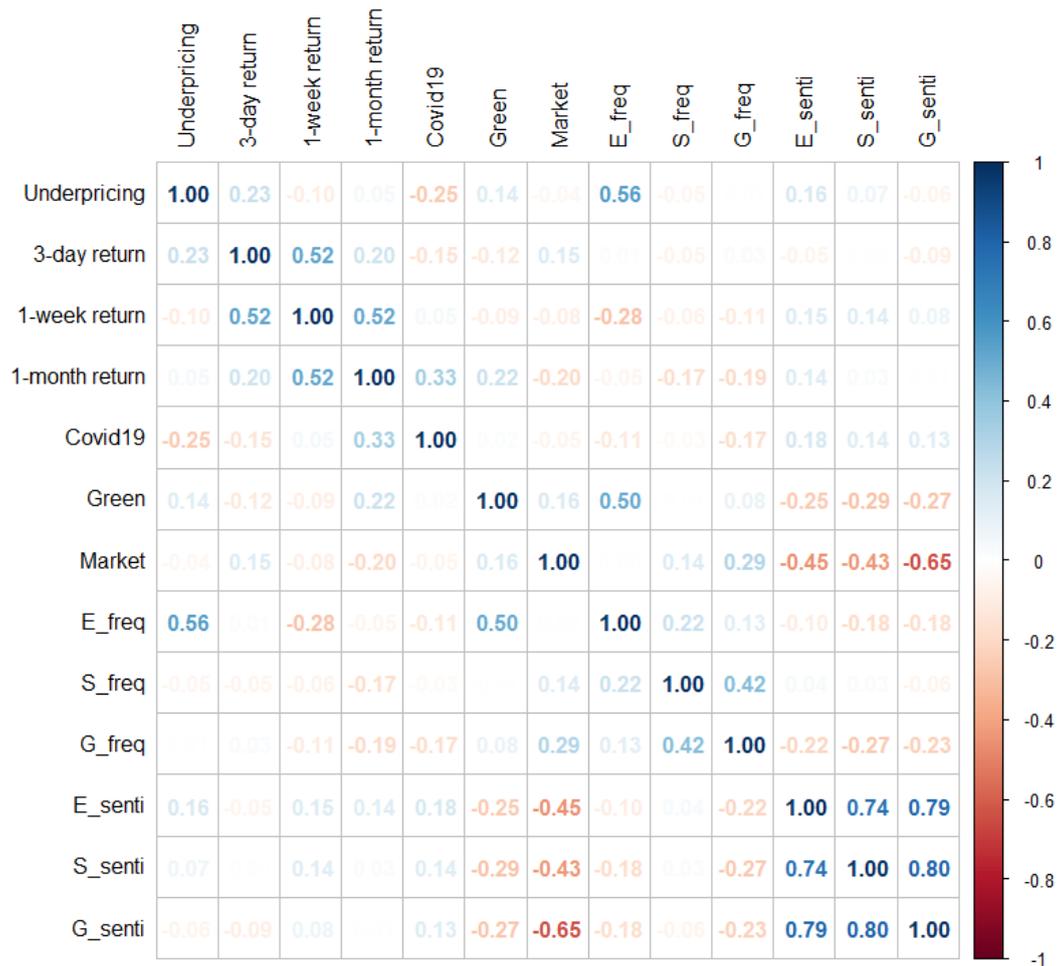


Figure 5.1: Correlation Matrix

5.2 Multiple regression

In this paper, we used four dependent variables (Underpricing, 3-day return, 1-week return, 1-month return) and two dimensions of independent variables (frequency and sentiment), leading us to do a large number of linear regressions. We were trying to avoid missing any information obtained from them, even though some of the models we have done are not statistically significant (e.g., insignificance, negative adjusted R-square). Therefore, in this section, we display the results of 32 linear regression models, the details of each model being as follows.

Table 5.2: Regression models description

Model Number	Description
1-4	Underpricing & ESG frequency
5-8	3-day return & ESG frequency
9-12	1-week return & ESG frequency
13-16	1-month return & ESG frequency
17-20	Underpricing & ESG sentiment
21-24	3-day return & ESG sentiment
25-28	1-week return & ESG sentiment
29-32	1-month return & ESG sentiment

After fitting the initial models, we performed OLS hypothesis validation for each model, and all models passed the tests of multicollinearity and autocorrelation, specifically VIF and DW tests. However, we noticed that some models did not pass the heteroskedasticity test, i.e., in the BP test, the output p-value was less than 0.05, implying that the original hypothesis of homoskedasticity of the model was rejected the conclusion that the models have heteroskedasticity. We can use the heteroskedasticity-robust standard error approach to solve this problem, so in the following regression model results, we show the regression results with robust standard errors.

5.2.1 ESG-term frequency

There is a total of 16 models have been conducted in the ESG-term frequency dimension. The first four models we fit are regression of underpricing on ESG-term frequency. From the regression table of models 1-4, we can see that environment frequency turns to be statistically significant at the 1% level. This means that Norwegian companies listed in the past 2020 years would have higher underpricing if they disclosed more about the environment in their prospectus filed before listing. However, social frequency and governance frequency are not significantly related to underpricing. Secondly, we observed the R-squared and found that the model with all ESG indicators has the highest R-squared, suggesting that the overall model has the best fit. Nevertheless, we also noted that the R-squared for the model with only social frequency or only governance frequency is very small, indicating that these two indicators have extremely low explanatory power for underpricing. In contrast, we saw that the model with only environmental frequency

has a higher R-squared, suggesting that the higher R-squared in the total model is the actual mainly contributed by the environment frequency indicator. The increase in R-squared from the model with only environmental frequency to the overall model is not necessarily due to the fact that social frequency and governance frequency explain more of the underpricing, but may simply to the fact that more noises make the R -squared is slightly higher. As for the control variables, none of them are significant except for covid19 measurement, which is significant at the 10% level in Model 2 and 3.

Table 5.3: Regression of underpricing on ESG freq

	<i>Dependent variable:</i>			
	Underpricing			
	(1)	(2)	(3)	(4)
Environment_frequency	28.726*** (6.573)			31.716*** (6.778)
Social_frequency		-2.650 (8.630)		-11.768 (8.060)
Governance_frequency			-1.970 (9.143)	-0.775 (8.259)
Covid19_measurement	-0.054 (0.036)	-0.078* (0.043)	-0.079* (0.044)	-0.053 (0.037)
Green_company	-0.146 (0.127)	0.138 (0.131)	0.140 (0.131)	-0.181 (0.128)
Market	-0.034 (0.177)	-0.098 (0.213)	-0.095 (0.220)	0.021 (0.183)
Constant	0.118 (0.093)	0.329* (0.171)	0.344 (0.284)	0.311 (0.237)
Observations	48	48	48	48
R ²	0.370	0.093	0.092	0.409
Adjusted R ²	0.312	0.008	0.007	0.322

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

Afterward, the following four models we fitted are regression of 3-day return on ESG-term frequency. From the regression results we observe that none of the variables are significant, and all four R-squared are remarkably low, representing that the explanatory power of these variables on the dependent variable is not different from zero. What is most striking

is that the adjusted R-squared is negative for all models. We have not encountered such results in our past statistical practice, and in fact, negative adjusted R-squared means the explanation towards response is extremely low or negligible, possibly due to our small sample size. In conclusion, we know that the ESG-term frequency has no explanatory power for the 3-day return.

Table 5.4: Regression of 3-day return

	<i>Dependent variable:</i>			
	3-day return on ESG freq			
	(5)	(6)	(7)	(8)
Environment_frequency	1.166 (2.100)			1.604 (2.220)
Social_frequency		-1.289 (2.296)		-1.783 (2.640)
Governance_frequency			-0.523 (2.439)	0.045 (2.705)
Covid19_measurement	-0.010 (0.012)	-0.011 (0.011)	-0.011 (0.012)	-0.009 (0.012)
Green_company	-0.045 (0.041)	-0.034 (0.035)	-0.033 (0.035)	-0.050 (0.042)
Market	0.064 (0.056)	0.066 (0.057)	0.065 (0.059)	0.071 (0.060)
Constant	0.029 (0.030)	0.057 (0.046)	0.051 (0.076)	0.054 (0.078)
Observations	48	48	48	48
R ²	0.067	0.067	0.061	0.079
Adjusted R ²	-0.020	-0.020	-0.026	-0.056
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01			

Models 9-12 are regression of the 1-week return on ESG-term frequency. We observe that models 10-12 still have negative adjusted R-squared, like the ones in previous models, and we argue that social frequency and governance frequency do not have any explanatory power for 1-week return. However, it is worthwhile to note that the adjusted R-squared is positive in the model that includes only environment frequency and is significant and negatively correlated with the independent variable at the 10% level, implying that for

Norwegian companies that go public in 2020, the more disclosures about environmental in the prospectus filed before the listing, the lower the 1-week return would be. However, the adjusted R-squared in the overall model is negative, and we believe that this result may be due to the inclusion of the social frequency and governance frequency indicators.

Table 5.5: Regression of 1-week return

	<i>Dependent variable:</i>			
	1-week return on ESG freq			
	(9)	(10)	(11)	(12)
Environment_frequency	-5.386*			-5.534*
	(2.817)			(2.990)
Social_frequency		-1.074		1.190
		(3.206)		(3.555)
Governance_frequency			-1.790	-1.629
			(3.388)	(3.643)
Covid19_measurement	0.001	0.005	0.004	-0.001
	(0.016)	(0.016)	(0.016)	(0.016)
Green_company	0.029	-0.025	-0.023	0.032
	(0.055)	(0.049)	(0.049)	(0.057)
Market	-0.048	-0.030	-0.022	-0.041
	(0.076)	(0.079)	(0.081)	(0.081)
Constant	0.017	0.002	0.037	0.046
	(0.040)	(0.064)	(0.105)	(0.104)
Observations	48	48	48	48
R ²	0.091	0.017	0.021	0.096
Adjusted R ²	0.007	-0.075	-0.071	-0.036
<i>Note:</i>		*p<0.1; **p<0.05; ***p<0.01		

The last four models in the ESG-term frequency dimension are regression of the 1-month return on ESG-term frequency. We can see that although all three indicators of ESG are not significant, the control variables of Covid19 measurement and Green company demonstrate significant correlations in all four models, in contrast to the non-significant correlations of the control variables in the previous models, such results may suggest that these two indicators do not manifest their effects until the 1-month time point after listing, which is understandable since the effects of a pandemic outbreak and whether a

company is a green company tend to be more far-reaching and may not show up in a very short period of time.

Table 5.6: Regression of 1-month return on ESG freq

	<i>Dependent variable:</i>			
	1-month return			
	(13)	(14)	(15)	(16)
Environment_frequency	-10.197 (8.575)			-8.519 (9.054)
Social_frequency		-9.372 (9.424)		-5.190 (10.766)
Governance_frequency			-7.585 (10.026)	-4.335 (11.032)
Covid19_measurement	0.102** (0.047)	0.109** (0.047)	0.104** (0.048)	0.099** (0.049)
Green_company	0.361** (0.166)	0.256* (0.143)	0.265* (0.144)	0.345* (0.171)
Market	-0.393* (0.230)	-0.333 (0.233)	-0.316 (0.241)	-0.341 (0.245)
Constant	-0.023 (0.121)	0.067 (0.187)	0.137 (0.312)	0.176 (0.316)
Observations	48	48	48	48
R ²	0.227	0.219	0.212	0.238
Adjusted R ²	0.155	0.146	0.138	0.127
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01			

5.2.2 Sentiment of ESG content

The following regression models 17-32 are based on the independent variables of the sentiment dimension. First of all, models 17-20 are regression of underpricing on ESG sentiment score. We see from the regression results that in the single indicator regression, only the environment sentiment score is significant, while social and governance sentiment score are not. However, in the multiple regression model, both environment and governance sentiment scores are significantly related to underpricing at the 5% level. The coefficient of environmental sentiment score is positive, while the coefficient of governance

sentiment score is negative, which means that a Norwegian company listing in 2020 would have a higher underpricing if its pre-listing prospectus is more positive in terms of environmental disclosure; conversely, a more positive disclosure in terms of governance may result in a lower underpricing. Secondly, we find that the model containing three indicators of ESG has the highest R-squared, indicating that it explains the variables best in these four models.

Table 5.7: Regression of underpricing on ESG senti

	<i>Dependent variable:</i>			
	Underpricing			
	(17)	(18)	(19)	(20)
Environment_sentiment_score	0.001* (0.0004)			0.001** (0.001)
Social_sentiment_score		0.0004 (0.0004)		0.001 (0.001)
Governance_sentiment_score			-0.0001 (0.0004)	-0.002** (0.001)
Covid19_measurement	-0.091** (0.042)	-0.083* (0.043)	-0.076* (0.044)	-0.091** (0.040)
Green_company	0.189 (0.130)	0.171 (0.134)	0.130 (0.135)	0.169 (0.125)
Market	0.063 (0.226)	-0.021 (0.229)	-0.160 (0.274)	-0.305 (0.257)
Constant	0.398*** (0.117)	0.325*** (0.109)	0.264** (0.126)	0.286** (0.119)
Observations	48	48	48	48
R ²	0.152	0.109	0.093	0.274
Adjusted R ²	0.074	0.026	0.008	0.168
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01			

When we fit the regression models for 3-day return and 1-week return, we encounter a problem similar to the previous regression results, i.e., the model's adjusted R-squared is negative, as we explained earlier: $\text{adjusted rsquared} = 1 - [(n-1)/(n-p)][\text{SSE}/\text{SSTO}]$ where n is observation size and p is prediction number. In our case, n equals to 48 and p equals to 3, so it is easy to calculate that the coefficient of the last part in this formula

is 47/45; if SSE/SSTO in our model is close enough to 1, then our adjusted R-squared could be negative, which suggests that we have too many predictors to chase too little information. Thus, we need to include more observations to avoid such problems. In summary, regression models 21-28 are all non-significant because none of the independent variables are significant, and their R-squared and adjusted R-squared tell us that these variables barely have any explanatory power over the independent variables.

Table 5.8: Regression of 3 day and 1 week return on ESG senti

	<i>Dependent variable:</i>							
	3-day return				1-week return			
	(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)
E_senti	0.00001 (0.0001)			0.00000 (0.0002)	0.0001 (0.0001)			0.0002 (0.0002)
S_senti		0.00004 (0.0001)		0.0001 (0.0002)		0.0001 (0.0002)		0.0002 (0.0003)
G_senti			-0.00001 (0.0001)	-0.0001 (0.0002)			0.00001 (0.0002)	-0.0002 (0.0003)
Covid19	-0.011 (0.012)	-0.011 (0.012)	-0.010 (0.012)	-0.011 (0.012)	0.003 (0.016)	0.004 (0.016)	0.005 (0.016)	0.003 (0.017)
Green	-0.033 (0.036)	-0.030 (0.036)	-0.034 (0.036)	-0.031 (0.037)	-0.017 (0.050)	-0.017 (0.050)	-0.023 (0.050)	-0.017 (0.051)
Market	0.063 (0.062)	0.070 (0.062)	0.059 (0.073)	0.050 (0.076)	-0.008 (0.086)	-0.013 (0.086)	-0.028 (0.102)	-0.050 (0.105)
Constant	0.037 (0.032)	0.040 (0.029)	0.035 (0.034)	0.032 (0.035)	0.002 (0.045)	-0.005 (0.041)	-0.013 (0.047)	-0.012 (0.049)
Observations	48	48	48	48	48	48	48	48
R ²	0.060	0.063	0.060	0.068	0.025	0.022	0.014	0.041
Adjusted R ²	-0.027	-0.024	-0.027	-0.069	-0.065	-0.068	-0.077	-0.099

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

The last four regression models are regression of the 1-month return on ESG-term frequency. From the regression results, we can find that the governance sentiment score is significant at the 10% level, and its coefficient is negative, which means that the more positive the disclosure about governance in the prospectus submitted before the listing, the lower the 1-month return is likely to be for Norwegian companies listed in 2020. The

model fitted by multiple linear regression has the largest R-squared, indicating that the fit is better compared to the other three.

Table 5.9: Regression of 1-month return on ESG senti

	<i>Dependent variable:</i>			
	1-month return			
	(29)	(30)	(31)	(32)
Environment_sentiment_score	0.0002 (0.0004)			0.001 (0.001)
Social_sentiment_score		-0.0002 (0.0005)		0.0002 (0.001)
Governance_sentiment_score			-0.0005 (0.0004)	-0.001* (0.001)
Covid19_measurement	0.106** (0.048)	0.112** (0.048)	0.117** (0.048)	0.107** (0.047)
Green_company	0.273* (0.148)	0.248 (0.149)	0.222 (0.147)	0.244 (0.146)
Market	-0.320 (0.257)	-0.399 (0.255)	-0.576* (0.299)	-0.657** (0.302)
Constant	-0.053 (0.133)	-0.098 (0.121)	-0.172 (0.137)	-0.148 (0.139)
Observations	48	48	48	48
R ²	0.204	0.203	0.223	0.278
Adjusted R ²	0.130	0.129	0.151	0.172
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01			

5.2.3 Summary of multiple regression results

Table 5.10: Multiple regression summary

	Underpricing	3-day return	1-week return	1-month return
E_freq	significant positive correlation	-	significant negative correlation	-
S_freq	-	-	-	-
G_freq	-	-	-	-
E_senti	significant positive correlation	-	-	-
S_senti	-	-	-	-
G_senti	significant negative correlation	-	-	significant negative correlation

We summarize and present all the regression results in a table, where "-" represents no significant correlation between the two variables. From the above table, we can find that underpricing is significantly correlated with three independent variables. The environmental indicator, both in terms of word frequency and sentiment score, has a positive effect on underpricing. In contrast, at the 3-day time point, none of the independent variables are significant; for 1-week time point, environment frequency is significant and negatively correlated with 1-week return; Finally, for 1-month time point, governance sentiment score is also significant and negatively correlated with the 1-month return.

5.3 Machine learning

From the discussion of multiple linear regression models in the previous section, it is not hard to see that our traditional linear regression models are restrained by many limitations: OLS assumptions, sample size, etc. These limitations may make our linear regression models meaningless and result in the independent variables having no explanatory power on the dependent variable. After getting such results, traditional

regression models can do nothing about it — this is our initial motivation for using machine learning models in our study. In this paper, we use the random forest and SVR (support vector regression). We also fitted the linear model during the machine learning process in order to compare their results.

Random forest is one of the most classic supervised algorithms in machine learning. It is an integrated learning algorithm with decision trees as the base learners. In the classification model, multiple decision trees are processed for voting; in the regression model, multiple decision trees results are averaged. Such averaging system makes random forest reduce the risk of overfitting. It is also a flexible and easy-to-use machine learning algorithm that brings good results in most cases, even without hyperparameter tuning. The most crucial point is that random forest is very stable; even if a new data point appears in the dataset, the whole algorithm would not be affected too much but only affect one decision tree.

SVR is a regression algorithm built on SVM, which is support vector machine. It maps vectors into a higher-dimensional space, where a maximum interval hyperplane is built. Two hyperplanes parallel to each other are on either side of the hyperplane, separating the data, and the separation hyperplane maximizes the distance between the two parallel hyperplanes. It is assumed that the larger the distance or gap between the parallel hyperplanes, the smaller the total error of the classifier. In our case, our model outputs negative values of adjusted R-squared because of the insufficient sample size, which makes our model meaningless. We use the SVR algorithm because it is a small-sample learning method that, by nature, avoids the traditional process of induction to deduction and achieves efficient "transductive inference" from training samples to prediction samples, simplifying the traditional regression problem, which is the same as our current dilemma.

We performed machine learning on all eight overall models and validated the models with 10-fold cross-validation. The table below clearly shows the results of each model for each combination of variables.

The Root Mean Square Error (RMSE) is a commonly used metric for quantifying the error of a model in predicting data. The RMSE tells us how far the predicted value is from the true value in general, thus allowing us to evaluate the model's accuracy. In addition, R-squared is another measure of goodness-of-fit that we use, the closer the

R-squared is to 1, the better the model.

Table 5.11: Machine Learning Results

	Linear Model		Random Forest		SVR	
	RMSE	Rsquared	RMSE	Rsquared	RMSE	Rsquared
Underpricing + freq	0.344	0.264	0.277	0.482	0.315	0.348
3-day return + freq	0.118	0.326	0.105	0.184	0.094	0.284
1-week return + freq	0.139	0.128	0.136	0.414	0.146	0.320
1-month return + freq	0.456	0.180	0.352	0.400	0.423	0.309
Underpricing + Senti	0.355	0.320	0.324	0.237	0.310	0.362
3-day return + Senti	0.102	0.305	0.102	0.193	0.100	0.262
1-week return + Senti	0.161	0.367	0.152	0.289	0.150	0.241
1-month return + Senti	0.441	0.176	0.391	0.305	0.449	0.293

The random forest has slightly more optimal results than linear model ones in R-squared, and SVR has the least optimal results. In terms of RMSE, five of the eight models have the best results from SVR, and the remaining three have the best results from the random forest algorithm. From the overall viewpoint, we use the average of the eight models to measure the overall algorithm superiority, and then we find that random forest has the best fit superiority results in both RMSE and R-squared.

6 Discussion

6.1 Discussion of textual analysis

This study is necessary to determine how pre-listed enterprises commit to ESG activities to compare the ESG performance of each company and continue our research. However, since no traditional ESG score is available for private enterprises, the only way forward is for private firms to develop their own ESG metrics. As previously stated, we will use two approaches of textual analysis in this research: term frequency and sentiment score. To begin, we will review the term frequency method's outcomes. Using our ESG dictionary, we can determine which terms exist in collecting prospectuses and admission documents.

In terms of the environmental context, waste and recycling seem to be heavily emphasized throughout the document collection. The terms commonly appear in the documents: pollution, hazardous, recycling, hazard, waste, emission, greenhouse, GHG, and plastic. This indicates that businesses are prepared to share how they manage garbage generated by their activities. The prevalence of the phrases "hydro," "wind," and "renewable" in a few documents demonstrate the tendency toward employing renewable resources. The social context in the prospectus or admission document generally focuses on labor standards and health, which is very important for corporations to present. Access to medication, education, and employment commonly occur in numerous documents. Surprisingly, the elements addressing diversity like the issues about gender and equality are featured in few prospectuses. It indicates that enterprises do not view these concerns as significant and do not need to be handled in the business setting. In the governance context, corporate governance, audit, and control are the criteria that every corporation needs to report under the stock markets' requirements. Several terminologies and elements are also addressed in documents, e.g., bribery, corruption, and transparency, because the Norwegian government emphasizes these issues. Overall, the goal of the social and governance elements is more into completing the requirements of the IPO prospectus or private placement admission documents, displaying the legal compliance of the businesses. Meanwhile, the environmental components tend to be volunteer disclosure compared to social and governance criteria. It could also be shown from the level of disclosure in different factors - G ranks the highest among three indicators, followed by S and, at last,

the E factor.

Sentiments more accurately reflect the quality of ESG operations. The descriptive study shows that all sentiment scores for ESG are mostly negative, particularly obvious for the governance factor in all prospectuses. Firms are nonetheless required to declare certain ESG practices that they cannot carry out, resulting in a negative tone in ESG settings. Meanwhile, the sentiment scores for Social and Environment are comparably better, indicating that a few businesses do better than others in terms of the related actions.

6.2 Discussion of regressions

This section discusses the study's primary objective: to examine the relations between a firm's ESG score and short-term return (underpricing). Thus, we can obtain two different indicators - term frequency and sentiment score through textual analysis.

In terms of frequency analysis, we assume that the more ESG activities a company engages in, the more ESG information the company wants to reveal to investors. Therefore, more ESG keywords appear in the IPO documents. We observe that the variable Environment frequency is statistically significant and positively relates to underpricing in the first three regressions. This actually violates the information revelation theory (Benveniste & Spindt, 1989) that investors perceive ESG firms as less risky, so they do not require more underpricing as compensation for offering higher demand, and firms are less underpriced compared to other firms in the market. The possible explanation is that if higher ESG disclosure attracts investor attention, underpricing should be greater for IPOs (Da et al., 2011; Liu et al., 2014). Additionally, according to the signaling hypothesis (Allen Faulhaber, 1989), companies may also want to underprice their listings by showing a better business value. Additionally, when investors become more concerned about ESG during the sustainability investment trend, their demand for equities increases, driving the price of the stocks higher on the first day and underpricing should be greater for IPOs (Da et al., 2011; Liu et al., 2014). Furthermore, Environment frequency is slightly negative and significant with the 1-week return. This means that, as time goes on, if corporations provide more information about the environment in their documents, their listings will be less underpriced. The findings are consistent with the hypothesis that

investors truly care about the businesses' environmental setting. The more information corporations provide, the less assurance there is regarding the firm's ex-ante value. As a result, corporations tolerate less underpricing, according to the Winner's curse argument (Rock, 1986). However, we recognize that investors do not seem to evaluate the social and governance contexts of listings since the Social frequency and Governance frequency are insignificant for all regression models, which have no impact on short-term returns. This may be plausible given that their primary emphasis is on labor standards and audit control, which seem consistent across all firms. It also means that investors may not feel less unsure since companies are expected to ensure that their activities adhere to the law to a sufficient degree.

Regarding sentiment analysis, which shows the attitude investors show to the ESG disclosure. This strategy elicits a more direct assessment of the ESG activity's quality. Contexts are classified into three categories, which correspond with three factors of ESG. The regression analysis reveals that the Environment sentiment score is significantly positive with underpricing while the Governance sentiment score is significantly negative. However, in other regressions, only the Governance sentiment score is negative and significant with the 1-month return. To account for the phenomena, we can connect it to the Winner's curse theory. As can be seen from the descriptive statistics that all three ESG sentiment scores have negative means, and since investors pay attention to environmental concerns, they may see non-environmental firms as risky. As a result, ex-ante uncertainty regarding the firm's valuation grows. Thus, corporations and underwriters need to underprice the IPO to create more demand. On the other hand, it seems as if the majority of prospectus and admission documents have a negative tone when discussing the governance framework since there are several regulations that corporations do not comply with.

Moreover, we also need to think about the control variables in our study, which are Covid19 measurement, Green company and Market. First, the pandemic has a weakly positive relationship with the 1-month return, which means the more companies are affected by the pandemic, the more abnormal the return they will get. This outcome is understandable given the small sample size and information provided from the covid period. As an exogenous emergency event, the pandemic has characteristics such as unpredictability and rapid spread, making it difficult for companies to respond promptly

when it hits. As a result, investors' uncertainty increases, which in turn raises the underpricing of companies' listings. The next control variable: Green company depicts the label of listed companies. We found that it is only significant for the 1-month return and explains a positive relationship, indicating that the "green impact" may take time to materialize completely. It is not contradicting previous scholars' research. Michael and Denis(2004) discover that technology businesses are more likely to be underpriced after adjusting for other factors in their sample of 1993–2000 IPOs. Walker et al.(2015) examine the relationship between lawsuit risk and underpricing for IPOs registered in the United States between 1996 and 2008 and discover that technology businesses are much more underpriced than non-technology enterprises, which they ascribe to increased ex-ante uncertainty. We suggest that it can be expanded to green firms as well since both are often composed of new enterprises depending on future development opportunities. As a result, we believe that green firms with higher ex-ante uncertainty would face larger underpricing, as Beatty and Ritter(1986) suggest. The last control variable, the two different markets corporations are listing in, shows an insignificant relation to almost all models. It means that the choice of IPO in Oslo Børs or private placement in Euronext Growth will not significantly differ the corporations' underpricing and short-term returns. Understandably, the private placement could be viewed as a simpler version of IPO. As such, the driven factors and effects can also be applied for both types of listings.

6.3 Implications

Our analysis combines two topics: short-term return (underpricing), a long-established concept in corporate finance, and ESG, a relatively new and developing concept. The amount of studies examining the link between ESG and listings, particularly with an emphasis on IPO underpricing, is small, and our work contributes to filling this research gap. The first hurdle that researchers must overcome when examining the ESG features of private enterprises is the lack of traditional ESG ratings. As a result, we suggest a novel technique based on textual analysis of IPO prospectuses and private placement admission documents. Additionally, we effectively use two textual analysis techniques: term frequency analysis for ESG keywords and sentiment analysis for the ESG context. The thesis has both contributed to the different listings and underpricing literature by

demonstrating how the market perceives certain ESG factors in a particular context. The findings imply that businesses' ESG features should be evaluated as predictors of IPO underpricing. Moreover, the findings demonstrate the market's integration of ESG into business. However, investors' consideration of social and governance parameters should be enhanced throughout the listing evaluation process.

Concerning underpricing, we utilized and described many control factors that have been stated before in the literature. We demonstrate that some control variables, including Covid19 measurement and Green company, have explanatory power for 1-month short term return, providing new possible affecting factors in the new era. Additionally, our analysis adds to the empirical evidence that the market does care about ESG factors, particularly in the context of listings. We have shown that investors consider the quality of a firm's environmental efforts when making IPO investment. This suggests that real-world examples of sustainable investing exist.

Our analysis not only addresses the research topic of whether ESG factors affect listing underpricing and short-term return, but also enables us to address a larger question: "How does the market view the ESG problems in a given setting?" We discovered evidence that market participants closely assess the environmental actions of corporations throughout the listing process.

On the other side, unlike the environmental criteria, the social and governance components of the company do not pique market interest. This demonstrates that the market should place a higher premium on social and governance considerations. Additionally, more rules should be included to force corporations to report their social and governance efforts separately and clearly in their listing documents. Because the market is aware of ESG issues, it encourages pre-listing enterprises, even if they are private, to have sustainable reporting. Furthermore, professional ESG rating companies may provide ESG ratings for private enterprises to raise their obligations on ESG activities.

6.4 Limitations

We accept the limits of our research and think carefully about how we might mitigate their impact on our findings. The study's first hurdle is generating an ESG score for

pre-listing enterprises, i.e., those that remain private and do not provide ESG information. Moreover, as previously stated, we depend on the IPO prospectus and the private placement admission document from a reliable data source. However, the market may get information about private enterprises from various data sources, including the news and the private information of other market players. As a result, our data source selection may be inadequate.

Second, we are aware of our textual analysis methodologies' limits. Both term frequency and sentiment analysis may involve removing the words from their context and underestimating the document's multidimensional nature (Loughran and McDonald, 2016). For instance, the ESG terms used in the thesis may have a different meaning in other settings that we are unaware of. The sentiment dictionary application is constrained because we used the LM dictionary without additional sentiment dictionaries tailored to the ESG scenario. For our own ESG dictionary, after using the word2vec algorithm to derive the results, we only solved the problem of repetition and partial word meanings, but in fact we did not do a more detailed evaluation of each word in the dictionary, which led to the fact that our dictionary may not measure the ESG metrics in a multidimensional way, but only based on our vector basis, which may be the main problem of textual analysis in this paper. In future research, we can further evaluate and filter the words in the lexicon, which may make the results of textual analysis more accurate.

Our small data sample is the third restriction of the thesis. The Euronext Growth market has become popular recently and in results do not give adequate companies in the time period for analysis. As we can see from the regression analysis, when we fit the regression models for 3-day return and 1-week return, our adjusted R-squared might be negative, which shows that we have too many factors to pursue too little information. Thus, we need to incorporate more observations to prevent such difficulties. Moreover, some firms were excluded from our final sample due to missing data or inapplicable with our analysis method. Both of these constraints may have resulted in skewed results. Additionally, it is worth noting that the sample of green firms is drawn from a condensed time period of a few months. The significantly greater underpricing of green firms may not hold with a larger sample size or over a more extended time period.

Finally, even if the control variables included in the regressions are carefully chosen, the

variable selection is restricted. Certain variables impacting underpricing are outside the scope of our investigation, such as the firm's age, investor's characteristics. However, there is compelling evidence that our model is implemented correctly, based on underpricing theories and other empirical investigations.

7 Conclusion

7.1 Summary of findings

The purpose of this thesis is to determine whether a firm's ESG disclosure affects its underpricing and short-term return. We researched by self-constructing ESG indicators for pre-listing corporations using two textual analysis techniques: term frequency analysis and sentiment analysis. We suggest that if businesses engage in ESG activities, they are ready to highlight them prominently in their listing documents. As a result, ESG terms are widely used in the context. Additionally, we use sentiment analysis to determine how investors perceive the ESG material in prospectuses and admission papers, providing additional insight into the quality of ESG operations.

To research the effect of ESG indicators and the extent to which they are underpriced, we conduct multiple regressions on each indicator, including the control variables. The first finding is that none of the variables relating to social and governance criteria are statistically significant, implying that investors may not consider their characteristics when deciding whether to invest in a company's listing. The second finding is that investors' concerns about environmental issues are expressed more sophisticatedly during the listing process. We can see that the more environmental information disclosed in a firm's prospectus, the more underpriced the firm's listing is, as determined by the frequency of environmental terms. And after one week, their listings become less underpriced supporting the theories that investors are really concerned about the environmental impact of the companies by more information revealed. Besides, we see that all sentiment scores for ESG factors are mostly negative. The sentiment expressed in the listing materials about the environment affects the underpricing. This raises the investor's ex-ante uncertainty about the firm's worth and results in a significant underpricing of the listing. Last, we notice that underpricing would be greater during the covid-19 pandemic period, and that green enterprises face much more underpricing than non-green ones.

In summary, we expand the literature on listing underpricing to a new section of the Norwegian market with our thesis. We may respond to our study question by stating that

some ESG characteristics do have an effect on a firm's underpricing and short-term return. On the one hand, investors are mostly unconcerned with a firm's social and governance initiatives, opting instead for environmental ones. Investors, on the other hand, place a higher premium on the quality of environmental operations. To get a high rating from investors, a business must have conspicuous and distinguishable environmental initiatives in comparison to other businesses. Additionally, we suggest that the other listing participants, issuers and underwriters, see all enterprises that are not environmentally sustainable as risky because they underprice their stocks to avoid potential future costs.

7.2 Further research

Our research objective is to evaluate the influence of ESG variables on the listing underpricing and the company's short-term return. As more firms are listed on the Norwegian stock market, it would be interesting to study if the green impact shown in this thesis holds true for a larger sample size and research the comparison of listing between Euronext Growth and Oslo Børs. We also encourage further research by duplicating the studies in terms of modifying the sample, for example, conducting the study for various geographical locations to assess whether, in other nations, the markets view the ESG criteria in different ways. To continue, future studies could explain why there is a difference between countries. For example, in Norway, we discovered that investors are concerned about a firm's environmental activities in the context of listings; however, Bollazzi and Risalvato(2017) conclude that ESG factors have no effect on IPO underpricing in the Italian market. The cause may be found in the regulations and the openness of the private firms' data. Another path for additional study is to uncover why investors prioritize the environmental element more than the social and governance. Besides, the longitudinal studies of the ESG influence will be beneficial to monitor the trend across the lengthy periods, whether the market has altered its assessment of ESG elements or not.

References

- Abrahamsen, T., Sveen, L. (2021). What you should know about a Euronext Growth Oslo listing. Thomessen. Available at: <https://www.thommessen.no/en/news/euronext-growth>
- Advokatfirmaet Wiersholm. (2021). Norway MA and IPO trends 2020. Available at: <https://www.wiersholm.no/en/newsletter/norway-ma-and-ipo-trends-and-insights-2020>
- Allen, F. Faulhaber, G.R. (1989). Signaling by underpricing in the IPO market. *Journal of Financial Economics*, 23 (2), 303-323.
- Amihud, Y., Mendelson, H. Pedersen, L. (1988). Liquidity and Asset Prices. *Foundations and Trends in Finance*, 1(4), 269-364.
- Andreou, P. C., Harris, T., Philip, D. (2020). Measuring firms' market orientation using textual analysis of 10-K filings. *British Journal of Management*, 31(4), 872-895.
- Arenas-Parra, M., Álvarez-Otero, S. (2020). CSR Disclosure: The IPO Case. *Sustainability*, 12(11), 4390.
- Baier, P., Berninger, M., Kiesel, F. (2020). Environmental, social and governance reporting in annual reports: A textual analysis. *Financial Markets, Institutions Instruments*, 29(3), 93-118.
- Beatty, R. P., Ritter, J. R. (1986). Investment Banking, Reputation, and the Underpricing of Initial Public Offerings. *Journal of Financial Economics*, 15, 213.
- Becchetti L, Ciciretti R, Hasan I, et al. (2012). Corporate social responsibility and shareholder's value. *Journal of Business Research*, 65(11), 1628-1635.
- Benveniste, L. M. Spindt, P. A. (1989). How investment bankers determine the offer price and allocation of new issues. *Journal of Financial Economics*, 24(2), 343-361.
- Black, B. S., Gilson, R.J. (1998). Venture capital and the structure of capital markets: Banks versus stock markets. *Journal of Financial Economics*, 47, 243-277.

- Bollazzi, F. Risalvato, G. (2017). IPO and CSR: An Analysis on Last Performance in Italian Stock Exchange. *China-USA Business Review*, 16(12), 588-600.
- Bryman, A. Bell, E. (2015). *Business Research Methods*. 4th edition. Oxford: Oxford University Press.
- Bui, T. M. A., Frongillo, A. (2020). How does the market perceive ESG in IPOs: Investigating how ESG factors affect IPO Underpricing in the U.S. market (Dissertation). Available at: <http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-172446>
- Chang, C.H. (2011). IPO underpricing: A social comparison perspective. *International Review of Economics Finance*, 20, 367-375.
- Chemmanur, T. J. Fulghieri, P. (1999). A Theory of the Going-Public Decision. *Review of Financial Studies*, 12, 249-79.
- Dolvin, S. D. (2012). IPO Underpricing: The Owners' Perspective. *Journal of Economics and Finance Education*, 11(2), 63-69.
- Doyle, T. M. (2018). Ratings that don't rate: The subjective world of ESG rating agencies. American Council for Capital Formation. Available at: https://accfcorgov.org/wp-content/uploads/2018/07/ACCF_RatingsESGReport.pdf
- Drobetz W, Schillhofer A, Zimmermann H. (2003). *Corporate governance and expected stock returns: Evidence from Germany*. John Wiley Sons.
- Durán-Santomil, P., Otero-González, L., Correia-Domingues, R.H. Reboredo, J.C. (2019). Does Sustainability Score Impact Mutual Fund Performance?. *Sustainability*, 11(10), 1-17.
- Eckbo, B. E., Masulis, R. W. Norli, O. (2008). *Handbook of Corporate Finance*. 1st Edition. Oxford: Elsevier Science Technology.
- Emilsen, N. H., Pedersen, K., Sættem, F. (1997). Børsintroduksjoner. *Beta-Tidsskrift for bedriftsøkonomi*, 1-13.
- Eun-Hee Kim, Thomas P. Lyon. (2014). Greenwash vs. Brownwash: Exaggeration and Undue Modesty in Corporate Sustainability Disclosure. *Organization Science*, 26(3), 705-723.

- Eurosif. (2018). European SRI Study 2018. Available at: <https://www.eurosif.org/wp-content/uploads/2018/11/European-SRI-2018-Study.pdf>
- European Commission. (2011). Communication from the commission to the European Parliament, The council, the European Economic and Social Committee and the Committee of regions: A renewed EU strategy 2011-14 for Corporate Social Responsibility. Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52011DC0681from=EN>
- Fenili, A., Raimondo, C. (2021). ESG and the Pricing of IPOs: Does Sustainability Matter. FEN: Institutions Financing Practices (Topic).
- Fernandez-Kranz D, Santalo J. (2010). When necessity becomes a virtue: The effect of product market competition on corporate social responsibility. *Journal of Economics Management Strategy*, 19(2), 453-487.
- Fjesme, S. L. (2011). Laddering in Initial Public Offering Allocations. Centre for Corporate Governance Research.
- 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, 210-233. Friedman, M. (1970). The Social Responsibility of Business Is to Increase Its Profits. *New York Times Magazine*, 13 September 1970, 122-126.
- Geddes, R. (2003). *IPOs and equity offerings*. 1st edition. Oxford: Elsevier Science Technology.
- Giese G, Lee L E, Melas D, et al. (2019). Foundations of ESG investing: How ESG affects equity valuation, risk, and performance. *The Journal of Portfolio Management*, 45(5), 69-83.
- Global Sustainable Investment Alliance. (2020). *Global Sustainable Investment Review 2020*. Available at: <http://www.gsi-alliance.org/wp-content/uploads/2021/08/GSIR-20201.pdf>
- Gregory A, Tharyan R, Whittaker J. (2013). Corporate social responsibility and firm value: Disaggregating the effects on cash flow, risk and growth. *Journal of Business Ethics*, 124(4), 633-657.

- Hertzel, Michael and Richard L. Smith. (1993). Market Discounts and shareholder Gains for placing Equity Privately. *Journal of Finance*, 48, 459-485.
- Holmström, B., Tirole, J. (1993). Market Liquidity and Performance Monitoring. *Journal of Political Economy*, 101(4), 678-709.
- Hong, Harrison Kacperczyk, Marcin. (2009). The price of sin: The effects of social norms on markets. *Journal of Financial Economics*, Elsevier, 93(1), 15-36.
- Huang, F. , Xiang, L., Liu, R., Su, S. Qiu, H. (2019). The IPO corporate social responsibility information disclosure: Does the stock market care?. *Accounting Finance*, 59, 2157-2198.
- Investopedia. (2021). IPO vs. Private Placement: What's the Difference?. Available at: <https://www.investopedia.com/ask/answers/021015/what-difference-between-ipo-and-private-placement.asp>
- KPMG. (2015). A Guide to GOING PUBLIC. Available at: <https://assets.kpmg/content/dam/kpmg/pdf/2015/06/KPMG-A-Guide-to-Going-Public-Interactive.pdf>
- Kumar, N. C. A., Smith, C., Badis, L., Wang, N., Ambrosy, P. Tavares, R. (2016). ESG factors and risk-adjusted performance: a new quantitative model. *Journal of Sustainable Finance Investment*, 6(4), 292-300.
- Lang, M., Stice-Lawrence, L. (2015). Textual analysis and international financial reporting: Large sample evidence. *Journal of Accounting and Economics*, 60(2-3), 110-135.
- Ljungqvist, A. (2008). Chapter 7: IPO Underpricing. In: Eckbo, B.E., ed. *Handbook of Corporate Finance*. 1st Edition. Oxford: Elsevier Science Technology.
- Ljungqvist, A., Nanda, V., Singh, R. (2006). Hot Markets, Investor Sentiment, and IPO Pricing. *The Journal of Business*, 79(4), 1667-1702.
- Lopez-de-Silanes, F., McCahery, J. A., Pudschedl, P. C. (2020). ESG performance and disclosure: A cross-country analysis. *Singapore Journal of Legal Studies*, (Mar 2020), 217-241.

- Loughran, T., McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of finance*, 66(1), 35-65.
- Maksimovic, V., and P. Pichler. (2006). Structuring the initial offering: Who to sell to and how to do it. *Review of Finance*, 10, 1-36.
- Mcpeak, C. and Tooley, N. (2008). Do Corporate Social Responsibility Leaders Perform Better Financially?. *Journal of Global Business Issues*, 2, 982-985.
- Michael, C. T., Denis, D. J. (2004). Do Initial Public Offering Firms Purchase Analyst Coverage with Underpricing? *The Journal of Finance*, 59(6), 2871-2901.
- Mohammad, W. M. W., Wasiuzzaman, S. (2021). Environmental, Social and Governance (ESG) disclosure, competitive advantage and performance of firms in Malaysia. *Cleaner Environmental Systems*, 2, 100015.
- Nollet, J., Filis, G. Mitrokostas, E. (2015). Corporate social responsibility and financial performance: A non-linear and disaggregated approach. *Economic Modelling*, 52, 400-407.
- Oslo Børs. (2020). Comparison between Oslo Børs, Oslo Axess and Merkur Market. Available at: https://www.oslobors.no/ob_eng/layout/set/print/Oslo-Boers/Listing/Shares-equity-certificates-and-rights-to-shares/Comparison-between-Oslo-Boers-Oslo-Axess-and-Merkur-Market
- Oslo Børs. (2018). Oslo Børs guidance on the reporting of corporate responsibility.
- Park, H.D. Patel, P.C. (2015). How Does Ambiguity Influence IPO Underpricing? The Role of the Signaling Environment. *Journal of Management*, 52, 796-818.
- PwC. (2017). Considering an IPO to fuel your company's future? Insight into the costs of going public and being public. Available at: <https://www.pwc.com/us/en/deals/publications/assets/cost-of-an-ipo.pdf>
- Ritter, J. R. (2021). IPO Data. Warrington College of Business. Available at: <https://site.warrington.ufl.edu/ritter/ipo-data/>
- Ritter, J.R. Welch, I. (2002). A Review of IPO Activity, Pricing, and Allocations. *The Journal of Finance*, 57, 1795-1828.

- Rock, K. (1986). Why new issues are underpriced. *Journal of Financial Economics*, 15, 187-212.
- Rydqvist, K. Högholm, K. (1995). Going public in the 1980s: Evidence from Sweden. *European Financial Management*, 1, 287-315.
- Shane, P. B., Spicer, B. H. (1983). Market Response to Environmental Information Produced outside the Firm. *The Accounting Review*, 58(3), 521-538.
- Sharpe, I., and L. Woo. (2005). Corporate control, expected underpricing, and the choice of issuance mechanism in unseasoned equity markets. *Journal of Corporate Finance*, 11, 716-35.
- Vance, S. (1975). Are Socially Responsible Corporations Good Investment Risks?. *Managerial Review*, 64, 18-24.
- Van Dijken, F. (2007). Corporate social responsibility: market regulation and the evidence. *Managerial Law*, 49(4), 141-184.
- Walker, T., Turtle, H. J., Pukthuanthong, K., Thiengtham, D. (2015). Legal oppourtunism, litigation risk, and IPO underpricing. *Journal of Business Research*, 68(2), 326-340.
- Welch. (1992). Sequential Sales, Learning, and Cascades. *Sequential Sales, Learning, and Cascades. The Journal of Finance*, 47(2), 695-732.
- Witold Henisz, Tim Koller, and Robin Nuttall. (2019). Five ways that ESG creates value.
- Wruck, Karen Hopper. (1989). Equity Ownership Concentration and Firm Value: Evidence from private equity Önancing. *Journal of Financial Economics*, 23, 3-28.
- Yilmaz, M. K., Aksoy, M., Tatoglu, E. (2020). Does the stock market value inclusion in a sustainability index? Evidence from Borsa Istanbul. *Sustainability*, 12(2), 483.
- Zingales, L. (1995). Insider ownership and the decision to go public. *Review of Economic Studies*, 62, 425-448.

Appendix

A1 Output of ESG initial word list by using Word2vec algorithm

S: Social (also presented in Figure 4.3)

	0	1	2	3	4	5	6	7	8	9
health	healthcare	hospital	population	welfare	payer	patient	age	connected	ehealth	medical
human	animal	routine	nutrition	food	computer	perceive	error	telecommuni	safety	technical
labor	wage	skilled	workforce	injury	gender	equality	litigation	death	certainty	skill
train	skilled	personnel	recruit	replacement	attract	experienced	shortage	competence	reliant	skill
society	transition	energy	europa	renewable	minimize	emission	climate	sector	utility	stakeholder
education	entertainment	elearning	chess	learning	career	nearly	child	learn	workforce	leadership
hire	personnel	recruit	skilled	continued	retain	skill	train	frame	key	klovning
employment	termination	severance	benefit	employee	terminate	salary	sverre	equality	options	gender
access	enable	help	opportunity	platform	sensitive	cloud	digital	easily	deliver	solution
public	council	repeal	parliament	incrimination	state	disqualify	review	offer	inquiry	expression

Figure A1.1: Social extended words

E: Environmental

	0	1	2	3	4	5	6	7	8	9
clean	mix	advanced	pure	environmentally	solar	light	cloudberry	approach	modern	air
environmental	safety	pollution	problem	strict	food	hazard	hazardous	welfare	animal	death
climate	greenhouse	emission	ghg	imo	fossil	goal	society	paris	gas	sustainable
renewable	energy	hydro	power	wind	nordics	fossil	transition	electricity	society	grid
carbon	capture	emission	fuel	acc	plastic	sustainable	cell	cleaning	gas	recovery
waste	plastic	recycling	chemical	agilyx	polystyrene	feedstock	crude	recycle	ps	heat
emission	greenhouse	ghg	shipping	fuel	imo	carbon	zero	climate	shift	sulphur
water	pollution	food	chemical	gas	climate	goal	carbon	fossil	air	heat
pollution	environmental	imo	hazardous	greenhouse	emission	air	strict	hazard	port	ghg
air	indoor	hazardous	pollution	rich	emission	environmentally	wave	nutrient	contamination	waste

Figure A1.2: Environmental extended words

G: Governance

	0	1	2	3	4	5	6	7	8	9
governance	practice	code	corporate	conflict	obligations	employees	requirements	implement	remuneration	recommendation
approve	authorise	authorize	rules	fsa	approval	inquiry	review	resolution	waive	/.
transparency	index	expand	acquirer	e&p	society	creation	constantly	diversify	sharing	børs
audit	consolidated	unconsolidated	audited	unaudited	interim	unaudite	end	ngaap	dgaap	december
control	detect	misstatement	override	collusion	influence	fraud	prevent	absolute	turn	perform
stakeholder	sustainable	society	ruche	ambition	chain	zero	green	geography	emission	approach
engagement	greve	services	role	consulting	experience	letter	marketing	ms.	microsoft	professional
compliance	regulation	comply	applicable	rule	requirement	ensure	acknowledge	reasonably	consummate	evidence
responsible	supervise	instruction	proper	budget	day	ensure	fit	investigation	manage	delegate
structure	chart	history	organisational	lei	subsidiary	ayfie	inc.	registered	materials	identifier

Figure A1.3: Governance extended words

A2 ESG dictionary

Table A2.1: Final ESG Dictionary

Environmental	clean, environmental, climate, renewable, carbon, waste, mix, greenhouse, energy, capture, indoor, advanced, emission, hydro, recycling, pure, problem, ghg, power, fuel, chemical, shipping, imo, wind, agilyx, rich, solar, fossil, nordics, plastic, polystyrene, light, hazard, goal, feedstock, environmentally, cloudberry, hazardous, transition, cell, crude, strict, wave, paris, electricity, cleaning, recycle, nutrient, modern, animal, gas, shift, port, contamination, sustainable, grid, recovery, heat, sulphur
Social	population, nutrition, workforce, recruit, europe, chess, benefit, opportunity, parliament, welfare, food, injury, replacement, learning, continued, employee, platform, incrimination, payer, computer, attract, minimize, career, retain, health, human, labor, train, society, education, hire, employment, access, public, healthcare, wage, skilled, entertainment, personnel, termination, enable, council, hospital, routine, elearning, severance, help, repeal, terminate, sensitive, state, patient, perceive, equality, experienced, nearly, skill, salary, cloud, disqualify, age, error, litigation, shortage, child, sverre, digital, connected, telecommunication, death, competence, sector, learn, frame, easily, ehealth, safety, certainty, reliant, utility, medical, key, options, deliver, technical, leadership, klovning, gender, solution, expression
Governance	obligations, approval, interim, influence, chain, experience, requirement, day, subsidiary, employees, inquiry, creation, unaudite, fraud, zero, letter, requirements, review, constantly, end, prevent, green, marketing, acknowledge, fit, inc., implement, resolution, diversify, ngaap, absolute, geography, reasonably, investigation, registered, remuneration, waive, sharing, dgaap, turn, microsoft, consummate, manage, materials, recommendation, december, perform, approach, professional, evidence, delegate, identifier, code, authorize, expand, unconsolidated, misstatement, governance, approve, transparency, audit, control, stakeholder, engagement, compliance, responsible, structure, practice, authorise, index, consolidated, detect, greve, regulation, supervise, chart, services, comply, instruction, history, corporate, rules, acquirer, audited, override, ruche, role, applicable, proper, organisational, conflict, fsa, e&p, unaudited, collusion, ambition, consulting, rule, budget

A3 Part of Python-code for textual analysis

```
import numpy as np
import pandas as pd
import os
```

```
def read_file(path):
    with open(path, 'r',encoding='utf-8') as file:
        content = file.read()
    return content

data_out = pd.DataFrame()

#read files
for root,dirs,files in os.walk(r"D:\NHH\Master_thesis\final_data\text_txt"):
    for file in files:
        print (os.path.join(root,file))
        path=os.path.join(root,file)
        str1=read_file(path)
        new = pd.DataFrame({'class': file, 'content': str1}, index=[0])
        data_out = data_out.append(new, ignore_index=True)

import pandas as pd
#data_out.to_excel(r'.\data.xlsx')

data_out
data_out['content'][0]
data_out.drop(data_out[data_out['content']==''].index,inplace=True)

#remove NA lines
data_out.dropna(subset=['content'],inplace=True)
data_out.reset_index(drop=True, inplace=True)
data_out

for i,j in zip(data_out['content'],data_out['class']):
    print(len(i),j)

from sklearn.feature_extraction import text
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.metrics import accuracy_score
from sklearn.base import TransformerMixin
from sklearn.pipeline import Pipeline
from sklearn.svm import LinearSVC

import string
import spacy
```

```

import re

punctuations = string.punctuation
punctuations=punctuations+'.1234567890'
parser = spacy.load('en_core_web_sm')

#stop words
stopwords = text.ENGLISH_STOP_WORDS

#tokens
def spacy_tokenizer(sentence):
    sentence = re.sub('[\u4e00-\u9fa5\u0030-\u0039]', '', sentence)
    #nlp objective
    tokens = parser(sentence)
    #word stemming
    tokens = [tok.lemma_.lower().strip() if tok.lemma_ != "-PRON-" else
               tok.lower_ for tok in tokens]
    #remove stop words, place names, etc.
    tokens = [tok for tok in tokens if (tok not in stopwords and tok not in
                                         punctuations)]
    return tokens

#create words bag
files=[]
for i in data_out['content']:
    i=spacy_tokenizer(i)
    files.append(i)

#####
#####Word2vec#####
#####

import gensim
sentences=files
model=gensim.models.Word2Vec(sentences,sg=1,window=20,min_count=20,hs=1,workers=8)
model.save("Word2vecmodel")

```

```
gensim.models.Word2Vec.load("Word2vecmodel")

import gensim
model2=gensim.models.Word2Vec.load("Word2vecmodel")

environmental=['clean', 'environmental', 'climate', 'renewable', 'carbon',
              'waste', 'emission', 'water', 'pollution', 'air']
print(len(environmental))

social=['health', 'human', 'labor', 'train', 'society', 'education', 'hire',
       'employment', 'access', 'public']
print(len(social))

governance=['governance', 'approve', 'transparency', 'audit', 'control',
            'stakeholder', 'engagement', 'compliance', 'responsible', 'structure' ]
print(len(governance))

aa=model2.wv.most_similar("clean",topn=10)
def w2v(ci):
    aa=model2.wv.most_similar("%s"%ci,topn=10) #top10 similar words
    word=[]
    sim=[]
    for i in aa:
        word.append(i[0])
        sim.append(i[1])
    df=pd.DataFrame(
        {
            "%s"%ci:word,
            "Similarity":sim
        }
    )
    df.to_excel(r'%sextendedtop10.xlsx'%ci)
    return df
df=w2v("clean")
```

```
df

data_environmental = pd.DataFrame()
for i in environmental:
    try:
        data_environmental = data_environmental.append(w2v(i).iloc[:, 0])
    except:
        pass

data_governance = pd.DataFrame()
for i in governance:
    try:
        data_governance = data_governance.append(w2v(i).iloc[:, 0])
    except:
        pass

data_social = pd.DataFrame()
for i in social:
    try:
        data_social = data_social.append(w2v(i).iloc[:, 0])
    except:
        pass

data_environmental.to_excel('data_environmental.xlsx')
data_governance.to_excel('data_governance.xlsx')
data_social.to_excel('data_social.xlsx')

import numpy as np
#transfer the matrix into list for the following calculation loop
environmental_extend=np.array(data_environmental).reshape(1,100)[0]
governance_extend=np.array(data_governance).reshape(1,100)[0]
social_extend=np.array(data_social).reshape(1,100)[0]
```

```
environmental_num=[]
governance_num=[]
social_num=[]
all=[]
for i in files:
    all.append(len(i))
    a=0
    b=0
    c=0
    for j in i:
        if j in environmental_extend:
            a=a+1
        if j in governance_extend:
            b=b+1
        if j in social_extend:
            c=c+1
    environmental_num.append(a)
    governance_num.append(b)
    social_num.append(c)
environmental_num,governance_num,social_num,all

data_out['environmental_num']=environmental_num
data_out['governance_num']=governance_num
data_out['social_num']=social_num

#calculate the ESG term frequency
data_out['Environment_ratio']=data_out['environmental_num']/all
data_out['Governance_ratio']=data_out['governance_num']/all
data_out['Social_ratio']=data_out['social_num']/all
data_out

import pysentiment2 as ps
lm = ps.LM()
#seperate the sentences
```

```
def sentence_split(str_sentence):
    list_ret = list()
    for s_str in str_sentence.split('.'):
        if '?' in s_str:
            list_ret.extend(s_str.split('?'))
        elif '!' in s_str:
            list_ret.extend(s_str.split('!'))
        else:
            list_ret.append(s_str)
    return list_ret

#calculate the character length of every factor
def text(content):
    Ethos_str = ''
    Pathos_str = ''
    Logos_str = ''
    for i in sentence_split(content):
        for j in environmental_extend:
            if j in i:
                #calculate the character length
                Ethos_str = Ethos_str + i
        for j in governance_extend:
            if j in i:
                #calculate the character length
                Pathos_str = Pathos_str + i
        for j in social_extend:
            if j in i:
                #calculate the character length
                Logos_str = Logos_str + i

    return Ethos_str, Pathos_str, Logos_str

e1 = []
e2 = []
e3 = []
for i in data_out['content']:
```

```
Ethos_str, Pathos_str, Logos_str = text(i)
#sentiment analysis
tokens = lm.tokenize(Ethos_str)
score = lm.get_score(tokens)
e1.append(score['Positive'] - score['Negative'])

tokens = lm.tokenize(Pathos_str)
score = lm.get_score(tokens)
e2.append(score['Positive'] - score['Negative'])

tokens = lm.tokenize(Logos_str)
score = lm.get_score(tokens)
e3.append(score['Positive'] - score['Negative'])

data_out['Environment_sentiment_score']=e1
data_out['Governance_sentiment_score']=e2
data_out['Social_sentiment_score']=e3
```
