Junior Management Science 7(3) (2022) 569-603

Junior Management Science

journal homepage: www.jums.academy

The Variety of CSR Disclosure and its Relationship with the Underlying Performance: A Textual Analysis

Niklas Winterberg

WHU – Otto Beisheim School of Management

Abstract

This study examines the relationship between corporate social responsibility (CSR) disclosure and the underlying CSR performance. The linguistic features of integrated and stand-alone sustainability reports from companies listed in the STOXX Europe 600 between 2010 and 2018 are investigated using computer-based textual analysis. The observed textual features are as follows: length, numeric content, horizon content, target orientation, readability, tone, topic-specific disclosure, and the number of topics covered. Additional variables include whether the report is following the framework of the Global Reporting Initiative (GRI), whether the CSR information is integrated into the annual report or prepared as a stand-alone report, and whether the company is defined as an early or late adopter. Concerning the relationship between the textual characteristics and CSR performance, the results support the hypotheses that length, target orientation, and the number of GRI topics covered are positively correlated with the performance. Concerning topic-specific disclosure, only the environmental and social dimensions are positively correlated with the corresponding performance. The results also reveal that companies with superior CSR performance tend to publish stand-alone reports under the GRI framework and started reporting before the announcement of the EU Directive in 2014.

Keywords: Corporate social responsibility; sustainability; sustainability reporting; textual analysis; GRI.

1. Introduction

In recent decades, companies have experienced an increasing demand from their stakeholders and shareholders to be informed about the company's impact on sustainability (Nazari, Hrazdil, & Mahmoudian, 2017). As climate change becomes an increasingly important issue, the public strives to understand how global companies are trying to solve this problem. Nonetheless, the demand is not only driven by the growing awareness of climate change. Human rights violations and other corporate scandals, such as accusations of child labor against Apple, Samsung, and other technology companies, have also led to external pressure for sustainability reporting (Wakefield, 2016). However, internal benefits equally caused that companies started reporting on their sustainability impact. These benefits are, for instance, that companies understand the underlying risks and improve efficiency of their operational processes (Global Reporting Initiative, 2016c). Consequently, a growing trend towards sustainability reporting has been observed in recent years (KPMG, 2017).

Due to this global trend, governments started to regulate sustainability reporting similarly to financial reporting. With reference to the latter, the European Union (EU) decided in 2002 that all listed companies in the EU must adopt the International Financial Reporting Standards (IFRS) (European Commission, n.da). As a sustainability counterpart to this regulation, the EU introduced the Directive 2014/95/EU. This regulation mandates that large firms have to prepare reports on their operations' environmental and social impacts (European Commission, n.db). A similar regulation has already come into force earlier in other countries, such as France, Denmark, and the United Kingdom (UK). These laws also mandate that companies disclose certain sustainability information (Fiechter, Hitz, & Lehmann, 2019; Hummel & Rötzel, 2019). In contrast to financial reporting, the EU Directive does not require companies to have their reports audited in detail or use a specific framework (European Commission, n.db).

Various frameworks and guidelines that support compa-

DOI: https://doi.org/10.5282/jums/v7i3pp569-603

nies during the preparation process of their sustainability reports emerged. These are designed to ensure that the reports' format and content are consistent and comparable among companies. According to a survey by KPMG (2017), the GRI framework is the most widely accepted framework for sustainability reporting. Nonetheless, there are also many other guidelines, such as the United Nations Global Compact (UNGC), the ISO 26000, and the OECD Guidelines for Multinational Enterprises.

The combination of a weak legal situation and a large number of guidelines allows companies to design their reports very individually. The variety of sustainability reports is further reinforced by the fact that companies are not required to disclose this information together with the annual report as long as it is "made publicly available within a reasonable period of time [...] after the balance sheet date, on the undertaking's website" (European Union, 2014, p. 5). Thus, they can independently decide whether they publish this information within their annual report or separately in a stand-alone sustainability report. As a result, this leads to a high diversity of sustainability reports.

The diversity of sustainability reports offers numerous research opportunities and is the fundamental motivation for this study. This study examines the differences between sustainability reports and how the corresponding textual characteristics are related to the underlying sustainability performance. For this purpose, a textual analysis based on statistical software is performed. This method allows to objectively and automatically retrieve textual characteristics from more than 2,500 reports. The observed textual characteristics are as follows: report length, readability, tone, horizon content, numeric content, target orientation, topic-specific CSR disclosure, and the number of GRI items covered. In addition, other essential variables include the use of the GRI framework, the report type, and whether the company is defined as an early or late adopter.

The first part deals with the differences in reports from specific groupings and with the textual development over time. Regarding the latter analysis, one could expect that the announcement of the EU Directive or its entry into force three years later might affect the reports' linguistic design. In addition to these events, there might be some other changes due to the general sustainability reporting trend. Moreover, reports from early and late adopters are compared. Similar to financial reporting, some companies have started to report on their sustainability responsibilities on their own initiative, while other companies have been moved to do so by regulatory pressure. This is the fundamental difference between early and late adopters. The study aims to examine the differences between the reports from both adopter types and to explore the underlying reasons. In addition, there are two other groups: integrated reports and stand-alone reports. The former are annual reports that include sustainability information, while the latter are separate reports that contain no further information.

The second part of the thesis deals with the relationship between sustainability disclosure and the corresponding performance. The motivation is to determine whether companies with a superior sustainability performance design their reports differently from companies with below-average performance. A regression is performed to investigate this relationship. The dependent variable is an economic, social, and governance (ESG) score, while the textual features and the other variables mentioned above are the independent variables. Moreover, additional analyses as well as different robustness tests are performed to confirm the key findings.

The remainder of this paper is structured as follows: Chapter two, *Corporate Social Responsibility Disclosure,* defines the concept of sustainability reporting and explains the EU Directive, which was announced in 2014. Afterwards, chapter three, *Frameworks & Guidelines,* introduces the different sustainability reporting frameworks. Thereby, it describes the GRI and the International Integrated Reporting Council (IIRC) as well as the corresponding frameworks. Chapter four, *Methodology,* explains how the study, in particular the textual analysis, was conducted. Afterwards, chapter five, *Literature Review & Hypothesis Development,* summarizes the current literature's findings and creates several hypotheses for the relationship between sustainability performance and the textual features. In the next step, chapter six, *Descriptive Statistics,* investigates the differences between integrated and stand-alone sustainability reports, the differences between reports from early and late adopters, and the textual development of sustainability reports over time. Furthermore, chapter seven, *Implications for CSR Performance,* contains the regression structure and tests the corresponding assumptions. Chapter eight, *Results,* includes the final results of the regression and relates to the formulated hypotheses. Moreover, it also elaborates on different robustness tests and additional analyses and includes managerial and theoretical implications. Finally, the thesis comes to a conclusion by addressing the limitations and summarizing the insights of this paper.

2. Corporate Social Responsibility Disclosure

2.1. Definition of Sustainability Reporting

The concept of sustainability is manifold and comprises various facets. In the current literature, it is often characterized as a three-pillar conception and includes the following three aspects: "economic", "social", and "environmental" (Purvis, Mao, & Robinson, 2019). In 1987, the United Nations Brundtland Commission defined sustainability as "meeting the needs of the present without comprising the ability of future generations to meet their own needs" (United Nations, 1987, p. 16). Thus, a major aspect of the concept is the long-term orientation and that economic, social, and ecological components can be beneficially combined. Transferred to a business setting, it relates to the companies' long-term survival by considering and combining the three perspectives (INTOSAI, 2013). The terms "sustainability", "corporate social responsibility", "corporate responsibility" (CR), or "economic, social, and governance" are often used

interchangeably to describe the same concept. Therefore, in this paper, they will not be further differentiated and utilized as synonyms (Ioannou & Serafeim, 2017).

Since traditional reporting and accounting is mainly based on financial disclosure, it is often criticized that this reporting type does not accurately represent the overall company due to missing information on environmental and social aspects (Aureli, Medei, Supino, & Travaglini, 2016). Internal as well as external stakeholders have pressured companies to start reporting on their sustainability responsibilities (Ballou, Heitger, Landes, & Adams, 2006). These reports are often declared as "Sustainability Report", "ESG Report", or "CSR Report". Often, non-financial information is also integrated into the annual report instead of being published in a separate report. In general, CSR or sustainability reporting can be defined as providing "nonfinancial information to key stakeholders, those people affected by a company's actions, on the company's operational, social, and environmental activities and its abilities to deal with related risks" (Ballou et al., 2006, p. 66). Thereby, it displays the linkage between the strategy of the company and its commitment to sustainable development. It also improves the companies' understanding of their sustainability performance and supports the company in measuring and communicating it (Global Reporting Initiative, n.db).

The development of sustainability reporting has started in the late 1980s when the first voluntary environmental reports have been disclosed. Large companies with highly pollutive operations were pressured by non-governmental organizations (NGOs) to report on these issues. In the following years, the reporting scope was widened, and CSR reporting was developed. Major reasons for this were concerns about human rights and labor conditions within the organizational structure and among the supply chain (INTOSAI, 2013). Another milestone of sustainability reporting was in 1998 when Elkington (1998) introduced the triple-bottom-line. This concept is based on the traditional bottom-line perspective and focuses not only on the economic value of a company but also on social and environmental values. A global trend has emerged during the recent decade with more and more companies starting to report on their sustainability activities. In 2011, around 44% of the G250 companies, which are the largest 250 corporations worldwide by revenue, disclosed CSR information within their annual reports or stand-alone reports. Six years later, the share of CSR reporting companies has increased to around 80% (KPMG, 2017).

Nowadays, financial investors are not only interested in financial performance but also in the management of sustainability issues. This trend can be seen in the global growth of sustainable investment funds or the introduction of specialized investment rating systems, such as the Dow Jones Sustainability Index (Pencle & Mălăescu, 2016). Hence, this trend underlines that social and environmental information is essential for various stakeholders' decision-making processes (Cho, Roberts, & Patten, 2010). The underlying reasons are not only moral or ethical issues but also the financial implications of sustainability issues. Enormous carbon dioxide (CO_2) emissions, for instance, might have a financial impact since the company has to buy emissions certificates or suffers a drop in sales due to a harmed company image (INTOSAI, 2013). Therefore, companies do not only have to follow the approach of maximizing shareholder value but also have to consider the interests of other key stakeholders (Ballou et al., 2006). In turn, this explains why an increasing number of companies have started reporting on their sustainability issues in recent years.

However, companies are not only extrinsically but also intrinsically motivated to pursue this reporting approach. On the one hand, for instance, sustainability reporting allows companies to better understand the underlying risks. On the other hand, employees might appreciate the efforts leading to higher job satisfaction. In turn, this might positively affect productivity and reduce employee turnover. Moreover, even if the reporting process requires time and effort, companies might benefit from it through resource and financial savings in the medium- and long-run. These, for instance, can be caused by reductions in energy or resource consumption (INTOSAI, 2013).

In contrast to financial reporting, sustainability reports do not only address shareholders but a wide range of stakeholders, for instance, local communities, environmental organizations, and the general public. These reports often cover the companies' philanthropic, environmental, operational, social, and economic objectives and how they manage the related risks (Ballou et al., 2006). However, the content is often very industry-specific. While automotive companies tend to focus on emissions and fuel consumption, textile companies tend to elaborate on topics like child labor and human rights (Liew, Adhitya, & Srinivasan, 2014).

All in all, sustainability reporting has become increasingly important in recent years and is used by companies to inform their stakeholders about their social and environmental responsibilities.

2.2. EU Directive

While financial reporting is rigorously controlled and regulated, the regulatory framework for disclosing non-financial information has been relatively weak. In 2013, the EU proposed the Directive 2014/95/EU, also called the non-financial reporting directive (NFRD). Thereby, the EU intended to improve the transparency of large European companies by mandating the disclosure of non-financial information. One year later, the directive was approved by the Council of the EU, and member states had to implement it within the following two years. The regulation has been effective since 2017. Thus, the first mandatory reports were published in 2018, covering the financial year 2017-2018 (European Commission, n.db; Global Reporting Initiative, n.df). Thereby, the directive led to a shift from voluntary to mandatory disclosure of non-financial information.

European companies of "public-interest", which employ more than 500 employees, must comply with the directive. In total, this amounts to approximately 6,000 entities within

the EU and includes banks, listed companies, insurance companies, or companies that are declared as "public-interest" entities by the authorities of the individual country (Global Reporting Initiative, n.df).

These companies must report on the following four buckets: environmental protection, anti-corruption, human rights, and social responsibility. While environmental protection comprises aspects like the usage of renewable energy and air pollution, social responsibility, for example, is related to working conditions and labor unions. Additionally, another compulsory part is the description of the business model, as well as a description of the diversity policies established for supervisory and management bodies. Companies must also explain the underlying risks and outcomes of policies that are implemented to tackle the four buckets.

However, if a company does not apply policies for one of those buckets, it is required to explain the reasons. Thereby, reporting flexibility is granted by the EU. Companies can either describe the applied policies or explain the underlying reasons for the non-existence of such policies. Furthermore, the directive neither includes an obligation to utilize a recognized reporting framework nor requires an audit company to accurately verify the information. Nevertheless, it encourages the use of frameworks, such as the UNGC or the GRI's Sustainability Reporting Guidelines (Global Reporting Initiative, n.df). With reference to assurance, it is only required that an independent audit firm checks the existence of non-financial information (European Union, 2014). Thus, the EU Directive does not impose any clear restrictions on the report design.

To sum up, the EU introduced the regulatory framework to enhance the transparency of major European enterprises. Thereby, it aims to improve the environmental and social performance of these companies. The EU also expects a positive long-term impact on economic growth and employment. Another objective is to increase the number of reporting companies and improve sustainability reports' information quality. Currently, the European Commission (EC) reviews the NFRD, which was decided at the end of 2019. In this way, the EC aims to continuously improve the directive and strengthen the sustainable development within the EU (Global Reporting Initiative, n.df). Even if the regulation was enormously criticized for its flexibility, it is considered as a milestone for sustainable development (Mittelbach-Hoermanseder, Hummel, & Rammerstorfer, 2019).

3. Frameworks & Guidelines

3.1. Overview of Frameworks & Guidelines

Referring to the regulatory framework for financial reporting in the EU, listed companies have to prepare their financial reports in accordance with the IFRS (European Commission, n.da). In contrast, there is no regulatory requirement for the use of a guiding framework referring to sustainability reporting (Nazari et al., 2017). As described above, even the implementation of the EU Directive did not lead to a mandatory reporting framework. As a result, companies can individually decide on their reporting methodology, which resulted in various reporting practices. Comparing the reports from different companies, they vary, for instance, regarding the format and the utilized performance indicators (Nazari et al., 2017). While some companies, for example, incorporate their sustainability information into their annual report, other companies publish biannually stand-alone reports (INTOSAI, 2013). In turn, according to a study by PwC (2014), the majority of financial investors worldwide are not satisfied with the heterogeneity of the current sustainability reporting practices.

Hence, numerous international and national initiatives came into existence, aiming to standardize and harmonize reporting practices. Depending on the industry, company, or strategic orientation, the different guidelines serve as a framework for a systematic reporting approach. The most commonly used framework is provided by the GRI and will be explained in the following chapter. Other frameworks include the OECD Guidelines for Multinational Enterprises or the UNGC initiated by the United Nations (UN). The latter comprises ten social and ecological standards, and around 12,000 companies from 170 countries are committed to the principles. Thereby, these firms have to publish an annual report called Communication on Progress (COP). On top of that, national guidelines like the Connected Reporting in the UK or the Sustainability Code in Germany are often applied by small enterprises with a short value chain or scarce financial resources (BMAS, n.d; INTOSAI, 2013). Since some companies integrate their sustainability information into their annual report, the IIRC provides guidelines for these report types.

3.2. Global Reporting Initiative

3.2.1. Overview & History

The Global Reporting Initiative is an independent international organization founded in 1997 in Boston and headquartered in Amsterdam (Global Reporting Initiative, n.da). It was initiated as a shared initiative of the UN Environmental Program and the Coalition for Environmentally Responsible Economies, a NGO based in the United States (US) (Clarkson, Li, Richardson, & Vasvari, 2008). Its global presence is supported through regional hubs in numerous countries, such as Colombia and India (Global Reporting Initiative, n.da). Furthermore, it is a non-profit foundation funded by various sources like partner organizations, businesses, and governments (Global Reporting Initiative, n.dc). The organization aims to improve the reporting quality of businesses worldwide. For this purpose, it supports governments and companies in understanding and communicating their sustainability impact. This attitude is reflected in their mission "to empower decisions that create social, environmental and economic benefits for everyone" (Global Reporting Initiative, n.da).

The organization's core product is the GRI sustainability reporting framework, which is called the GRI Standards. The

first version was released in 2000. It is the first framework for sustainability reporting and is also the most widely adopted one. Since the first release, the guidelines have been further developed and improved and rely on global best practices (Global Reporting Initiative, n.da). In 2017, 75% of all reports published by G250 companies were in line with the GRI framework (KPMG, 2017). Furthermore, the guidelines are universally applicable and can be applied by any organization independent of its size, location, sector, or whether it is a public or a private company. Thereby, GRI aims to ensure that stakeholders receive comparable data from different companies (Global Reporting Initiative, n.db).

The guidelines are based on a multi-stakeholder approach by bringing different stakeholders together to participate in a dialogue and jointly decide on changes and improvements. This approach should guarantee that different perspectives are considered. Moreover, during the continuous development and improvement process, a consensus-seeking approach is utilized. This procedure ensures that the interests of various stakeholders, for instance, businesses, governments, and labor unions, are taken into account (Global Reporting Initiative, n.dc).

Apart from developing its reporting framework, GRI also actively promotes the implementation at the company level. In the course of this, the organization advises market regulators, stock exchanges, and governments during the development process of sustainability policies. The collaborative approach aims to create a beneficial environment for CSR reporting (Global Reporting Initiative, n.da). As a result, GRI is referenced in the sustainability policies of 35 countries. Moreover, the organization also collaborates with other issuers of reporting guidelines like the UNGC and the International Organization for Standardization (ISO). In cooperation, GRI published guidelines on how to combine different frameworks to harmonize the reporting process (Global Reporting Initiative, n.dd). Moreover, it cooperates with these organizations to avoid content duplication (Global Reporting Initiative, n.de).

3.2.2. Structure of GRI Standards

In 2016, GRI published the latest version of the GRI Standards, which replaced the predecessor framework G4. Apart from content clarifications and the usage of a more straightforward language, the overall structure was changed. Thereby, the content has been restructured into a modular structure consisting of six interrelated sets of standards. Three of these sets are universal standards, which are applied by every reporting company. Additionally, three topic-specific buckets cover economic, environmental, and social issues. The modular structure grants higher flexibility since individual standards within a set can be modified without revising the whole set. Hence, the framework is more responsive to changes and can be regularly updated (GRI Secretariat, 2016).

In general, each standard includes the following three aspects: requirements, recommendations, and guidance. Requirements are all mandatory instructions. These have to be fulfilled by the reporting company to prepare a sustainability report that is in accordance with the GRI Standards. While requirements are labeled with the modal verb “shall”, recommendations are indicated with the modal verb “should”. The latter one represents measures which are advised but not mandatory. Lastly, the guidance part provides explanations and background information to support businesses in understanding the different guidelines. Thereby, this section also includes examples that act as an orientation for the reporting company (Global Reporting Initiative, 2016a).

The first universal standard is “GRI 101 – Foundation”. This standard acts as a starting point for the reporting process and includes the ten reporting principles. These principles cover topics like reliability, comparability, and completeness, ensuring reporting quality and sufficient reporting content. Moreover, it explains how to use the various standards and also how to reference them. In the same part, companies have to conduct the materiality assessment to identify the topics with the most significant impact on stakeholders (GRI Secretariat, 2016). Furthermore, GRI 101 also includes the requirements for publishing a report, which is in accordance with the GRI Standards, or for publishing a report, which contains only selected standards (Global Reporting Initiative, 2016a).

In the second universal standard, “GRI 102 – General Disclosures”, businesses have to report contextual information. This standard should ensure that stakeholders are able to understand the business model as well as the environment of the reporting company. Hence, this set covers topics like the organizational profile, the company’s strategy, and the reporting practice. In this standard, companies can decide on reporting in accordance with the GRI Standards core or comprehensive option. For the latter option, companies have to report on all disclosure items. In contrast, they only have to report fewer mandatory items in the core option, as it will be described later (Global Reporting Initiative, 2016b).

The third universal standard, “GRI 103 – Management Approach”, has to be utilized with the topic-specific buckets. Thereby, it should explain the reasons for the topic’s materiality and the organization’s approach to tackle this issue. If the organization has not implemented a particular management approach, it has to explain the underlying reasons. This is in line with the “report or explain” approach of GRI. Moreover, the company is required to evaluate the management approach with regard to its effectiveness (Global Reporting Initiative, 2016c).

The three topic-specific buckets are the economic, social, and environmental standards series. While the economic standards comprise topics like procurement practices (GRI-204) and anti-corruption issues (GRI-205), the environmental standards focus on topics like water usage (GRI-303), biodiversity (GRI-304), and waste management (GRI-306). With regard to the social standards, the content is, for instance, related to occupational health and safety (GRI-403), child labor (GRI-408), and customer privacy (GRI-418). In total, all three topic-specific buckets include 34 individual standards (GRI Secretariat, 2016). However, the range of

these topics is not exhaustive. If a particular material topic does not accurately match one of the standards, the company still has to report "other appropriate disclosures" (Global Reporting Initiative, 2016a, p. 19).

Moreover, there are two different approaches on how to use the GRI Standards. Companies can utilize the guidelines as a set, or they can only use selected items to report on specific topics. Regardless of the reporting method, businesses can individually decide on publishing a stand-alone sustainability report or an integrated report. Moreover, they can also reference information that is disclosed in other electronic or paper-based locations.

The first approach is the extensive version to present a complete picture of the company's sustainability responsibilities and the underlying management approach. Following this approach, companies prepare a report *in accordance with the GRI Standards*. Each reporting company, which wants to pursue this reporting approach, must publish a GRI index that contains the items on which the company has reported. Furthermore, these companies need to conduct a materiality assessment and report on every topic with a material impact on sustainability (Global Reporting Initiative, 2016a).

This approach is sub-divided into a comprehensive and a core option. The latter one is a downsized version, including the minimum amount of information to comprehend the organization and its impact on stakeholders. The comprehensive option is an extended version that contains additional information on the company's strategy, governance, and ethics. In addition, the company has to report on every topic-specific item within a material topic. Referring to the topic-specific item "GRI 303 – Water", for instance, the company would have to report on all three disclosure items (303-1; 303-2; 303-3) to fulfill the requirements of the comprehensive option. The two options refer only to the application of the GRI Standards but not to the reporting quality (Global Reporting Initiative, 2016a).

The second approach includes the utilization of only specific standards, if the company does not want to report on all of its material topics. This is called the *GRI-referenced* claim and allows businesses to report on specific information without providing a complete overview of its material topics (Global Reporting Initiative, 2016a).

3.3. Integrated Reporting

In 2002, a Danish biotechnology company called Novozymes published the first annual report that included information on environmental and social aspects. The former CEO of the company, Steen Riisgaard, stated that Novozymes decided to combine the information in one report since "various stakeholders [were] asking for a wider overview of the business" (Eccles, Krzus, & Solano, 2019, p. 2). Nowadays, these reports are called integrated reports since traditional financial reporting is combined with reporting on sustainability responsibilities. In the following years, more and more companies, such as BASF in 2007, have adopted this reporting type and issued integrated reports. However, these firms did not apply a common framework during the preparation process. Hence, the comparability of these reports was relatively low (Eccles et al., 2019).

To ensure uniform and comparable integrated reports, the Prince's Accounting for Sustainability (A4S) Project and GRI jointly launched the International Reporting Committee in 2010. This initiative developed the first reporting framework for the preparation of integrated reports. The organization was later rebranded into the International Integrated Reporting Council. This global alliance comprises investors, company representatives, and regulators, among others. The target of this initiative is to refine and disseminate the integrated reporting framework, which was published at the end of 2013 (Eccles et al., 2019). In addition, the IIRC wants to achieve superior information quality, which is accessible for "providers of financial capital" (International Integrated Reporting Council, 2013, p. 4). Proponents of this reporting method argue that it allows stakeholders to better understand the interconnections between the traditional reporting content and sustainability issues (Melloni, Caglio, & Perego, 2017).

To sum up, the integrated reporting framework is a frequently used guideline for integrated reporting and provides businesses with instructions on merging financial and sustainability reporting. Additionally, it also includes the Guiding Principles, which act as the basic concept for preparing and presenting the reporting content, and it also includes the Content Elements. The latter are questions, for example, concerning the business model and the governance structure to provide businesses with reporting guidelines. However, according to the International Integrated Reporting Council (2013), they should not serve as a rigid reporting structure since the content depends on the individual company situation. In this thesis, all annual reports, which contain information on the companies' sustainability issues, are defined as integrated reports. However, for the purpose of this study, it is not relevant whether these reports are in accordance with the integrated reporting framework.

4. Methodology

4.1. Introduction to Textual Analysis

In this paper, computer-based textual analysis of standalone and integrated reports is conducted using R. The reports are from companies that are a part of the STOXX Europe 600. R is a free programming language and environment with regards to the areas of graphics and statistical computing. In general, textual analysis can be defined as the "notion for parsing text for patterns" and is also known as natural language processing or computational linguistics (Loughran & Mcdonald, 2016, p. 1). This method comprises numerous techniques like sentiment analysis or the measurement of document similarity. Thereby, it allows researchers to extract information from unstructured data types, such as annual reports or other company statements (Liu, Wu, Yang, & Yu, 2020; Loughran & Mcdonald, 2016).

At the beginning of the 20th century, manual textual analysis was utilized. The motivation was, for example, to analyze works of Shakespeare or political speeches during the world wars (Loughran & Mcdonald, 2016). However, this manual method is time-consuming and prone to subjectivity. Hence, it is not recommended for the analysis of large-scale data samples (Yang, Dolar, & Mo, 2018). Computer-based textual analysis overcomes these problems since it is based on algorithms, which are not affected by subjective judgments (Mittelbach-Hoermanseder et al., 2019).

Moreover, in recent years, computing power has rapidly increased through technological innovation. In turn, this expanded the application opportunities of computer-based textual analysis. Technological progress, combined with the online availability of accounting- and finance-related documents, has led to the increasing use of this research method in the fields of accounting and finance (Loughran & Mcdonald, 2016). Thus, previous analyses were mainly concerned with, for instance, the degree of CSR disclosure in annual reports (Mittelbach-Hoermanseder et al., 2019), textual analysis of CSR reports in the US (Clarkson et al., 2020), and disclosure of sustainability information in annual reports (Hummel & Rötzel, 2019). All in all, textual analysis is a reliable and replicable form of qualitative analysis, which will be used in this paper to examine stand-alone and integrated sustainability reports, respectively.

4.2. Data Sample & Data Pre-Processing

The sample includes all firms that were part of the STOXX Europe 600 index at the beginning of June 2020. One has to consider changes in the index's composition, which came into effect on 22nd June 2020. These were caused by a regular review conducted by the index provider. Hence, companies such as Hugo Boss and EasyJet are still in the sample of observed companies, although they were excluded from the STOXX Europe 600 index through the last review (STOXX, 2020). In addition, as not all reports from 2019 have been published yet, the period between 2010 and 2018 is observed.

To find all relevant sustainability and integrated reports, the ASSET4 database, a division of Thomson Reuters, has been utilized. ASSET4 is a specialized ESG database, which contains around 900 indicators per year per company and provides information on the CSR performance of numerous businesses. This information is retrieved from publicly accessible sources and is predominantly related to the following four dimensions: economic, social, environmental, and corporate governance. Research analysts evaluate this information and create annual ESG scores for each firm (Clarkson et al., 2020).

A list indicating whether a company issued a sustainability report was retrieved from the ASSET4 database in the first step. The list also included an indicator of whether the reports are in accordance with the GRI framework. According to the database, 4,061 reports have been published between 2010 and 2018. Nevertheless, if all companies had reported for each year, 5,400 reports would be available (Appendix 1).

In the next step, the reports were retrieved as .pdf documents. Following the approach of Clarkson et al. (2020), the number of reports per year per was restricted to one observation per company. In some cases, for example, companies published an additional summary report of their sustainability report. In this case, the document with the highest number of pages was selected if more than one report per year was available. The reports were downloaded from the following sources: GRI database, company websites, and www.corporateregister.com. The latter one is the leading organization that provides CSR reports in its database (Dhaliwal, Li, Tsang, & Yang, 2011).

However, only 3,187 reports could be downloaded (Appendix 1). The primary issue was that many companies remove their sustainability reports after a certain time period. Furthermore, other reasons, which led to the exclusion, will be described subsequently. While some companies published only an executive summary instead of the full report (e.g., IHG's report in 2017), other companies published the report in a language other than English (e.g., Gecina's bilingual reports). Moreover, in some cases, false reports were published on the corporate website. Kesko, for instance, attached the sustainability report of 2012 instead of their report of 2010. Furthermore, some companies pursue a biennial reporting approach and, hence, published one report covering two years (e.g., Voestalpine). Since both years are covered, the ASSET4 database indicates that two reports are available. In turn, this led to a distortion of the initial sample size, which seems to be smaller than predicted. In the case of a company merger, the reports prior to the merger were excluded from the sample as, otherwise, the sustainability reports of either of the merging parties would have to be selected (e.g., EssilorLuxottica). Lastly, some companies published only a web version of their sustainability report, which could not be downloaded (e.g., Tele2 2013-2016).

During the gathering of the sustainability reports, it was noticed that there are some deviations from the ASSET4 database. Not all reports that were declared as GRI reports are indeed in accordance with the GRI framework. These errors were manually revised. However, these incorrectly declared reports account for only ∼1.2% of all sustainability reports.

Before the actual text analysis, one had to pre-process the integrated reports since they contain not only sustainability information but also other operational and financial information like balance sheets and income statements. To increase the comparability with the stand-alone reports, the sustainability part was manually extracted from the integrated reports. However, some reports could not be edited since the reports were protected (e.g., Coca Cola and Thyssenkrupp). Hence, these reports were excluded from the overall sample. This also applies to reports for which it was not possible to identify an extractable sustainability part because the individual sustainability components were spread across the entire integrated report. Afterwards, the sample comprised 2,619 reports (Appendix 1).

Nevertheless, the data sample still contained reports from

companies that did not report over the entire period or for which not all reports were available. Some companies have started their reporting activity in later years, for instance, due to the pressure of the EU Directive. However, the inclusion of reports from companies that have published only one or two reports might bias the analysis. Hence, to consider both aspects, the minimum number of reports per company was set at four. This limit guarantees that companies that have started reporting no later than 2015 can be included in the analysis. In the end, this resulted in a final sample of 2,431 reports (Appendix 1). Finally, the reports were imported into R. For this purpose, the package *readtext* was utilized.

In the next step, the data had to be pre-processed to accurately analyze numerous reports and determine the different textual characteristics. However, not all steps are relevant for the analysis of all observed variables. Thus, all steps will be described in the following part, and individual deviations for specific variables will be mentioned later. Moreover, the *quanteda* package was utilized. This package in R is predominantly determined for the quantitative analysis of textual data.

Firstly, the text was tokenized. During this process, the original text is broken down into a sequence of its individual components. These components are called tokens. There are different tokenization approaches. In this thesis, the text body was tokenized at non-letters. As a result, at every space or punctuation, a new token was created.

Secondly, specific characters were removed from the text body to improve the accuracy of the analysis. For this purpose, hyphens, punctuation, and symbols were filtered out. On top of that, in some cases, numbers were removed from the text.

Thirdly, all tokens were converted to lower cases. In turn, this ensures that identical words are identified as one word. Otherwise, words in different cases are not equivalent to each other and will be treated separately.

Fourthly, English stopwords were filtered out. These words are generic terms that have a low informational value and are irrelevant for the analysis. The package *quanteda* contains a list of 175 predefined stopwords. Examples for this word category are words like “that”, “would”, or “the”. Moreover, this reduces the number of tokens, which diminishes complexity and optimizes the processing time of the analysis.

Lastly, a stemming procedure was performed. This process is a type of morphological analysis, which aims to transform words into their root form, for instance, removing suffixes like “ed” and “ing” (Liew et al., 2014). Thereby, terms like “went” and “going” are converted into “go”. This procedure has a positive impact on precision since different versions of the same word are recognized as the same term. In addition, it further consolidates the dataset by reducing the number of tokens and, thus, positively affects processing time.

4.3. Observed Variables

As the data has been collected, imported, and pre-processed, the different observed variables can be analyzed. In the following sections, the variables will be defined, and the underlying methodology will be explained.

4.3.1. Length

The first observed variable is the length of disclosure. According to the current literature, the report length provides numerous insights. For example, the informational content might be derived from the length of CSR reports since longer reports can contain more information. This underlying argumentation has been utilized in various studies about financial and non-financial reports (e.g., Li, 2008; Muslu, Mutlu, Radhakrishnan, & Tsang, 2019). In contrast, it is also argued that the length might act as an indicator of the complexity of the companies’ CSR activities as these have to be more extensively described (Muslu et al., 2019). Moreover, Neu, Warsame, and Pedwell (1998) show that companies seek to influence the perception of stakeholders about the related CSR performance by disclosing environmental data. Hence, they argue that the length of a report also acts as a proxy for the management of stakeholders’ impressions.

In this thesis, three indicators are utilized to measure report length: number of characters, number of total words, and number of unique words. The first indicator does not require any pre-processing steps and counts all elements, including letters and numbers. The second indicator is the number of total words, which is the most common measurement type of report length in the current literature (e.g., Clarkson et al., 2020; Hummel & Rötzel, 2019; Muslu et al., 2019). Lastly, the number of unique words indicates how many different words are utilized in one report. The latter two indicators require all pre-processing steps except for the removal of stopwords. Following the approach of Hummel and Rötzel (2019), the logarithm of the measurement values is used to reduce skewness.

4.3.2. Readability

Readability is a parameter for the complexity of the language and determines “how easily the reader can grasp the content of a text” (Mittelbach-Hoermanseder et al., 2019, p. 27). Researchers often observed this variable in financial and non-financial studies. Managers might utilize the readability of corporate documents as an instrument to obfuscate a low performance since a difficult readability increases the burden of understanding. Thereby, it might diminish the resulting negative reaction of investors and analysts. Concerning CSR disclosure, the role of readability is especially pronounced due to the narrative and poorly regulated nature of CSR reports (Wang, Hsieh, & Sarkis, 2018).

Various indices have been adopted in previous studies for the measurement of readability (e.g., Clarkson et al., 2020; Hummel & Rötzel, 2019; Mittelbach-Hoermanseder et al., 2019; Muslu et al., 2019; Nazari et al., 2017; Wang et al., 2018). In this thesis, the Flesch-Kincaid, the Flesch Reading

Ease, and the Fog Index are utilized. According to Li (2008), these are the most reliable instruments. The calculation of these indicators is based on the average number of syllables per word and the average number of words per sentence. The Fog Index and the Flesch-Kincaid Grade Level estimate how many years of formal US education are required to understand the content of a text (Mittelbach-Hoermanseder et al., 2019; Wang et al., 2018). In addition, the Flesch Reading Ease score ranges from 0 (*Professional*) to 100 (5^{th} *grade*) (Flesch, n.d). Hence, there is an inverse relationship between both Flesch indices.

The calculation of the indices does not require any pre-processing steps. For the calculation, sentences with a minimum and maximum length of three and 75 tokens were considered, respectively. These limits ensure that incorrect sentence structures are excluded, and precision is optimized. Following the approach of Franco, Hope, Vyas, and Zhou (2015), the three indices were combined into one aggregate measure. This was conducted by computing the average of the indices' percentile ranks and dividing the result by 100. Thereby, higher values reflect a lower degree of readability.

4.3.3. Tone

By conducting a sentiment analysis, the tone of a report is determined. This variable indicates how positive or negative the document's language is, but it is not directly related to the content. Hence, it does not indicate, for example, whether good or bad news are published. However, even if the tone is not directly related to the content, sentences with a negative tone are likely to pertain negative information (Muslu et al., 2019). Due to the narrative nature of sustainability reports in contrast to financial reports, managers tend to utilize tone to communicate information, which is hard-to-quantify (Du & Yu, 2020). Referring to CSR disclosure, Muslu et al. (2019) argue that businesses with a more negative tone are more transparent since the negative tone relates to the publication of negative aspects. Hence, these companies are more willing to also inform the public about their negative impact on sustainability.

In order to conduct the sentiment analysis, a list of positive and negative words was utilized. This list was developed by Loughran and Mcdonald (2011) and is especially created for accounting research. All pre-processing steps of the text were conducted except for the removal of stopwords and the stemming process. Afterwards, all positive and negative words were counted. Finally, the tone was calculated by dividing the difference between positive and negative words by the total number of words (Figure 1). Thus, the value ranges between minus one and one.

$$\frac{Positive\ Words - Negative\ Words}{Total\ Words}$$

Figure 1: Sentiment Analysis – Formula

4.3.4. Topic-Specific CSR Disclosure

Many researchers investigated the degree of CSR content in company reports (Melloni et al., 2017; Muslu et al., 2019; Nazari et al., 2017). Some studies argue that CSR disclosure is utilized as a legitimizing tool leading to an increasing disclosure of sustainability information of companies with poor CSR performance (Cho & Patten, 2007). In contrast, Al-Tuwaijri, Christensen, and Hughes (2004) argue that companies with a superior CSR performance tend to disclose more CSR information. The underlying reason is that these companies want to convince market participants of their superior performance.

This thesis aims to assess the disclosure of specific CSR topics. In particular, the three dimensions measured are the three dimensions of the GRI Standards: economic, social, and environmental. Some researchers utilized the frequency of specific keywords as a measurement method of CSR disclosure. However, in this paper, the approach of Mittelbach-Hoermanseder et al. (2019) was applied. Thereby, one word window per dimension was created. Each word window contains twenty terms related to the specific dimension. Afterwards, the cosine similarities between the predefined word windows and the various reports were calculated. Cosine similarity is calculated as the inner product of two vectors: one vector refers to the topic-specific standard, and the other refers to a sustainability report. Thereby, the relative word frequencies were compared. This calculation resulted in the similarities between the reports' vocabulary and the topic-specific word windows (Mittelbach-Hoermanseder et al., 2019). The outcome ranges between zero and one. The latter result indicates that both documents have equal proportions, and the former result means that the documents do not share any similarities (Lang & Stice-Lawrence, 2015). Hence, a high value indicates that a report has a high similarity with the topic-specific vocabulary and that the company discloses a high degree of topic-specific information.

Every word window contains twenty words, including the respective search term and nineteen additional terms. However, "environment" is not included in the corresponding word window due to the term's ambiguity and is replaced by the term "ecology". The word windows are based on the ones defined by Mittelbach-Hoermanseder et al. (2019). Since Mittelbach-Hoermanseder et al. (2019) measured the five dimensions of the EU Directive, the word windows were consolidated and complemented by other terms. The additional terms were retrieved from the different standards. For example, the word-window of the economic dimension contains terms such as "corruption" (GRI-205), "monopoly" (GRI-206), and "tax" (GRI-207). The entire set of terms of the word windows and the corresponding GRI Standards are provided in Appendix 5.

4.3.5. Numeric Content

Numeric content in corporate statements allows the reader to gain insights into whether companies only talk about their activities or substantiate this with quantitative content and key performance indicators (KPIs). Researchers

argue that numerical information, in comparison to qualitative information, tends to be "more accurate, objective, comparable, and verifiable" (Hummel & Rötzel, 2019, p. 30). This information makes it easier for investors and analysts to understand the content of the reports and compare it with the ones of other companies. This is, for example, shown in the study of Huang, Nekrasov, and Teoh (2012). This study provides evidence that numbers in the title of earnings press releases cause a stronger reaction of investors.

In this paper, the indicator representing numeric content is the quantity of Arabic numbers per 1,000 words. Hence, it is measured as the ratio of the quantity of Arabic numbers to the total sum of words (Figure 2). However, it must be stressed that not the amount of digits is measured, but the amount of numbers. Thus, 100 is counted as one number and not as three digits. In addition, no pre-processing steps are necessary.

$$\left(\frac{Arabic\ Numbers}{Total\ Words}\right) * 1000$$

Figure 2: Numeric Content – Formula

4.3.6. Horizon Content

The degree of horizon-related information reflects the future orientation of corporate reports. Muslu et al. (2019) argue that sustainability reports, which elaborate on future trends, tend to be more informative for market participants. Referring to the research on financial reports, researchers provide evidence that forward-looking Management Discussion and Analysis (MD&A) statements are more informative and help market participants to consider the company's future performance in their analysis (Muslu, Radhakrishnan, Subramanyam, & Lim, 2015).

The approach of Muslu et al. (2019) was applied to calculate the degree of horizon content. The number of all horizon-related words was counted using the predefined list of Muslu et al. (2019). This list includes short- and long-term related phrases like "next period", "subsequent quarter", and "upcoming month" (Appendix 6). Lastly, the number of horizon words was divided by the total number of words and multiplied by 1,000 to receive the amount of horizon content per 1,000 words (Figure 3). Regarding the pre-processing steps, all of them were conducted except for the removal of hyphens, numbers, punctuation, and stopwords, and the stemming process. The underlying reason is that the list of Muslu et al. (2019) includes hyphens and numbers and is not stemmed.

$$\left(\frac{Horizon\ Words}{Total\ Words}\right) * 1000$$

Figure 3: Horizon Content – Formula

4.3.7. GRI Index

As described in the introductory part, every company that publishes a report following the GRI Standards has to release a GRI index. This document contains all topics mentioned by the company in its report. The total number of included topics acts as an indicator of the breadth of the report. Many researchers have already investigated the GRI index during their studies: Hummel and Schlick (2016) created a measurement scheme based on the GRI index to measure reporting quality, while Clarkson et al. (2008) developed a content index and incorporated components of the GRI index. For this study, the GRI index for each report is approximated to estimate the number of GRI items covered. One could also manually retrieve the reported items. However, this approach is relatively time-consuming. In addition, the approximation approach allows the estimation of a GRI index of companies that did not release a report in accordance with the GRI Standards.

In the first step, a dictionary for each GRI item was developed (Appendix 7; Appendix 8; Appendix 9). These act as search terms to identify whether a company has reported on the specific issues. Since generic terms such as "supplier" are mentioned in many buckets, for instance GRI-204 and GRI-308, the terms had to be very specific. In the next step, the only pre-processing steps were the removal of symbols and the transformation to lower tokens. Afterwards, the terms for each report were counted. As the single mention of a word does not mean that the company has reported on this item, thresholds were set. For items having more than five words in their dictionary, the limit was set at five. Therefore, the words had to occur a total of five times to exceed the threshold and confirm the item. For the remaining categories, a single mention is sufficient, as these are usually very specific. This becomes clear if one looks at the integrated report of Zalando in 2017. Even though they reported on "GRI-418 Customer Privacy", the report includes the terms of the corresponding dictionary ("customer privacy", "customer data") only four times. Finally, the total number of items confirmed was counted, ranging from zero to 34 and approximating the reporting breadth.

4.3.8. Target Orientation

Target setting is a fundamental element of management control systems (Malmi & Brown, 2008). These systems help companies break down their strategy into clear objectives for the different management levels and facilitate the execution of the company's strategy (Kaplan & Norton, 2008). Hence, target setting is a steering instrument and allows the management team to align the activities of the company with the desired organizational outcome (Malmi & Brown, 2008). Therefore, it is interesting to investigate whether companies formulate clear objectives in their reports. A precise formulation of sustainability objectives might indicate a higher commitment to a sustainability strategy.

For this paper, the degree of target orientation is measured by the number of target-related words per 1,000 words. Firstly, a list was created containing fifteen terms, which are

associated with target orientation. For this purpose, the first step was to search for synonyms for "target". As a result, terms such as "goal" and "objective" were found. In addition, for a test sample of ten reports (Appendix 10), the *keyword-in-context* function was used. This function allows searching for terms that are often mentioned in connection with the terms above. Furthermore, the 100 most frequent terms were searched for in the same test reports. All values that matched from a personal assessment were added to the list. In the last step, the set was supplemented by additional words that matched from personal experience resulting in the final dictionary (Appendix 11). Afterwards, the number of target-related words per report was counted and divided by the number of total words. To receive the number of target-related terms per 1,000 words, it was multiplied by 1,000 (Figure 4).

$$\left(\frac{Target\ Words}{Total\ Words}\right) * 1000$$

Figure 4: Target Orientation – Formula

5. Literature Review & Hypothesis Development

The following part delivers a comprehensive analysis of earlier results obtained from the existing literature. After a thorough evaluation of the different perspectives, the hypotheses concerning the relationship between the different variables and CSR performance will be formulated. The hypotheses relate to the textual features and the following additional factors: GRI framework, report type, and adopter type.

Referring to the length of disclosure, Clarkson et al. (2020) provide evidence that companies with superior CSR performance tend to disclose more comprehensive sustainability reports. The reason for this might be that well-performing companies can include more content since they pursue more CSR activities (Clarkson et al., 2020).

However, Li (2008) examined the relationship between the length of annual reports and financial performance and came to a different conclusion. The researcher mentions that longer reports lead to higher information-processing costs. Therefore, market participants need to invest more resources to understand the content of these reports accurately. Hence, managers might use the length as a strategic measure to decrease transparency and conceal information from market participants (Li, 2008). This strategy is also emphasized by Aureli et al. (2016), who state that companies can merely repeat information to increase the length of reports without providing additional and meaningful content. This insight underlines that longer reports can serve the purpose of disguising poor performance. Furthermore, this finding can be easily applied to CSR reporting. Companies could also increase the length of their sustainability reports to hide information about their poor CSR performance in an enormous amount of less relevant information. With regard to financial reporting, the US Securities and Exchange Commission (SEC) even recommends keeping sentences and documents as short and concise as possible to avoid this problem (SEC, 1998).

Current literature also often refers to the legitimacy theory (e.g., Clarkson et al., 2008; Clarkson, Overell, & Chapple, 2011; Patten, 2002). This theory states that the social legitimacy of businesses is monitored through a process of public policy. In case a company assumes that its social legitimacy is endangered, it is incited to approach the problem proactively. Thus, companies with a low CSR performance might utilize sustainability reporting as a legitimizing tool to positively affect the public perception about the company's performance. As a result, they report in detail on their sustainability issues to justify their legitimacy, which leads to longer reports. In addition, managers can employ this tool to inform the public about actual performance changes (Patten, 2002). Hence, this might lead to longer reports from companies with poor CSR performance.

Moreover, researchers also often refer to the voluntary disclosure theory (e.g., Al-Tuwaijri et al., 2004; Clarkson et al., 2008, 2011). In contrast to the legitimacy theory, it states that businesses with superior CSR performance have an incentive to disclose a high level of information. Thereby, these companies can highlight their performance and differentiate themselves from the competition. Competitors with poor performance cannot easily imitate this. This condition can alleviate the problem of adverse selection since it allows companies to reveal their actual performance level, which is not directly accessible to market participants. These companies might expect that market participants will react positively to this (Clarkson et al., 2008). In turn, this fact might lead to longer sustainability reports from well-performing companies. Hence, the voluntary disclosure theory suggests that CSR performance positively correlates with the length of sustainability reports.

The findings of the current literature about the relationship between length and CSR performance are inconclusive. Nonetheless, based on the insights gained through the voluntary disclosure theory, the hypothesis is formulated as follows:

> *Hypothesis 1: The length of sustainability reports is positively correlated with CSR performance.*

Next, the relationship between CSR performance and readability is examined. In general, businesses have an incentive to incorporate negative and positive sustainability information into their reports to preserve a positive company image. Otherwise, the disclosure of merely positive information would weaken the report's credibility and damage the company's reputation. Therefore, managers might utilize readability as an instrument to mitigate the response to negative information (Wang et al., 2018). This strategy would mean that reports of underperforming companies are less readable.

For this, Nazari et al. (2017) refer to the Efficient Market Hypothesis (EMH). This concept states that the market price incorporates all publicly available information. However, the researchers argue that the EMH is constrained by market participants' cognitive capabilities to retrieve all available information. Hence, they propose the Incomplete Revelation Hypothesis (IRH) of Bloomfield (2002). This theory states that market participants do not immediately react to complex information due to their cognitive limitations. Hence, the cognitive limitation of market participants allows managers to diminish the negative impact of bad news by worsening the readability since this makes it more difficult to accurately comprehend the information (Nazari et al., 2017; Wang et al., 2018).

Concerning financial reports, this obfuscation strategy can even inhibit the ability of small investors to process the reports since it requires too many resources (Miller, 2010). Thus, the usage of readability as a tool to obfuscate a poor performance was also observed in financial reports. Li (2008) found that poor readability is associated with low financial performance. In this way, companies try to hide their poor performance.

Referring to the voluntary disclosure theory, one could also assume that companies with superior CSR performance aim to make the report as comprehensible as possible. Hence, they tend to utilize plain language, which can be easily processed to underline their superior performance (Wang et al., 2018).

However, the study of Clarkson et al. (2020) came to the opposite conclusion. Their study revealed that reports from companies with good sustainability performance are less readable. The researchers argue that reports from these companies include "more sophisticated analyses", which are more difficult to understand (Clarkson et al., 2020, p. 21).

Nevertheless, the majority of current literature supports the hypothesis that reports of companies with poor performance are less readable. Since a high value of the variable indicates a low level of readability, the second hypothesis is the following:

> *Hypothesis 2: The readability of sustainability reports is negatively correlated with CSR performance.*

Referring to the literature on tone, Davis and Tama-Sweet (2012) investigated the role of this linguistic feature in earnings press releases. They found a positive correlation between the tone indicator and the subsequent ROA. In addition, a further study revealed that a more pessimistic tone in the MD&A section is associated with lower future ROA (Davis & Tama-Sweet, 2012). Hence, financial research revealed a positive correlation between tone and the subsequent financial performance.

Concerning sustainability performance, the study of Cho et al. (2010) indicates that firms aim to manage the perception of stakeholders by biasing the verbal tone in their sustainability reports. The reports from bad CSR performers reveal a high level of optimistic language compared to the opposite group. This finding shows that underperforming firms tend to focus on the good news while blurring the bad ones (Cho et al., 2010). This phenomenon is especially pronounced for sustainability reporting since it is still unregulated and difficult to verify (Du & Yu, 2020). Thereby, these findings support a negative correlation between both factors.

Nevertheless, the study of Clarkson et al. (2020) came to the opposite conclusion that the tone in reports from companies with poor CSR performance tends to be more negative. This result is not in line with the predicted greenwashing behavior. Greenwashing can be defined as "a discrepancy between words and deeds, which combines poor environmental performance and positive communication about the environmental performance" (Pizzetti, Gatti, & Seele, 2019, p. 2). Clarkson et al. (2020) argue that their finding, which contradicts the expected greenwashing behavior, might be caused by the attempt of managers to enhance credibility. Alternatively, they state that this might be a result of an unconscious process to fend off criticism. Nonetheless, they point out the need for further research (Clarkson et al., 2020). Moreover, Du and Yu (2020) found that improvements in tone indicate an improvement in subsequent sustainability performance. As managers employ positive and negative words to inform stakeholders about their expectations of future performance, an improvement in tone acts as an indicator for a higher future CSR performance. One condition for this assumption is that sustainability reports contain relevant information about future performance (Du & Yu, 2020).

The findings of the various studies are divergent. Nonetheless, the third hypothesis follows the concept of greenwashing, meaning that well-performing firms aim to shape stakeholders' perception by biasing the tone in their reports:

> *Hypothesis 3: The tone of sustainability reports is negatively correlated with CSR performance.*

The observed indicators *GRI Index* and *Topic-Specific CSR Disclosure* refer to the content of sustainability reports. For this reason, the literature review is performed for both categories together. Ingram and Frazier (1980) investigated the relation between environmental disclosure content and the underlying environmental performance. While content is measured by twenty categories like "Regulatory Compliance" or "Environmental Control", the environmental performance is measured by an index of the Council of Economic Priorities (CEP). The results imply that the content does not correlate with environmental performance (Ingram & Frazier, 1980). Other researchers came to the same conclusion (Wiseman, 1982). Nonetheless, it could be argued that the issue did not have the importance that it has today. This would explain diverging results.

Moreover, Patten (2002) identified several flaws in the previous studies; for example, no other control variables were considered, and the indicators for the environmental performance were not appropriate. Hence, his study considered these issues. The study revealed a negative correlation between the content of disclosure and environmental performance. This finding corresponds to the study of Cho and

Patten (2007), who concluded that firms utilize disclosure as a legitimizing tool. Clarkson et al. (2008) also refer to the legitimacy theory. This theory suggests that firms include more content in their sustainability reports to justify their poor performance and change the public's perception of their actual performance (Clarkson et al., 2008).

However, the study of Clarkson et al. (2008) is in line with the voluntary disclosure theory. As described previously, firms with superior CSR performance could incorporate more content to differentiate themselves from the competition. This proactive strategy allows well-performing companies to disclose their true performance, which, otherwise, is not directly accessible to market participants. These companies expect a positive market response and believe that this will lead to future benefits. Hence, this theory suggests a positive relation between CSR performance and the amount of content (Clarkson et al., 2008). The study of Al-Tuwaijri et al. (2004) revealed similar findings suggesting that firms with superior CSR performance tend to cover more topics in their sustainability reports.

To sum up, the various studies provide mixed results about the relationship between CSR performance and CSR disclosure. Moreover, one has to underline that many studies only observed the relationship between environmental disclosure and the associated performance. Nevertheless, the environmental dimension is only one element of sustainability and does not reflect the overall concept. In turn, the following hypotheses are formulated in line with the voluntary disclosure theory mentioned previously:

> *Hypothesis 4a: The number of GRI items per sustainability report is positively correlated with CSR performance.*

> *Hypothesis 4b: The degree of topic-specific CSR disclosure is positively correlated with CSR performance.*

The literature on the relationship between CSR performance and numerical content in sustainability reports is limited. Clarkson et al. (2011) examined the relation between the nature of disclosure and CSR performance. According to this study, the nature of disclosure can be characterized as "soft" or "hard". "Hard" disclosure is defined as information that is "objective and externally verifiable" (Clarkson et al., 2011, p. 2). This definition can also be applied to the concept of numerical content since, in comparison to narratives, it is more objective and easier to verify. In their study, a sample of 51 listed Australian companies is observed. While environmental performance is measured by the quantity of toxic releases weighted by annual sales, the degree of "hard" disclosure is measured by the ratio of "hard" disclosure items to the total number of disclosure items. The latter is based on an index developed by Clarkson et al. (2008). Although the researchers expected a positive relationship between "hard" disclosure items and environmental performance, the study revealed a negative relationship (Clarkson et al., 2011).

Moreover, one can also refer to the voluntary disclosure theory, as well as to the legitimacy theory, to formulate a hypothesis. Concerning the former one, this theory predicts that companies with superior CSR performance disclose CSR information as a unique selling proposition to differentiate themselves from underperforming companies. Hence, they might substantiate their reports with more numerical content, which is more credible, objective, and verifiable than narratives. This content cannot be imitated by firms with poor performance and underlines the outstanding performance. Therefore, this theory suggests a positive correlation between numerical content and CSR performance (Clarkson et al., 2011).

With reference to the legitimacy theory, underperforming companies publish CSR reports to maintain their legitimacy and positively affect the public's perception of the company. Therefore, this theory predicts that companies disclose less numeric content since this could harm the legitimacy. In turn, they prefer to publish a high degree of narratives because they aim to shape the public's perception and not communicate their actual performance (Clarkson et al., 2011).

Moreover, Hummel and Schlick (2016) examined the relationship between the quality of sustainability reports and the associated performance. High-quality disclosure is defined as the "disclosure of numerical data on a company-wide level that fulfill[s] or exceed[s] the minimum requirements derived from the GRI guidelines" (Hummel & Schlick, 2016, p. 460). In contrast, low-quality disclosure does not fulfill the requirements or provides any other information. In particular, the latter aspect demonstrates that low-quality disclosure tends to contain less numeric content when companies provide any other information than required. Moreover, the researchers argue that high-quality disclosure fulfills the following criteria: verifiability, comparability, and reliability. Numeric content can be easily compared among companies, is more reliable than narratives, and can be better verified. These characteristics also underline that numeric content can be defined as high-quality disclosure. The results of the study indicate that high-quality disclosure is positively related to CSR performance (Hummel & Schlick, 2016). Hence, one can assume that the relation between numeric content and CSR performance is also positive. In this paper, the hypothesis states as follows:

> *Hypothesis 5: The numeric content of sustainability reports is positively correlated with CSR performance.*

Current literature has not yet investigated the relationship between horizon content and sustainability performance. However, Muslu et al. (2019) examined the impact of CSR report narratives on the accuracy of analyst forecasts. Their results reveal that the quality of CSR reports, measured by a disclosure score, positively affects the analyst forecasts' precision. The researchers developed the associated disclosure score based on the following criteria: tone, readability, length, numerical content, and horizon content. Concerning the latter component, they argue that reports tend to

be more informative when they elaborate on the future outlook (Muslu et al., 2019). Muslu et al. (2015) support this assumption since they found that MD&A disclosures with a high degree of horizon content help market participants to forecast financial performance. Moreover, Hussainey and Walker (2009) examined the relation of forward-looking statements in annual reports and market participants' ability to predict future earnings. Their results also indicate that forward-looking disclosures improve the precision of analyst forecasts. Hence, it can be concluded that reports with a high degree of horizon content tend to be more informative.

This assumption can be utilized for the interpretation of the voluntary disclosure theory and the legitimacy theory. Regarding the former one, well-performing companies might aim to maximize the informativeness of their reports. The argumentation is based on the same train of thoughts as in the previous sections. Hence, reports from well-performing firms are likely to discuss the future CSR strategy and the associated activities. Thereby, these companies aim to differentiate themselves from underperforming firms. This theory suggests that horizon content is positively related to CSR performance.

Moreover, the implications of the legitimacy theory are inconclusive. On the one hand, one could argue that underperforming companies tend to keep the informativeness as low as possible since they do not want to disclose their misconduct. This behavior might be especially pronounced for underperforming companies, which do not want to change their sustainability strategy in the future. On the other hand, Asay, Libby, and Rennekamp (2018) studied the relation between firm performance and language choices in narrative publications. They argued that there are two arguments why these companies could incorporate more future-related content in their reports: Firstly, companies with a poor CSR performance might elaborate on future initiatives to distract from the past or use it as a measure of defense. Secondly, they might aim to affect the perception of the market positively. Therefore, these companies focus on what they will change during the next periods and how they will improve (Asay et al., 2018). This behavior would tend to apply to companies that want to improve their CSR performance in the future. Consequently, this theory does not propose a clear relationship between the current CSR performance and horizon content. However, since there is more support for a positive relationship, the hypothesis is as follows:

> *Hypothesis 6: The horizon content of sustainability reports is positively correlated with CSR performance.*

The amount of literature elaborating on the relationship between CSR performance and target orientation of sustainability reports is similarly limited. Targets are an essential part of management control in most organizations and act as a decision-making tool (Arnold & Artz, 2015). To develop a hypothesis about the relationship, one can refer to the goal-setting theory developed by Latham and Locke (1979). The researchers argue that challenging but attainable objectives, which are clear and specific, positively impact performance. This effect results from four mechanisms. Firstly, objectives direct the behavior and effort of employees towards measures, which serve the purpose of fulfilling the objective. Secondly, objectives motivate employees through the so-called "energizing function" (Locke & Latham, 2002, p. 706). Thirdly, objectives, in particular difficult ones, improve the persistence of employees and, fourthly, they indirectly affect employee behavior by generating knowledge and arousal (Locke & Latham, 2002). Thus, this theory suggests that the formulation of sustainability goals has a positive impact on performance. However, the mere number of words does not measure the content or the formulation of objectives since the theory is based on specific and attainable objectives. Moreover, this argumentation refers to a causal effect from the report design on CSR performance, while the opposite effect is observed.

As in the previous paragraphs, one can refer to the voluntary disclosure theory and the legitimacy theory. The first theory predicts a positive correlation between CSR performance and the degree of target orientation. Well-performing companies might tend to formulate clear objectives to show the market that they have met their targets in the following period. This action would enable them to demonstrate their superior CSR performance and set themselves apart from the competition. In contrast, companies with poor performance would formulate fewer or no targets at all, as they are unlikely to meet them.

From the legitimacy theory perspective, one could argue that underperforming companies might set targets to convince market participants of their legitimacy. According to the arguments provided previously, they might even set themselves more targets to distract from the current situation. As a consequence, the legitimacy theory does not provide a consistent prediction in line with the goal-setting theory or the voluntary disclosure theory. Nonetheless, the present hypothesis is based on the latter perspectives:

> *Hypothesis 7: The target orientation of sustainability reports is positively correlated with CSR performance.*

In addition to the textual characteristics, the effect of other variables is observed. An additional factor is whether the company is an early or late adopter of sustainability reporting. Early adopters are companies that pursued sustainability reporting before the announcement of the EU Directive in 2014. In contrast, late adopters are defined as companies that started reporting on their sustainability issues after the announcement until the regulation came into force. In the current literature on IFRS adoption, there is also a third group, the so-called resisters, which are businesses that started reporting after the application of the IFRS became mandatory (Christensen, Lee, Walker, & Zeng, 2015). In terms of the sustainability reporting environment, this would include companies that started reporting in 2017. At this

point in time, the first mandatory reports were published covering the financial year 2017-2018. However, this group is excluded from the current sample since the minimum number of reports per company is four. The resister group has only published two reports during the observation period for the years 2017 and 2018.

Bhimani, Silvola, and Sivabalan (2016) conducted interviews and surveys to examine the relation between the two reporter types and CSR embeddedness. The latter is defined as "alignment between the contents reported and their actual manifestation" (Bhimani et al., 2016, p. 82). Thereby, it acts as an indicator to what extent the sustainability reports' content is integrated into the overall strategy. Even if it is not directly related to CSR performance, embeddedness might be positively correlated with CSR performance when companies do not only talk about their CSR strategy but, in turn, implement it. The researchers argue that early adopters' motivation is related to their genuine ambition for CSR activities and their intention to talk about the underlying performance. In contrast, late adopters are only motivated to keep up with early adopters and merely improve their reputation. Hence, sustainability practices are more integrated into the strategy of early adopters leading to a higher sustainability performance (Bhimani et al., 2016).

This argumentation is consistent with the theory of voluntary disclosure. Firms with superior CSR performance aim to inform the public about their positive impact. As a result, they started reporting on sustainability issues without regulatory pressure. In contrast, companies with poor CSR performance do not want to share this information with the public and only start reporting in response to regulations. Without this pressure, they would probably hesitate to adopt the practice of sustainability reporting. Hence, this supports the hypothesis that early adopters obtain a higher CSR performance than late adopters.

Nonetheless, given the legitimacy theory, it could be argued that companies with poor CSR results started reporting even before the announcement of the directive. Since these companies are striving to positively shape the perception of their stakeholders and convince them of their legitimacy, they might have an incentive to start reporting without regulatory pressure. In contrast, high-performing companies do not face external pressure to justify their legitimacy and only start reporting in response to upcoming regulations. Thus, this supports the hypothesis that early adopters tend to be companies with poor CSR performance to justify their legitimacy. Nonetheless, since the voluntary disclosure theory is supported by the study of Bhimani et al. (2016), the hypothesis is as follows:

> *Hypothesis 8: Companies with a higher CSR performance tend to be early adopters of sustainability reporting.*

This paper also examines whether there is a relationship between reporting under the GRI framework and CSR performance. Bernard, Abdelgadir, and Belkhir (2015) conducted a sector-specific analysis to examine the relationship between the two factors. They utilized CO_2 emissions as a measure of sustainability performance in the period between 2007 and 2011. In their study, they found no performance differences between GRI and non-GRI reporting firms. Therefore, this study does not identify any significant effect. However, the study does not include any control variables, and it is questionable whether CO_2 emissions are a representative instrument for measuring sustainability performance.

Referring to the voluntary disclosure theory, one could assume that high-performing companies tend to commit themselves to the GRI framework to help stakeholders better understand the company's CSR activities. Moreover, underperforming companies probably avoid using the GRI guidelines, as they can be better compared to high-performing companies. Thus, companies with superior CSR performance can distinguish themselves even better from the competition through reporting under the GRI framework. Lastly, well-performing companies make their intentions even more apparent to stakeholders by committing to this framework since the preparation of these reports requires effort and resources.

With reference to the legitimacy theory, one could also argue that companies report under the GRI framework to compensate for their poor sustainability performance. In this way, they might try to signal their stakeholders that they want to work on themselves. These firms expect that a commitment to one of the leading reporting guidelines would compensate for the poor performance and has a positive effect on the stakeholders' attitude towards the company. Even if they cannot deliver sufficient CSR results, they justify their legitimacy by producing high-quality reports in line with the GRI framework (Bernard et al., 2015).

However, the literature on this relationship is limited, and there is little empirical evidence provided by studies. For this reason, it is only hypothesized that both factors are correlated, but not the direction:

> *Hypothesis 9: The CSR performance affects whether companies prepare their sustainability reports in accordance with the GRI framework.*

The last factor observed concerns whether companies publish a separate report or integrate the sustainability section into their annual report. However, no empirical study dealing with the differences among stand-alone or integrated reports could be identified. Therefore, the underlying argumentation is based on the voluntary disclosure theory and the legitimacy theory, respectively.

With reference to the former one, the line of argumentation is not clear. On the one hand, one could argue that high-performing companies publish a separate sustainability report highlighting their superior CSR performance. One report, which solely focuses on the company's impact on sustainability, underlines the company's efforts and raises stakeholders' awareness. Moreover, other companies might aim to hide their poor performance within their annual reports. Thereby, they can distract from their sustainability failure by focusing on financial performance. On the other

hand, one could argue that well-performing companies publish an integrated report highlighting the linkage between their CSR strategy and the operational business. These companies might want stakeholders to fully understand the company's impact and, hence, want to present financial and sustainability performance together.

Concerning the legitimacy theory, the argumentation builds upon the discussion on the GRI framework, previously. Companies with poor CSR performance might prepare a stand-alone report to protect their legitimacy. Even if they cannot deliver satisfactory results, they show their stakeholders that they are aware of their problems and wish to improve. For this reason, they prepare a stand-alone report underlining the company's efforts. However, one could also argue that these companies publish an integrated report. In this report, they can show the connection between their business model and their sustainability performance. Thereby, they might justify their performance with the firm's business model and claim that the poor performance, for instance, is due to the financial result or the operational complexity. By showing the linkage between business model and sustainability performance, they might try to defend their legitimacy.

To sum up, both theories predict that the CSR performance has a particular impact on the reporting method. However, the direction of the effect is unclear. For this reason, the hypothesis states as follows:

> *Hypothesis 10: The CSR performance affects whether companies publish integrated or stand-alone reports.*

6. Descriptive Statistics

Before the regression analysis is conducted, the descriptive statistics are presented subsequently. In addition to the results of the total dataset, the results for the following groupings will be compared: integrated vs. stand-alone reports and reports from early adopters vs. reports from late adopters. Besides, the changes over time will be examined. Referring to the length indicators, the absolute values and not the logarithmic values are presented since this facilitates the corresponding interpretation. Mann-Whitney-U tests are conducted to test whether the differences between reports published by early and later adopters, as well as the differences between integrated and stand-alone reports, are significant. This test is a non-parametric test for two independent samples, which checks whether they have an equal distribution. It is also often called the Wilcoxon-Mann-Whitney test.

6.1. Overview

The minimum, maximum, and mean values of all variables for the total dataset are listed in Table 1. Looking at the length indicators, one can detect enormous differences. While some companies report in great detail, other companies spend only a few hundred words on their sustainability reports. In general, an average sustainability report comprises around 27,000 words. Furthermore, the tone variable predicts that the reports are generally more positive than negative but almost neutral. This can be compared to annual reports, which usually have a lower value (Mittelbach-Hoermanseder et al., 2019). However, the dataset mainly comprises stand-alone sustainability reports (Appendix 2). Since annual reports are legal documents, they require a more neutral language, which might cause the difference between the two studies (Mittelbach-Hoermanseder et al., 2019).

The readability index cannot directly be interpreted since it represents the average of percentile ranks. The corresponding values cover the full range from zero to one. Looking at the individual readability indices, the Fog Index, as well as the Flesch-Kincaid Index, are above the normal range. Li (2008) mentions that the range of the Fog Index is from 8 (*childish*) to 18 (*difficult*), while all texts with even higher values are unreadable. For this reason, the average Fog score of this sample (21) seems to be out of range. This finding can also be observed for the Flesch-Kincaid Index. Since both scores represent the required years of formal education, it also underlines that the mean values might be too high. However, referring to a study of Caglio, Melloni, and Perego (2020), their results also reveal a high average Fog score of approximately 23. Hence, the results indicate that a high academic level is required to understand the content of sustainability reports. This may be based on the use of several technical terms related to the topic of sustainability, which might not refer to individual's common way of speaking. Concerning the Flesch Reading Ease, the results can be interpreted. The mean value of this readability index corresponds to a required educational level of a college graduate (Flesch, n.d). This insight is in line with the other two readability indices, which indicate that a high educational level is required.

The three indicators for measuring the disclosure of specific CSR topics reveal that social issues are the most prevalent topic in the various reports of the dataset. This is in line with the findings of Mittelbach-Hoermanseder et al. (2019), who also utilized the cosine similarity. The researchers observed that social is one of the predominant issues in annual reports. However, their study revealed that the cosine similarities of the social and environmental topics are lower in absolute values. Since their research solely investigated the content of annual reports and not sustainability reports, the diverging results may be caused by sample differences. The degree of CSR topics in annual reports is probably lower because the main focus of these reports is on financial and operational issues.

Referring to the GRI index, some companies cover all GRI items, while others do not even cover a single item. Since the sample includes reports, which are not in accordance with the GRI framework, this might explain the finding. However, on average, companies report on approximately 50% of the GRI items.

Table 1: Descriptive Statistics – Overview

Variable	Range	Minimum	Mean	Maximum
#Characters	0 – ∞	4,081	319,472	2,534,315
#Total Words	0 – ∞	417	27,065	205,362
#Unique Words	0 – ∞	169	2,247	6,859
Numeric Content	0 – 1,000	12.0	87.1	461.4
Horizon Content	0 – 1,000	0.0	1.0	5.9
Target Orientation	0 – 1,000	0.0	6.4	28.4
Tone	(1) – 1	(0.018)	0.006	0.035
Readability Index	0 – 1	0.0	0.5	1.0
Flesch Reading Ease	0 – 100	4	23	73
Fog	yrs. of edu.	15	21	26
Flesch-Kincaid	yrs. of edu.	12	17	22
Economic	0 – 1	0.000	0.066	0.178
Environment	0 – 1	0.000	0.146	0.385
Social	0 – 1	0.000	0.174	0.316
GRI Index	0 – 34	0	16	34

6.2. Development over Time

In recent years, it has been observed that more and more companies have begun to report on their sustainability responsibilities (KPMG, 2017). For this reason, it is of great interest to investigate how sustainability reports and their contents have developed over time. In particular, events such as the announcement of the EU Directive in 2014 and its subsequent entry into force in 2017 could have had an impact. Table 2 shows the mean values for all textual characteristics between 2010 and 2018.

Referring to the three length indicators, one can observe an increase in the average length after the EU Directive announcement. This effect is most pronounced for the number of characters and the total number of words. The corresponding plots can be seen in Appendix 12 to Appendix 14. Hummel and Rötzel (2019) investigated the impact of the introduction of the Companies Act 2006 Regulations 2013 in the UK on annual reports. This regulation demands the disclosure of various CSR information relating to issues like human rights and gender distribution. Their study also reveals that the implemented law led to an increase in the report length. The researchers argue that this can be explained by additional disclosure required by the regulation (Hummel & Rötzel, 2019). Hence, the increase in report length could be caused by the announcement of the EU Directive since businesses have started to report on additional topics not covered previously. This is further illustrated by the fact that the directive mandates the disclosure of the following four issues: environmental protection, human rights, anti-corruption, and social responsibility (European Union, 2014). However, one has to consider that the regulation was only announced but not implemented at that point in time. Looking at the literature on the mandatory IFRS adoption, the finding of this study is consistent with the study of Lang and Stice-Lawrence (2015). This study found that compulsory IFRS adoption led to an increase in the amount of disclosure.

In contrast, Chauvey, Giordano-Spring, Cho, and Patten (2015) argue that the increase in report length could be caused by more companies publishing stand-alone reports instead of integrated reports. However, in this study, the proportion of integrated reports increases over time, and these report types tend to be shorter (Appendix 24 – Appendix 27). Therefore, the proportion of integrated reports cannot be the underlying reason for this phenomenon. Moreover, the share of reports in accordance with the GRI framework is also almost constant, being no decisive factor, as well (Appendix 28). Even the emergence of late adopters due to the announcement in 2014 cannot be regarded as an underlying reason for the increasing length of reports, since their reports appear to be shorter than the reports from early adopters (Appendix 29 – Appendix 32). In summary, it can be assumed that the requirements of the EU Directive have led to an increase in the length of sustainability reports.

Regarding the horizon content, there is little movement over the observed period, and no specific trend can be identified (Appendix 16). The marginal increase in numeric content, shown in Appendix 15, may be explained by public demands for more performance indicators rather than narrative descriptions of the implemented policies. This is based on the fact that, in comparison to narrative descriptions, numeric content allows stakeholders to better assess the companies' CSR performance (Bhimani et al., 2016). Furthermore, Hummel and Rötzel (2019) provide evidence that the introduction of a similar regulation in the UK has also led to an increase in numeric content. This might be caused by the requirements of the directive.

On top of that, following the announcement of the EU Directive, the degree of words related to target orientation has increased (Appendix 17). As already mentioned in chapter five, the literature on this subject is relatively limited. However, the EU Directive mandates that reports must include tar-

gets for the implemented diversity policies (European Union, 2014). This regulation will inevitably have led to companies reporting more about the corresponding objectives. In addition, it could also have triggered that these companies started reporting on the goals of other CSR policies. Since legal regulations required them to set targets for their diversity policies, they might also have set targets for other policies. In turn, this might lead to an increase in target orientation.

Besides, the changes in tone are marginal and inconsistent (Appendix 18). This finding is consistent with the study of Hummel and Rötzel (2019), who similarly found no change in tone. Furthermore, it can be seen that readability has become increasingly difficult since 2014 (Appendix 19). In contrast, Hummel and Rötzel (2019) observed that the UK's regulatory intervention led to an improvement in readability. One reason for the observed negative impact in this study could be that the EU Directive does not provide clear reporting guidelines but leads to more disclosure of technical topics. In turn, this could negatively affect readability (Lang & Stice-Lawrence, 2015). However, as the requirements of the EU Directive are not very technical, it is questionable whether it indeed increased the complexity. Nonetheless, the EU Directive does not require reports to be based on particular frameworks but requires the disclosure of additional content. The combination of these aspects could be detrimental to readability if companies need to include more content in their reports without having appropriate guidelines for creating such a corporate document (European Union, 2014). Another reason could be the emergence of late adopters in response to the announcement. As these companies have not previously reported on their sustainability issues, their incentives for preparing a high-quality report could be relatively low. Another reason might be that late adopters are inexperienced in preparing such a report, thereby, harming readability.

With reference to the thematic disclosure indicators, it can be noted that the degree of economic content has continuously risen since the beginning of the observation period (Appendix 20). In contrast, environmental content has increased from 2014 onwards (Appendix 21), while the level of social content varies over the period and does not give a clear direction (Appendix 22). The findings are partially consistent with the results of Mittelbach-Hoermanseder et al. (2019), who found annual increases for the different topic-specific CSR disclosures.

Concerning the social dimension, one could argue that the EU Directive requires that companies report on social matters (European Union, 2014). Hence, this should have led to an increase in social content. Nonetheless, the average cosine similarity of the social dimension is already high in 2010, exceeding the economic and environmental values for the entire period (Table 2). Therefore, it is not surprising that the social value has not risen further since it was already a prevalent topic. Another reason, which could explain this insight, is the "comply or explain" clause (Ioannou & Serafeim, 2017, p. 3). The EU Directive allows businesses to justify why they do not pursue specific policies (European Union, 2014). Thereby, companies can avoid implementing such policies as well as the corresponding reporting, if they explain the underlying reasons. Hence, the announcement and the introduction of the EU Directive did not necessarily lead to a higher level of disclosure (Ioannou & Serafeim, 2017).

Furthermore, it is interesting that the mean value of the environmental dimension experiences an increase after the announcement of the EU Directive. This increase could be caused by the mandate to report on environmental matters like "the use of renewable [...] energy, greenhouse gas emissions, water use and air pollution" (European Union, 2014, p. 2). A similar insight was generated by the study of Hummel and Rötzel (2019), who observed an increase of information on required topics after a similar regulation became effective in the UK. Due to the fear of being caught to be non-compliant or due to perceived benefits of being compliant, companies might include more CSR information in their reports as a reaction to the EU Directive (Ioannou & Serafeim, 2017).

Concerning the constant increase of economic disclosure from 2010 onwards, no literature explains the effect before 2014. The economic dimension is related to topics like bribery, taxes, and minimum wages. On top of that, it can be observed that there is a particular trend towards integrated reporting since the share of integrated reports has continuously risen from 2010 onwards (Appendix 24). Since these report types combine business topics with sustainability issues, the CSR sections might include more content related to the economic dimension. Hence, the trend towards integrated reporting could be an underlying reason for the increase in economic content. Nonetheless, the subsequent chapter reveals that integrated reports contain less economic content than stand-alone sustainability reports. Since there is no specific explanation for this observation, this might be caused by the recognized trends towards sustainability reporting. Companies may have expanded their reporting scope by including other aspects than only environmental and social information.

The number of GRI items has increased after the directive entered into force in 2017 (Appendix 23). In the first step, the number of reports complying with the GRI framework is observed. Even if the GRI index is approximated for non-compliant companies, companies actively applying this framework might have a tendency to cover more GRI topics. However, the share of reports under the GRI framework remained almost constant during the observed period (Appendix 28). Therefore, this effect is not caused by an increasing number of companies reporting under the GRI framework. Another reason might be that companies have increased the content of their reports regardless of the framework following previous arguments provided on topic-specific CSR disclosure.

To sum up, it can be said that even if the EU Directive did not come into force until 2017, the mere announcement of the regulation was accompanied by significant changes to the textual characteristics of sustainability reports. Thus, the present study provides valuable insights on the development

of the key components of sustainability reports.

6.3. Integrated vs. Stand-Alone Reports

Another revealing observation is the comparison of integrated reports and stand-alone reports. Similar to the findings of Mittelbach-Hoermanseder et al. (2019), the sample indicates a trend towards integrated reporting since the proportion of integrated reports has increased over time, while the absolute number of reports also has risen (Appendix 3, Appendix 24). This trend is also consistent with the findings of Lukomnik, Kwon, and Welsh (2018), who observed that the number of reports in accordance with the integrated reporting framework doubled between 2013 and 2018. However, it should be noted that even if researchers have already conducted textual analyses on stand-alone and integrated sustainability reports, no study examined the differences between the two types of reports (e.g., Caglio et al., 2020; Mittelbach-Hoermanseder et al., 2019; Nazari et al., 2017).

Concerning the length indicators, stand-alone reports are generally longer than integrated reports (Table 3). This difference is significant at the 1% significance level for all three variables. In fact, they tend to be even twice as long if considering the number of characters and words only. Caglio et al. (2020) conducted a textual analysis on integrated reports in South Africa, and their results revealed that integrated reports include, on average, around 31,000 words. These integrated reports are twice as long as the integrated reports from this sample, which obtain, on average, approximately 15,000 words. Nonetheless, the researchers analyzed the overall report and extracted not only the sustainability part. Hence, the underlying methodological differences may explain the variations in the results.

One can refer to the audience of the different reports to explain the variance in length between the two reporting types. Yusof (2018) argues that the main target group of sustainability reports are stakeholders, while shareholders are the main target group of integrated reports. Furthermore, one could assume that companies tailor their reports to the needs of their audience. The study of Cohen, Holder-Webb, and Zamora (2015), based on a survey of more than 200 professional investors, revealed that investors prefer concise non-financial information. As a result, companies might keep their integrated reports short since the key audience are investors, who prefer brief rather than extensive reports. The integrated reporting framework also encourages this format by stating that “an integrated report should be concise” (International Integrated Reporting Council, 2013, p. 34). Thus, businesses that utilize this framework to produce their integrated report might follow this guiding principle and keep their reports short and concise. Another potential explanation is introduced by Yusof (2018), who investigated the change in sustainability disclosure when companies move from publishing stand-alone reports to publishing integrated reports. The study shows that after the integration, sustainability sections also tend to be shorter. The researcher argues that more information has to be included in a report and, hence, companies have to short financial as well as non-financial information to combine both parts (Yusof, 2018). This trade-off could also be a reason for this phenomenon.

Concerning horizon content and target orientation, there are only minor differences between the two reporting types. While the difference for horizon content is not statistically significant, the difference concerning target orientation is significant at 1%. However, it might be challenging to explain this phenomenon since stand-alone reports contain, on average, 0.3 target-related words more per 1,000 words. In addition, the tone of both report types is similar, and the marginal difference is not statistically significant.

Moreover, stand-alone reports tend to contain a higher degree of numeric content, which is significant at 1% (Table 3). Current literature does not provide any explanations for this observation. Nonetheless, it may be caused by the data extraction methodology. Since only the CSR sections from integrated reports were extracted, the tables might not be considered during the process. This is the case when the corresponding tables are located in subsequent chapters, such as the Appendix. Thus, the potential exclusion of tables from integrated reports might explain the difference in numeric content.

In terms of readability, integrated reports, on average, seem to be less readable than stand-alone sustainability reports (Table 3). This applies to the readability index as well as to the individual components. All of these differences are statistically significant at 1%. This finding is inconsistent with the study of Lueg, Lueg, Andersen, and Dancianu (2016), who found out that preparing an integrated report allows companies to use less technical language. In turn, this should lead to more readable reports. However, the current literature does not provide any reasons for the finding of this study. One possible explanation may be based on the combination of financial as well as sustainability issues. On the one hand, combining both parts leads to a certain level of complexity and, on the other hand, financial reports may require the use of more technical terms. Hence, these reasons might lead to a decrease in readability. Finally, this finding could also be potentially explained by the GRI framework. In Appendix 2, it can be seen that around 80% of the stand-alone reports comply with the GRI guidelines, while only about 47% of the integrated reports comply with the GRI framework. Therefore, integrated reports might be less readable due to the lack of guidance.

Concerning topic-specific disclosure, stand-alone reports are characterized by marginally higher economic content (Table 3). Besides, these reports also contain, on average, more content related to the environmental dimension. Both differences are statistically significant at 1%. This observation is consistent with the findings of Yusof (2018) and follows the argumentation that companies have to balance financial and sustainability issues in an integrated report. In turn, this can lead to a lower proportion of economic and environmental content. This insight could also be potentially explained by the study of Marx and Mohammadali-Haji (2014). The researchers examined integrated reports in South Africa and

Table 2: Descriptive Statistics – Development over Time

Variable	2010	2011	2012	2013	2014	2015	2016	2017	2018
#Characters (in k)	293	303	299	309	325	313	324	341	355
#Total Words (in k)	26	27	26	26	27	26	27	28	29
#Unique Words (in k)	2.3	2.3	2.2	2.2	2.2	2.2	2.2	2.3	2.3
Numeric Content	83.4	87.7	87.2	85.5	87.6	86.0	87.3	87.8	90.0
Horizon Content	1.0	1.0	1.0	1.1	1.0	1.0	1.0	1.0	1.0
Target Orientation	6.2	6.2	6.3	6.1	6.2	6.4	6.6	6.7	7.0
Tone	0.006	0.006	0.006	0.005	0.006	0.006	0.006	0.005	0.005
Readability Index	0.48	0.48	0.48	0.49	0.50	0.50	0.52	0.54	0.54
Economic	0.059	0.061	0.061	0.063	0.065	0.066	0.068	0.071	0.073
Environment	0.136	0.140	0.141	0.141	0.143	0.148	0.150	0.152	0.157
Social	0.175	0.172	0.174	0.172	0.174	0.174	0.173	0.176	0.173
GRI Index	15.4	16.0	15.7	15.7	15.7	15.6	15.8	17.2	17.4

found out that some businesses merely declared their annual reports as integrated reports without changing the reports' content. The underlying motivation is to pretend that they are on the cutting edge of sustainability reporting. As a result, this might lead to a low disclosure level of CSR topics in integrated reports (Marx & Mohammadali-Haji, 2014).

However, the results show that integrated reports tend to include more information on the social dimension than stand-alone sustainability reports. This difference is also significant at 1%. This contradicts the findings of Yusof (2018), who states that companies disclose approximately 70% less social information in their integrated reports. Nonetheless, his study focuses on reports from companies that have switched from stand-alone to integrated reports, while this study covers all report types (Yusof, 2018). Consequently, there is no empirical or theoretical finding explaining this observation.

Lastly, stand-alone reports cover around 50% more GRI topics than integrated reports (Table 3). Similar to the other observations, this difference is also statistically significant at the 1% level. As described above, most of the stand-alone reports follow the GRI guidelines. Companies that prepare their sustainability reports without relying on a specific framework might choose the focal points individually. In contrast, GRI-compliant companies probably focus on particular GRI topics. In turn, this might lead to the disclosure of more GRI topics and explains the differences in the reporting scope.

In summary, various textual differences between integrated and stand-alone reports can be detected. The main findings are related to the differences in length, readability, and topic-specific content.

6.4. Early Adopter vs. Late Adopter

The announcement of the EU Directive in 2014 led to the emergence of late adopters. As described above, these companies anticipated that the regulation would come into force in the following years and started reporting between 2014 and 2016 in response to the announcement of the EU Directive. All reports from companies that published fewer than four reports were excluded from the sample. Thus, in this study, the group of late adopters compromises all companies that started reporting for the financial years 2014 and 2015. The percentage of reports published by early and late adopters over time can be found in Appendix 29.

Concerning the average length of reports, it can be seen that the reports from early adopters tend to be more as twice as long as those from late adopters. This observation is statistically significant at the 1% level for all three variables (Table 4). It is also in line with the study of Stent, Bradbury, and Hooks (2013), who observed the impact of the transition to mandatory IFRS reporting in New Zealand. They also found that early adopters utilized about twice as much space for their annual reports compared to late adopters.

To explain this insight, one can refer to the motivational factors for voluntary reporting. The disclosure of CSR information can lead to various benefits such as lower cost of capital. In contrast, it is also associated with different costs, for instance, preparation and proprietary costs. If the benefits exceed the corresponding costs, firms are encouraged to voluntarily disclose CSR information (Ioannou & Serafeim, 2017; Leuz, 2010). Consequently, one can assume that the costs exceed the benefits for late adopters and, thus, they have not reported until the regulation was announced. This explanation is in line with the study of Christensen et al. (2015) about the mandatory adoption of the IFRS. The researchers argue that early adopters have strong financial reporting incentives, while late adopters adopt the framework in anticipation of the transition to compulsory reporting (Christensen et al., 2015). For this reason, one can assume that voluntary reporters obtain more significant reporting incentives to disclose CSR information. Since the perceived net benefit of sustainability reporting is minimal or even negative, late adopters might tend to keep their reports as short as possible. In this way, they could prepare for the incoming regulation and fulfill the upcoming minimum requirements. However, they keep their reports short to avoid the occurrence of additional costs.

Table 3: Descriptive Statistics – Integrated & Stand-Alone

Variable	Integrated	Stand-Alone	Sig. Level
#Characters	168,364	376,727	1%
#Total Words	15,106	31,596	1%
#Unique Words	1,581	2,500	1%
Numeric Content	80.7	89.5	1%
Horizon Content	1.0	1.0	n.s.
Target Orientation	6.2	6.5	1%
Tone	0.006	0.006	n.s.
Readability Index	0.60	0.47	1%
Flesch Reading Ease	22	24	1%
Fog	18	17	1%
Flesch-Kincaid	22	21	1%
Economic	0.061	0.067	1%
Environment	0.135	0.150	1%
Social	0.182	0.170	1%
GRI Index	11.5	17.8	1%

Mann-Whitney-U-Test (1%, 5%, 10%, or not significant (n.s.))

Moreover, it can be seen that early adopters, on average, incorporate more numbers into their sustainability reports (Table 4). This difference is significant at 5%. To explain this finding, one can refer to the companies' compliance with the GRI Standards. In Appendix 33, it can be seen that less than 50% of the reports published by late reporters are in accordance with the GRI framework, while about 75% of the reports published by early reporters comply with these guidelines. Bhimani et al. (2016) provide evidence that early adopters pursue a differentiation strategy to distinguish themselves from their competitors. This strategy could lead to a tendency for them to report under the GRI framework to set themselves further apart. Looking at the individual GRI disclosure items' requirements, these often demand the publication of numerical data. The GRI-303 standard, for example, requires the disclosure of the amount of water extracted as well as the amount of water recycled (Global Reporting Initiative, 2016d). Moreover, plotting the data in a diagram with the number of GRI items on the x-axis and numeric content on the y-axis, the trend line shows a tendency of the number of reported GRI positions to positively correlate with the degree of numeric content (Appendix 34). This correlation underlines that reports with many covered GRI topics include, on average, more numerical data than non-compliant reports. Hence, the early adopters' differentiation strategy might lead to a higher proportion of companies reporting under the GRI framework, and, thus, these companies incorporate a higher level of numerical content into their sustainability reports.

Moreover, it could also be assumed that the collection of numerical data is relatively expensive. As the net benefit of sustainability reporting for late reporters is minimal or even negative, they are unwilling to invest in numerical data collection. Therefore, these companies might seek to produce their reports in a way that minimizes costs. Hence, the differentiation strategy, as well as the cost intensity of numerical data, could be the underlying reasons for the observation.

Furthermore, the difference in horizon content between the two types is marginal but statistically significant at 1% (Table 4). The effect could potentially be explained by the survey conducted by Bhimani et al. (2016). The results suggest that early rather than late adopters tend to incorporate sustainability aspects into their three-or five-year plan as well as in their business vision. As described above, these companies have real incentives to report on their sustainability issues. For this reason, they might also include these topics in their future strategy and incorporate it into the statements mentioned above. This behavior could explain why reports from early adopters tend to contain more horizon content. In addition, there is also a small difference with regards to target orientation that reports from early adopters tend to include more target-related words. Nevertheless, this difference is not statistically significant.

Referring to the readability, the corresponding index indicates that early adopters' reports are less readable (Table 4). Nonetheless, this is statistically insignificant. Looking at the individual components, there are only minor differences in the decimal places, which are also not statistically significant. Consequently, one could assume that the different reporting types and the associated incentives have no impact on readability. However, even if there are no differences, readability is still at a very high level, which increases the difficulty for stakeholders to comprehend the content of the reports accurately.

With reference to the economic, environment, and social dimension, the reports from early adopters tend to contain more content from all three dimensions. The difference in social content is statistically significant at 1%. Even if the re-

ports from late adopters are characterized by a lower degree of social content, it is still higher than the economic and environmental content of reports from early adopters. Bhimani et al. (2016) explain that early adopters report on their sustainability responsibilities to create value, while late adopters report on these issues to deflect threats to their brand image. Hence, one explanation could be that late adopters report only the minimum amount of content to satisfy stakeholders and maintain their brand image. In contrast to the other group, they are not as intrinsically motivated and, thus, they tend to report relatively fewer contents than early adopters.

However, the difference in economic and environmental content between reports from early and late adopters is marginal and not even significant at 10%. This finding seems to contradict the previous insights. Nonetheless, the study conducted by Bhimani et al. (2016) also showed that late adopters try to keep up with their competitors by imitating them. For this purpose, they use similar reporting methods, resulting in reports with similar topics (Bhimani et al., 2016). As a result, this imitation strategy could explain why late adopters obtain a similar level of economic and environmental content, although they publish shorter reports and receive a marginal net benefit from reporting.

Referring to the number of reported GRI buckets, early reporters include an average of about sixteen buckets, while late adopters report an average of about thirteen buckets. This difference is significant at 1%. One possible explanation is based on the differentiation strategy mentioned above. Early reporters authentically pursue CSR reporting to differentiate themselves from their competitors. Therefore, they might use the opportunity to report under the GRI framework to stand out further. It turns out that relatively more reports from early adopters are in line with the framework compared to late adopters. If companies report according to the guidelines, there might be a high probability that they use the related topics as a guideline, resulting in a greater coverage of GRI topics.

To sum up, there are significant differences between the reports from early and late adopters. The current literature indicates that the reasons for those could be related to the early adopters' differentiation strategy and the underlying differences in reporting incentives.

7. Implications of CSR Performance

7.1. Empirical Model

As the data has been collected and the overall dataset, as well as the characteristics of the different groups, have been described, the formulated research hypotheses can be tested. The following empirical model is utilized to test the hypotheses:

$$
\begin{aligned}
&CSR\ Performance \\
&\quad = \beta_0 + \beta_1 LGTH + \beta_2 NC + \beta_3 HC + \beta_4 TO \\
&\quad + \beta_5 TN + \beta_6 RDB + \beta_7 ECON + \beta_8 ENV \\
&\quad + \beta_9 SOC + \beta_{10} ITMS + \beta_{11} SZ + \beta_{12} ROA \\
&\quad + \beta_{13} CURRAT + \beta_{14} CAPSP + \beta_{15} PNEW \\
&\quad + \beta_{16} LEV + \beta_{17} INST + \beta_{18} ADT + \beta_{19} GRI \\
&\quad + \beta_{20} INTG + \sum_{j=1}^{9} \mu_j Year + \sum_{k=1}^{11} \gamma_k Industry
\end{aligned}
$$

Based on this empirical model, a linear regression is conducted using R and the corresponding *caret* package. The model allows forecasting the relation between different input variables and an independent output variable.

7.1.1. Dependent Variable

The dependent variable of the regression is the individual CSR performance. Consistent with the study of Clarkson et al. (2020), the corresponding performance data was retrieved from the ASSET4 database. The performance score is an ESG score ranging between 0 (*weak performance*) and 100 (*strong performance*) and is based on 178 indicators. Thereby, it comprises the following dimensions: resource use, emission, innovation, workforce, human rights, community, product responsibility, management, shareholders, and CSR strategy (Eikon, 2017). As described in the beginning, ASSET4 offers extensive information on the sustainability aspects of companies worldwide. Moreover, researchers argue that using this data is beneficial since the underlying evaluation of companies' CSR performance is more comprehensive and objective compared to other sustainability databases (Clarkson et al., 2020). In addition, the corresponding scores are not solely based on the companies' sustainability reports, but also incorporate information from "stock exchange filings, [...] annual reports, non-governmental organizations' websites, and various news sources" (Ioannou & Serafeim, 2012, p. 21). Therefore, this underlines that the validity of the data is ensured.

In contrast to Clarkson et al. (2020), the combined ESG score of the ASSET4 database is utilized, and no dimension is excluded. This score is based on social, environmental, and corporate governance factors. Thereby, the ESG performance score reflects the overall concept of sustainability.

7.1.2. Independent Variables

The variables of interest are the textual characteristics as well as additional company- and report-specific characteristics. *LGTH* reflects the length of reports. Due to the risk of multicollinearity, only one of the length indicators is considered in the regression. For this reason, the logarithm of the number of characters serves as a proxy for the report length. *NC*, *HC*, and *TO* are the variables representing the degree of numeric content, horizon content, and target-related words

Table 4: Descriptive Statistics – Early & Late Adopter

Variable	Early Adopter	Late Adopter	Sig. Level
#Characters	323,956	136,032	1%
#Total Words	27,449	11,349	1%
#Unique Words	2,265	1,510	1%
Numeric Content	87.3	77.8	5%
Horizon Content	1.0	0.8	1%
Target Orientation	6.4	5.8	n.s.
Tone	0.006	0.005	n.s.
Readability Index	0.50	0.45	n.s.
Flesch Reading Ease	23	23	n.s.
Fog	17	17	n.s.
Flesch-Kincaid	21	21	n.s.
Economic	0.065	0.063	n.s.
Environment	0.146	0.138	n.s.
Social	0.174	0.157	1%
GRI Index	16.2	12.6	1%

Mann-Whitney-U-Test (1%, 5%, 10%, or not significant (n.s.))

per 1,000 words. On top of that, *TN* represents the tone indicator, while *RDB* is the readability index. The individual readability scores are excluded from the regression due to the risk of multicollinearity. The thematic CSR disclosure dimensions are *ECON* (economic), *ENV* (environment), and *SOC* (social). *ITMS* reflects the number of GRI topics covered per report. In addition to the textual characteristics, dummy variables are included. *ADT* is a dummy variable, whether the company is an early or late reporter. Furthermore, *GRI* reflects whether the report is in accordance with the GRI framework or not. Lastly, the dummy variable *INTG* indicates whether the report is an integrated or stand-alone report.

7.1.3. Control Variables

Control variables are included in the regression to control other factors that might influence the underlying performance. These variables comprise both financial as well as non-financial variables. The choice of control variables is based on previous studies (e.g., Clarkson et al., 2008, 2011; Nazari et al., 2017). The first control variable is *SZ*, which serves the purpose of reflecting the size of the company. For this purpose, it is measured by the natural logarithm of total assets. Another control variable is *ROA*, which is the income of the enterprise divided by its total assets. *CURRAT* is measured by the ratio of total current assets to total current liabilities. *CAPSP*, which reflects the degree of capital spending, is the ratio of total spending to total sales revenues. Moreover, *PNEW* is the amount of net property, plant, and equipment divided by gross property. *LEV*, which reflects the company's leverage ratio, is calculated by dividing total debt by total assets.

Furthermore, *INST* is the percentage of institutional ownership. Since the database does not provide such information, other measures were used as a proxy for institutional ownership. As the database provides information on the percentage of strategic ownership differentiated by different owner types, INST equals the sum of strategic ownership held by institutions, investment banks, and pension/endowment funds. All of these data points were retrieved from the WorldScope databank via the financial information service of Thomson Reuters. WorldScope provides fundamental financial data on leading companies worldwide. For each year, the value at the end of the year was taken to ensure consistency.

Lastly, dummy variables for the industry and year were added. These variables control for specific temporal and industry effects. Nonetheless, the database does not include the corresponding control data for all companies in all periods. Thus, the observations with missing control variables were excluded from the regression. This procedure led to the exclusion of 532 reports resulting in a final regression sample of 1,899 reports.

7.2. Assumptions of Linear Regression

Before one can interpret the results, the four assumptions of a linear regression have to be tested. These assumptions include linearity (linear relationship), normality, homoscedasticity, and the avoidance of multicollinearity. The corresponding tests will be conducted with the statistical software R and are mainly based on the visualization of the data. If one of these assumptions does not hold, the scientific findings of the regression are inefficient or even strongly biased. Hence, the regression assumptions are necessary to accurately interpret the results (Statistics Solutions, n.d).

Firstly, multicollinearity means that at least two independent variables are highly correlated and demonstrate a linear relationship. Hence, this causes that one of the different

independent variables is redundant. The occurrence of multicollinearity has to be avoided since it leads to somewhat unstable parameter estimates. There are various methods to test for multicollinearity. In this thesis, the Variance Inflation Factor (VIF) is utilized. Multicollinearity causes that the variance of the regression increases and, hence, makes it unreliable. The VIF makes use of this phenomenon and measures how much of the inflated variance is due to multicollinearity. A corresponding value above four indicates that multicollinearity might exist and that a further analysis should be conducted. If the VIF score is higher than ten, this is a strong indication for multicollinearity. In this case, this has to be corrected (Corporate Finance Institute, n.d). In Appendix 35, the different VIF values for each predictor are shown. The indicator for length, which is the logarithmic number of characters, and the number of reported GRI buckets obtain the highest VIF values. These values are slightly above four. If one of the variables is excluded from the regression, the VIF values tend to be lower. This insight indicates that these predictors are somewhat correlated. This correlation is evident because when companies report on more GRI topics, the reports inevitably become longer. Since the VIF values are above the limit of four, this indicates that multicollinearity might exist. Thus, the variable representing the number of GRI items reported is excluded from the regression to solve this problem. In turn, the VIF value of report length decreases to around 2.5. The new values can be seen in Appendix 36.

Secondly, the second assumption, which is linearity, means that the relationship between the dependent variable and the numerous predictors is linear. This assumption implies that the outcome variable is a straight-line function of the different predictors. Moreover, it indicates that the regression slope does not depend on the value of the other independent variables. To test this assumption, one can plot the residuals against the fitted values. The assumption holds if one can detect a horizontal line with no specific pattern (Nau, n.d). The corresponding plot can be seen in Appendix 37, which provides evidence that a linear relationship between the numerous predictors and the CSR performance exists.

Thirdly, the assumption of normality means that the residual errors are assumed to be normally distributed. A violation of this assumption has a negative effect on the significance level of the coefficients and the corresponding confidence intervals. This assumption can be tested using a QQ-plot with the theoretical quantiles on the x-axis and the standardized residuals on the y-axis (Nau, n.d). This plot is shown in Appendix 38. Since the data is on the diagonal dashed line, the data is assumed to be normally distributed. Hence, the normality assumption also holds.

Fourthly, the last assumption is the presence of homoscedasticity. Homoscedasticity means that the error terms of all values of the predictors are on the same level. If the assumption does not hold and heteroscedasticity is present, this leads to biased standard errors. In turn, this results in false conclusions about the significance of the numerous regression coefficients. To test for homoscedasticity, one can use the scale-location plot, which shows the fitted values on the x-axis and the root of standardized errors on the y-axis (Statistics Solutions, 2013). This plot is shown in Appendix 39. It demonstrates a horizontal line with a small downward tendency on the right side. This shows that the associated points tend to be equally spread. Hence, the fitted values' variances tend to be constant among all values except a small deviation on the right side. Therefore, it can be assumed that the assumption of homoscedasticity holds.

Moreover, besides the analysis of the regression assumptions, one should look at the existence of outliers. With regard to the dataset of this study, outliers are reports with strongly deviating characteristics. Using the statistical software R, these values are highlighted in the different plots. Three outliers can be identified. In the next step, one should check how realistic the values of these outliers are. For this purpose, the values are compared to the median values of the overall dataset. The textual characteristics of the three outliers identified seem to be reliable. However, their ESG scores are relatively low since these values are close to or even zero. For this reason, one could consider excluding the outliers from the dataset. However, as the ASSET4 database publishes these values, the three corresponding reports will be retained in the analysis. Nonetheless, a more in-depth analysis concerning outliers will be conducted in the robustness section.

In summary, the underlying assumptions of a linear regression are fulfilled. In addition, a few outliers are identified. Since the outliers and the corresponding characteristics seem to fit the dataset except for the CSR performance deviations, the reports remain in the dataset. However, this will be further considered during the robustness tests of the regression. All in all, the regression results can be interpreted since the underlying regression assumptions are fulfilled.

8. Results

8.1. Hypotheses Testing

After conducting the textual analysis, collecting further data, and testing the regression assumptions, the regression can be performed, and its results can be interpreted. Before the individual hypotheses are analyzed, the explanatory power of the regression is observed. The adjusted R^2, which adjusts for the number of independent variables, amounts to 0.551. Hence, 55.1% of the variability of the CSR performance is explained by the different predictors. The study of Patten (2002) examined the relationship between environmental disclosure and the corresponding performance, and its model obtained an adjusted R^2 of 0.38. The researcher states that the "explanatory power is relatively high" (Patten, 2002, p. 770). In contrast, the study of Clarkson et al. (2020) observed a similar relation between sustainability performance and disclosure and obtained an adjusted R^2 of around 0.72. Therefore, the corresponding value of this study is between the values of the current literature. This comparison emphasizes that the explanatory power of this

study is acceptable. Moreover, the F-statistic indicates that the regression model is statistically significant at 0.1%. The regression table is shown in Table 5.

The first hypothesis predicts that the length of sustainability reports is positively correlated with CSR performance. The regression coefficient is positive and statistically significant at 0.1%. This coefficient implies that companies obtaining higher sustainability performance tend to produce more comprehensive sustainability reports. This insight is consistent with the finding of Clarkson et al. (2020) that companies with superior CSR performance prepare longer reports. The underlying reason could be the fact that they can report on more topics due to a larger number of pursued CSR activities. Moreover, this is also in line with the voluntary disclosure theory and contradicts the legitimacy theory. All in all, the hypothesis about the relationship between report length and sustainability performance is supported.

The second hypothesis states that the readability of CSR reports negatively correlates with the corresponding performance. This hypothesis means that reports from well-performing companies are assumed to be more readable. Looking at the associated regression coefficient, it contradicts the hypothesis. The coefficient implies that companies that prepare reports with a low level of readability have a higher CSR performance. This finding is consistent with the study of Clarkson et al. (2020), although the researchers expected a different relationship. The researchers argue that more sophisticated analyses might cause poor readability. However, the coefficient is not statistically significant. To sum up, the regression analysis does not support the second hypothesis.

The third hypothesis forecasts that the tone of sustainability reports is negatively correlated with the associated performance. However, the regression generated a positive coefficient, which is statistically significant at the 1% level. This outcome indicates that companies with a poor (good) CSR performance publish reports with a more negative (positive) tone. As described in the literature review, this can be explained by the behavior of managers of poor performing companies, who include more negative phrases to fend off criticism and/or build up credibility (Clarkson et al., 2020). For these reasons, they might use a more negative tone in their sustainability reports. In summary, the third hypothesis is not supported.

Hypothesis 4a predicts a positive relationship between the number of reported GRI items and CSR performance. Due to multicollinearity, this variable was excluded from the regression. However, in order to examine this hypothesis, the length indicator was excluded from the regression, and the indicator for the number of GRI items was included. In this way, the occurrence of multicollinearity is avoided. The remaining regression assumptions are also tested for this regression equation. The regression reveals a positive coefficient of around 0.49, which is significant at 0.1%. This finding supports the hypothesis that companies, which tend to report on more GRI topics, have a higher CSR performance.

Hypothesis 4b states that topic-specific CSR disclosure positively correlates with sustainability performance. Referring to the environmental and social dimensions, the regression results imply that the relationship between these dimensions and the performance is positive. The coefficient of environmental disclosure is significant at 0.1%, while the social disclosure coefficient is significant at 1%. This is in line with the voluntary disclosure theory and the findings of the current literature (e.g., Al-Tuwaijri et al., 2004; Clarkson et al., 2008). In contrast, the coefficient of the economic dimension is negative, which contradicts the hypothesis. One could argue that this is based on the fact that the dependent variable is an ESG score, which does not consider economic factors (Eikon, 2017). Moreover, the associated coefficient is not statistically significant. As the economic dimension is not part of the ESG score, this might explain that the degree of economic disclosure has no significant effect on the dependent variable. In conclusion, the hypothesis is widely supported, except for the relationship between economic disclosure and CSR performance.

The fifth hypothesis predicts that companies with superior CSR performance tend to include more numerical data into their reports. The related coefficient is negative, but close to zero (-0.002). As shown in Table 1, the difference between the minimum and maximum numeric content for the overall dataset is around 450 numbers per 1,000 words. According to the regression, this difference in the extreme values would result in a 0.72 difference in CSR performance. This underlines that the impact is marginal. Moreover, the coefficient is not statistically significant. Therefore, the stated hypothesis is not supported by the regression analysis.

Moreover, the sixth hypothesis forecasts a positive relationship between the degree of horizon content and sustainability performance. However, the corresponding coefficient indicates that the relationship is the other way around. This deviation might be caused by poor performing companies that focus on future performance to distract from the current one. These companies might also incorporate future-related content to emphasize their ambition to improve their performance in the future (Asay et al., 2018). Nonetheless, the coefficient is not statistically significant. Thus, the hypothesis is not supported.

The seventh hypothesis deals with the relationship between the degree of target orientation and CSR performance. Based on the goal-setting theory and the voluntary disclosure theory, it was assumed that firms with a superior CSR performance tend to formulate more targets in their sustainability reports (Latham & Locke, 1979). The former allows managers to motivate employees and direct their behavior. In contrast, the latter allows these firms to show the market that they are able to meet their stated objectives. As a result, they incorporate more target-related words into their sustainability reports. However, a necessary condition is that the formulated objectives are clear, specific, and attainable (Latham & Locke, 1979). The regression results confirm this hypothesis since the corresponding coefficient is positive. Furthermore, it is statistically significant at 0.1%. In summary, the

hypothesis about the relation between the degree of target orientation and CSR performance is confirmed.

As the hypotheses concerning the textual characteristics have been analyzed, the hypotheses concerning the other characteristics will be observed in the next step. The eighth hypothesis is related to whether a company is defined as an early adopter or a late adopter. The associated coefficient shows that being a late adopter has a negative effect on sustainability performance. According to the regression, this effect amounts to around 9.8 performance points. Considering that the ESG score ranges between 0 and 100, this is an enormous difference. In addition, the coefficient is significant at the 0.1% level. This outcome is consistent with the study of Bhimani et al. (2016), which states that late adopters are only motivated to pursue sustainability reporting to improve their reputation. Thus, late adopters only want to keep up with their competitors and do not authentically pursue CSR activities. In turn, these companies tend to obtain a lower CSR performance. The results are also in line with the voluntary disclosure theory. To sum up, the hypothesis about the relationship between the adopter type and CSR performance is confirmed.

Concerning the hypothesis dealing with the impact of reporting under the GRI framework, the results reveal that non-complying companies, on average, obtain a score that is around 5.8 performance points lower. The coefficient is significant from zero at the 0.1% level. This outcome contradicts the study of Bernard et al. (2015) and the legitimacy theory. On the contrary, the results emphasize that companies with a superior sustainability performance commit themselves to the GRI framework to differentiate themselves from the competition. However, the stated hypothesis does not determine a specific direction but merely assumes that there is a certain relationship. Therefore, the hypothesis is supported. In addition, the impact of GRI reporting seems to be positive.

The last hypothesis states that the firm's sustainability performance has an impact on the reporting method. Nonetheless, it does not predict a specific direction. The regression analysis shows that the coefficient of preparing a stand-alone sustainability report is positive. The coefficient is statistically significant at 1%. This finding demonstrates that companies that have superior sustainability performance tend to publish stand-alone reports. The underlying reason might be that these firms want to highlight their superior performance. Therefore, they tend to publish a report that solely focuses on this aspect. In contrast, firms with poor performance might publish integrated reports to hide their sustainability failure in their financial results. To sum up, the hypothesis is supported. Moreover, the effect of integrated reporting seems to be negative.

All in all, the regression analysis confirms the hypotheses concerning length, number of GRI items, target orientation, adopter type, GRI framework, and the reporting method. Moreover, the hypothesis about the degree of topic-specific CSR disclosure is widely supported. The coefficients relating to numerical content, horizon content, and readability are insignificant, while contrary to expectations, the tone positively correlates with CSR performance. It is noteworthy that most of the hypotheses can be explained by the voluntary disclosure theory and not by the legitimacy theory. These insights underline that firms with superior CSR performance tend to be more engaged in sustainability reporting. As a result, these companies prepare, on average, longer stand-alone reports with a high degree of formulated targets and a high degree of topic-specific disclosure. In addition, these reports tend to be in accordance with the GRI framework and cover many of the related topics. In contrast, underperforming companies tend to prepare short reports integrated into their annual reports and contain a few formulated targets. Moreover, these reports are rarely prepared using the GRI framework.

8.2. Robustness Tests

Different robustness tests are performed to check the strength of the regression model. These tests should confirm the model as well as the corresponding findings by applying different conditions. The first robustness test refers to the definition of early and late adopters. As described in the literature review, late adopters are defined as companies that have started reporting in response to the announcement of the EU Directive in 2014. However, according to Fiechter et al. (2019), some EU companies might have anticipated the regulation and started reporting before the announcement. Therefore, in the first robustness test, late adopters are defined as companies that started reporting in 2013 or later. Thereby, the number of reports from late adopters increases from 58 to 169 reports. The significance levels of the individual predictors before and after the robustness test, as well as the corresponding coefficients, can be seen in Appendix 40.

For most predictors, the results remain unchanged. While the significance level of the tone value coefficient increased from 1% to 0.1%, the significance level of the coefficient for publishing an integrated report decreased from 1% to 5%. However, the results remain constant and are only marginally affected. In addition, the coefficient of horizon content becomes slightly significant at 10%. As the correlation is negative, companies with a poor CSR performance focus more on future aspects in their reports. The underlying reasons might be that these companies want to distract from the past or positively shape stakeholders' perception by emphasizing future initiatives (Asay et al., 2018). This outcome contradicts the originally formulated hypothesis about the relationship between CSR performance and horizon content. However, this hypothesis was not confirmed in the main analysis, either. All in all, a different definition of early and late adopters leads to only marginal differences in the various variables and, hence, confirms the previous findings.

The second robustness test involves another dependent variable instead of the ASSET4 ESG score. Other studies use different sustainability scores, such as the performance data from Kinder, Lydenberg, and Domini (KLD) or the Bloomberg ESG Score (e.g., Clarkson et al., 2020; Nazari et al., 2017).

Table 5: Regression Results

Variable	Coefficient	Relationship	Sig. Level
LGTH	10.100	+	0.1%
NC	-0.002	-	n.s.
HC	-0.776	-	n.s.
TO	0.481	+	0.1%
TN	133.274	+	1%
RDB	0.431	+	n.s.
ECON	-9.339	-	n.s.
ENV	42.929	+	0.1%
SOC	18.742	+	1%
ITMS[1]	0.488	+	0.1%
GRI (No)	-5.823	-	0.1%
ADT (Late)	-9.832	-	0.1%
INTG (No)	1.747	+	1%
SZ	3.205	+	0.1%
ROA	0.079	+	5%
CURRAT	-1.113	-	0.1%
CAPSP	-0.018	-	n.s.
PNEW	-8.456	-	0.1%
LEV	9.827	+	0.1%
INST	7.270	+	10%
Temporal Effects		Yes	
Industry Effects		Yes	

1 = Separate regression without LGTH due to multicollinearity

However, this data is not available in the database of Thomson Reuters. Thus, the ESG score from the ASSET4 database is adjusted. Following the approach of existing studies (e.g., Clarkson et al., 2020; Ioannou & Serafeim, 2012), the governance dimension is excluded from the overall score. In the ASSET4 database, the individual scores for each ESG dimension are available. The average score of the environmental and social dimensions is computed to exclude the governance dimension. This score acts as a new dependent variable to test the robustness.

Looking at the results in Appendix 41, the coefficient of the tone variable became insignificant. However, in the main analysis, the corresponding hypothesis was not supported either. Furthermore, the significance of social disclosure's coefficient decreased from 1% to 5%, while the negative coefficient of economic disclosure became significant at 1%. However, the hypothesis regarding the latter coefficient was not supported before. More interesting is the change in the readability coefficient. In the main analysis, the coefficient was not significant and positive. In this robustness test, the coefficient became negative and significant. This outcome supports the stated hypothesis that firms with poor CSR performance produce less readable reports. An underlying reason might be that they aim to obfuscate their poor CSR performance. Lastly, the effect of producing an integrated report became insignificant. This change contradicts one of the critical findings of the main analysis. To sum up, the main findings remain constant except for the impact of producing an integrated report.

The third robustness test manages the occurrence of outliers. To identify all outliers, the approach of John Tukey, a researcher who invented the boxplot, is applied. For this purpose, the interquartile range (IQR) has to be calculated. The IQR is the difference between the 25^{th} and 75^{th} percentile. All data points that are one and a half times the IQR higher (lower) than the 75^{th} percentile (25^{th} percentile) are defined as outliers (Purplemath, n.d). The outlier identification process is performed for each numeric variable. Instead of removing these values from the dataset, values outside the lower limit are replaced by the value of the 5^{th} percentile. Moreover, all values that lie outside the upper limit are replaced by the corresponding value of the 95^{th} percentile. In this way, the dataset is cleared for outliers. The new levels of significance, as well as the new coefficients, are listed in Appendix 42.

In addition to various changes in the control variables, the coefficient of numeric content becomes significant at 1%. The results indicate that there is a negative relationship between this variable and CSR performance. This finding demonstrates that companies with poor CSR performance include more numerical data in their reports. Even if a positive relationship was predicted, the hypothesis was not supported in the main analysis. Furthermore, the significance levels from tone and target orientation decrease slightly. All

in all, the results remained mostly unchanged.

The fourth robustness test is related to future and past sustainability performance. The analysis examined the relationship between textual characteristics and CSR performance in the same period. However, one could argue that companies have an excellent performance in period $t-1$ but incorporate this into their reports for period t. A potential explanation could be that they were not aware of their excellent performance. This phenomenon might be particularly prevalent if, for example, they publish the report shortly after the end of the financial year. In this case, they might not be aware of their excellent CSR performance and incorporate this into the subsequent report. This might lead to a time-delayed effect. Moreover, one could also imagine a similar effect in the opposite direction. For example, due to upcoming initiatives, firms are aware that they will obtain a superior CSR performance in the period $t+1$. For this reason, they design their reports differently in period *t,* highlighting their upcoming performance. Therefore, robustness tests that examine the relationship between textual characteristics and future (past) performance are conducted.

The corresponding regression tables are shown in Appendix 43 and Appendix 44. Most of the changes are only marginal. The only remarkable effect concerns the social disclosure coefficient. This coefficient becomes insignificant if past performance acts as the dependent variable. However, the remaining coefficients remain mostly constant.

For the fifth robustness test, an additional control variable that captures the reporting experience is introduced. The variable *EXP* reflects the number of reports under the GRI framework since 2005. For instance, if a company has published eight reports from which six reports are according to the GRI framework since 2005, the corresponding value is equal to six. All reports before 2005 are not relevant for this analysis. The regression results are listed in Appendix 45. There are only minor changes in the significance levels, but the key insights remain robust. The coefficient *EXP* is positive and significant at 1%, indicating that reporting experience positively affects CSR performance.

Finally, following the approach of Mittelbach-Hoermanseder et al. (2019), an additional control variable was introduced taking into account the effect of national culture. This aspect is particularly important since some countries, such as France, Denmark, and the UK, have already introduced regulations on the disclosure of sustainability information before the announcement of the EU Directive (Fiechter et al., 2019; Hummel & Rötzel, 2019). This could have an impact on the results. The first two letters of the International Securities Identification Number (ISIN) indicate the company's country. These letters were extracted and used as an additional dummy variable to account for national differences.

After performing the regression with the new control variable, the results demonstrate some interesting changes. These can be seen in Appendix 46. Similar to the first robustness test, the coefficient of the horizon content became significant. This result contradicts the initial hypothesis and supports the theory that CSR performance and horizon content are negatively correlated. Moreover, the coefficient of publishing an integrated report became insignificant, similar to the second robustness test. The other variables remained mostly unchanged. Therefore, this robustness test confirms the key findings of this thesis illustrated previously except for the relationship between report type and sustainability performance.

In summary, the four robustness tests caused only marginal changes for the different variables. Most of the changes concerned variables whose hypothesis had not been confirmed before. Nevertheless, the hypothesis regarding the report type was not always supported. When choosing a different ESG score or controlling for national differences, the coefficient became insignificant. These robustness tests underline that this variable is sensitive to the conditions and question the relationship between this variable and CSR performance. Nonetheless, the robustness tests underline that the overall model is largely robust, and the main findings remain mostly unchanged.

8.3. Additional Analysis

So far, the study confirms the hypotheses concerning the relationship between CSR performance and report length, environmental content, social content, number of GRI items covered, target orientation, adopter type, GRI framework, and reporting method. Robustness tests in the previous sections also supported these hypotheses. However, the causal direction of the regression goes from textual characteristics to CSR performance. This structure allows to analyze the relationship between textual characteristics and CSR performance in a single regression. Nevertheless, the underlying reasoning of the hypotheses is based on the effect of CSR performance on textual characteristics. However, one could argue that regression results only reflect a correlation and not a causal relationship. This argumentation would mean that the underlying regression is valid.

However, additional analyses are performed. These are intended to show that the findings are valid with CSR performance as the independent variable and text characteristics as the dependent variables. Therefore, a regression is performed for each variable of the confirmed hypotheses. This approach results in a total of eight additional regression equations. The regression structure is similar to the previous analysis. The control variables remain the same. The eight text characteristics are each used as dependent variable, while CSR performance is the predictor. In addition, the dummy variables concerning the report type, the GRI framework, and the adopter type are also included as control variables. The individual variables are not included if they are used as the dependent variable themselves. In this case, the variable is omitted because, otherwise, it would act as both a dependent and an independent variable.

Concerning report length, target orientation, number of GRI items, and environmental and social disclosure, a linear regression is performed. However, this is not possible for the other three variables, as they are binary variables. In this

case, a logistic regression is performed that is capable of modeling a binary dependent variable. The regression equation, for example, for the report length, is as follows:

$$\begin{aligned} LGTH = \beta_0 + \beta_1 CSR\,Performance + \beta_2 SZ + \beta_3 ROA \\ + \beta_4 CURRAT + \beta_5 CAPSP + \beta_6 PNEW + \beta_7 LEV \\ + \beta_8 INST + \beta_9 ADT + \beta_{10} GRI + \beta_{11} INTG \\ + \sum_{j=1}^{9} \mu_j Year + \sum_{k=1}^{11} \gamma_k Industry \end{aligned}$$

The regression tables are shown in Appendix 47 – Appendix 54. For each regression, the coefficient for CSR performance is significant at 1% or even 0.1%. Furthermore, the direction of the coefficient in each regression is consistent with the hypotheses. For instance, the CSR performance coefficient is negative when the GRI variable is the dependent variable. A regression outcome close to one predicts that the report is not in accordance with the GRI framework. Hence, this outcome shows that the higher the ESG score is, the more the model predicts that the corresponding report is in line with the GRI framework. This prediction is consistent with the hypothesis that firms with superior CSR performance tend to report under this framework. All in all, the additional analyses confirm the previous findings.

8.4. Managerial Implications

The findings of the current analysis provide manifold implications for businesses, regulators, analysts, and other market participants. Firstly, the study shows that narrative characteristics (e.g., readability, report length) and additional characteristics (e.g., adopter type, GRI compliance) allow stakeholders to assess businesses' sustainability performance. The associated performance is an essential factor for the decision-making process of consumers and investors. As a result, managers have to consider this during the creation process of their sustainability report. Instead of only considering which contents they include in their reports, managers also have to consider how they communicate the corresponding contents. This approach allows managers to actively shape stakeholders' perception and convince them of their sustainability efforts. Reporting under the GRI framework or incorporating a high degree of environmental content, for example, can lead to CSR performance being rated higher than it genuinely is. If they do not consider these aspects, stakeholders might perceive the firm's sustainability performance differently.

Hence, managers should invest more time and resources into the preparation process of their sustainability reports to optimally design the textual narratives. Managers could set up a dedicated department as well as internal guidelines to ensure a sufficient reporting quality. As a result, the creation of such reports receives full attention. This investment guarantees that the textual design is optimized. To check the format and the quality of their sustainability reports, firms could utilize textual analysis. Thereby, they could analyze their past reports, which helps them to identify their weaknesses. Afterwards, they can tackle these weaknesses and publish an improved report for the next period.

Furthermore, companies should design their reports according to the findings of the study. One possibility would be to commit themselves to the GRI framework voluntarily. As shown in the regression analysis, this signals market participants a high level of sustainability performance. Even if they cannot deliver this level of performance, they can positively shape their stakeholders' perceptions.

Moreover, in particular for private investors, it is often challenging to assess the actual sustainability performance of companies. This is based on the fact that the access to databases such as the ASSET4 is often subject to a fee. Hence, investors often have to rely on the contents of sustainability reports to assess companies' CSR performance. However, the study expands the toolkit of private investors since they can utilize the linguistic features as an additional proxy for a firm's sustainability performance. Since the latter factor is an increasingly important decision-making criterion, this offers an immense benefit for private investors. Furthermore, this is also a vast advantage for institutional investors and analysts. Different databases do not always cover small and private companies. Thus, these investors also have difficulties assessing the actual CSR performance and can use linguistic features as an additional dimension for their analysis.

Lastly, as shown in chapter six, *Descriptive Statistics*, there are substantial differences in the design of sustainability reports. While some companies publish only a few pages within their annual report, other companies intensively describe their sustainability efforts substantiated by numerical data and in accordance with the GRI framework. Even if stakeholders can derive the corresponding performance from such characteristics, it is challenging to compare the reports among different companies. This is a major finding for regulators in Europe. Even if the EU Directive mandates that companies report on their sustainability issues, companies are not obliged to use a specific framework. This increases the pressure on regulators to tighten the regulatory framework to make sustainability reporting more comparable and consistent. Thereby, they should consider following the approach of financial reporting and introducing a specific reporting framework.

8.5. Theoretical Implications

From a theoretical point of view, this thesis contributes to the literature in several respects. Firstly, it is the first study examining the differences between integrated and stand-alone reports as well as the differences between reports from early and late adopters by conducting a textual analysis. While many researchers focused solely on integrated or stand-alone reports (e.g., Clarkson et al., 2020; Muslu et al., 2019; Nazari et al., 2017), no prior study already investigated the differences between these reports and explained the underlying reasons. Referring to integrated and stand-alone reports, this study explores significant differences, for instance, concerning report length, content, and readability. Moreover, the re-

sults also reveal significant differences between the reports from early and late adopters.

Secondly, this study examines textual characteristics, which have not been observed by other researchers. Mittelbach-Hoermanseder et al. (2019) measured topic-specific CSR disclosure by the cosine similarity. However, they measured the topics of the EU Directive, while this study measures the three dimensions of the GRI Standards. In addition, another textual feature that was previously unobserved is the degree of target orientation. Targets are a fundamental element in the field of management accounting and control. Therefore, it is an interesting finding that target orientation also plays an essential role in sustainability reporting. This study is also the first to utilize textual analysis to approximate a GRI index for each report. This approach allows to determine how many topics are covered, even if the reports are not following the GRI framework. To sum up, the results imply that target orientation, topic-specific CSR disclosure, as well as the number of GRI topics covered, significantly correlate with CSR performance.

Thirdly, the study is one of the first observing the impact of the EU Directive entering into force in 2017. The observation period allows to detect any changes in response to the shift to mandatory sustainability reporting. Since the shift happened in 2017, previous literature could not examine this impact. Therefore, they could only observe the effect of the announcement of the EU Directive in 2014, which was not yet binding this year. The results imply that more GRI topics have been covered after the regulation came into force (Appendix 23).

Fourthly, some studies also examined the relationship between textual features of sustainability reports and the associated performance (e.g., Clarkson et al., 2020; Hummel & Schlick, 2016; Nazari et al., 2017; Patten, 2002). However, this study focuses not only on textual features but also on other characteristics such as the report type (integrated or stand-alone), adopter type, and the use of the GRI reporting framework. The results of the regression imply that all of these characteristics significantly correlate with sustainability performance, which display relevant insights expanding current literature.

Moreover, this study also has some implications for future research. Since the literature about the companies' motivations for preparing an integrated or stand-alone report is limited, future research can build on this research opportunity. These insights would enable researchers to precisely explain the differences between integrated and stand-alone sustainability reports. Furthermore, future research can investigate the differences between reports from voluntary adopters and resisters. The latter group postponed the reporting process until the EU Directive came into force in 2017. Due to the observation period of the study until 2018, this grouping could not be observed separately. Lastly, since the sample is based on the STOXX Europe 600, the relation between sustainability disclosure and the corresponding performance in other regions might be an interesting subject for future research. Since the analysis is based on a code in R, it can be performed easily and time efficiently for other regions. To do so, only the reports to be analyzed have to be collected. Therefore, the same study could be replicated, for example, with reports from US companies to discover regional differences.

8.6. Limitations

This thesis also has some limitations, which need to be considered. Firstly, the number of reports from late adopters is relatively small. Therefore, one could question the representativeness of this group. While the sample comprises 2,373 reports from early adopters, there are only 58 reports from late adopters. The underlying reason for this is the utilized sampling rule, which states that only reports from companies with a minimum number of four reports are considered in the study. This limit ensures that companies with only one or two reports do not distort the results in the corresponding years. Since the emergence of late adopters was initiated by the EU Directive's announcement in 2014, late adopters could have published a maximum of five reports covering the fiscal years between 2014 and 2018. Late adopters, which started reporting even later in 2016, were excluded from the regression due to the sampling rule mentioned above.

In addition, as described in the methodology part, individual reports were sometimes not accessible due to various problems. In some cases, for instance, the reports were removed from the corporate website, or only a web version was available. Hence, a single missing report from a late adopter could lead to the exclusion of all other reports from that company, as they could have published a maximum of five reports during that period. Consequently, the minimum number of reports per company could have been reduced to counteract this problem. However, this would have called into question the reliability of the entire dataset. For this reason, it was decided to retain the sampling rule. As a result, the small sample size of late adopters may not be truly representative of the entire group of late adopters.

Secondly, both the Fog and Flesch-Kincaid readability values appear to be above the indices' regular score range. This deviation makes it more challenging to interpret the corresponding values accurately. Both indices refer to the required years of education. Therefore, index values above twenty seem difficult to be interpreted since this number of years of education is rather unlikely. However, the Flesch Reading Ease scores are within their normal value range. This index score is based on the same textual components as the other indices. The calculation of all three indices is based on the number of syllables per word and the number of words per sentence (Li, 2008). This shows that the data on which the calculation is based has to be correct since the Flesch Reading Ease, as well as the results of all other textual characteristics, appear plausible. For this reason, the results of the Flesch-Kincaid index, as well as the Fog index, have to be valid. Therefore, a limitation of the present study is the difficulty to interpret the corresponding results because they exceed the indices' regular range. Nonetheless, conclusion can still be drawn from the differences in readability.

Thirdly, some companies, which report under the GRI framework, publish the GRI index separately and do not incorporate it into their sustainability report. Hence, this can lead to deviations in the results of the textual analysis. The GRI indices contain additional information on sustainability topics and also comprise different required KPIs. Thus, this can impact the different textual variables, such as the indicators for thematic disclosure and numeric content. Moreover, this might have implications for the approximated GRI index. The derivation process of the GRI index is based on specific search terms. These terms are included in the text as well as in the index itself. Hence, the separate publication of the GRI index could result in some topics not being recognized. Nevertheless, this only applies to less than 1% of the sample.

Fourthly, another limitation is the lack of current literature. There is only limited literature on the impact of the EU Directive on the various textual characteristics. This problem has been solved by referring to literature examining the effects of similar regulations in the UK and France. In addition, no literature deals with the differences between integrated and stand-alone reports. Most of the papers focuses on only one of these two reporting types. This problem is further enhanced by the fact that the literature presents varying definitions of the concept of integrated reporting. While some researchers define integrated reports as annual reports with a CSR section, other researchers define them as reports in accordance with the integrated reporting framework. Due to the lack of literature, the argumentation is mainly based on theories, such as the legitimacy and voluntary disclosure theory.

Regarding the reports from early and late adopters, conclusions could be drawn from literature on the introduction of mandatory IFRS reporting. In this setting, researchers have already studied the differences between early and late adopters, as well as their motivations and characteristics. However, the extent to which there are similarities and differences between sustainability and financial reporting is questionable. In addition, when mandatory IFRS reporting was introduced, textual analysis was not as advanced as it is today. As a result, researchers have often observed features other than the textual characteristics of this study.

Fifthly, the argumentation is often based on the assumption that companies, which have a superior (poor) CSR performance, design their reports differently. Hence, the causal direction goes from the CSR performance to the different characteristics. However, one could also argue that the causal direction is the other way around. For example, this study argues that companies with a superior CSR performance report under the GRI framework to differentiate themselves. Nonetheless, one could also argue that companies, which report under the GRI framework, are able to reach a higher CSR performance since the framework allows them to structurally pursue their CSR activities. Therefore, reverse causality cannot be completely ruled out.

Lastly, another limitation concerns the sample of integrated reports. While stand-alone sustainability reports were retrieved as a whole, only the sustainability sections from integrated reports were considered for this study. For this purpose, the reports had to have a clear sustainability section, which can be extracted. Integrated reports, where the CSR information was distributed throughout the entire report, were excluded from the regression. Nonetheless, even for reports with a clear sustainability section, it is still possible that companies have also reported on sustainability issues in one of the other chapters. In turn, this might have influenced the results.

9. Conclusion

More and more companies have started reporting on their sustainability issues. The underlying reasons are manifold. A milestone in Europe was the EU's announcement in 2014 that individual companies are obliged to report on their sustainability issues from 2017 onwards. Due to the high number of reporting guidelines and the weak legal framework, companies can create their sustainability reports individually. This results in a great variety of sustainability reports. For this reason, this offered an interesting and unique research opportunity.

This study examined the relationship between sustainability disclosure and the underlying performance. Sustainability disclosure was represented by numerous variables such as the length of the reports as well as the corresponding readability. It was hypothesized that longer reports, which are easily readable with a more negative tone covering many topics and having a high degree of numeric, horizon, and target content, indicate a superior sustainability performance. Moreover, it was predicted that early adopters obtain a higher CSR performance. Simultaneously, the relationship with reporting under the GRI framework and issuing a stand-alone report was unclear.

The study was conducted for all STOXX Europe 600 companies, for which the sustainability reports were available. Besides stand-alone reports, the sample also included integrated reports that were embedded in the annual report. A textual analysis was performed using the statistical program R to preserve the textual features of the numerous reports. This methodology allows to easily replicate the study with another sample, for example, from another region.

In the first part, descriptive statistics were shown. Concerning the development over time, the results indicate that the announcement of the EU Directive led to an increase in length, target orientation, and topic-specific disclosure. In addition, readability became more difficult. Furthermore, three years later, an increase in covered GRI topics could be observed when the regulation entered into force. Moreover, integrated and stand-alone reports were identified to differ significantly from each other. Stand-alone reports tend to be longer, contain more numerical data, and include more targets, while integrated reports are less readable and tend to incorporate less content. Referring to the reports from early and late adopters, the results also revealed significant differences. The reports from early adopters tend to be longer and contain more numerical, horizon, and topic-specific content.

The underlying argumentation for these differences is based on the differentiation strategy of early adopters and the differences in reporting incentives.

In the second part, a regression was performed to examine the relationship between CSR disclosure and the associated performance. The hypotheses concerning numerical content, horizon content, tone, and readability were not supported. However, the results confirmed the hypothesis that companies with a superior sustainability performance produce longer stand-alone reports with a higher degree of target orientation and a high number of GRI topics. Moreover, the results supported the hypotheses that these companies tend to be early adopters and report under the GRI framework. The prediction concerning the topic-specific disclosure was supported to a great extent except for the relation between economic content and performance. The voluntary disclosure theory can explain most of these correlations, stating that well-performing companies voluntarily disclosure sustainability information to diminish information asymmetry between the company and investors and demonstrate its extraordinary performance.

All in all, the quantitative analysis offers various insights into the field of sustainability reporting. On the one hand, it underlines that there are significant differences between the sustainability reports of companies. On the other hand, it highlights that the expression of several textual features correlates with the underlying sustainability performance. In turn, these insights provide essential implications for academics, investors, regulators, and other market participants.

References

Al-Tuwaijri, S. A., Christensen, T. E., & Hughes, K. (2004, jul). The relations among environmental disclosure, environmental performance, and economic performance: a simultaneous equations approach. *Accounting, Organizations and Society*, *29*(5-6), 447–471. doi: 10.1016/s0361-3682(03)00032-1

Arnold, M. C., & Artz, M. (2015, jan). Target difficulty, target flexibility, and firm performance: Evidence from business units' targets. *Accounting, Organizations and Society*, *40*, 61–77. doi: 10.1016/j.aos.2014.12.002

Asay, H. S., Libby, R., & Rennekamp, K. (2018, apr). Firm performance, reporting goals, and language choices in narrative disclosures. *Journal of Accounting and Economics*, *65*(2-3), 380–398. doi: 10.1016/j.jacceco.2018.02.002

Aureli, S., Medei, R., Supino, E., & Travaglini, C. (2016, jan). Sustainability disclosure after a crisis. *International Journal of Social Ecology and Sustainable Development*, *7*(1), 35–49. doi: 10.4018/ijsesd.2016010102

Ballou, B., Heitger, D. L., Landes, C. H., & Adams, M. (2006). The future of corporate sustainability reporting. *Journal of Accountancy*, *202*(6), 70–74.

Bernard, S., Abdelgadir, S., & Belkhir, L. (2015, nov). Does GRI reporting impact environmental sustainability? an industry-specific analysis of CO2 emissions performance between reporting and non-reporting companies. *Journal of Sustainable Development*, *8*(9), 190. doi: 10.5539/jsd.v8n9p190

Bhimani, A., Silvola, H., & Sivabalan, P. (2016, mar). Voluntary corporate social responsibility reporting: A study of early and late reporter motivations and outcomes. *Journal of Management Accounting Research*, *28*(2), 77–101. doi: 10.2308/jmar-51440

Bloomfield, R. J. (2002). The 'incomplete revelation hypothesis' and financial reporting. *SSRN Electronic Journal*. doi: 10.2139/ssrn.312671

BMAS. (n.d). *Standards der CSR-berichterstattung.* https://www.csr-in-deutschland.de/DE/Unternehmen/CSR-Berichterstattung/Standards/standards.html. Retrieved 2022-05-19, from https://www.csr-in-deutschland.de/DE/CSR-Allgemein/CSR-in-der-Praxis/CSR-Berichterstattung/csr-berichterstattung.html

Caglio, A., Melloni, G., & Perego, P. (2020, oct). Informational content and assurance of textual disclosures: Evidence on integrated reporting. *European Accounting Review*, *29*(1), 55–83. doi: 10.1080/09638180.2019.1677486

Chauvey, J.-N., Giordano-Spring, S., Cho, C. H., & Patten, D. M. (2015, mar). The normativity and legitimacy of CSR disclosure: Evidence from france. *Journal of Business Ethics*, *130*(4), 789–803. doi: 10.1007/s10551-014-2114-y

Cho, C. H., & Patten, D. M. (2007, oct). The role of environmental disclosures as tools of legitimacy: A research note. *Accounting, Organizations and Society*, *32*(7-8), 639–647. doi: 10.1016/j.aos.2006.09.009

Cho, C. H., Roberts, R. W., & Patten, D. M. (2010, may). The language of US corporate environmental disclosure. *Accounting, Organizations and Society*, *35*(4), 431–443. doi: 10.1016/j.aos.2009.10.002

Christensen, H. B., Lee, E., Walker, M., & Zeng, C. (2015, jan). Incentives or standards: What determines accounting quality changes around IFRS adoption? *European Accounting Review*, *24*(1), 31–61. doi: 10.1080/09638180.2015.1009144

Clarkson, P. M., Li, Y., Richardson, G. D., & Vasvari, F. P. (2008, may). Revisiting the relation between environmental performance and environmental disclosure: An empirical analysis. *Accounting, Organizations and Society*, *33*(4-5), 303–327. doi: 10.1016/j.aos.2007.05.003

Clarkson, P. M., Overell, M. B., & Chapple, L. (2011, mar). Environmental reporting and its relation to corporate environmental performance. *Abacus*, *47*(1), 27–60. doi: 10.1111/j.1467-6281.2011.00330.x

Clarkson, P. M., Ponn, J., Richardson, G. D., Rudzicz, F., Tsang, A., & Wang, J. (2020, mar). A textual analysis of US corporate social responsibility reports. *Abacus*, *56*(1), 3–34. doi: 10.1111/abac.12182

Cohen, J. R., Holder-Webb, L., & Zamora, V. L. (2015, jun). Nonfinancial information preferences of professional investors. *Behavioral Research in Accounting*, *27*(2), 127–153. doi: 10.2308/bria-51185

Corporate Finance Institute. (n.d). *What is the variance inflation factor (vif)?* https://corporatefinanceinstitute.com/resources/knowledge/other/variance-inflation-factor-vif/.

Davis, A. K., & Tama-Sweet, I. (2012, jan). Managers' use of language across alternative disclosure outlets: Earnings press releases versus MD&a∗. *Contemporary Accounting Research*, *29*(3), 804–837. doi: 10.1111/j.1911-3846.2011.01125.x

Dhaliwal, D. S., Li, O. Z., Tsang, A., & Yang, Y. G. (2011, jan). Voluntary nonfinancial disclosure and the cost of equity capital: The initiation of corporate social responsibility reporting. *The Accounting Review*, *86*(1), 59–100. doi: 10.2308/accr.00000005

Du, S., & Yu, K. (2020, apr). Do corporate social responsibility reports convey value relevant information? evidence from report readability and tone. *Journal of Business Ethics*, *172*(2), 253–274. doi: 10.1007/s10551-020-04496-3

Eccles, R. G., Krzus, M. P., & Solano, C. (2019). A comparative analysis of integrated reporting in ten countries. *SSRN Electronic Journal*. doi: 10.2139/ssrn.3345590

Eikon. (2017). *Thomson reuters esg scores*. Thomson Reuters.

Elkington, J. (1998). Partnerships fromcannibals with forks: The triple bottom line of 21st-century business. *Environmental Quality Management*, *8*(1), 37–51. doi: 10.1002/tqem.3310080106

European Commission. (n.da). *Financial reporting.* https://ec.europa.eu/info/business-economy-euro/company-reporting-and-auditing/company-reporting/financial-reporting_en.

European Commission. (n.db). *Non-financial reporting.* https://ec.europa.eu/info/business-economy-euro/company-reporting-and-auditing/company-reporting/non-financial-reporting_en.

European Union. (2014). *Directive 2014/95/eu.* https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32014L0095&from=DE.

Fiechter, P., Hitz, J.-M., & Lehmann, N. (2019). Real effects of disclosure regulation: Evidence from the european union's CSR directive. *SSRN Electronic Journal*. doi: 10.2139/ssrn.3033883

Flesch, R. (n.d). *How to write plain english.* https://web.archive.org/web/20160712094308/http://www.mang.canterbury.ac.nz/writing_guide/writing/flesch.shtml.

Franco, G. D., Hope, O.-K., Vyas, D., & Zhou, Y. (2015, jul). Analyst report readability. *Contemporary Accounting Research*, *32*(1), 76–104. doi: 10.1111/1911-3846.12062

Global Reporting Initiative. (2016a). *Gri 101: Foundation.*

Global Reporting Initiative. (2016b). *Gri 102: General disclosures.*

Global Reporting Initiative. (2016c). *Gri 103: Management approach.*

Global Reporting Initiative. (2016d). *Gri 303: Water.*

Global Reporting Initiative. (n.da). *About gri.* https://www.globalreporting.org/information/about-gri/Pages/default.aspx.

Global Reporting Initiative. (n.db). *About sustainability reporting.* https://www.globalreporting.org/information/sustainability-reporting/Pages/default.aspx.

Global Reporting Initiative. (n.dc). *Gri and sustainability reporting.* https://www.globalreporting.org/information/sustainability-reporting/Pages/gri-standards.aspx.

Global Reporting Initiative. (n.dd). *Gri's history.* https://www.globalreporting.org/information/about-gri/gri-history/Pages/GRI's%20history.aspx.

Global Reporting Initiative. (n.de). *The gri standards: The global standards for sustainability reporting.* https://www.globalreporting.org/standards/media/2458/gri_standards_brochure.pdf.

Global Reporting Initiative. (n.df). *Sustainability reporting in the european union.* https://www.globalreporting.org/information/policy/Pages/EUpolicy.aspx.

GRI Secretariat. (2016). *Introducing the gri standards.* https://www.youtube.com/watch?time_continue=2&v=RDe8Kkh1BrU&feature=emb_title.

Huang, X., Nekrasov, A., & Teoh, S. H. (2012). Headline salience and over- and underreactions to earnings. *SSRN Electronic Journal*. doi: 10.2139/ssrn.2101887

Hummel, K., & Rötzel, P. (2019, feb). Mandating the sustainability disclosure in annual reports—evidence from the united kingdom. *Schmalenbach*

Business Review, *71*(2), 205–247. doi: 10.1007/s41464-019-00069-8

Hummel, K., & Schlick, C. (2016, sep). The relationship between sustainability performance and sustainability disclosure – reconciling voluntary disclosure theory and legitimacy theory. *Journal of Accounting and Public Policy*, *35*(5), 455–476. doi: 10.1016/j.jaccpubpol.2016.06.001

Hussainey, K., & Walker, M. (2009, jan). The effects of voluntary disclosure and dividend propensity on prices leading earnings. *Accounting and Business Research*, *39*(1), 37–55. doi: 10.1080/00014788.2009.9663348

Ingram, R. W., & Frazier, K. B. (1980). Environmental performance and corporate disclosure. *Journal of Accounting Research*, *18*(2), 614. doi: 10.2307/2490597

International Integrated Reporting Council. (2013). *The international <ir> framework*. https://integratedreporting.org/wp-content/uploads/2015/03/13-12-08-THE-INTERNATIONAL-IR-FRAMEWORK-2-1.pdf.

INTOSAI. (2013). *Sustainability reporting: Concepts, frameworks and the role of supreme audit institutions*. https://www.wgea.org/media/2935/2013_wgea_sustainability-reporting.pdf.

Ioannou, I., & Serafeim, G. (2012, oct). What drives corporate social performance? the role of nation-level institutions. *Journal of International Business Studies*, *43*(9), 834–864. doi: 10.1057/jibs.2012.26

Ioannou, I., & Serafeim, G. (2017). *The consequences of mandatory corporate sustainability reporting*. Harvard Business School.

Kaplan, R. S., & Norton, D. P. (2008). Mastering the management system. *Harvard Business Review*, *86*(1), 62–77.

KPMG. (2017). *The road ahead–the kpmg survey of corporate responsibility reporting 2017*. https://home.kpmg/content/dam/kpmg/xx/pdf/2017/10/kpmg-survey-of-corporate-responsibility-reporting-2017.pdf.

Lang, M., & Stice-Lawrence, L. (2015, nov). Textual analysis and international financial reporting: Large sample evidence. *Journal of Accounting and Economics*, *60*(2-3), 110–135. doi: 10.1016/j.jacceco.2015.09.002

Latham, G. P., & Locke, E. A. (1979, sep). Goal setting—a motivational technique that works. *Organizational Dynamics*, *8*(2), 68–80. doi: 10.1016/0090-2616(79)90032-9

Leuz, C. (2010, jan). Different approaches to corporate reporting regulation: How jurisdictions differ and why. *Accounting and Business Research*, *40*(3), 229–256. doi: 10.1080/00014788.2010.9663398

Li, F. (2008, aug). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics*, *45*(2-3), 221–247. doi: 10.1016/j.jacceco.2008.02.003

Liew, W. T., Adhitya, A., & Srinivasan, R. (2014, apr). Sustainability trends in the process industries: A text mining-based analysis. *Computers in Industry*, *65*(3), 393–400. doi: 10.1016/j.compind.2014.01.004

Liu, M., Wu, K., Yang, R., & Yu, Y. (2020). Textual analysis for risk profiles from 10-k filings. *CPA Journal*, *90*(6), 36–41.

Locke, E. A., & Latham, G. P. (2002, sep). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, *57*(9), 705–717. doi: 10.1037/0003-066x.57.9.705

Loughran, T., & Mcdonald, B. (2011, jan). When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of Finance*, *66*(1), 35–65. doi: 10.1111/j.1540-6261.2010.01625.x

Loughran, T., & Mcdonald, B. (2016, jun). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, *54*(4), 1187–1230. doi: 10.1111/1475-679x.12123

Lueg, K., Lueg, R., Andersen, K., & Dancianu, V. (2016, feb). Integrated reporting with CSR practices. *Corporate Communications: An International Journal*, *21*(1), 20–35. doi: 10.1108/ccij-08-2014-0053

Lukomnik, J., Kwon, S., & Welsh, H. (2018). State of sustainability and integrated reporting 2018. In (chap. IRRC Institute. Retrieved from IRRC Institute website: https://www.weinberg.udel.edu/IIRCiResearchDocuments/2018/11/2018-SP-500-Integrated-Reporting-FINAL-November-2018-1.pdf). Retrieved from IRRC Institute website: https: IRRC Institute.

Malmi, T., & Brown, D. A. (2008, dec). Management control systems as a package—opportunities, challenges and research directions. *Management Accounting Research*, *19*(4), 287–300. doi: 10.1016/j.mar.2008.09.003

Marx, B., & Mohammadali-Haji, A. (2014, apr). Emerging trends in reporting: An analysis of integrated reporting practices by south african top 40 listed companies. *Journal of Economic and Financial Sciences*, *7*(1), 231–250. doi: 10.4102/jef.v7i1.138

Melloni, G., Caglio, A., & Perego, P. (2017, may). Saying more with less? disclosure conciseness, completeness and balance in integrated reports. *Journal of Accounting and Public Policy*, *36*(3), 220–238. doi: 10.1016/j.jaccpubpol.2017.03.001

Miller, B. P. (2010, nov). The effects of reporting complexity on small and large investor trading. *The Accounting Review*, *85*(6), 2107–2143. doi: 10.2308/accr.00000001

Mittelbach-Hoermanseder, S., Hummel, K., & Rammerstorfer, M. (2019). Information content of corporate social responsibility disclosures in europe: An institutional perspective. *SSRN Electronic Journal*. doi: 10.2139/ssrn.3365521

Muslu, V., Mutlu, S., Radhakrishnan, S., & Tsang, A. (2019, jan). Corporate social responsibility report narratives and analyst forecast accuracy. *Journal of Business Ethics*, *154*(4), 1119–1142. doi: 10.1007/s10551-016-3429-7

Muslu, V., Radhakrishnan, S., Subramanyam, K. R., & Lim, D. (2015, may). Forward-looking MD& a disclosures and the information environment. *Management Science*, *61*(5), 931–948. doi: 10.1287/mnsc.2014.1921

Nau, R. (n.d). *Regression diagnostics: Testing the assumptions of linear regression*. http://people.duke.edu/~rnau/testing.htm.

Nazari, J. A., Hrazdil, K., & Mahmoudian, F. (2017, aug). Assessing social and environmental performance through narrative complexity in CSR reports. *Journal of Contemporary Accounting & Economics*, *13*(2), 166–178. doi: 10.1016/j.jcae.2017.05.002

Neu, D., Warsame, H., & Pedwell, K. (1998, apr). Managing public impressions: Environmental disclosures in annual reports. *Accounting, Organizations and Society*, *23*(3), 265–282. doi: 10.1016/s0361-3682(97)00008-1

Patten, D. M. (2002, nov). The relation between environmental performance and environmental disclosure: a research note. *Accounting, Organizations and Society*, *27*(8), 763–773. doi: 10.1016/s0361-3682(02)00028-4

Pencle, N., & Mălăescu, I. (2016, sep). What's in the words? development and validation of a multidimensional dictionary for CSR and application using prospectuses. *Journal of Emerging Technologies in Accounting*, *13*(2), 109–127. doi: 10.2308/jeta-51615

Pizzetti, M., Gatti, L., & Seele, P. (2019, dec). Firms talk, suppliers walk: Analyzing the locus of greenwashing in the blame game and introducing 'vicarious greenwashing'. *Journal of Business Ethics*, *170*(1), 21–38. doi: 10.1007/s10551-019-04406-2

Purplemath. (n.d). *Interquartile ranges & outliers*. https://www.purplemath.com/modules/boxwhisk3.htm.

Purvis, B., Mao, Y., & Robinson, D. (2019, sep). Three pillars of sustainability: in search of conceptual origins. *Sustainability Science*, *14*(3), 681–695. doi: 10.1007/s11625-018-0627-5

PwC. (2014). *Sustainability goes mainstream: Insights into investor views*. https://www.pwc.com/us/en/pwc-investor-resource-institute/publications/assets/pwc-sustainability-goes-mainstream-investor-views.pdf.

SEC. (1998). *A plaing english handbook: How to create clear sec disclosure documents*. U.S. Securities and Exchange Commission, Washinton, DC.

Statistics Solutions. (2013). *Homoscedasticity*. https://www.statisticssolutions.com/homoscedasticity/.

Statistics Solutions. (n.d). *Assumptions of linear regression*. https://www.statisticssolutions.com/assumptions-of-linear-regression/.

Stent, W., Bradbury, M., & Hooks, J. (2013, sep). What firms' discretionary narrative disclosures reveal about the adoption of international financial reporting standards. *Australian Accounting Review*, *23*(3), 252–263. doi: 10.1111/auar.12002

STOXX. (2020). *Changes in composition of stoxx benchmark indices*. https://www.stoxx.com/press-releases-details?articleId=1846984593.

United Nations. (1987). *Report of the world commission on environment and*

development: Our common future.

Wakefield, J. (2016). *Apple, samsung and sony face child labour claims [news website]*. https://www.bbc.com/news/technology-35311456.

Wang, Z., Hsieh, T.-S., & Sarkis, J. (2018, jul). CSR performance and the readability of CSR reports: Too good to be true? *Corporate Social Responsibility and Environmental Management*, *25*(1), 66–79. doi: 10.1002/csr.1440

Wiseman, J. (1982, jan). An evaluation of environmental disclosures made in corporate annual reports. *Accounting, Organizations and Society*, *7*(1), 53–63. doi: 10.1016/0361-3682(82)90025-3

Yang, F., Dolar, B., & Mo, L. (2018, mar). Textual analysis of corporate annual disclosures: A comparison between bankrupt and non-bankrupt companies. *Journal of Emerging Technologies in Accounting*, *15*(1), 45–55. doi: 10.2308/jeta-52085

Yusof, S. M. (2018). *Social environmental disclosure between gri-sustainability reporting and iirc – integrated reporting among european companies.*