Measuring the impact of the transition to mandatory CSR reporting in Europe

Abstract

We examine the transition to mandatory corporate social responsibility (CSR) reporting by large European listed companies around Directive 2014/95/EU. The new Directive defines quality reporting principles establishing minimum topic coverage and setting the same rules for all member states. However, the Directive's framework leaves much room for managerial discretion. We conduct our study in two stages. We start with evaluating the effect of the first mandatory CSR regulation in Europe on firms' disclosure strategies, followed by the analysis of the Directive's impact on the financial market. We resort to automated textual analysis as our primary tool which allows us to capture quality dimensions explicitly targeted by the Directive. Using a difference-in-difference model where countries already having mandatory reporting rules serve as controls, we show that the Directive increased the number of reporters, the volume of disclosures and led to significant changes in various qualitative dimensions for those that did not. Some qualitative changes are positive (less biased, broader coverage of CSR-specific topics, and more long-term oriented discussions). Others are difficult to interpret before analyzing the impact on financial users. We next analyze the market's perception by combining textual attributes into an index. Overall, bid-ask spreads are negatively related to the index over the entire pre- and post-periods. The change in our index indicates a decrease in the bid-ask spread, greater transparency, but only as of 2017 the year of implementation.

Keywords— textual analysis, annual report, corporate social responsibility, international, disclosure regulation, Directive 2014/95

JEL— M14, M41, G38, Z13

1.   Introduction

Over the past years, an increasing number of companies have adopted CSR reporting practices to respond to market and social pressures concerning sustainability (KPMG, 2017). However, until 2017, CSR disclosures in Europe mainly remained a voluntary practice. The lack of regulation left much room for managerial discretion, raising questions about the credibility and usefulness of such disclosures (Cho et al., 2012; Cho et al., 2014; Michelon et al., 2015). An assessment by the European Commission (European Commission, 2013) establishes that 94% of the 42,000 large European companies did not disclose non-financial information at the time and only 36% of reporters published at least one report per three years. Concerning quality, the EU Commission concludes that companies tend to release mostly positive information omitting negative material performance. Moreover, reports are inconsistent over time and, therefore, difficult to compare. Finally, they often skip relevant topics such as human rights or corruption matters and lack external verification.

The survey conducted by the EU Commission echoes findings in the literature investigating CSR reporting practices under voluntary regimes. For example, Hummel and Schlick (2016) conclude that poor CSR performers are using CSR disclosures as a means of impression management and advocate the necessity to implement a "binding regulatory framework for contents of sustainability reports" (p.473). However, they point to inconclusive evidence concerning the effectiveness of mandatory sustainability regulation, which often results in low compliance (Larrinaga et al., 2002). Therefore, Hummel and Schlick (2016) add that there is a need to "investigate both pre-regulation adaptations of reporting behavior and post-regulation sustainability disclosure quality to determine the effectiveness of new regulatory frameworks" (p.474).

The lack of high-quality voluntary CSR disclosures is often blamed on costly implementation and the failure of market forces (e.g., rating agencies, audits, and assurances) to efficiently detect "greenwashing". Moreover, according to Christensen et al. (2019), firms do not consider aggregated benefits for the whole economy in their reporting choices. It leads to a situation where the public value of disclosed information is higher than the private one, which can only be addressed by a mandate. Therefore, regulators can favor reporting mandates "to induce more CSR" and avoid "dead-weight losses" to the economy "resulting from no or bad disclosures" (Christensen et al., 2019).

Starting from 2017, the European Commission mandates large European companies to publish CSR disclosures in annual or stand-alone CSR reports. According to the new Directive, 2014/95/EU, four topics (environmental, social, and employee matters; respect for human rights; anti-corruption; and bribery matters) become compulsory to disclose. The Directive states that its primary goal is to increase the transparency of non-financial information. The option chosen for the implementation of the Directive requires reporting on a "report or explain basis". The Directive allows companies to select the most appropriate reporting framework and does not impose an assurance requirement. The EU Commission issued principle-based guidelines as a point of reference for good reporting practice.

We conduct our study in two stages. We start with evaluating the effect of the first mandatory CSR regulation in Europe on firms' disclosure strategies, followed by the analysis of the Directive's impact on financial markets. According to Christensen

et al. (2019), there is still scarce evidence of the effectiveness of CSR mandates. However, it offers better identification strategies than voluntary regimes. Direct tracking of disclosure changes adds more certainty to the observed outcome, making attribution of results to regulation more robust.

Our first-stage results indicate that the Directive triggered a reaction from companies. It increased the number of reporters, the volume of disclosures and led to significant changes in various qualitative dimensions. However, the interpretation of the qualitative changes is twofold, complicating the overall assessment of change. Some qualitative changes are positive (less biased, broader coverage of CSR-specific topics, and more long-term oriented discussions). Others are difficult to interpret before analyzing the impact on financial users. Namely, as we will further discuss, the lower similarity with higher complexity can signify a positive as much as a negative development. As a result, it is not clear how these changes affect users' perceptions when aggregated.

In our second-stage analysis, we explore the implications of this transition on the transparency of financial markets. First, we conduct a preliminary test to validate that CSR disclosures collected from annual and stand-alone reports are taken into consideration by financial users. We confirm a significant impact of CSR disclosures on reducing the bid-ask spread - our proxy for transparency of information. We select the bid-ask spread to assess capital effects of the regulation as it is "very responsive to corporate disclosures and is probably one of the capital-market outcomes that we understand the best" (Christensen et al., 2019; Leuz & Wysocki, 2016). The link occurs when CSR disclosures concentrate in one report (either annual or a stand-alone) or are mixed in both.

Further, we analyze the relationship between individual textual properties such as topic coverage, sentiment bias, readability, similarity and length, and bid-ask spreads. Our results indicate that better transparency is associated with: lower wordiness, captured by the number of sentences in disclosure; broader coverage of specific topics such as environmental, social, bribery, human rights, and long term orientation; lower complexity captured by readability indexes; lower bias captured as the absolute difference between positive and negative words; and higher similarity from year to year.

Since the interpretation of individual parameters is insufficient to assess the overall impact of changes, we combine all textual characteristics in one score and examine its relationship with the bid-ask spread. We analyze the perception of textual attributes by users in aggregate, as some improvements can overweight quality degradation or vice versa. A combined score is highly significantly associated with a bid-ask spread with one standard deviation increase in score associated with a 8.9% decline in the bid-ask spread, which is considered economically significant. Significance also persists after controlling for CSR performance with the Asset4 rating. Our results indicate the incremental usefulness of lexical properties for market participants beyond CSR ratings, supporting our choice of direct textual analysis over non-transparent rating scores to capture CSR disclosures.

Finally, we analyze how the shift in textual attributes initiated by the transition to mandatory reporting translates into economic consequences. We only consider those companies that published reports in both periods. Going from zero reports to having a disclosure will mechanically inflate our combined score without necessarily indicating quality improvement. Our results suggest that better reporting significantly improves

bid-ask spread only from 2017 - the year of the Directive's implementation.

We contribute to the literature in the following ways. First, most of the prior studies apply a limited set of linguistic properties. These studies mainly focus on stand-alone CSR reports or integrated reports as isolating CSR topics in annual reports was challenging. Focusing on stand-alone CSR reports leaves aside many disclosures made in annual reports, which can bias the full picture as companies sometimes switch from one source to another over the years. Further, studies concentrating on integrated reports do not isolate CSR disclosures, impeding proper CSR narrative analysis. We also notice a gap in the number of textual attributes analyzed, where we mostly find such metrics as readability and sentiment. Second, we observe that none manages to isolate CSR discussions while controlling for a whole set of fixed effects relevant for a longitudinal large-scale study (year, industry, country fixed effects) and other required controls.

Third, our review reveals that papers mostly cover US companies (Jegadeesh & Wu, 2013; Lawrence, 2013; Loughran & McDonald, 2011, 2015; Nazari et al., 2017). However, the European context of the Directive also presents an exciting area of research, particularly for CSR, as the European economic space consists of various institutional settings and is more stakeholder-oriented. Consumers, employees, governments, and communities in stakeholder-oriented countries are more concerned about firms conforming to legal standards. They can influence firms' operational decisions, leading to a higher CSR disclosure level (Dhaliwal et al., 2012; Hummel et al., 2018).

Fourth, capturing the capital market benefits of increased disclosures with alternative measures using natural language processing techniques was urged by several academics (Beyer et al., 2010; Core, 2001). According to Beyer et al. (2010), currently used "proxies for disclosures costs are correlated with capital market benefits of increased disclosures, making it difficult to infer how increased disclosures affect capital markets benefits" (p.314). Our study contributes to this strand of literature by covering the entire set of CSR disclosures in a European setting with a broad set of textual attributes based on guidelines during the transition from voluntary to mandatory reporting.

This study also contributes to the literature in many other ways. First, it analyzes regulatory impact in two-stages: from isolating exact changes initiated by the mandate to establishing an economic link between these changes and financial users. Second, it validates the textual channel to be informative and representative of such an impact. Third, it adds to the literature focusing on corporate reporting's textual properties and their role in financial users' decision making. Fourth, we validate the importance of including annual reports in such analysis, which contain an essential share of CSR disclosures. This study opens new opportunities for future research focusing on CSR disclosures and CSR in general by suggesting a direct way of measuring this phenomenon.

Our results may be of interest to different parties. It provides evidence for regulators regarding the efficiency of costly change. It indicates to companies which disclosure attributes are essential for users and how to shape their reports better. Finally, it offers direct users of this information new tools for assembling and analyzing CSR disclosures by their proper means and the importance of including these factors in their decision making.

The rest of the paper is organized as follows. First, we present a review of the

literature justifying our research motivation. Next, we describe the institutional setting and theoretical background with hypotheses development, followed by textual channel's operationalization. Then we describe each stage's research design, sampling, and descriptive statistics. Then we present our main findings and conclusions.

2. Literature review

The effect of the transition to compulsory CSR reporting was explored before for pioneers of mandatory regimes such as China, Denmark, Malaysia, and South Africa (Ioannou & Serafeim, 2017). Ioannou and Serafeim (2017) examine the impact of the transition to compulsory reporting on Bloomberg Disclosure score and its association with firm valuation. They confirm a significant increase in reporting scores following regulation and a positive impact on Tobin's Q. However, they argue that "it could be the case that sustainability disclosure regulations affect firms differently in other countries where institutions vary in terms of the extent to which organizations already make ESG disclosures and in terms of how they compete on the basis of their ESG performance" (p.31).

Lack of generalizability of results for countries with weakly enforced regulations onto countries with more substantial enforcement also emerges from a series of other studies (Cahan et al., 2016; De Villiers & Marques, 2016; Hummel et al., 2018; Mittelbach-Hörmanseder et al., 2020). For example, Cahan et al. (2016) show that CSR disclosures are more value relevant in countries with weak institutions, while De Villiers and Marques (2016) find more substantial value relevance in countries with sound governance mechanisms.

The first studies analyzing the reaction to the Directive lead to mixed conclusions. On the one hand, the Directive initiated an increase in CSR activities (Fiechter et al., 2018). On the other hand, the Directive provoked an adverse market reaction to its adoption (Grewal et al., 2019), indicating that investors see this change as rather costly.

Regarding literature in textual analysis, a large body of literature provides evidence of the economic importance of financial reporting's lexical properties captured with automated textual analysis (Bonsall IV et al., 2017; Huang et al., 2014; Lang & Stice-Lawrence, 2015; Li, 2008, 2010a; Loughran & McDonald, 2014, 2015, 2016).

Thanks to highly regulated reporting in the US, US firms' financial disclosures serve as a good starting point for automated analysis of corporate disclosures' textual properties. Li (2008) publishes one of the first studies investigating the role of lexical features of financial texts. Using a readability metric of 10-K reports, he demonstrates that corporate narratives could be opportunistically structured as poor financial performers appear to have more obfuscating disclosures. Right after that, studies examining corporate reporting through textual analysis enlarged a set of relevant textual properties by starting to explore vocabulary-based sentiment dimensions.

First studies examining the role of optimistic and pessimistic language confirmed its value-relevance (Davis et al., 2012; Loughran & McDonald, 2011). A positive tone associates with positive investor reactions and contains incremental information content in predicting future performance (Kearney & Liu, 2014). However, investors appear to react more strongly to negative text (Huang et al., 2014), recognizing firms' opportunistic behavior. Beyond corporate filings, the tone provides investors with

additional information through other channels, such as earnings releases, conference calls, or analyst reports (Huang et al., 2014; Lehavy et al., 2011).

Moving forward from readability and tone attributes, researchers applied additional natural language processing techniques. They capture comparability using cosine similarity (Brown & Tucker, 2011; Lang & Stice-Lawrence, 2015), identify topics using Latent Dirichlet Allocation (Ball et al., 2013) or explore information content using Naive Bayes classification (Li, 2010b).

The US-setting allows the examination of different parts and types of financial disclosures (e.g., MD&A, press releases, earnings announcements) due to easy access and a standardized structure. The first studies in non-US settings appeared later and could only process entire annual reports without segmenting them into subtopics of interest (Lang & Stice-Lawrence, 2015). To the best of our knowledge, the first attempt to isolate topics and analyze their textual properties using automated textual treatment was made by Athanasakou et al. (2019), who specifically focus on strategy discussions in annual reports.

Overall, papers covering textual attributes in financial context demonstrate that textual disclosures are informative for financial users (Brown & Tucker, 2011; Li, 2010a; Loughran & McDonald, 2016). Although, studies of linguistic properties in the CSR context remain scarce.

Prior research suggests that CSR disclosures are value relevant for financial users (see Appendix 5), yet they are different from financial disclosures on many levels, motivating separate investigation of their linguistic properties. According to Ioannou and Serafeim (2017) and Dhaliwal et al. (2012), the main distinguishing characteristic of CSR reporting consists of a broader set of stakeholders. Another difference between CSR and financial disclosures includes their verifiability and credibility, given high levels of discretion inherent to CSR disclosures even after the transition to mandatory reporting (Radhakrishnan et al., 2018). Christensen et al. (2013) also complement this discussion by pointing towards a broader diversity of topics, measurements, and activities. Since these differences may influence the interpretation of textual attributes, it requires validation in the CSR context. Moreover, studying CSR reporting's linguistic properties is a promising area for identifying impression management and capturing useful features of disclosures for financial users. Finally, according to Hummel et al. (2018), studying CSR disclosures with natural language processing techniques has the "potential to overcome some of the major limitations in empirical CSR disclosure research, in particular problems surrounding replicability and inter-coder reliability of manual CSR disclosure measurements and lack of generalizability due to small sample sizes" (p.18).

Among the first studies, we find the one by Cho et al. (2010) based on US 10-K environmental disclosures, which states that worse environmental performers use language and verbal tone to influence their public image. Some older studies also explored CSR disclosures using automatic language processing but mainly focused on quantity rather than more sophisticated qualitative linguistic properties. For example, Cormier et al. (2011) find that poor environmental performers are verbose, and web-based publications are not relevant for financial users. More recent publications find value-relevance of some qualitative linguistic properties in CSR reports such as readability, tone, and several topics. Under this strand, Melloni et al. (2017) provide evidence that poor social performers have less readable reports. Nazari et al. (2017) suggest that more readable stand-alone CSR reports increase transparency regarding

firms' social and environmental performance while using less-readable language in CSR reports increases obfuscation. They also add that those complexity indices used as obfuscation measures in prior finance and accounting research can help shareholders, financial analysts, and investors determine CSR disclosure's credibility.

The most comprehensive analysis of textual properties is made by Muslu et al. (2019). They construct a CSR disclosure score and show that it is positively associated with analyst forecast accuracy [1].

Several studies analyze CSR topics' value-relevance in different countries, including Europe (Hummel et al., 2018; Mittelbach-Hörmanseder et al., 2020) and confirm that CSR disclosures are affected by the institutional environment. Current studies measuring CSR disclosures with natural language processing techniques, however, appear to be incomplete. For example, research by Mittelbach-Hörmanseder et al. (2020) omits CSR disclosures published in stand-alone CSR reports in its analysis of value-relevance of Directive's mandated topics. Moreover, their proposed methodology for the extraction of CSR disclosures from annual reports does not validate their selection criteria and risks CSR topics' misspecification. Based on the methods of Mittelbach-Hörmanseder et al. (2020), we can only observe the amount of CSR specific words in selected documents. Still, their occurrences do not necessarily appear in the CSR section. For example, counting such a term as "environment" in the whole annual report can also capture discussions of the economic environment in the financial part resulting in erroneous attribution of these counts to CSR and, therefore, inaccurate specification of CSR topics.

Others study linguistic properties of integrated reports and confirm incremental informativeness of readability, length, and bias beyond firms' fundamentals and CSR scores (Caglio et al., 2020). They explore integrated reporting as a proxy for CSR disclosures. However, without isolating CSR discussions from the financial information, it is difficult to conclude the role of CSR information since these findings reflect the aggregated impact of financial and CSR information published together. Finally, relevant studies mostly cover voluntary regimes, except for Mittelbach-Hörmanseder et al. (2020).

3.   Institutional setting

Among the first attempts to regulate CSR reporting by European companies was Account Modernization Directive 2003/51/EC. It required some reporting on environmental and employee affairs. Then followed Accounting Directive 2013/34/EU, which imposed a "corporate governance statement" in companies' management reports. Neither of the existing acts covering non-financial reporting had standards or guidelines.

On a member state level, a few European countries introduced mandatory disclosure rules before 2015. Among the most prominent adopters of mandatory regimes were Scandinavian countries (from 2009), France (from 2001), and the United Kingdom (2006). Below we describe the timeline and scope of concerned companies for corresponding countries.

1. Denmark 2009: state-owned companies and companies with total assets of more than EUR 19 million, revenues more than EUR 38 million, and more

---

[1]The score proposed by Muslu et al. (2019) is a linear combination of tone, readability, length, numerical, and horizon content of stand-alone CSR reports

than 250 employees (stricter conditions than in Directive), must report on their responsibility to society (CSR) and are encouraged to do so using GRI Sustainability Reporting Guidelines;

2. UK 2006: British Companies Act mandates listed on London Stock Exchange companies to disclose, in their annual Business Review, information on the environmental, workplace, social and community matters "to the extent that they are important to understanding company's business";

3. Sweden 2009: all state-owned companies required to produce an annual sustainability report following GRI G3 guidelines;

4. France 2012: Grenelle II Act requiring all companies with over 500 employees to issue a yearly "social and environmental report."

The rest of the European states regarded CSR as a purely voluntary practice up to the new Directive.

New Directive 2014/95/EU determines the content of disclosed information and imposes a framework with a more structured presentation and specific requirements applicable to all European countries. The Directive sets a benchmark for its transposition in member state legislation. According to this regulatory act, organizations must disclose if they cumulatively have: 1) an average number of over 500 employees during the financial year and 2) total assets exceeding EUR 20 million or a net turnover exceeding EUR 40 million.

The covered topics must at minimum consist of human rights, environmental, social, employee, anti-corruption, and bribery matters. Member states are entitled to self-determine a set of undertakings falling under the Directive, penalties for non-compliance, and audit requirements. We present a more detailed comparison of member state requirements vs. the benchmark determined by the Directive in Table 1. On the whole, we can see that member states either maintain the same conditions or apply a stricter framework.

(**Table 1 about here**)

According to the Directive, member states mostly mandate auditors to verify disclosures' existence in annual or stand-alone reports. Nonetheless, several countries include content checks in the audit procedure. However, difficulties in measuring CSR impede the development of proper frameworks to realize such audits and question their efficiency.

Concerning legal enforcement, article 33.1 of Directive stipulates that members of administrative, management and supervisory bodies have collective responsibility for management report, including both annual and stand-alone CSR reports. Consequences of non-compliance differ across countries and vary from purely administrative and monetary penalties to imprisonment. Further analysis of financial penalties reveals that they are insignificant for large corporations and that imprisonment measures are in place only in countries with a small number of affected companies. Our analysis indicates that the Directive's setting lacks vigorous enforcement defined by Christensen et al. (2019) as "the product of likelihood that non-compliance is detected and the magnitude of the imposed penalties for non-compliance" (p.23).

4.   Theoretical background and hypotheses development

First-stage: companies reaction

First, we assess the Directive's effectiveness by measuring its impact on the quantity and quality of reporting. It includes capturing change in the number of reporters, the volume of disclosures, and the qualitative characteristics of texts.

The fundamental mechanism that would drive such change is the Directive's impact on the cost-benefit tradeoff regarding firms' decision to report on CSR. Such tradeoff depends on two conditions formally stated by signaling theory. According to Zerbini (2017), companies signal the "true type" if: 1) the returns of signaling "true type" are greater than its costs or 2) the costs of falsification are greater than returns. If either of the conditions is not satisfied, a company chooses to falsify or avoid disclosing information.

Let's examine factors that may or may not lead to change following the new Directive, noting that there is no stringent enforcement mechanism. First, mandates serve as a commitment tool for the companies (Christensen et al., 2019), which improves comparability for the users. However, the question arises concerning the drivers of a company's engagement to CSR reporting. The Directive clearly states the scope of companies affected by CSR issues. It also defines topics, which clarifies the definition of CSR and simplifies benchmarking. Moreover, the set of mandatory items includes CSR issues relevant to most large corporations. As a result, a clear scope and definition improve stakeholders' understanding and awareness about CSR, making falsification more easily detectable or "costly" and leaving less room to justify non-compliance for the reporters.

Another source of costs for non-disclosure can arise from the threat of future stricter enforcement in the absence of reaction to this mandate initiated by the highest European institution. The first indication of such a risk is the European Commission's obligation to assess the Directive's implementation within four years and develop further recommendations for the legislator. Therefore, the Directive is a "first step to provide insights about whether a more stringent EU approach is needed" (Aureli et al., 2018). According to Leuz and Verrecchia (2000), if companies can commit to a quality level reporting, the regulator is less inclined to intervene. Finally, the new mandate can also increase the benefits of disclosing CSR by reducing credibility issues proper to a voluntary regime and widespread CSR perception as greenwashing.

On the opposite side of the cost-benefit tradeoff lie increased costs of disclosing the information. First, to initiate such reporting, companies incur direct costs of preparing CSR disclosures and creating procedures to implement CSR activities. A recent study by Fiechter et al. (2018) shows a significant increase in CSR expenditures by companies falling under the Directive. It can also increase proprietary costs. Finally, companies risk provoking adverse reactions by disclosing information that could have been relevant to their stakeholders before a mandatory regime but was not communicated.

The EU Commission in its pre-Directive analysis had already noted that 1) a significant number of companies did not disclose CSR information; 2) those that disclosed information did so selectively; and 3) the quality of the disclosure was questionable. The literature broadly supports the fact that companies are generally selective about their CSR disclosures (Kozlowski et al., 2015). Michelon et al. (2015)

also document the low quality of voluntary CSR disclosures before the Directive. Given the cost-benefit tradeoffs do not guarantee a change following the Directive, we state the three following hypotheses in null form:

HYPOTHESIS 1. *The Directive will not affect the number of CSR reporters.*

HYPOTHESIS 2. *The Directive will not affect the volume of CSR disclosures.*

HYPOTHESIS 3. *The Directive will not lead to the change of qualitative characteristics of reports.*

Overall, the voluntary regime's failure to provide investors with quality CSR disclosures was mostly due to the low costs of falsification. The new Directive provides a unique setting that could make the falsification costly despite the lack of formal enforcement. We summarize new costs and benefits related to the Directive in the table below. However, we can only assess them qualitatively as there is not enough information to measure the costs and benefits directly, and " the new effects of a CSR mandate are not a priori obvious" (Christensen et al., 2019).

Reaction if		No reaction if
"Doing nothing" is costly	Reaction is beneficial	Change is costly

↑ threat of stricter regulation
↑ falsification costs
  • ↑ comparability
  • ↑ benchmarking
  • ↑ awareness about CSR
	↑ credibility of CSR	↑ direct reporting costs
		↑ proprietary costs
		  • ↑ stakeholder's scrutiny
		  • ↓ attractiveness
		  • ↑ exposure of selective disclosure

Second-stage: economic consequences

Several arguments motivate our choice of the bid-ask spread to capture the economic consequences of mandatory reporting. First, the bid-ask spread is a traditional proxy of disclosure transparency. According to the Directive (Aureli et al., 2018), transparency is defined by the "sharing of a certain quantity and quality of information" (p.6). The Directive links transparency with quantity, comparability, accessibility, accuracy, and reliability of disclosed information.

Lee and Yahn (1997) present a detailed explanation of the link between bid-ask spread and transparency of accounting information. More specifically, they show that the bid-ask spread includes three types of costs incurred by professional dealers: order-processing costs, inventory costs, and adverse-selection costs. The first two elements occur due to the costs of arranging transactions and keeping an inventory

of stock. The third is associated with the informational environment[2]. Dealers set selling (ask) and buying (bid) prices based on their estimation of information asymmetries between them and other investors. To avoid losing to more informed investors, dealers set higher ask and lower bid prices to compensate these losses by a gain of negotiating with uninformed investors. Costly private information search increases the gap between informed and unaware investors resulting in higher bid-ask spreads. Verrecchia (1982) concludes that increased public information alleviates these asymmetries. Based on the literature review, Lee and Yahn (1997) first find that accounting disclosures reduce the adverse selection component of the bid-ask spread. Referring to Greenstein and Sami (1994), they argue that "by improving quality and content of accounting information, accountants may be able to help improve transactional efficiency of stock markets" (p.55).

According to Dhaliwal et al. (2011), more transparent disclosures mitigate informational asymmetries if they are value-relevant. An increasing amount of studies find value - relevance of CSR disclosures (see Appendix 5). CSR disclosures may affect firm value through different channels. For example, firms initiating CSR activities' disclosure attract dedicated institutional investors, raise more equity capital, and attract greater analyst coverage. CSR disclosures may also change investors' perceptions of firm value, help build political connections, or mitigate negative CSR events and provide legitimacy (Cahan et al., 2016).

Several studies explore the link between CSR disclosures and bid-ask spreads (Caglio et al., 2020; Cho et al., 2013; Cormier et al., 2011; Cui et al., 2018; Egginton & McBrayer, 2019). Although they suggest a relationship between CSR disclosure transparency and equity market liquidity, they have some drawbacks that could bias and limit our understanding of this link. For example, studies exploring the value-relevance of CSR disclosures mostly use CSR scores provided by rating agencies. Analysis of such ratings reveals several inconsistencies and questions their quality (Bouten et al., 2018).

To summarize, the main arguments justifying our choice of bid-ask spreads as our dependent variable are the following: 1) theoretical link with a better informational environment; 2) empirical indications of the existence of a relationship between bid-ask spreads and CSR disclosures; 3) an explicit statement of transparency improvement as the main objective of the Directive declared in the first paragraph of the Directive. [3]

Preliminary hypotheses

Before answering our research questions, we focus on a primary concern about CSR disclosures' information content. According to Li (2008), studies analyzing

---

[2]Following Christensen et al. (2019), we define a better informational environment by lower costs of information search and processing, which is an expected outcome of the standardized reporting as it "should make it easier and less costly for investors to acquire and interpret CSR information." (p.41)

[3]"In its communication entitled 'Single Market Act — Twelve levers to boost growth and strengthen confidence — "Working together to create new growth"', adopted on 13 April 2011, the Commission identified the need to raise to a similarly high level across all member states the transparency of the social and environmental information provided by undertakings in all sectors. This is fully consistent with the possibility for member states to require, as appropriate, further improvements to the transparency of undertakings' non-financial information, which is by its nature a continuous endeavour."(European Commission, 2014)

textual properties should first examine the informativeness of their sources. On one side, there is prior evidence suggesting that CSR reports are informative (Dhaliwal et al., 2012).On the other hand, greenwashing practices could have changed its perception over time. Our null hypothesis for this preliminary test is formulated as follows:

HYPOTHESIS 4a. *CSR disclosure is not linked to the bid-ask spread.*

The second part of preliminary tests consists of establishing informativeness of different sources of CSR disclosures (annual report, stand-alone report, or mixed) as it has been argued that different communication devices have different relevance for market participants (Cormier et al., 2011). Discussion by Mittelbach-Hörmanseder et al. (2020) indicates the irrelevance of stand-alone CSR reports, suggesting that one may omit these observations. However, their evidence is contradictory to studies that only focused on CSR stand-alone reports and confirmed their informativeness (Dhaliwal et al., 2012). Our second null hypothesis of this preliminary analysis is formulated as follows:

HYPOTHESIS 4b. *The source of CSR reporting does not influence the link with the bid-ask spread.*

Main hypotheses

To answer our research questions, we first establish an empirical link between our textual attributes and bid-ask spreads. We rely on existing literature to justify our expectations of associations between individual textual parameters and better transparency. Even though empirical evidence indicates the existence of such a link, it is mostly limited to length, readability, and tone (Caglio et al., 2020; Cormier et al., 2011). However, prior literature often suffers from incomplete settings (lack of proper control variables, unaccounted fixed effects, missing observations), which we remedy.

Style attributes change user beliefs independent of content (Muslu et al., 2019). Therefore, such characteristics as tone, readability, document length can be used strategically to manage user perception of the firm.

Regarding tone, managers have opportunistic incentives to disclose more positive aspects, which seems to be the case in the CSR context (Cho et al., 2010). Consistent with this literature and Directive associating good quality CSR reporting with the unbiased presentation, we expect reports with a lower bias to be more transparent. On the other hand, sentiment can be informative about underlying performance. In that case, its absence would decrease transparency.

Concerning readability, since less readable language can be a way of hiding poor performance (Li, 2008), we expect better readability to be associated with higher transparency. Alternatively, given the specificity of CSR disclosure, simple language can be perceived as "cheap talk," which would have an opposite impact on transparency.

According to Loughran and McDonald (2014), another way of obfuscating the reader is by using lengthier reports, based on which we expect more concise reporting to be associated with higher transparency. However, it is possible that longer statements only reflect the complexity of the underlying business (Li, 2008), suggesting

an inverse relationship. The long-term orientation of CSR information is likely to be informative as it reflects information about future trends and targets (Muslu et al., 2019). On the other hand, discussing future projects rather than concentrating on the accomplishment of current objectives could be less informative for investors.

Concerning environmental matters, prior studies find a positive relationship between their coverage and firm value (Clarkson et al., 2013; Cormier et al., 2011). Rare studies focusing on social aspects (Richardson & Welker, 2001) find a negative effect on firm value. To the best of our knowledge, the only study analyzing disclosures on human rights and bribery matters was conducted by Mittelbach-Hörmanseder et al. (2020). They find that these topics are value relevant, but sign depends on the observed period with a positive relationship before Directive and negative afterwords.

Finally, following Lang and Stice-Lawrence (2015), we associate similarity with better transparency as it is positively related to greater liquidity. Here again, the direction can be opposite as the similarity metric can also capture uninformative boilerplate disclosure.

Given the scarcity of prior research on textual attributes in CSR context, this investigation is, to some extend, descriptive. However, we link our textual attributes to outcomes controlling for a broad set of fixed effects and various control variables, including the CSR performance score provided by Asset4. As expectations for our textual attributes are twofold, we formulate our hypotheses in a null form.

> HYPOTHESIS 5a. *Reports with higher topics coverage, more long-term orientation, lower bias, lower complexity, and better comparability do not affect the bid-ask spread.*

To test the overall perception of individual changes of textual attributes, we aggregate all characteristics in one disclosure index and analyze its relationship with the bid-ask spread, which brings us to the next null hypothesis.

> HYPOTHESIS 5b. *Aggregated CSR disclosure score does not affect the bid-ask spread.*

Commitment to enhanced reporting should improve the informational environment by reducing the adverse - selection component of the bid-ask spread (Leuz & Verrecchia, 2000). However, this outcome can only be observed in the market if it is not saturated with information. If the transition to mandatory regime results in the spread of uninformative boilerplate disclosures, it will also undermine the effectiveness of the Directive. This leads to our second null hypothesis.

> HYPOTHESIS 6. *Change in reporting initiated by Directive will not affect the bid-ask spread.*

5. Measuring CSRD quality

The guidelines clarify the EU's understanding of non-financial reporting quality. It relies on six principles suggesting disclosing information, which is:

1. Material

2. Fair, balanced and understandable

3. Comprehensive but concise

4. Strategic and forward-looking

5. Stakeholder orientated

6. Consistent over time and coherent

We focus on principles assessable with NLP techniques. Concerning materiality, article one of the Directive stipulates that information should be provided "to the extent necessary for an understanding of [...] impact of (company's) activity" (European Commission, 2014). It is unclear from who's perspective it needs to be "necessary for an understanding" (Jeffwitz & Gregor, 2017). Difficulties for understanding and capturing the materiality principle are not uncommon and received attention in prior literature (Christensen et al., 2019).

The same lack of clarity impedes the analysis of the fifth principle. The Guidelines comment that "companies should provide relevant, useful information on their engagement with relevant stakeholders, and how their information needs are taken into account," leaving aside the definition of relevant stakeholders. Thus, due to an unclear meaning and vague extent of information, we omit two above-enumerated principles: materiality and stakeholder orientation of disclosures.

Guidelines reiterate numerous times that reporting should avoid the use of boilerplates: "non-financial statement is also expected to be concise and avoid immaterial information. Disclosing immaterial information may make the non-financial statement less easy to understand since it would obscure material information. Generic or boilerplate information that is not material should be avoided" (p.8). In the Directive's context, one way to ensure conciseness is by avoiding boilerplates. Lang and Stice-Lawrence (2015) define boilerplate as an uninformative standardized disclosure. They use four-word phrases, which are extremely common among documents to identify boilerplates. They exclude sentences with common tetragrams (combinations of four words) occurring from 30% to 75% of documents. The main assumption is that duplication of other firms' disclosures is unlikely to communicate firm-specific information. Using this methodology, we identify a set of boilerplates occurring in our set of texts. All extracted tetragrams were not related to CSR. Thus, we did not further proceed with the computation of this attribute and excluded boilerplates from further analysis.

The final property, which is not measurable with NLP techniques, is coherence. Going back to the original meaning of the word "coherence", we can see that it originates from Latin, meaning "to stick together." This definition implies a logical and ordered relation of words. Given that we implement a "bag of words" approach that disregards grammar and word order, our methods do not allow us to assess this property.

For the rest of the principles, we provide our reasoning for NLP proxy choices summarized below:

	Guidelines principles	Corresponding NLP tools
1	Fair, balanced	Tone
2	Understandable	Readability
3	Comprehensive	Report length, topics coverage
4	Strategic and forward-looking	Horizon and strategic vocabulary
5	Consistent over time	Similarity of texts over time

Fair, balanced (Tone)

Guidelines state that "non-financial statement should give fair consideration to favourable and unfavourable aspects, and information should be assessed and presented in an unbiased way" (p.7). Prior literature has shown that tone changes users' appreciation of financial narratives independently of content. Tone can either signal future performance (Li, 2010b) or be a result of strategic manipulations of users' perceptions (Huang et al., 2014). In a CSR context, disclosures can be optimistically biased to manage stakeholders' impressions (Cho et al., 2010). We evaluate the tone of CSR disclosures and analyze the impact of the Directive on this attribute.

We rely on a dictionary-based approach using classification provided by Loughran and McDonald (2015) as it is explicitly developed for economic context. Following Loughran and McDonald (2011), we multiply term frequency by inverse document frequency to attribute more weight to unusual words. According to Loughran and McDonald (2011) term weighting alleviates the problem of attributing too much importance to frequent words as "word occurring 10 times more frequently is most likely not 10 times more informative" (p.18). 'Term frequency-inverse document frequency' score (TF-IDF score) computation has the following formula:

$$TfIDfscore = n_{ij} / \sum_j n_{ij} \times \log N/df_i \tag{1}$$

where:

n_{ij} - number of occurrences of term i in document j

$\sum_j n_{ij}$ - document length

N - total number of documents

df_i - number of documents containing term i

We apply TF-IDF weighting using "TfidfVectorizer" from the "sklearn" Python module.

Our final parameter of interest is bias computed as the difference between corresponding TF-IDF scores of positive and negative words. Since users perceive negative CSR disclosures as more credible than positive CSR narratives, we will interpret a decrease of bias as a positive change.

According to Loughran and McDonald (2016), the main caveat of the dictionary approach is the negligence of negations. This problem mainly occurs when counting positive words: "negative words seem unambiguous – rarely does management negate a negative word to make a positive statement" (p.35). Lack of a pre-trained database for financial reporting makes Naive Bayes approach non-applicable when dealing with sophisticated negations. We are only limited to the straightforward detection of negations surrounding positive words. Namely, if negations ("not",

"no," "none," "nothing," "nowhere") precede a positive term, we count it as negative. However, our analysis demonstrates that the negations in sentiment analysis do not significantly change the values. Therefore, we pursue our study using tone counts without correction for negations.

Understandable (Readability)

The guidelines (European Commission, 2017) suggest that "information may also be made more understandable by using plain language and consistent terminology." (p.7)). The textual analysis captures understandability with readability indexes. The readability of documents is associated with timely information processing (You & Zhang, 2009). There is evidence that managers use more complex language to obfuscate users in case of poor performance (Ajina et al., 2016; Li, 2008). Impression management through language obfuscation was also confirmed in CSR context (Nazari et al., 2017). The most common metric is FOG index (Li, 2008; Miller, 2010) calculated as follows:

$$FOG = (words/sentences + \%complex\ words) \times 0.4 \tag{2}$$

Some argue (Loughran & McDonald, 2014) that this metric is not appropriate to assess financial texts known for their sophisticated vocabulary while still well understood among financial users. CSR texts address a broader set of users, making use of the FOG index more justified in this context. However, we also compute readability with other indexes, such as SMOG or Flesch-Kincaid grade level (FLE). They use the same components as FOG, but instead of binary classification of complex words (words with more than three syllables), relies on an explicit count of syllables:

$$FLE = (11.8 \times syllables/words) + (0.39 \times words/sentences) - 15.59 \tag{3}$$

Finally, the SMOG index was developed as a more accurate substitute for the FOG index. The corresponding formula of SMOG index (SMG) is the following:

$$SMOG = 1.0430\sqrt{\%complex\ words \times 30/number\ of\ sentences} + 3.1291 \tag{4}$$

We use the "textstat" Python module for the computation of readability indexes. The resulting values correspond to grades of education required to understand the text. The level of complexity increases with an increasing level of education.

Increased complexity can be associated with more obfuscation or a more complex underlying business, requiring further controls to avoid mixing two interpretations.

Comprehensive (Report length, Topic coverage)

Guidelines refer to the comprehensiveness of texts as the breadth of information disclosed (i.e., coverage and dispersion of different topics, (Beretta & Bozzolan, 2008)). We capture this attribute with report length and topic coverage.

Report length can have a twofold interpretation. The provision of more information is likely to signify higher transparency but can also go along with increased disclosure complexity due to information processing costs (Li, 2008). The ambiguous

nature of this parameter requires additional information to be associated with quality. We follow Lang and Stice-Lawrence (2015) and measure report length as a natural logarithm of words contained in the annual report. To split our texts into words, we apply the "nltk" Python module.

Topic coverage is evaluated based on word count using lexicons corresponding to each topic, as presented in Appendix 2. Similar to sentiment analysis, we apply TF-IDF weighting using "TfidfVectorizer" from the "sklearn" Python module. A significant increase in the length of CSR discussions and topic coverage will indicate that companies are concerned by CSR matters but were not disclosing this information before the Directive. In contrast, the addition of small boilerplates stating that companies are not involved will result in an insignificant change.

Strategic and Forward-looking (Horizon and Strategic vocabulary)

The suggestion to present strategic and forward-looking information echoes recent studies' findings, which confirm that horizon content increases informativeness of CSR disclosures and associates with better quality disclosures (Muslu et al., 2019). Brochet et al. (2015) analyzes the time horizon of conference calls' narratives and reveals that managers who emphasize short-term demonstrate opportunistic behavior at the expense of long-term performance. Using lists of horizon terms developed in previous studies (see Appendix 2), we assess the fourth principle of CSR reporting. Similar to other vocabulary-based variables, we apply TF-IDF weighting using "TfidfVectorizer" from the "sklearn" Python module. We interpret a broader use of horizon and strategic lexicon after Directive as a positive change.

Consistent over time (Similarity of texts over time)

The final principle states: "content of the non-financial report should be consistent over time. This enables users of information to understand and compare past and present changes in a company's development" (p.9). We analyze consistency over time using cosine similarity, which measures the closeness of two sentences in terms of the word content and word position.

According to Brown and Tucker (2011), the raw measure of cosine similarity is mechanically related to document length[4]: "longer a pair of documents, more probable a word is included in both documents, leading to a lower likelihood that document will differ" (p.317). We remove this mechanical relationship using Brown and Tucker (2011) methodology. Using Python "sklearn" module, we compute raw scores of cosine similarity that we regress on the first five polynomials of length. Values before adjustments range from zero to one, where higher parameters are associated with higher similarity. Final cosine is lower and can be negative as we remove the fitted score from raw metrics.

More similarity between the same company's texts from year to year could either signify better comparability, which simplifies information processing or boilerplate disclosures, which do not convey specific information about CSR activities.

Given the ambiguous interpretation of similarity, we do not associate it with one-directional expectation but mostly keep it to capture the impact of the Directive on this dimension.

---

[4]Readability indexes suffer from the same problem, requiring length adjustment as well

Section conclusion

The enumerated principles of CSR disclosure quality, as well as chosen NLP proxies, are likely to capture quality construct for several reasons. According to IASB (2013), understandability, comparability, relevance, and reliability are essential quality attributes, reflecting good practice. We capture these concepts as follows:

- understandability - with readability indexes;
- comparability - with cosine similarity metric;
- relevance - with relevant CSR vocabulary (topic coverage, horizon);
- reliability - with unbiased tone.

Second, evidence suggests that readability, tone, topic coverage, and report length reflect the informativeness of financial disclosures beyond numeric data. According to a literature review by Li (2010a), more complex financial reports associate with management obfuscation and require considerable user's effort to process. Future-oriented economic narratives are more informative, and tone associates with future performance. Informativeness of CSR disclosures is still an emerging topic. A recent study by Muslu et al. (2019) analyzes CSR disclosures with a directly computed textual score. They associate higher CSR disclosure scores with fewer optimistic and more pessimistic keywords, better readability, more length, and more numeric and horizon content. They find that CSR reporters with high disclosure scores have more accurate forecasts and better CSR performance ratings issued by KLD. Another study by Nazari et al. (2017) finds that within the context of voluntary CSR disclosure, the increased size of CSR reports indicates better CSR performance and confirms the tendency to use more complex language in case of poor CSR performance.

Scarce literature establishing a link between CSR textual attributes and the market makes interpretation of the direction of change ambiguous. Here we apply the Directive's perspective to interpret impact and associate an improvement of textual quality with increased topic coverage, greater use of forward-looking and long-term oriented vocabulary, and a decrease of optimistic bias, lower complexity, and better similarity. However, further validation is required from the users' perspective to ensure that our proxies capture what we intend and to understand their perception when assessed in aggregate. The summary of operationalized variables is presented in Appendix 4.

6. CSR Disclosure score

Our index consists of the following components:

1. Coverage of dimensions such as environmental, social, bribery, human rights, and long-term orientation

2. Sentiment bias

3. Similarity of texts

4. Complexity (captured by readability and length)

We rank each attribute from one to ten, where the highest rank corresponds to the best value of the textual parameter. For consistency of interpretation, we reverse readability, length, and bias, ensuring that most obfuscating, wordiest, and most

biased reports get the lowest rank. We group all parameters capturing the same concept into one measure by taking their average to avoid amplifying their weight in the index. Our final score is formally computed as follows:

$$CSRQ = mean(TOP + BIAS + READ + COS + lnLEN) \tag{5}$$

where:

TOP - topic coverage, which equals the average of word counts of environmental, social, bribery, human rights, and long term oriented vocabulary;

$BIAS$ - sentiment bias, which equals to the absolute difference between positive and negative words;

$READ$ - readability, which equals the average value of FOG, $SMOG$, and FLE readability indexes;

$lnLEN$ - length, which equals to the count of sentences in disclosure defined using nltk package;

COS - cosine similarity, which equals to the similarity between texts from year to year.

By excluding the effect of length from each component, we ensure that each attribute obtains an equal weight, and length does not solely drive our results. However, equal weighting may also limit our understanding of the relationship with attributes perception by market users. Another possibility is the existence of a non-linear relationship. Here we assume a simplistic linear relationship, which is a first step in validating the importance of these attributes for the market.

7. Research design

First-stage research design

The impact of regulation is evaluated using a difference-in-difference (DID) design with the control group consisting of pre-treatment compliers (Atanasov & Black, 2016; Dahya & McConnell, 2007). The DID groups are schematized below:

Groups	Pre 2015	Post 2014
Control	Mandatory reporting	Mandatory
Treatment	Voluntary reporting	Mandatory

As we have seen above, mandatory regimes were already implemented by the UK, Sweden, France, and Denmark before the Directive. First, we exclude Sweden as it is the only country with a narrower scope compared to the Directive, mandating CSR reporting only for state-owned companies. Second, we exclude 2011, 2012 from the sample because several legislations concerning CSR occurred during this period that could potentially affect local reporting practices leading to the violation of the DID parallel trend assumption.

Danish, French, and UK undertakings operate in environments where CSR disclosures were mandated long before 2015. We assign companies from UK, France, and Denmark to the control group, while firms from countries with voluntary frameworks compose the treatment group.

The experience of Danish companies served as a model for the Directive (Szabó & Sørensen, 2015), while the UK has a higher degree of compliance with new requirements (Aureli et al., 2018) leading to a smallest post-Directive change in these countries. UK and French companies have historically paid more attention to issues of long-term environmental and social risk (Aureli et al., 2018; Chen & Bouvain, 2009), while such aspects as human rights and corruption received less attention (Aureli et al., 2018).

Among the EU member states, the UK, Denmark, and France are considered to be among the leaders in CSR (Mullerat, 2013). For a certain time, UK has even appointed a CSR Minister, and France created an inter-ministerial CSR coordination committee. Meanwhile, Denmark is called a "welfare state" as it has a higher socially oriented political regime and has attained a high CSR performance over the past years.

To conclude, the control group's choice depended on the existence of mandatory regimes before the Directive, the scope of companies obliged to publish CSR, and the level of CSR activities in the corresponding countries. We select countries with the mandatory regime, the same range of companies required to report compared to the Directive, and a high CSR activity as our control group. Such a selection ensures minor changes in these countries brought by the Directive.

One of the potential issues for the selected control group is compliance under established mandatory regimes. On the one hand, low compliance in the control group will leave room for changes under the European mandate. In such a scenario, DID will not capture the change's amplitude as both groups will change at the same pace. However, we reason that as the mechanism to force out companies to publish CSR was already activated by the mandatory local regimes and the requirements were very similar, the new legislation will not bring new elements to the cost-benefit tradeoff in these countries. France, Denmark, and the UK have already undergone the same change and reached their plateau of the "comply and explain" regulation. Our data collection also confirms the high coverage of CSR in these countries. Therefore, we expect that the new legislation will not affect our control group. However, the trend evolution of CSR disclosures in these countries can still occur without undermining our analysis as long as the parallel trend assumption holds. Descriptive statistics presented later add more weight to our choice of the control group.

Further, we discuss the specifications of our model. The difference-in-difference method consists of analyzing a regression of the form presented in equation 6.

$$y = \alpha + \beta_{1-3}Time * Treat + \beta_4 Controls + \beta_5 FE + e \qquad (6)$$

where:

y - textual outcome

β_1 - expected mean change from before to after the shock among the control group (UK, France, Denmark)

β_2 - estimated mean difference in y between the treatment and control groups before the shock

β_3 - difference-in-difference estimator, shows whether the mean differences between the treatment and control groups after the shock are significantly different from the gap that existed before the shock between those same groups

β_4 - control variables isolating effects of companies' characteristics (such as growth (relative change in revenues), size (market value), financial performance (ROA) and A4IR CSR performance (ESG^5 on the content and presentation of reports

β_5 - country, industry, year fixed effects

The parallel trend assumption requires that, in the absence of treatment, the difference between the "treatment" and "control" group is constant over time. We verify this condition based on the trends presented in Figure 1.

A more formal approach to test the parallel trend assumption is presented in Figure 2. This Figure indicates the significance of the difference between the annual change of the dependent variable for the control and treatment groups. After visual and formal inspections, we keep all variables since none suggests a clear violation of the parallel trend assumption.

Second-stage research design

We start by conducting a preliminary test as the first indication of informativeness of CSR disclosures in our setting (Report REP equal to one if CSR disclosure exists, zero otherwise) using equation 7. This model allows us to explore the relationship between the existence of CSR disclosure and bid-ask spread ($lnBA_t$).

$$lnBA_t = \alpha + \beta_1 REP_{t-1} + \beta_2 Controls_{t-1} + \beta_3 FE + e \tag{7}$$

Then, we explore the role of the document where it is published ($SOURCE$ equal to one (S_1) if issued in annual report, $SOURCE$ equal to two (S_2) if published in both reports, $SOURCE$ equal to three (S_3) if published in stand-alone CSR report) (equation 8).

$$lnBA_t = \alpha + \beta_1 S_{1,t-1} + \beta_2 S_{2,t-1} + \beta_3 S_{3,t-1} + \beta_4 Controls_{t-1} + \beta_5 FE + e \tag{8}$$

To answer our first research question, we analyze the relationship between textual attributes (topic coverage (TOP), sentiment bias ($BIAS$), readability ($READ$), similarity (COS), length ($lnLEN$) and bid-ask spreads ($lnBA_t$) (equation 9).

$$lnBA_t = \alpha + \beta_1 TOP_{t-1} + \beta_2 BIAS_{t-1} + \beta_3 READ_{t-1} + \beta_4 COS_{t-1} \\ + \beta_5 lnLEN_{t-1} + \beta_6 Controls_{t-1} + \beta_7 FE + e \tag{9}$$

We further explore the relationship between the aggregated measure ($CSRQ$) and the bid-ask spread ($lnBA_t$) using equation 10.

$$lnBA_t = \alpha + \beta_1 CSRQ_{t-1} + \beta_2 Controls_{t-1} + \beta_3 FE + e \tag{10}$$

---

[5]This rating belonging to the Thomson Reuters database widely relies on public disclosures as a basis of the overall expert's judgment about the company's CSR performance. Even though it is not a measure of disclosure quality, it is likely highly correlated with the extent of CSR disclosures (Bouten et al., 2018) as they form the primary source for CSR performance definition and affect the way raters perceive CSR information. We include this score to control for underlying CSR performance and capture incremental informativeness of textual parameters. Textual attributes can offer a more transparent tool to assess CSR than ratings and ensure broader coverage not limited by the rating database. Therefore, introducing textual scores may help overcome the main drawbacks of the ratings, such as quality, transparency, coverage, reverse causality. However, using it as an alternative tool to capture CSR, we enter a "horse race, " requiring testing if textual attributes communicate additional information omitted by the rating experts.

Following Leuz (2010), we obtain daily closing bid and ask prices and compute spread by taking the difference between ask and bid prices divided by their average. Yearly value is obtained by taking an average of all daily spreads within the corresponding year and applying logarithmic transformation after following prior literature. Logarithmic transformation is applied to correct the skewness of this variable ($lnBA_t$).

The information covering year t in the report appears only in year $t+1$, a few months after the end of the fiscal year. Since the real dates of publication are not known we use a conventional adjustment, which supposes that the reports become public three months after the end of the fiscal year as schematized below:

We match values of yearly averages of bid-ask spreads with a year of public disclosure, and not the year of coverage. Therefore, textual values observed for reporting of year t are regressed on the bid-ask spread in $t+1$. We adjust the bid-ask spread to the assumed dates of public disclosure. It is identified by adding three months to the fiscal year-end, which is indicated in the report [6].

We conduct this analysis for the period from 2013 to 2017, controlling for country, industry, and year fixed effects (FE). Since we are interested in the incremental usefulness of textual metrics over existing CSR scores, namely, in our case Asset4 ESG performance, we include it as a control in our analysis. We also control for trading volume, price variability, and company size as these parameters are shown to affect bid-ask spreads (Leuz & Verrecchia, 2000). Finally, we control for revenue growth, return on assets, and book to market as they also influence the overall level of disclosures according to prior studies (Athanasakou et al., 2019; Caglio et al., 2020; Egginton & McBrayer, 2019). All variables are defined in Appendix 4.

The model presented in equation 11 helps to determine if, in aggregate, there was a significant change in textual disclosures around the Directive, based on which we can conclude if it entailed economic consequences. Change equals the difference between means of post and pre Directive values for each variable of regression.

$$\Delta lnBA_t = \alpha + \beta_1 \Delta CSRQ_{t-1} + \beta_2 \Delta Controls_{t-1} + e \tag{11}$$

We also explore the changes in individual textual components for the same sample of reporters to understand the main drivers, as indicated in equation 12.

$$\begin{aligned}\Delta lnBA_t = \alpha &+ \beta_1 \Delta TOP_{t-1} + \beta_2 \Delta BIAS_{t-1} + \beta_3 \Delta READ_{t-1} + \beta_4 \Delta COS_{t-1} \\ &+ \beta_5 \Delta lnLEN_{t-1} + \beta_6 \Delta Controls_{t-1} + e\end{aligned} \tag{12}$$

---

[6]We conduct robustness tests and check the sensitivity to the selection of public disclosure dates, but it does not affect our results.

8.   Sample selection and Descriptive statistics

Sample selection

Using the Reuters database, we preselect a set of companies required to disclose non-financial information based on country-level criteria published in the overview of the Directive prepared by the Global Reporting Initiative (GRI, 2017). The total amount of listed companies falling under the Directive, excluding the financial sector, is 1,164. A detailed breakdown by country is presented in Table 2.

(**Table 2 about here**)

CSR disclosures exist in different formats. Some companies choose to publish it in a stand-alone report. Others include it in the annual report or both. The Directive acknowledges non-financial disclosures in both reports. Therefore, to analyze its impact, we need to cover both channels as companies can also change the source of disclosure over time. Another reason for including both sources of CSR disclosures is the recurrent problem of CSR studies consisting of insufficient datasets. By including annual reports, we add a significant amount of observations to our dataset. However, there is a risk of capturing overlapping disclosures between the two reports. We address this issue by comparing these reports using the cosine similarity metric. We do not find "copy-paste" disclosures as none of the comparison have a perfect similarity.

We start by exploring databases already containing links to CSR reports, such as Corporate Register and GRI database. However, they appear to be unsatisfactory for this task. Regarding Corporate Register, it claims to gather most of the reports (annual and stand-alone) but limits free access. Concerning the GRI, their database contains an important share of outdated web links. As a result, we end up collecting the links manually from companies' websites. We assemble web links of .pdf files for the period ranging from 2011 to 2017 and download them with Python. We include only English written reports. Since we also include countries where English is not a national language, we recognize that some information can be "lost in translation", however, we partially alleviate this limitation by controlling for country fixed effects. After obtaining all the links per year per company, we convert .pdf files to .txt format. One of the impediments at this stage are .pdf documents saved as an image. The tools for recognizing texts in such formats are at the initial stage of development and offer very messy optical recognition, which does not allow us to include these observations. Our final database consists of active links, referring to either annual or CSR reports per year. The sample obtained at this stage consists of 614 firms.

Since we solely focus our analysis on CSR information, the second stage of data collection consists of the extraction of CSR sections from annual reports using Python. Our algorithm consists of proper identification of entire CSR sections based on the distribution of CSR words in documents using vocabulary built for that purpose (see Appendix 2). The vocabulary used for the identification of CSR parts relies on CSR standards and frameworks such as AA1000, UNGC, GRI, ISO14001, SA8000, and others (see Appendix 1). These documents allowed us to build a list of words and word collocations (up to 4-words, also called tetragrams) that were then independently analyzed by two master students, who selected CSR items from the list. The final set of words is presented in Appendix 2. Our algorithm reaches a

92% rate of successful extraction of CSR disclosures based on a set of 50 randomly selected and manually classified reports (Author, 2019).

Finally, only those companies with at least one observation, consisting of either CSR section from an annual report or stand-alone CSR report, are kept, reducing the sample size to 385 firms resulting in a 1,925 firm-year observations. The average values of main variables computed for our final dataset, along with other descriptive statistics, are presented in Table 3.

(**Table 3 about here**)

Panel A summarizes descriptive statistics for the quantity, quality, and control variables: number of reporters, CSR reports length, sentiment, readability, topic coverage, horizon vocabulary, the similarity of texts over time. We identified skewness of textual length and market value distributions during the check of variables' normality that we address with a logarithmic transformation.

The number of reporters (REP) reflects CSR disclosure independent of its size for a given firm-year observation. It is a dummy variable with values equal to zero for the absence of CSR disclosure (in the annual report or stand-alone CSR report) and one otherwise. 81 % of our sample has CSR disclosures, which reinforces our assumption that there is room for more publications. Most of the reports come from the United Kingdom, where CSR reporting is mandatory.

The textual part of average CSR disclosure has about 170 sentences ranging from zero to 10,000 sentences (before the transformation of LEN with natural logarithm). The topic coverage descriptive statistics consist of direct word counts (before transformation with TF-IDF weighting), significantly varying depending on the topic. The environmental (ENV) and social (SOC) issues get higher weight in CSR discussions, with 31 and 50 words per document, respectfully. The bribery (BRB) and human rights (HR) topics receive a smaller coverage with only about three to four words per document. Bribery vocabulary is the most modest, which may be misleading and create an impression of a lower discussion. It is the drawback of using a vocabulary approach for topics describable with only a small set of words. However, it is enough to capture the topic's presence and the extent to which it changes after the Directive.

Sentiment statistics illustrate the dominance of positive tone in CSR discussions, which is in line with the literature, finding CSR discussions to be overpositive (Cho et al., 2010) as a means of "greenwashing".

The readability indexes are around 10, meaning that for the understanding of those texts the reader requires 10 years of education slightly lower than the values found in literature analyzing 10-K reports (Lang & Stice-Lawrence, 2015; Loughran & McDonald, 2014): 18 years for FOG, 19 years for SMOG (Muslu et al., 2019); and 14 years for Flesch (De Franco et al., 2015). The latter is consistent with the idea that CSR discussions aim to satisfy a broader range of stakeholders' informational needs.

CSR subsection's average similarity (COS) before Taylor's adjustment is around 0.75, while the range of value lies between zero and one. Values close to one correspond to higher similarity. The sample size for this parameter is smaller because its computation requires at least two consecutive observations, which is not always the case since we also have companies that publish CSR disclosures sporadically.

Mean statistics breakdown by treatment / control and pre/post

Table 3, panel B, reports mean statistics for pre/post periods contrasting control and treatment groups. Both groups have a higher number of reporters after the Directive. However, the treatment group has a higher growth (an increase of 14% vs. an increase of 3% for the control group). Changes in the control group do not invalidate the benchmark choice as long as it satisfies the parallel trend assumption. The number of reporters is significantly higher for the control group, which matches our expectation of CSR prevalence in the control group.

The control group does not significantly change disclosures' length (from 170 to 220 sentences), while the treatment group has substantially longer CSR disclosures after 2015 (from 80 to 220 sentences). Interestingly, both groups converge to a similar length after the Directive.

The interpretation of topic coverage depends on whether we present it in absolute or weighted values. In our principal analysis, we apply TF-IDF weighting. However, we also show absolute values before the transformation. It helps to interpret and understand the impact of the weights better. Following our discussion above, TF-IDF transformation takes into account frequent words, which obtain lower weight, as the same things do not convey more information. Analysis of topic coverage in absolute measures shows that it is more widespread in our treatment group. However, if we reduce the weight of most common CSR words and give more weight to unique CSR words, our control group has higher coverage of CSR topics. This observation suggests that in a voluntary environment, the content of CSR discussion is more generic than in the control group with mandatory regimes. However, the weighting does not affect our interpretation of the trend, which is positive for both groups.

The control group appeals less to sentiment vocabulary, while the treatment group appears to use more positive and negative words. In both groups, the bias is positive, and it decreases over time.

The readability indexes go up. Companies in the control group with mandatory regimes implemented before the Directive have a higher level of complex CSR reports. The transition of the treatment group to compulsory reporting goes together with an increase in complexity. Depending on the interpretation of this measure in CSR context, it can be due to the more complex nature of underlying CSR activities when we switch from the "cheap talk" to real actions, or on the contrary to another form of hiding the actual performance and obfuscating the reader on purpose.

Reporting becomes more similar over time, attaining the same level of similarity in the control and treatment groups. Finally, companies also start using more long-term oriented vocabulary.

To summarize, our choice of the control group is consistent with a better level of quantitative (number of reporters, reports' length) and qualitative (topics coverage, sentiment, horizon) attributes observed thanks to descriptive statistics' analysis. The treatment group experiences significant changes after shock compared to previous years, while the control group has less pronounced time changes consistent with its status as a benchmark.

Convergence over time

Table 3, panel B, also allows us to analyze the general trend in CSR disclosures after the transition to a mandatory regime. Looking at the difference between the

control and treatment group in pre vs. post Directive periods, we can infer whether companies converge in their reporting practices or the new regulatory regime leads to more discrepancies. The extent of the gap between these two groups reduces for the number of reporters, the coverage of social issues, the use of sentiment and long-term oriented vocabulary, textual complexity, and similarity. At the same time, it disappears entirely for the length of reports and human rights topic coverage. The differences seem to intensify only for coverage of the environmental topic, with the latter receiving broader attention in the treatment group.

This evidence suggests that there is a tendency towards convergence in disclosure practices after a mandatory regime. This observation is consistent with the institutional theory, which "focuses on understanding the imitation of practices – termed as "isomorphism" (which often leads to an increasing homogenization across organizations) – as the outcome of a process through which companies seek to acquire legitimacy by conforming to regulative, normative and cultural-cognitive pressures" (Ioannou & Serafeim, 2019).

9.   Results

First-stage main results

Table 4 presents results capturing effect of the Directive on CSR disclosures. Panels A differ from panels B by the introduction of ESG rating (Asset4) as a control variable. Other control variables are similar for both panels, with market value transformed to the natural logarithm, and all variables winsorized at 5-95 level[7].

(**Table 4 about here**)

We adjust for length for such attributes as cosine similarity and readability indexes, to avoid a mechanical relationship with documents' length. We do not make such adjustments for vocabulary-based metrics as by taking out length components, we would loose information about channels through which changes in volume occur.

We disclose these two specifications to show the incremental information of our parameters beyond the Asset4 rating and directly trace the change of textual attributes. As we can see, the Asset4 score entirely captures such dimensions as social topic coverage, long-term orientation, number of reports, and similarity. As for the other parameters, the ESG rating only slightly reduces the coefficients and misses the bias. It also reduces sample size by a third.

Our control group (*1.TIME*) goes through some changes during the five years. It increases the length of reports, mainly through broader bribery and human rights coverage, the similarity of texts from year to year, and the use of sentiment vocabulary. These changes reflect the general time trend without the impact of the Directive. As long as the general trend satisfies the parallel trend assumption, these significant increases are not a concern.

Looking at the contrast between the treatment and control groups before the Directive (*1.TREAT*), we confirm weaker CSR engagement of the treatment group, which echoes our descriptive statistics overview. There are significantly fewer reports in the treatment group before the Directive. Disclosures are also considerably smaller.

---

[7]We test with other levels of winsorization (1-99), and the results stay the same

The treatment group has less informative topic coverage except for environmental issues. Reporters from the treatment group include less long-term oriented, and sentiment vocabulary.

Our main parameter of interest is the interaction of the treatment and time variables (*1.TREAT#1.TREAT*). The Directive significantly impacts several aspects of CSR disclosures of the treatment group compared to the control. First, we validate our underlying assumption that there is room for new reporters as the number of reporters increases after 2015. Second, treatment companies dedicate significantly more space to CSR discussions with the main focus on social, environmental and bribery issues. The latter validates our assumption that companies were not publishing CSR reports at their full potential and that non-reporters were doing so before not because they were not concerned but because of the high costs of disclosing this information. The reports of treatment firms also become more conservative with the increase of negative vocabulary. Reports become more long-term oriented, less similar, and more complex after the Directive. These latter results suggest that companies report more specific information about more complex underlying activities or, on the contrary, that they are inconsistent in their reporting and try to obfuscate the reader.

After including the ESG rating in our analysis, a few results become insignificant, meaning that the score already captures this information. It is the case for social and bribery topic coverage, long-term orientation, number of reports, and similarity. Finally, with ESG rating as a control variable, sentiment bias becomes significant, indicating that there is a more substantial decrease of bias after the Directive by the treatment group.

We include country and year fixed effects for robustness tests, as main effects capture it only partially. Panels A2 and B2 present the results with fixed effects which only slightly affect the coefficients.

Christensen et al. (2019) make a distinction between early adopters and low reporters in the context of transition to mandatory reporting. They state that "the comparability benefits would not only accrue to firms with currently low levels of CSR disclosure but also affect best-practice firms that already disclose information on most of the relevant topics" (p.31). We analyze such breakdowns of our dataset (results are not reported) and discover that our results are driven by the change in a subgroup of low reporters. In contrast, all-time reporters only have a significant increase in readability.

Second-stage results

Table 5 shows the final set of variables further used in tests. Exclusion of missing values for the financial variables leaves us with a sample of 1,849 firm-year observations. Further inclusion of control for ESG performance reduces our sample to 970 observations.

(**Table 5 about here**)

Preliminary tests

Table 6 presents distribution of CSR disclosures across different sources from 2013 to 2017. The overall trend indicates higher coverage of CSR over time, with the biggest share published in annual reports.

<div align="center">(**Table 6 about here**)</div>

Our initial test in Table 7 reveals a highly significant relationship between CSR disclosure existence and bid-ask spread. Consistent with prior findings (Cho et al., 2013; Cormier et al., 2011; Cui et al., 2018; Egginton & McBrayer, 2019; Radhakrishnan et al., 2018), we can see that existence of CSR disclosures reduce information asymmetries in financial markets. This finding is an indication of CSR disclosures' informativeness as they are taken into account by market participants (Hypothesis 4a).

<div align="center">(**Table 7 about here**)</div>

After the rejection of Hypothesis 4a, we investigate the role of each source of CSR publications. Table 7 displays results for our second preliminary test (Hypothesis 4b) and shows that annual reports are also an essential source of CSR disclosures for investors as are stand-alone CSR reports or a mix of the two. Moreover, contrary to the finding of Mittelbach-Hörmanseder et al. (2020), we show that CSR disclosures published in stand-alone reports are significant for the market and cannot be omitted from further analysis.

Main tests

Table 8 presents results for equations 9 and 10 testing our first central hypotheses (Hypothesis 5a and Hypothesis 5b). The sample size is smaller for the model with individual textual attributes because of the similarity metric. We find a significant association between our index and the natural logarithm of bid-ask spread beyond performance captured by Asset4 rating, which indicates that our metric matters beyond this score. This finding is economically meaningful, as one standard deviation in $CSRQ$ score reduces bid-ask spread by 8.9% [8]. It is similar to a bid-ask spread reduction of 5.39% associated with reporting changes found by Egginton and McBrayer (2019). When main components are disclosed, we can see that this result is driven by complexity (captured by length and readability), topic coverage and similarity of texts, which significantly affect bid-ask spread.

<div align="center">(**Table 8 about here**)</div>

For robustness, we further investigate our sample by reducing it only to a set of companies that produced at least one CSR disclosure every year. According to Table 9, the effect for index is maintained. However, the decomposition of the index indicates that for this sub-sample, topic coverage loses its importance.

<div align="center">(**Table 9 about here**)</div>

Finally, to address our second central hypothesis (Hypothesis 6), we analyze the change due to the Directive (see Table 10). First, we look at the economic consequences of reporting changes for early adopters. For that purpose, we choose as a cut-off 2015 - year after the Directive's adoption, which took place in 2014.

---

[8]We obtain this measure by multiplying coefficient of $CSRQ$ (-0.08) by the standard deviation of $CSRQ$ (1.057839) divided by the standard deviation of $lnBA$ (0.9417851). Standard deviations are obtained using Stata.

Comparing pre and post-adoption periods, we can see that CSR disclosures are not associated with a change in the bid-ask spread, which implies that it did not have economic consequences straight after adoption.

(**Table 10 about here**)

Our second analysis focuses on subsequent relevant signals for the market, which is the actual implementation of the Directive effectively entering into force in 2017. We repeat prior analysis by comparing periods before and after implementation of the Directive (see Table 11). This test indicates a significant contribution of the aggregated textual index to the bid-ask spread change. However, due to the small sample size, it is difficult to tell which of its components drives results as none is significant.

(**Table 11 about here**)

10. Conclusion

Non-trivial extraction of European CSR disclosures limited prior studies to analyzing only incomplete settings. Analyzing significant changes in reporting, such as the transition of all large European companies to the mandatory regime, requires treating large data sets and develop transparent metrics. Such an examination is needed to assess the outcomes of a costly change. Among a panoply of available methods satisfying these conditions, natural language processing techniques offer the most promising tools. However, their application is scarce and incomplete in the CSR context. We contribute to this strand of literature, aiming to uncover the role of different textual attributes obtained using the automated textual treatment. However, further research is required to better understand the link between textual characteristics of disclosures and underlying CSR performance. It is a promising area to uncover managerial incentives beyond such disclosures under the mandatory regime.

Our two-stage analysis confirms that a change occurs at the company level, starting to prepare for a new regime. These changes are also relevant for market participants because they affect transparency but only when the mandate enters into force. Despite the prior unsuccessful implementation of mandatory regimes in distinct localities (Larrinaga et al., 2002), the vast scope of the new initiative seems to send a more reliable signal about the commitment of regulators and generates a significant response by companies. More generally, our results contribute to the literature on the effectiveness of disclosure regulations.

First, we question if the new Directive increases the number of CSR reporters. Our analysis confirms that the Directive had enough power to initiate a reaction from companies since the number of reports increased significantly.

The next question consists of verifying the extent of this impact. Do companies increase their reporting substantially or add a few statements? For that purpose, we look at the rise in volume. Results confirm a substantial change, showing that companies had new things to tell about CSR, instead of just adding a paragraph about the Directive's irrelevance for their business, which would not have resulted in significant change.

To the extent that textual parameters capture underlying concepts, texts become more long-term oriented, cover CSR topics more broadly, and become less biased, indicating an improvement for users. On the other hand, they also become more complex and less similar.

Our analysis indicates that the "comply or explain" model of the mandate, when implemented at such a broad scope, initiates a reaction despite weak enforcement. We argue that the impact occurs due to the shift in a cost-benefit tradeoff with an increased falsification cost, the threat of future stricter regulations, and a more credible CSR image.

However, the interpretation of this reaction is twofold. Higher complexity (readability) and lower similarity analyzed together could signify more specific and detailed disclosures. Indeed, it makes sentences longer and less comparable from year to year, which is a good sign when aiming to reflect strategic CSR instead of "one size fits all". On the other hand, higher complexity could appear because they do not have an option for selective disclosure anymore and have to disclose even weak CSR activities but do it in an obfuscating way, which is a sign of bad quality.

We explore the economic role of CSR disclosures and their linguistic properties. Prior literature documents the value relevance of CSR disclosures but suffers from some limitations. Some use non-transparent disclosure scores or incomplete observations reduced to stand-alone CSR reports. Others explore the US setting or only a voluntary regime covering a limited set of linguistic properties. Our study tries to alleviate these concerns. First, we make sure to validate the CSR narrative's informativeness for financial users and highlight their most relevant attributes. Second, we assess the effectiveness of the transition to a mandatory CSR reporting regime from the financial users' perspective.

Using NLP techniques, we capture CSR disclosures beyond their mere existence to answer questions about its economic role during the transition to a mandatory regime. From the regulator's perspective, understanding these implications helps to evaluate the effectiveness of new measures. From the academic view, it allows exploring informativeness of CSR report quantity (as measured by length and reporting frequency) and quality (as measured by coverage of environmental, social, bribery, human rights, and long-term oriented vocabulary, readability, sentiment bias, and similarity) for financial markets. It is particularly interesting in the transition to mandatory reporting. While under voluntary regimes, the role of these disclosures is heavily discredited by greenwashing practices, the transition to mandatory reporting involves new mechanisms that prevent former shortcomings. However, the transition to mandatory reporting can also have undesired consequences such as the convergence of practices, which is not always wanted as investors seek CSR solutions individually tailored to the entity's business model. We find that the changes occurring after the Directive's adoption do not improve financial users' informational environment. During the transition period (from 2015 to 2017), only the variation in topic coverage was significant. However, there is an indication of a relevant change for investors in 2017, after the Directive's implementation. Our results show that variation in reporting is informative for the market as it significantly improves transparency.

Additionally, we provide evidence for the role of textual attributes in capturing CSR disclosures. First, we find that it is incrementally informative compared to Asset4 rating. We also disclose the main drivers of our aggregated textual index. It seems that investors pay attention to the level of verbal complexity, preferring lower

obfuscation and less wordiness, while also appreciating consistent reporting from year to year.

Finally, our dataset allows us to explore the role of different sources of CSR disclosures. According to Lewis and Young (2019), "tracking information across multiple reporting channels and evaluating the consistency of content and messaging is challenging" (p.596). Based on this premise, we posit that CSR reports simultaneously covered in multiple channels do not improve the informational environment. On the other hand, prior literature argues that CSR disclosures in stand-alone reports do not affect market participants (Hummel et al., 2018). Our analysis indicates that both sources matter individually.

It is essential to acknowledge a series of limitations to our study. Even though CSR disclosures mostly consist of narratives, capturing only "soft" attributes may omit important information. Developing more hard metrics isolating numeric CSR data is a promising area for improvement of CSR analysis based on publicly available information. Moreover, future research is needed to assess if the beneficial effects of the new regulation are long-lasting. Only by analyzing post-mandatory periods can we state whether this regime reduces the informativeness of CSR disclosures or provide a better instrument for companies to differentiate themselves based on their CSR activities. Finally, more insights can be gained by looking into member states' individual experiences and finding the main drivers of successful implementation. There are also technical limitations that may affect our data. Although English is a generally accepted business language in European countries, some information can still be lost in translation, as it is not a native language for all the participants. there is also a significant loss of information occurring because of reports uniquely published in foreign languages and those that do not convert to .txt format.

Overall, this study contributes to CSR literature investigating economic consequences of disclosure regulations (Ioannou & Serafeim, 2019), summarizing CSR narratives with alternative measures besides non-transparent ratings (Muslu et al., 2019) and extending the scope of CSR disclosure analysis by including annual reports (Mittelbach-Hörmanseder et al., 2020).

References

Ajina, A., Laouiti, M., & Msolli, B. (2016). Guiding through the fog: Does annual report readability reveal earnings management? *Research in International Business and Finance*, 38, 509–516.

Atanasov, V. A., & Black, B. S. (2016). Shock-Based Causal Inference in Corporate Finance and Accounting Research. *Critical Finance Review*, 5, 207–304. https://doi.org/10.1561/104.00000036

Athanasakou, V., El-Haj, M., Rayson, P., Walker, M., & Young, S. (2019). Annual Report Management Commentary articulating strategy and bsuiness models. https://doi.org/10.17635/lancaster/researchdata/232

Aureli, S., Salvatori, F., & Magnaghi, E. (2018). The Transposition of the Non-Financial Reporting Directive in the UK, France and Italy. *Symphonya. Emerging Issues in Management*, 1, 48–67. https://doi.org/10.4468/2018.1.04aureli.magnaghi.salvatori

Author. (2019). *Vocabulary based information retrieval: a look at CSR reports by European listed companies* (Working paper).

Ball, C., Hoberg, G., & Maksimovic, V. (2013). *Disclosure and Firm Separation: A Text-Based Examination.* (Working paper). https://doi.org/10.2139/ssrn.2260371

Barth, M. E., Cahan, S. F., Chen, L., & Venter, E. R. (2017). The economic consequences associated with integrated report quality: Capital market and real effects. *Accounting, Organizations and Society*, 62, 43–64. https://doi.org/10.1016/j.aos.2017.08.005

Beretta, S., & Bozzolan, S. (2008). Quality versus Quantity: The Case of Forward-Looking Disclosure. *Journal of Accounting, Auditing & Finance*, 23 (3): 333–376. https://doi.org/10.1177/0148558X0802300304

Beyer, A., Cohen, D. A., Lys, T. Z., & Walther, B. R. (2010). The financial reporting environment: Review of the recent literature. *Journal of accounting and economics*, 50 (2-3): 296–343. https://doi.org/10.1016/j.jacceco.2010.10.003

Bonsall IV, S. B., Leone, A. J., Miller, B. P., & Rennekamp, K. (2017). A plain english measure of financial reporting readability. *Journal of Accounting and Economics*, 63 (2-3): 329–357. https://doi.org/10.1016/j.jacceco.2017.03.002

Bouten, L., Cho, C. H., Michelon, G., & Roberts, R. W. (2018). CSR Performance Proxies in Large-Sample Studies: 'Umbrella Advocates', Construct Clarity and the 'Validity Police'. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3107182

Brochet, F., Loumioti, M., & Serafeim, G. (2015). Speaking of the short-term: disclosure horizon and managerial myopia. *Review of Accounting Studies*, 20 (3): 1122–1163. https://doi.org/10.1007/s11142-015-9329-8

Brown, S. V., & Tucker, J. W. (2011). Large-Sample Evidence on Firms' Year-over-Year MD&A Modifications. *Journal of Accounting Research*, 49 (2): 309–346. https://doi.org/10.1111/j.1475-679X.2010.00396.x

Caglio, A., Melloni, G., & Perego, P. (2020). Informational Content and Assurance of Textual Disclosures: Evidence on Integrated Reporting. *European Accounting Review*, 29 (1): 55–83. https://doi.org/10.1080/09638180.2019.1677486

Cahan, S. F., De Villiers, C., Jeter, D. C., Naiker, V., & Van Staden, C. J. (2016). Are CSR Disclosures Value Relevant? Cross-Country Evidence. *European*

Accounting Review, 25 (3): 579–611. https://doi.org/10.1080/09638180.2015.1064009

CDSB. (2019). *The Climate Disclosure Standards Board Reporting Framework*. CDSB. https://www.cdsb.net/sites/default/files/cdsb_framework_2019_v2.2.pdf

Chen, S., & Bouvain, P. (2009). Is corporate responsibility converging? a comparison of corporate responsibility reporting in the USA, UK, Australia, and Germany. *Journal of business ethics*, 87 (1): 299–317. https://doi.org/10.1007/s10551-008-9794-0

Cho, C. H., Guidry, R. P., Hageman, A. M., & Patten, D. M. (2012). Do actions speak louder than words? An empirical investigation of corporate environmental reputation. *Accounting, Organizations and Society*, 37 (1): 14–25. https://doi.org/10.1016/j.aos.2011.12.001

Cho, C. H., Michelon, G., Patten, D. M., & Roberts, R. W. (2014). CSR report assurance in the USA: An empirical investigation of determinants and effects. *Sustainability Accounting, Management and Policy Journal*. https://doi.org/10.1108/SAMPJ-01-2014-0003

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

Cho, S. Y., Lee, C., & Pfeiffer Jr, R. J. (2013). Corporate social responsibility performance and information asymmetry. *Journal of Accounting and Public Policy*, 32 (1): 71–83. https://doi.org/10.1016/j.jaccpubpol.2012.10.005

Christensen, H. B., Hail, L., & Leuz, C. (2013). Mandatory IFRS reporting and changes in enforcement. *Journal of Accounting and Economics*, 56 (2-3): 147–177. https://doi.org/10.1016/j.jacceco.2013.10.007

Christensen, H. B., Hail, L., & Leuz, C. (2019). Adoption of CSR and Sustainability Reporting Standards: Economic Analysis and Review. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3427748

Clarkson, P. M., Fang, X., Li, Y., & Richardson, G. (2013). The relevance of environmental disclosures: Are such disclosures incrementally informative? *Journal of Accounting and Public Policy*, 32 (5): 410–431. https://doi.org/10.1016/j.jaccpubpol.2013.06.008

Conceptual Framework of the Sustainability Accounting Standards Board. (2013). https://www.sasb.org/wp-content/uploads/2013/10/SASB-Conceptual-Framework-Final-Formatted-10-22-13.pdf

Core, J. E. (2001). A review of the empirical disclosure literature: Discussion. *Journal of accounting and economics*, 31 (1-3): 441–456. https://doi.org/10.1016/S0165-4101(01)00036-2

Cormier, D., Ledoux, M.-J., & Magnan, M. (2011). The informational contribution of social and environmental disclosures for investors. *Management Decision*.

Cui, J., Jo, H., & Na, H. (2018). Does Corporate Social Responsibility Affect Information Asymmetry? *Journal of Business Ethics*, 148 (3): 549–572. https://doi.org/10.1007/s10551-015-3003-8

Dahya, J., & McConnell, J. J. (2007). Board Composition, Corporate Performance, and the Cadbury Committee Recommendation. *Journal of financial and quantitative analysis*, 535–564. https://doi.org/10.1017/s0022109000004099

Davis, A. K., Piger, J. M., & Sedor, L. M. (2012). Beyond the Numbers: Measuring the Information Content of Earnings Press Release Language. *Contemporary*

Accounting Research, 29 (3): 845–868. https://doi.org/10.1111/j.1911-3846.2011.01130.x

De Franco, G., Hope, O.-K., Vyas, D., & Zhou, Y. (2015). Analyst report readability. *Contemporary Accounting Research*, 32 (1): 76–104. https://doi.org/10.1287/mnsc.2017.2751

De Villiers, C., & Marques, A. (2016). Corporate social responsibility, country-level predispositions, and the consequences of choosing a level of disclosure. *Accounting and Business Research*, 46 (2): 167–195. https://doi.org/10.1080/00014788.2015.1039476

Dhaliwal, D. S., Li, O. Z., Tsang, A., & Yang, Y. G. (2011). Voluntary Nonfinancial Disclosure and the Cost of Equity Capital: The Initiation of Corporate Social Responsibility Reporting. *The Accounting Review*, 86 (1): 59–100. https://doi.org/10.2308/accr.00000005

Dhaliwal, D. S., Radhakrishnan, S., Tsang, A., & Yang, Y. G. (2012). Nonfinancial disclosure and analyst forecast accuracy: International evidence on corporate social responsibility disclosure. *The accounting review*, 87 (3): 723–759. https://doi.org/10.2308/accr-10218

Egginton, J. F., & McBrayer, G. A. (2019). Does it pay to be forthcoming? Evidence from CSR disclosure and equity market liquidity. *Corporate Social Responsibility and Environmental Management*, 26 (2): 396–407. https://doi.org/10.1002/csr.1691

European Commission. (2013). Commission staff working document impact assessment, 1–88. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52013SC0127&from=EN

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

European Commission. (2017). European Commission: Non-Financial Reporting. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52017XC0705(01)&from=EN

Fiechter, P., Hitz, J.-M., & Lehmann, N. (2018). Real Effects in Anticipation of Mandatory Disclosures: Evidence from the European Union's CSR Directive. *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.3033883

Greenstein, M., & Sami, H. (1994). The impact of SEC's segment disclosure requirement on bid-ask spreads. *Accounting review: A quarterly journal of the American Accounting Association*, 69 (1): 179–199.

Grewal, J., Riedl, E. J., & Serafeim, G. (2019). Market Reaction to Mandatory Nonfinancial Disclosure. *Management Science*, 65 (7): 3061–3084. https://doi.org/10.2139/ssrn.2657712

GRI. (2013). *G4 Sustainability Reporting Guidelines - Part 1: Reporting Principles and Standard Disclosures.* GRI. https://www.globalreporting.org/resourcelibrary/G3-Guidelines-Incl-Technical-Protocol.pdf

GRI. (2017). *Member State Implementation of Directive 2014/95/EU.* GRI. https://www.accountancyeurope.eu/wp-content/uploads/1711-NFRpublication-GRI-CSR-Europe.pdf

Huang, A. H., Zang, A. Y., & Zheng, R. (2014). Evidence on the information content of text in analyst reports. *The Accounting Review*, 89 (6): 2151–2180. https://doi.org/10.2308/accr-50833

Hummel, K., Mittelbach-Hoermanseder, S. A., Cho, C. H., & Matten, D. (2018). *Implicit Versus Explicit Corporate Social Responsibility Disclosure: A Textual Analysis.* https://doi.org/10.2139/ssrn.3090976

Hummel, K., & Schlick, C. (2016). 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. https://doi.org/10.1016/j.jaccpubpol.2016.06.001

IASB, A. (2013). A review of the conceptual framework for financial reporting. *London: IASB.*

Ioannou, I., & Serafeim, G. (2017). *The Consequences of Mandatory Corporate Sustainability Reporting* (Working paper No. [11-100]:). https://doi.org/10.2139/ssrn.1799589

Ioannou, I., & Serafeim, G. (2019). *Corporate Sustainability: A Strategy?* (Working Paper No. [19-065]:). https://doi.org/10.2139/ssrn.3312191

ISO. (2015). *Introduction to ISO 14001:2015.* ISO. https://doi.org/10.1108/S0749-742320160000019022

Jeffwitz, C., & Gregor, F. (2017). Comparing the implementation of the EU non-financial reporting directive. *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.3083368

Jegadeesh, N., & Wu, D. (2013). Word power: A new approach for content analysis. *Journal of Financial Economics*, 110 (3): 712–29. https://doi.org/10.1016/j.jfineco.2013.08.018

Kearney, C., & Liu, S. (2014). Textual sentiment in finance: A survey of methods and models. https://doi.org/10.1016/j.irfa.2014.02.006

Kozlowski, A., Searcy, C., & Bardecki, M. (2015). Corporate sustainability reporting in the apparel industry an analysis of indicators disclosed. *International Journal of Productivity and Performance Management.* https://doi.org/10.1108/IJPPM-10-2014-0152

KPMG. (2017). The road ahead: The KPMG Survey of Corporate Responsibility Reporting 2017. https://integratedreporting.org/wp-content/uploads/2017/10/kpmg-survey-of-corporate-responsibility-reporting-2017.pdf

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

Larrinaga, C., Carrasco, F., Correa, C., Llena, F., & Moneva, J. (2002). Accountability and accounting regulation: the case of the Spanish environmental disclosure standard. *European Accounting Review*, 11 (4): 723–740. https://doi.org/10.1080/0963818022000001000

Lawrence, A. (2013). Individual investors and financial disclosure. *Journal of Accounting and Economics*, 56 (1): 130–47. https://doi.org/10.1016/j.jacceco.2013.05.001

Lee, C. M. C. C. M., & Yahn, T. L. (1997). Accounting Information and Bid-Ask Spreads. *Accounting horizons*, 11 (4): 50–60. https://doi.org/10.1111/j.1475-679X.2007.00238.x

Lehavy, R., Li, F., & Merkley, K. (2011). The effect of annual report readability on analyst following and the properties of their earnings forecasts. *The Accounting Review*, 86 (3): 1087–1115. https://doi.org/10.2308/accr.00000043

Leipziger, D. (2017). *The corporate responsibility code book*. https://doi.org/10.9774/gleaf.9781783530670{_}12

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

Leuz, C., & Verrecchia, R. E. (2000). The Economic Consequences of Increased Disclosure. *Journal of accounting research*, 91–124. https://doi.org/10.2307/2672910

Leuz, C., & Wysocki, P. D. (2016). The Economics of Disclosure and Financial Reporting Regulation: Evidence and Suggestions for Future Research. *Journal of accounting research*, 54 (2): 525–622. https://doi.org/10.1111/1475-679X.12115

Lewis, C., & Young, S. (2019). Fad or future? Automated analysis of financial text and its implications for corporate reporting. *Accounting and Business Research*, 49 (5): 587–615. https://doi.org/10.1080/00014788.2019.1611730

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

Li, F. (2010a). Survey of the Literature. *Journal of accounting literature*, 29, 143–165.

Li, F. (2010b). The information content of forward- looking statements in corporate filings-A naïve bayesian machine learning approach. *Journal of Accounting Research*, 48 (5): 1049–1102. https://doi.org/10.1111/j.1475-679X.2010.00382.x

Loughran, T., & McDonald, B. (2011). When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *Journal of Finance*, 66 (1): 35–65. https://doi.org/10.1111/j.1540-6261.2010.01625.x

Loughran, T., & McDonald, B. (2014). Measuring readability in financial disclosures. *the Journal of Finance*, 69 (4): 1643–1671. https://doi.org/10.1111/jofi.12162

Loughran, T., & McDonald, B. (2015). The Use of Word Lists in Textual Analysis. *Journal of Behavioral Finance*, 16 (1): 1–11. https://doi.org/10.1080/15427560.2015.1000335

Loughran, T., & McDonald, B. (2016). Textual Analysis in Accounting and Finance: A Survey. *Journal of Accounting Research*, 54 (4): 1187–1230. https://doi.org/10.1111/1475-679X.12123

Matsumura, E. M., Prakash, R., & Vera-Munoz, S. C. (2014). Firm-value effects of carbon emissions and carbon disclosures. *The accounting review*, 89 (2): 695–724. https://doi.org/10.2308/accr-50629

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

Michelon, G., Pilonato, S., & Ricceri, F. (2015). CSR reporting practices and the quality of disclosure: An empirical analysis. *Critical perspectives on accounting*, 33, 59–78. https://doi.org/10.1016/j.cpa.2014.10.003

Miller, B. P. (2010). The effects of reporting complexity on small and large investor trading. https://doi.org/10.2308/accr.00000001

Mittelbach-Hörmanseder, S., Hummel, K., & Rammerstorfer, M. (2020). Information Content of Corporate Social Responsibility Disclosures in Europe: An

Leipziger, D. (2017). *The corporate responsibility code book*. https://doi.org/10.9774/gleaf.9781783530670{_}12

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

Leuz, C., & Verrecchia, R. E. (2000). The Economic Consequences of Increased Disclosure. *Journal of accounting research*, 91–124. https://doi.org/10.2307/2672910

Leuz, C., & Wysocki, P. D. (2016). The Economics of Disclosure and Financial Reporting Regulation: Evidence and Suggestions for Future Research. *Journal of accounting research*, 54 (2): 525–622. https://doi.org/10.1111/1475-679X.12115

Lewis, C., & Young, S. (2019). Fad or future? Automated analysis of financial text and its implications for corporate reporting. *Accounting and Business Research*, 49 (5): 587–615. https://doi.org/10.1080/00014788.2019.1611730

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

Li, F. (2010a). Survey of the Literature. *Journal of accounting literature*, 29, 143–165.

Li, F. (2010b). The information content of forward- looking statements in corporate filings-A naïve bayesian machine learning approach. *Journal of Accounting Research*, 48 (5): 1049–1102. https://doi.org/10.1111/j.1475-679X.2010.00382.x

Loughran, T., & McDonald, B. (2011). When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *Journal of Finance*, 66 (1): 35–65. https://doi.org/10.1111/j.1540-6261.2010.01625.x

Loughran, T., & McDonald, B. (2014). Measuring readability in financial disclosures. *the Journal of Finance*, 69 (4): 1643–1671. https://doi.org/10.1111/jofi.12162

Loughran, T., & McDonald, B. (2015). The Use of Word Lists in Textual Analysis. *Journal of Behavioral Finance*, 16 (1): 1–11. https://doi.org/10.1080/15427560.2015.1000335

Loughran, T., & McDonald, B. (2016). Textual Analysis in Accounting and Finance: A Survey. *Journal of Accounting Research*, 54 (4): 1187–1230. https://doi.org/10.1111/1475-679X.12123

Matsumura, E. M., Prakash, R., & Vera-Munoz, S. C. (2014). Firm-value effects of carbon emissions and carbon disclosures. *The accounting review*, 89 (2): 695–724. https://doi.org/10.2308/accr-50629

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

Michelon, G., Pilonato, S., & Ricceri, F. (2015). CSR reporting practices and the quality of disclosure: An empirical analysis. *Critical perspectives on accounting*, 33, 59–78. https://doi.org/10.1016/j.cpa.2014.10.003

Miller, B. P. (2010). The effects of reporting complexity on small and large investor trading. https://doi.org/10.2308/accr.00000001

Mittelbach-Hörmanseder, S., Hummel, K., & Rammerstorfer, M. (2020). Information Content of Corporate Social Responsibility Disclosures in Europe: An

Institutional Perspective. *European Accounting Review*, 1–40. https://doi.org/10.2139/ssrn.3365521

Mullerat, R. (2013). Corporate Social Responsibility: A European Perspective. *The Jean Monnet/Robert Schuman Paper Series*, 13 (6): 1–22.

Muslu, V., Mutlu, S., Radhakrishnan, S., & Tsang, A. (2019). Corporate Social Responsibility Report Narratives and Analyst Forecast Accuracy. https://doi.org/10.1007/s10551-016-3429-7

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

OECD. (2008). *Guidelines for multinational enterprises*. OECD. https://doi.org/10.1787/mne-2002-en

Plumlee, M., Brown, D., Hayes, R. M., & Marshall, R. S. (2015). Voluntary environmental disclosure quality and firm value: Further evidence. *Journal of Accounting and Public Policy*, 34 (4): 336–361. https://doi.org/10.1016/j.jaccpubpol.2015.04.004

Radhakrishnan, S., Tsang, A., & Liu, R. (2018). A corporate social responsibility framework for accounting research. *The International Journal of Accounting*, 53 (4): 274–294. https://doi.org/10.1016/j.intacc.2018.11.002

Reverte, C. (2016). Corporate social responsibility disclosure and market valuation: evidence from Spanish listed firms. *Review of Managerial Science*, 10 (2): 411–435. https://doi.org/10.1007/s11846-014-0151-7

Richardson, A. J., & Welker, M. (2001). Social disclosure, financial disclosure and the cost of equity capital. *Accounting, organizations and society*, 26 (7-8): 597–616. https://doi.org/10.1016/S0361-3682(01)00025-3

Soyka, P. A. (2013). The International Integrated Reporting Council (IIRC) integrated reporting framework: toward better sustainability reporting and (way) beyond. *Environmental Quality Management*, 23 (2): 1–14. https://doi.org/10.1002/tqem.21357

Standard, A. S. E. (2015). Institute of Social and Ethical Accountability: 896 London. *United Kingdom*, 897.

Szabó, D. G., & Sørensen, K. E. (2015). *New EU Directive on the Disclosure of Non-Financial Information (CSR)* (Working paper No. [15-01]:). https://doi.org/10.1515/ecfr-2015-0307

United Nations. (2000). The ten principles of global compact. https://doi.org/10.1002/npr.l0067

United Nations. (2019). *The sustainable development goals report 2019*. United Nationss. https://unstats.un.org/sdgs/report/2019/The-Sustainable-Development-Goals-Report-2019.pdf

Verbeeten, F. H., Gamerschlag, R., & Möller, K. (2016). Are CSR disclosures relevant for investors? Empirical evidence from Germany. *Management Decision*. https://doi.org/10.1108/MD-08-2015-0345

Verrecchia, R. E. (1982). The Use of Mathematical Models in Financial Accounting. *Journal of Accounting Research*, 1–42. https://doi.org/10.2307/2674672

You, H., & Zhang, X.-j. (2009). Financial reporting complexity and investor underreaction to 10-k information. *Review of Accounting studies*, 14 (4): 559–586. https://doi.org/10.1007/s11142-008-9083-2

Zerbini, F. (2017). CSR Initiatives as Market Signals: A Review and Research Agenda. *Journal of Business Ethics*, 146 (1): 1–23. https://doi.org/10.1007/s10551-015-2922-8

TABLE 1

Member States requirements after the Directive (Source: GRI 2017)

Country	Staff	Balance Sheet	Net Turnover	Currency	Comparison	Penalties	Audit
Directive	500	20 mio	40 mio	EUR	benchmark	not specified	existence check
Austria	500	20 mio	40 mio	EUR	=	Monetary	=
Belgium	500	17 mio	34 mio	EUR	stricter	Monetary	=
Bulgaria	500	38 mio	76 mio	BGN	=	Monetary	= +content check
Croatia	500	15 mio	30 mio	HRK	stricter	Monetary	=
Cyprus	500	20 mio	40 mio	EUR	=	Monetary	= +content check
Czech Republic	500	na	1 bn	CZK	=	Monetary	=
Denmark	250	156 mio	313 mio	DKK	=	=	= +content check
Estonia	500	na	na	na	stricter	=	=
Finland	500	20 mio	40 mio	EUR	=	Monetary	=
France	500	20 mio	40 mio	EUR	=	Monetary	= +content check
Germany	500	20 mio	40 mio	EUR	=	Monetary	not mandatory
Greece	500	na	na	na	stricter	Monetary	=
Hungary	500	6 bn	12 bn	HUF	=	Monetary	=
Iceland	250	3 bn	6 bn	ISK	stricter	+ Imprisonment	= +content check
Italy	500	20 mio	40 mio	EUR	=	Monetary	= +content check
Latvia	500	20 mio	40 mio	EUR	=	Monetary	= +content check
Lithuania	500	20 mio	40 mio	EUR	=	Monetary	=
Luxembourg	250	20 mio	40 mio	EUR	stricter	=	=
Malta	500	20 mio	40 mio	EUR	=	Monetary	=
Netherlands	500	20 mio	40 mio	EUR	=	=	=
Norway	500	20 mio	40 mio	EUR	=	+ Imprisonment	=
Poland	500	85 mio	170 mio	PLN	=	Monetary	=
Portugal	500	na	na	na	stricter	Monetary	=
Republic of Ireland	500	20 mio	40 mio	EUR	=	+ Imprisonment	=
Romania	500	na	na	na	stricter	Monetary	= +content check
Slovak Republic	500	20 mio	40 mio	EUR	=	Monetary	=
Slovenia	500	20 mio	40 mio	EUR	=	Monetary	=
Spain	500	20 mio	40 mio	EUR	=	=	=
Sweden	250	175 mio	350 mio	SEK	stricter	Monetary	=
United Kingdom	500	na	na	na	stricter	Monetary	= +content check

Notes: This table compares implementation requirements formulated in Directive and their transposition into the Member States legislation. We compare the scope of entities falling under Directive, penalties, and auditor's level of involvement. The comparison is made after conversion to EUR.

=: same level of requirements as determined by the Directive.

na: not applicable

TABLE 2
Number of companies by country

Austria	28
Denmark	48
Greece	29
Lithuania	5
Poland	115
Slovenia	5
Belgium	14
Estonia	2
Hungary	3
Luxembourg	3
Portugal	19
Spain	65
Croatia	5
Finland	47
Iceland	5
Malta	2
Republic of Ireland	7
Sweden	42
Cyprus	2
France	82
Italy	47
Netherlands	44
Romania	4
United Kingdom	345
Czech Republic	5
Germany	175
Latvia	2
Norway	10
Slovak Republic	4
Total	**1,164**

Notes: This table shows the distribution of companies falling under the Directive by country.

TABLE 3
Full sample

Panel A: Descriptive Statistics

	count	mean	sd	min	p50	max
Quantity variables						
REP	1,925	0.81	0.39	0	1	1
lnLEN	1,925	5.13	2.72	0	5.82	9.22
Quality variables						
– Topics						
ENV	1,925	30.92	46.31	0	12	366
SOC	1,925	53.87	71.39	0	27	563
BRB	1,925	2.88	5.66	0	0	55
HR	1,925	4.29	13.09	0	1	225
– Sentiment						
NEG	1,925	134.98	195.94	0	56	1347
POS	1,925	194.83	266.73	0	84	2136
BIAS	1,925	59.85	123.7	-531	21	946
– Readability						
FOG	1,925	8.88	4.59	0	9.96	18.05
FLE	1,925	10.43	5.15	0	12.4	24
SMG	1,925	12.11	5.88	0	14.6	19.8
– Horizon						
HOR	1,925	47.44	70.96	0	19	598
– Similarity						
COS	1,564	0.76	0.27	0	0.88	0.99
Control variables						
REV	1,925	0.18	5.08	-0.86	0.03	221.21
ROA	1,925	5.94	9.13	-70.08	5.25	217.76
lnMV	1,925	7.6	2	1.65	7.5	12.67
ESG	1,200	72.8	28.21	0	85.97	96.41

Panel B: Mean statistics by treatment / control, pre/post

	Treated			Control vs. Treated		Control		
	Before	Diff	After	Before Diff	After Diff	Before	Diff	After
REP	0.67	0.14***	0.81	-0.20***	-0.09**	0.87	0.03	0.9
lnLEN	4.42	0.98***	5.40	-0.71***	-0.03	5.13	0.30	5.43
– Topics								
ENV	0.62	0.11**	0.73	0.01	0.13***	0.62	-0.02	0.6
SOC	0.92	0.21***	1.12	-0.31***	-0.17***	1.22	7.00	1.3
BRB	0.04	0.04***	0.08	-0.04***	-0.04***	0.08	0.04***	0.12
HR	0.08	0.05***	0.13	-0.03*	-0.02	0.12	0.03*	0.15
– Sentiment								
NEG	0.1	0.04***	0.14	-0.06***	-0.05***	0.16	0.02***	0.18
POS	0.13	0.02***	0.15	-0.06***	-0.05***	0.19	0.01	0.2
BIAS	0.03	-0.01**	0.01	0	-0.01	0.03	-0.01	0.02
– Readability								
FOG	7.15	1.45***	8.6	-2.89***	-1.53***	10.04	0.1	10.13
FLE	8.67	1.91***	10.58	-2.47***	-0.80**	11.14	0.24	11.38
SMG	10.05	2.22***	12.27	-2.91***	-0.99**	12.96	0.3	13.26
– Horizon								
HOR	1.45	0.30***	1.75	-0.35***	-0.14**	1.8	0.09	1.89
– Similarity								
COS	0.53	-0.06***	0.6	-0.11***	-0.07***	0.65	0.06***	0.71
Control variables								
REV	0.09	-0.03	0.05	-0.65	-0.02	0.74	-0.67	7.00
ROA	5.5	-0.04	5.46	-1.64*	-0.78	7.13	-0.89	6.24
lnMV	7.51	0.2	7.71	0.06	0.06	7.45	0.2	7.65
ESG	64.12	8.56***	72.68	-7.27**	-8.03***	71.39	9.32***	80.71
N	450		675			320		480

Notes: Panel A shows descriptive statistics for a sample of companies with at least one CSR disclosure during 2013 - 2017. The definition of variables is presented in Appendix 4. Panel B contrasts the mean values of observations for different sub-samples: control vs. treated and prior vs. post Directive. The Control group consists of firms from Denmark, UK, France. The treatment group consists of firms from the rest of the EU. The year 2015 is a pre-post cutoff. The TF-IDF weighting attributes lower weights to common words and higher weights to rare terms. The definition of variables is in Appendix 4.
*** p<0.01, ** p<0.05, * p<0.1

TABLE 4
Difference-in-difference

Panel A1: without *ESG*

Variables	Existence	Volume	Topic Coverage				Similarity	Readability			Sentiment			Horizon
	REP	*LEN*	*SOC*	*ENV*	*BRB*	*HR*	*COS*	*FOG*	*FLE*	*SMG*	*NEG*	*POS*	*BIAS*	*HOR*
Difference in difference														
1.TIME#1.TREAT	0.51*	0.67***	0.13*	0.12***	0.01*	0.02	-0.08**	0.16***	0.27***	0.29***	0.02**	0.01	-0.01	0.20**
	(0.27)	(0.18)	(0.06)	(0.04)	(0.01)	(0.01)	(0.03)	(0.05)	(0.08)	(0.05)	(0.01)	(0.01)	(0.01)	(0.08)
Main effects														
1.TIME	0.29	0.24**	0.07	-0.02	0.03***	0.02***	0.03*	-0.09	-0.12	-0.13*	0.02**	0.01**	-0.00	0.07
	(0.23)	(0.10)	(0.05)	(0.04)	(0.01)	(0.01)	(0.01)	(0.06)	(0.11)	(0.06)	(0.01)	(0.00)	(0.01)	(0.06)
1.TREAT	-1.25***	-0.82***	-0.30***	-0.01	-0.03***	-0.04**	0.01	-0.14	0.09	0.05	-0.07***	-0.06***	0.01	-0.36***
	(0.21)	(0.14)	(0.05)	(0.04)	(0.01)	(0.01)	(0.02)	(0.09)	(0.06)	(0.06)	(0.01)	(0.01)	(0.01)	(0.06)
Control variables														
REV	-0.73	-1.06*	-0.10	-0.11	-0.01	-0.09**	-0.14***	0.12	-0.68	-0.83	-0.01	-0.03	-0.02	-0.22
	(0.60)	(0.48)	(0.07)	(0.09)	(0.03)	(0.04)	(0.04)	(0.95)	(0.96)	(1.10)	(0.02)	(0.03)	(0.05)	(0.20)
ROA	0.03**	0.00	0.01	-0.00	-0.00**	-0.00	0.00	0.08**	0.06*	0.07*	-0.00	0.00**	0.00***	0.01
	(0.01)	(0.02)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.03)	(0.03)	(0.03)	(0.00)	(0.00)	(0.00)	(0.01)
MV	0.29***	0.50***	0.05***	0.07***	0.01***	0.02***	0.01**	0.06	0.42***	0.50***	0.01***	0.01***	-0.00	0.13***
	(0.04)	(0.05)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)	(0.08)	(0.10)	(0.11)	(0.00)	(0.00)	(0.00)	(0.02)
Constant	1.61**	1.48**	0.77***	0.09	-0.01	-0.02	-0.08**	0.25	0.04	0.09	0.08***	0.09***	0.01	0.81***
	(0.67)	(0.46)	(0.10)	(0.07)	(0.01)	(0.03)	(0.03)	(0.18)	(0.23)	(0.14)	(0.02)	(0.02)	(0.02)	(0.19)
Observations	1,925	1,925	1,925	1,925	1,925	1,925	1,925	1,925	1,925	1,925	1,925	1,925	1,925	1,925
R-squared	logistic	0.14	0.07	0.08	0.07	0.05	0.02	0.07	0.07	0.08	0.14	0.13	0.02	0.10
Taylor length adjustment	NO	NO	NO	NO	NO	NO	YES	YES	YES	YES	NO	NO	NO	NO
Industry FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
Country FE	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO

Panel A2: robustness test with country, year, industry fixed effects, without *ESG*

Variables	Existence	Volume	Topic Coverage				Similarity	Readability			Sentiment			Horizon
	REP	*LEN*	*SOC*	*ENV*	*BRB*	*HR*	*COS*	*FOG*	*FLE*	*SMG*	*NEG*	*POS*	*BIAS*	*HOR*
Difference in difference														
1.TIME#1.TREAT	0.58**	0.68***	0.13**	0.12***	0.01	0.02	-0.08**	0.16***	0.27**	0.28***	0.02**	0.01	-0.01	0.21**
	(0.28)	(0.19)	(0.06)	(0.04)	(0.01)	(0.01)	(0.03)	(0.05)	(0.08)	(0.05)	(0.01)	(0.01)	(0.01)	(0.08)
Control variable														
REV	-0.63	-0.86*	-0.14	-0.04	-0.04	-0.12**	-0.09**	0.17	0.17	0.23	-0.02	-0.03	-0.01	-0.21
	(0.62)	(0.40)	(0.09)	(0.09)	(0.04)	(0.04)	(0.04)	(0.24)	(0.25)	(0.21)	(0.03)	(0.02)	(0.04)	(0.21)
ROA	0.03**	-0.01	0.01	-0.00	-0.00**	-0.00	0.00	0.00	0.00	0.00	-0.00	0.00	0.00**	0.00
	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
MV	0.26***	0.50***	0.07***	0.05***	0.00	0.02***	0.00*	0.02	0.02	0.00	0.01***	0.01***	0.00	0.14***
	(0.05)	(0.04)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)	(0.02)	(0.04)	(0.02)	(0.00)	(0.00)	(0.00)	(0.02)
Constant	2.29***	2.26***	0.76***	0.55***	-0.00	-0.05	-0.01	0.16	-0.31	0.42	0.04	0.00	-0.03	0.88***
	(0.83)	(0.53)	(0.09)	(0.14)	(0.02)	(0.05)	(0.03)	(0.26)	(0.48)	(0.30)	(0.03)	(0.03)	(0.06)	(0.24)
Observations	1,925	1,925	1,925	1,925	1,925	1,925	1,925	1,925	1,925	1,925	1,925	1,925	1,925	1,925
R-squared	logistic	0.24	0.17	0.18	0.21	0.09	0.07	0.13	0.08	0.13	0.20	0.23	0.07	0.17
Taylor length adjustment	NO	NO	NO	NO	NO	NO	YES	YES	YES	YES	NO	NO	NO	NO
Industry FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Country FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Panel B1: with *ESG*

Variables	Existence	Volume	Topic Coverage				Similarity	Readability			Sentiment			Horizon
	REP	*LEN*	*SOC*	*ENV*	*BRB*	*HR*	*COS*	*FOG*	*FLE*	*SMG*	*NEG*	*POS*	*BIAS*	*HOR*
Difference in difference														
1.TIME#1.TREAT	0.43	0.40*	0.06	0.11**	0.01	0.01	-0.03	0.11	0.24***	0.20***	0.02***	-0.00	-0.02**	0.09
	(0.40)	(0.19)	(0.06)	(0.04)	(0.01)	(0.01)	(0.02)	(0.07)	(0.06)	(0.04)	(0.01)	(0.01)	(0.01)	(0.08)
Main effects														
1.TIME	0.24	0.04	0.01	-0.03	0.02**	0.00	0.02	-0.06	-0.09	-0.09	0.01	0.01	0.01	0.01
	(0.33)	(0.19)	(0.05)	(0.05)	(0.01)	(0.01)	(0.01)	(0.06)	(0.09)	(0.05)	(0.01)	(0.01)	(0.01)	(0.09)
1.TREAT	-0.56*	0.01	-0.16***	0.14*	-0.02*	-0.00	0.02	-0.01	0.26***	0.20**	-0.05***	-0.04***	0.01	-0.08
	(0.30)	(0.32)	(0.04)	(0.07)	(0.01)	(0.02)	(0.02)	(0.08)	(0.07)	(0.07)	(0.01)	(0.01)	(0.01)	(0.12)
Control variables														
ESG	0.02***	0.03***	0.01***	0.00**	0.00	0.00*	0.00	0.01	0.03***	0.03***	0.00**	0.00***	-0.00	0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
REV	0.18	-0.28	0.23	-0.14	0.04	-0.08	-0.11	1.38	1.11	1.09	0.03	-0.01	-0.05	0.12
	(0.95)	(0.75)	(0.21)	(0.15)	(0.05)	(0.06)	(0.07)	(1.32)	(1.35)	(1.51)	(0.03)	(0.04)	(0.04)	(0.33)
ROA	0.06***	0.01	0.01*	-0.00	-0.00**	0.00	0.00	0.08**	0.07**	0.08**	-0.00	0.00**	0.00*	0.01
	(0.02)	(0.02)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.03)	(0.03)	(0.03)	(0.00)	(0.00)	(0.00)	(0.00)
MV	-0.18**	0.10	-0.06**	0.04**	0.01**	0.03***	0.00	-0.51**	-0.33*	-0.34	0.01*	-0.00	-0.01***	-0.02
	(0.08)	(0.08)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.16)	(0.17)	(0.20)	(0.00)	(0.00)	(0.00)	(0.03)
Constant	3.27***	2.66***	1.29***	0.15*	-0.04	-0.13**	-0.02	0.61*	0.39	0.30	0.09***	0.19***	0.10**	1.51***
	(0.87)	(0.51)	(0.19)	(0.07)	(0.03)	(0.05)	(0.05)	(0.30)	(0.26)	(0.21)	(0.02)	(0.04)	(0.03)	(0.23)
Observations	1,200	1,200	1,200	1,200	1,200	1,200	1,200	1,200	1,200	1,200	1,200	1,200	1,200	1,200
R-squared	logistic	0.14	0.11	0.09	0.07	0.06	0.01	0.08	0.05	0.06	0.12	0.10	0.03	0.08
Taylor length adjustment	na	na	na	na	na	na	YES	YES	YES	YES	na	na	na	na
Industry FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
Country FE	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO

Panel B2: robustness test with country, year, industry fixed effects, with *ESG*

Variables	Existence	Volume	Topic Coverage				Similarity	Readability			Sentiment			Horizon
	REP	LEN	SOC	ENV	BRB	HR	COS	FOG	FLE	SMG	NEG	POS	BIAS	HOR
Difference in difference														
1.TIME#1.TREAT	0.47	0.40*	0.06	0.11**	0.01	0.01	-0.03	0.11	0.24***	0.20***	0.02***	-0.00	-0.02**	0.09
	(0.41)	(0.19)	(0.06)	(0.04)	(0.01)	(0.01)	(0.02)	(0.07)	(0.06)	(0.04)	(0.01)	(0.01)	(0.01)	(0.08)
Control variables														
ESG	0.02***	0.03***	0.00***	0.00***	0.00*	0.00	-0.00	-0.00	-0.00*	-0.00	0.00**	0.00	-0.00	0.01***
	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
REV	0.14	-0.33	0.11	-0.04	-0.01	-0.10	-0.07	0.26	0.58*	0.49*	0.02	-0.03	-0.05	0.04
	(1.08)	(0.69)	(0.22)	(0.17)	(0.05)	(0.06)	(0.07)	(0.25)	(0.26)	(0.24)	(0.04)	(0.02)	(0.04)	(0.34)
ROA	0.06***	0.01	0.01	0.00	-0.00**	0.00	0.00	0.01	0.01	0.00	-0.00	0.00	0.00	0.01**
	(0.02)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
MV	-0.27***	0.09	-0.03	0.00	0.00	0.03***	-0.00	-0.02	-0.02	-0.03	0.01**	0.00	-0.01*	-0.03
	(0.10)	(0.07)	(0.03)	(0.02)	(0.00)	(0.01)	(0.00)	(0.04)	(0.04)	(0.02)	(0.00)	(0.00)	(0.00)	(0.03)
Constant	3.90***	3.32***	1.17***	0.65***	-0.02	-0.17	0.07	0.59	0.75	0.72**	0.01	0.08*	0.08	1.61***
	(1.01)	(0.37)	(0.09)	(0.19)	(0.02)	(0.10)	(0.05)	(0.43)	(0.42)	(0.29)	(0.04)	(0.04)	(0.08)	(0.16)
Observations	1,200	1,200	1,200	1,200	1,200	1,200	1,200	1,200	1,200	1,200	1,200	1,200	1,200	1,200
R-squared	logistic	0.20	0.18	0.20	0.21	0.09	0.04	0.21	0.16	0.19	0.20	0.21	0.10	0.11
Taylor length adjustment	na	na	na	na	na	na	YES	YES	YES	YES	na	na	na	na
Industry FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Country FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Notes: Table 4 presents DID analysis results regarding the impact of the Directive on metrics capturing quantitative and qualitative attributes of CSR disclosures. The Control group consists of firms from Denmark, UK, France. The treatment group consists of firms from the rest of the EU. The year 2015 is a pre-post cutoff. Panel A presents results without controlling for the external CSR rating. In contrast to panel A1, panel A2 controls for country and year fixed effects, which leads to the omission of main effects due to collinearity. We adjust for length only cosine similarity and readability indexes to avoid associating their change with simple existence or higher volume of reports. Our vocabulary-based variables are adjusted with TF-IDF, which attributes lower weights to common words. We do not adjust these variables for length as they indicate channels through which the change in volume occurs. Panel B presents results with controlling for the external CSR rating. In contrast to panel B1, panel B2 controls for country and year fixed effects. The definition of variables is presented in Appendix 4.

Robust standard errors in parentheses, clustered at firm level

*** p<0.01, ** p<0.05, * p<0.1

TABLE 5
Second-stage descriptive statistics

Variables	count	min	p25	mean	p75	max	sd
Control variables							
lnBA	1849	-3.23	-2.17	-1.03	-0.13	2.12	1.25
lnVLT	1849	2.54	3.01	3.23	3.45	3.94	0.31
lnVOL	1849	-2.43	2.81	4.63	6.71	9.41	2.76
BM	1849	-1.80	1.10	2.56	3.08	17.00	2.55

Notes: This table displays descriptive statistics for firm-year observations for the second-stage analysis for the period 2013-2017. We apply natural logarithm to correct for skewness and winsorize at (1-99) level to adjust for outliers. The definition of variables is presented in Appendix 4.

TABLE 6

Distribution of CSR disclosures by year by source

Source	2013	2014	2015	2016	2017
No disclosure	53%	57%	50%	48%	42%
CSR in annual report	27%	30%	28%	31%	31%
CSR in stand-alone report	13%	13%	14%	14%	17%
CSR in both	8%	0	8%	8%	10%

Notes: This table shows the share of CSR disclosures published in different sources such as CSR stand-alone reports and annual reports, or both from 2013 to 2017. It is build based on a sample of companies falling under the Directive and having at least one report during the whole period. The biggest share of CSR disclosures comes from annual reports, while stand-alone reports cover only a third.

TABLE 7

Economic role of CSR disclosure existence and its source

Variables	(1) lnBA	(2) lnBA
Issuance of CSR report		
1.REP	-0.12***	
	(0.04)	
CSR report type		
1.SOURCE		-0.11**
		(0.04)
2.SOURCE		-0.11**
		(0.05)
3.SOURCE		-0.19***
		(0.05)
Control variables		
lnMV	-0.32***	-0.32***
	(0.01)	(0.01)
lnVLT	0.36***	0.36***
	(0.06)	(0.06)
lnVOL	-0.22***	-0.22***
	(0.01)	(0.01)
ROA	-0.02***	-0.02***
	(0.00)	(0.00)
REV	0.03	0.03
	(0.12)	(0.12)
BM	-0.00	-0.00
	(0.01)	(0.01)
Constant	1.61***	1.58***
	(0.25)	(0.25)
Observations	1,849	1,849
R-squared	0.75	0.75
Industry FE	YES	YES
Year FE	YES	YES
Country FE	YES	YES

Notes: This table contrasts the effect of CSR disclosures existence (*REP*=1) on bid-ask spread vs. absence of CSR coverage (*REP*=0). It further decomposes *REP* disclosures according to the document in which it has been published. *SOURCE* = 1 if only in Annual report, *SOURCE* = 2 if present in Annual report and stand-alone CSR report, *SOURCE* = 3 if only presented in stand-alone CSR report. The test is performed using a sample from 2013 to 2017, including non-reporters of CSR and contrasting them to reporters. The definition of variables is presented in Appendix 4.

Robust standard errors in parentheses, clustered at firm level

*** p<0.01, ** p<0.05, * p<0.1

TABLE 8

CSR Disclosure index, its components and the bid-ask spread for all the reporters

Variables	(1) lnBA	(2) lnBA	(3) lnBA
Aggregated score			
CSRQ	-0.08***	-0.09***	
	(0.01)	(0.01)	
Score components			
TOP_10			-0.04*
			(0.02)
BIAS_10			-0.01
			(0.01)
READ_10			-0.02**
			(0.01)
COS_10			-0.02***
			(0.01)
lnLEN_10			-0.03***
			(0.01)
Control variables			
lnMV	-0.35***	-0.33***	-0.34***
	(0.04)	(0.04)	(0.04)
lnVLT	0.29***	0.27***	0.26***
	(0.09)	(0.08)	(0.07)
lnVOL	-0.09***	-0.09***	-0.09***
	(0.03)	(0.02)	(0.02)
ROA	-0.01***	-0.01***	-0.01***
	(0.00)	(0.00)	(0.00)
REV	-0.35*	-0.40**	-0.40**
	(0.18)	(0.17)	(0.17)
BM	0.01***	0.01***	0.01***
	(0.00)	(0.00)	(0.00)
ESG		-0.00***	-0.00***
		(0.00)	(0.00)
Constant	1.42***	1.68***	1.86***
	(0.39)	(0.48)	(0.55)
Observations	970	970	970
R-squared	0.56	0.57	0.57
Industry FE	YES	YES	YES
Year FE	YES	YES	YES
Country FE	YES	YES	YES

Notes: These regressions indicate a significant negative relationship between CSR index (*CSRQ*) and bid-ask spread. Results are presented with and without controlling for Asset4 performance score (*ESG*). Our index is incrementally informative to the latter. The decomposition of the index into its components indicates that main drivers are readability (*READ_10*), similarity (*COS_10*), and length (*LEN_10*). These analyses are conducted based on a sample of observations with an existing CSR disclosure and Asset4 score from 2013 to 2017. The sample size in column (3) is limited by available values of similarity, which require at least two years of consecutive CSR disclosures. The definition of variables is presented in Appendix 4.

Robust standard errors in parentheses, clustered at firm level

*** p<0.01, ** p<0.05, * p<0.1

TABLE 9

CSR Disclosure index, its components and the bid-ask spread for regular reporters

Variables	(1) $lnBA$	(2) $lnBA$
Aggregated score		
$CSRQ$	-0.06***	
	(0.02)	
Score components		
TOP_10		-0.02
		(0.02)
$BIAS\_10$		-0.00
		(0.01)
$READ\_10$		-0.02*
		(0.01)
COS_10		-0.02**
		(0.01)
$lnLEN\_10$		-0.02***
		(0.01)
Control variables		
$lnMV$	-0.27***	-0.28***
	(0.02)	(0.02)
$lnVLT$	0.42***	0.42***
	(0.12)	(0.11)
$lnVOL$	-0.12***	-0.12***
	(0.04)	(0.04)
ROA	-0.01***	-0.01***
	(0.00)	(0.00)
REV	-0.52**	-0.52**
	(0.24)	(0.24)
BM	0.01*	0.01
	(0.01)	(0.01)
ESG	-0.00**	-0.00**
	(0.00)	(0.00)
$Constant$	0.87**	0.98***
	(0.38)	(0.32)
$Observations$	833	833
$R\text{-}squared$	0.55	0.56
$Industry$ FE	YES	YES
$Year$ FE	YES	YES
$Country$ FE	YES	YES

Robust standard errors in parentheses, clustered at firm level
*** p<0.01, ** p<0.05, * p<0.1

Notes: The results of these regressions indicate a significant negative relationship between CSR index (*CSRQ*) and bid-ask spread, which is incremental to the Asset4 performance score (*ESG*). The decomposition of the index into its components indicates that the main drivers are readability (*READ_10*) and length (*LEN_10*). These analyses are conducted based on a sample of observations with a CSR disclosure published every year from 2013 to 2017 and an available Asset4 score. The sample size in column (3) is limited by available values of similarity, which require at least two years of consecutive CSR disclosures. The definition of variables is presented in Appendix 4.
Robust standard errors in parentheses, clustered at firm level
*** p<0.01, ** p<0.05, * p<0.1

TABLE 10

The contribution of change in reporting to the change in bid-ask spread before/ after 2015

Variables	(1) $lnBA$	(2) $lnBA$
Aggregated score		
$CSRQ$	-0.04	
	(0.03)	
Score components		
TOP_10		-0.04
		(0.03)
$BIAS\_10$		0.01
		(0.01)
$READ\_10$		-0.01
		(0.01)
COS_10		-0.02
		(0.02)
$lnLEN\_10$		0.00
		(0.01)
Control variables		
$lnMV$	-0.21**	-0.21**
	(0.11)	(0.11)
$lnVLT$	0.32	0.27
	(0.27)	(0.27)
$lnVOL$	-0.13**	-0.13**
	(0.05)	(0.05)
ROA	-0.02**	-0.02**
	(0.01)	(0.01)
REV	-0.36	-0.40
	(0.25)	(0.25)
BM	-0.02	-0.02
	(0.02)	(0.02)
ESG	-0.00	-0.00
	(0.00)	(0.00)
$Constant$	-0.02	-0.01
	(0.03)	(0.03)
Observations	249	249
R-squared	0.19	0.20

Notes: Δ values are obtained by subtracting corresponding variable's means computed over the period before 2015 from means computed over the period starting from 2015. Where 2015 is the year of first reporting after the adoption of Directive by the European Commission. The results for the aggregated index are insignificant. The sample consists of a number of firms with a CSR disclosure published every year from 2013 to 2017 and an available Asset4 score. The definition of variables is presented in Appendix 4.

Robust standard errors in parentheses, clustered at firm level

*** p<0.01, ** p<0.05, * p<0.1

TABLE 11

The contribution of change in reporting to the change in bid-ask spread before/after 2017

Variables	(1) $lnBA$	(2) $lnBA$
Aggregated score		
$CSRQ$	-0.05** (0.02)	
Score components		
TOP_10		-0.01 (0.02)
$BIAS\_10$		-0.00 (0.01)
$READ\_10$		-0.01 (0.01)
COS_10		-0.02 (0.01)
$lnLEN\_10$		-0.01 (0.01)
Control variables		
$lnMV$	-0.38*** (0.10)	-0.39*** (0.10)
$lnVLT$	-0.01 (0.30)	0.00 (0.30)
$lnVOL$	-0.16*** (0.05)	-0.16*** (0.05)
ROA	-0.01 (0.01)	-0.01 (0.01)
REV	-0.04 (0.19)	-0.06 (0.20)
BM	0.00 (0.01)	0.00 (0.01)
ESG	-0.00 (0.00)	-0.00 (0.00)
$Constant$	0.14*** (0.04)	0.14*** (0.04)
$Observations$	249	249
$R\text{-}squared$	0.18	0.18

Notes: Δ values are obtained by subtracting corresponding variable's means computed over the period before 2017 from means computed over the period starting from 2017. Where 2017 is the year of mandatory implementation of Directive by European companies. The results for the aggregated index are significant. The sample consists of a number of firms with a CSR disclosure published every year from 2013 to 2017 and an available Asset4 score. The definition of variables is presented in Appendix 4.

Robust standard errors in parentheses, clustered at firm level

*** p<0.01, ** p<0.05, * p<0.1

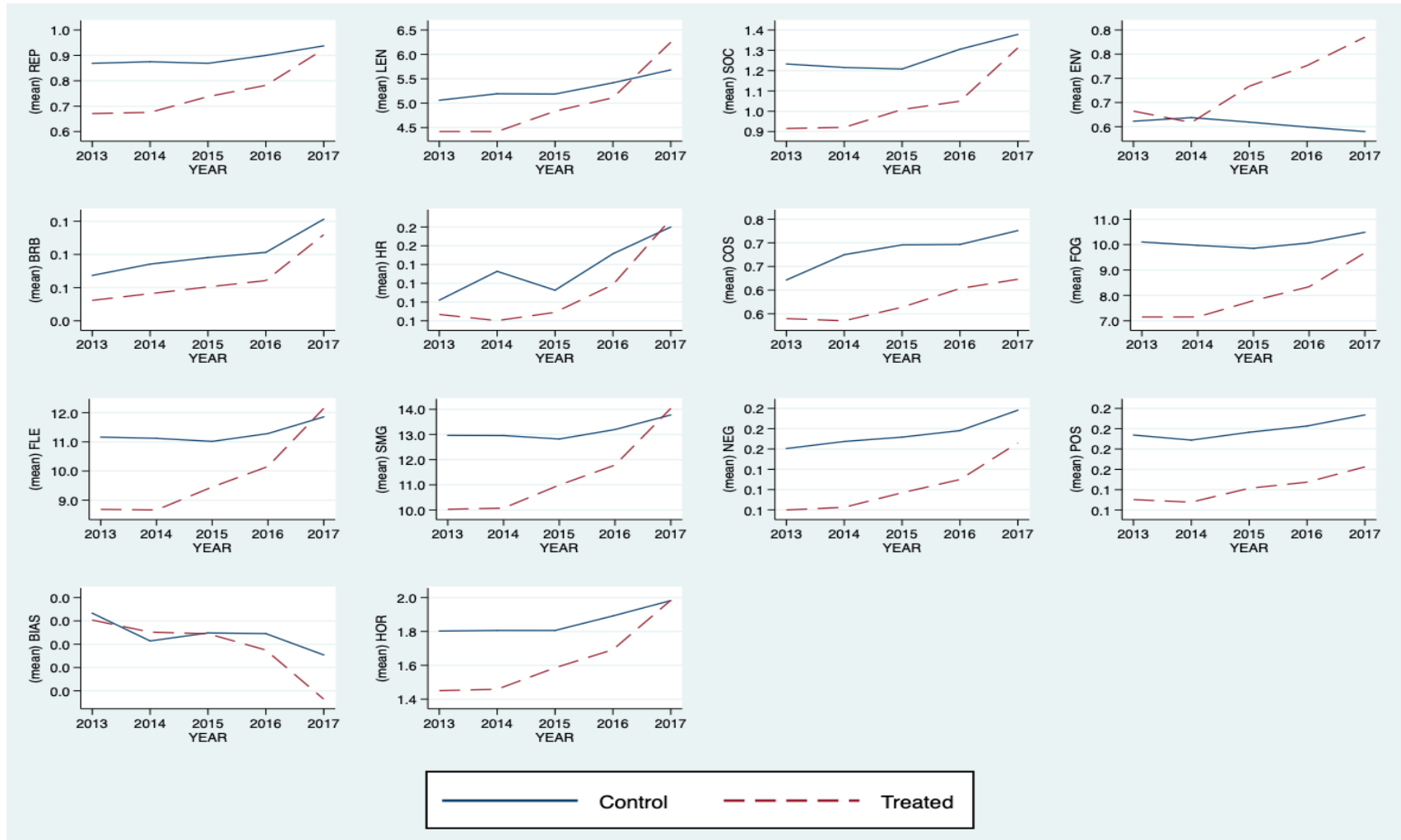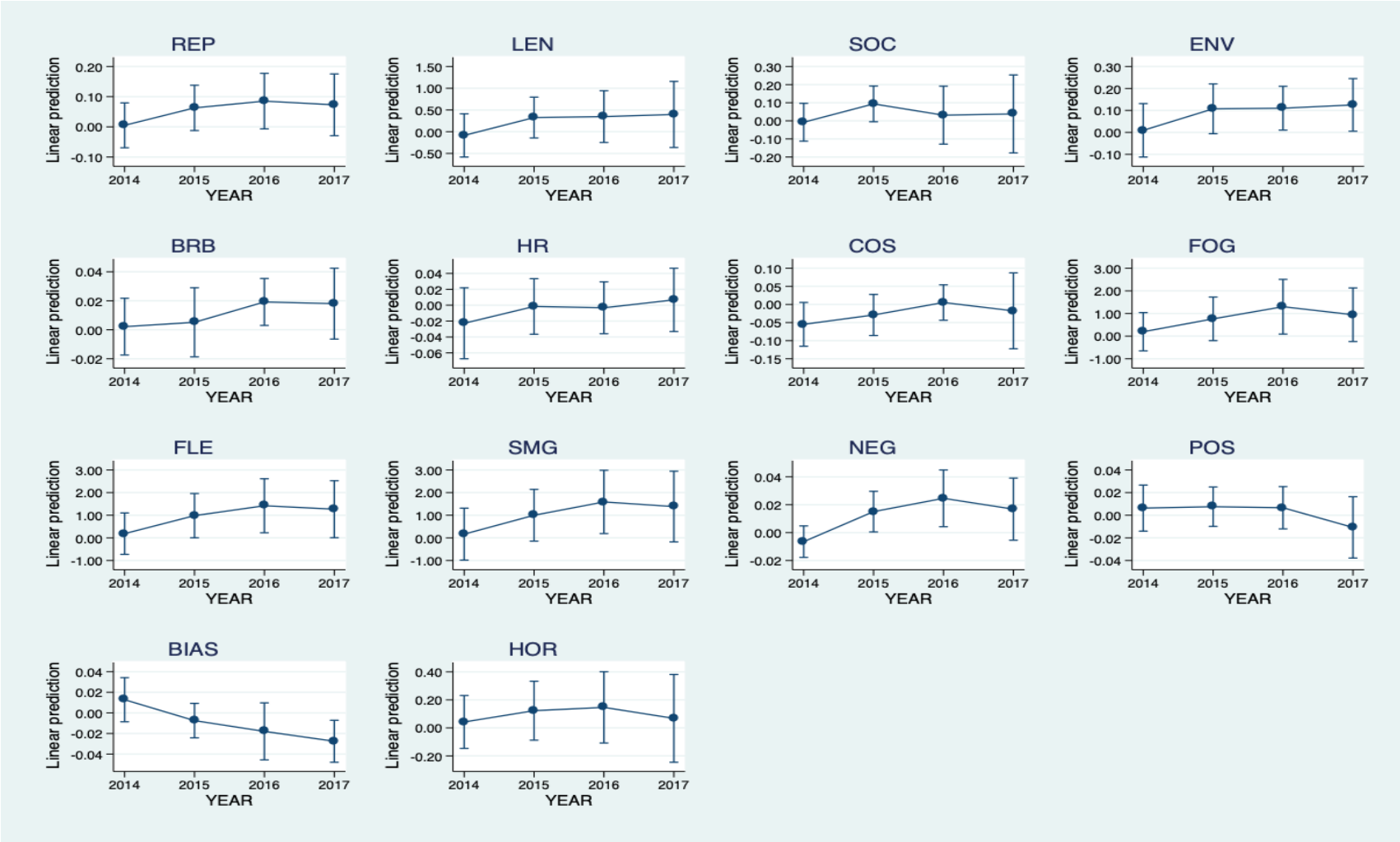**Figure 1** Parallel Trend Assumption: visual inspection

Figure 2 Parallel Trend Assumption: margins plot, formal testing

APPENDIX 1

CSR frameworks used for CSR vocabulary

Name	Description	Reference
AA1000	internationally accepted, principles-based framework and guidance that organizations can use to identify, prioritize and respond to sustainability challenges to improve long-term performance	Standard, 2015
UNGC	non-binding United Nations pact to encourage businesses worldwide to adopt sustainable and socially responsible policies, and to report on their implementation	United Nations, 2000
GRI	first global standards for sustainability reporting	GRI, 2013
ISO14001	specifies the requirements for an environmental management	ISO, 2015
SA8000	auditable certification standard that encourages organizations to develop, maintain, and apply socially acceptable practices in the workplace	Leipziger, 2017
SASB	sustainability accounting standards in disclosing material sustainability issues for the benefit of investors and the public	"Conceptual Framework of the Sustainability Accounting Standards Board", 2013
SDG	are a collection of 17 global goals designed to be a "blueprint to achieve a better and more sustainable future for all"	United Nations, 2019
Directive 2014/95/EU	EU law requiring large companies to disclose certain information on the way they operate and manage social and environmental challenges	European Commission, 2014
OECD	non-binding principles and standards for responsible business conduct in a global context consistent with applicable laws and internationally recognised standards	OECD, 2008
IIRC	framework for the representation of a company's performance in terms of both financial and sustainability information	Soyka, 2013
CDSB Framework	framework for reporting environmental and climate change information in mainstream reports, such as annual reports, 10-K filing, or integrated report	CDSB, 2019

Notes: This table presents a set of documents used for the selection of the CSR related words and word collocations.

APPENDIX 2

CSR Vocabulary

General
2014/95/eu
aa1000
ability to create value
accountability principle
corporate social opportunity
corporate social responsibility
corporate sustainability
creating sustainable value
csr
dialogue
directive (2014/95)
djsw
environmental and social
environmental social
eu non financial reporting
eu non-financial reporting
ftse4good
g4
g4 guidelines
global compact
gri
impact assessment
indicators g4
international ir framework
material non financial issue
material non-financial issue

multi stakeholder
multi-stakeholder
non financial
non-financial
non financial information
non-financial information
non financial statement
non-financial statement
non profit
non-profit
report sustainability
responsible business
responsible research
sasb
sasb standards
sdg
stakeholder
stakeholder engagement
sustainability
sustainability accounting standards board
sustainability context
sustainability issues
sustainability report
sustainability reporting
sustainability topics
sustainable business

sustainable development
un convention
un declaration
un guiding principles
ungc
voluntary
volunteer

Bribery
antibribery
anticorruption
bribe
bribery
collusion
corrupt
corruption
money laundering
sa800
whistleblowing

Environment
air emission
biodiversity
biomass
biosphere
climate change
climate related financial disclosures

co2
co2 emissions
conservation
deforestation
deplete
eco
eco friendly
ecofriendly
ecological
ecosystem
endangered
energy consumption
energy indirect scope
environmental impact
environmental information
environmental liabilities
environmental management
environmental policy
environmental protection
environmental responsibility
environmentally
environmentally friendly
environmentally responsible
extinction
fauna
flora
footprint

freshwater
ghg emissions
ghg protocol
global warming
green technology
greenhouse gas
hazardous waste
high biodiversity value
iso 14001
natural capital
nature friendly
nature-friendly
negative environmental
non renewable
non-renewable
oil spill
ozone depletion
ozone depleting
ozone-depleting
palm oil
pollutant
polluting
pollution
protected areas
protecting the environment
purification
radiation
rational use
recycle
recycled
recycling

renewable energy
renewable resources
resource conservation
resources institute wri
soil erosion
species extinction
task force on climate
toxic
user friendly
user-friendly
waste water
wastewater

Human Rights
asylum
bullying
business and human right
child labor
child labour
coercion
compulsory labor
cruel
data protection
dignity
discrimination
equal opportunities
equal rights
ethnic
ethnic minority
exile
female talent

forced labour
freedom
freedom of association
good labour practices
harassment
human rights
human rights impact
humanity
indigenous peoples
inequality
inhuman
intimidation
migrant
negative human rights
nondiscrimination
oppression
persecution
poverty
privacy
protect respect and remedy
racial
racial minority
refugees
servitude
sexual harassment
sexual orientation
slave
slavery
torture
tribal
tyranny

Social
absentee
absenteeism
abuse
attracting people
charitable
charitable activity
charity
community relationships
cultural
culture
disability
donate
donation
educational
educational institutions
employee engagement
employee matters
employee safety
employee turnover
employer
ethic
ethical
exit interview
family
fatalities
future generations
health and safety
health safety environment
hr management

human resources	safe working environment	outlook	14 year
ilo convention	safety	strategic	15 year
immoral	safety at work	objective	16 year
impact on society	schooling	target	17 year
injury	skills of employees	benchmark	18 year
international labour organization	social commitment	commitment	19 year
labor practices	social impact	prospect	20 year
labour convention	social policy	forward looking	one year
labour practices	social program	forward-looking	two year
local communities	social responsibility	coming year	three year
maternity leave	socially	following year	four year
moral	socially responsible	incoming year	five year
motherhood	staff training	next year	six year
motivated worker	talented workforce	subsequent year	seven year
ohsas 18001	training	upcoming year	eight year
organisational social responsibility	unemployment	long run	nine year
our people	unethical	looking ahead	ten year
participation committee	well being	1 year	eleven year
paternity leave	well-being	2 year	twelve year
personnel	widowhood	3 year	thirteen year
personnel attraction	young people	4 year	fourteen year
personnel management	young workers	5 year	fifteen year
philanthropic		6 year	sixteen year
philanthropy	**Horizon**	7 year	seventeen year
professional training	foreseeable	8 year	eighteen year
promoting talent	future	9 year	nineteen year
qualified personnel	long term	10 year	twenty year
retain talent	long-term	11 year	
reward and recognition	long run	12 year	
safe place to work	long-run	13 year	

APPENDIX 3

Details on parsing process with Python

1. Report length

   Apply "nltk sent_tokenize", to split text into sentences, remove punctuation and digits

2. Tone

   Apply "TfidfVectorizer" from sklearn to a list of all documents that splits texts into words, transforms to lower case and removes punctuation. This module assigns a weight to each token which depends on its frequency in a document and in the entire corpora

3. Readability

   Apply "textstat" from textstat to each document splitting text into words and transforming to lower case

4. Topic coverage

   Apply "TfidfVectorizer"

5. Forward-looking and long-term oriented vocabulary

   Apply "TfidfVectorizer"

6. Similarity

   Apply "CountVectorizer" from sklearn to a list of all documents that splits texts into words, transforms to lower case and removes punctuation. This module assigns a weight to each token equal to its frequency in a document

APPENDIX 4

Nomenclature

Symbol	Adjustment	Definition	Computation/Source
REP	none	Existence of CSR disclosure	Dummy variable, 1 when CSR disclosure exists, 0 otherwise
LEN*	logarithm	Length	number of sentences obtained using *nltk* package in Python
ENV	Taylor adj.	Environmental words	word count (see Appendix 10), obtained using *TfidfVectorizer* from the sklearn Python module
SOC	same	Social words	same
BRB	same	Bribery words	same
HR	same	Human rights words	same
HOR	same	Horizon words	same
TOP	same	Topics	mean(*ENV,SOC,BRB,HR,HOR*)
FOG	same	FOG readability index	see eq. 2
FLE	same	Flesh readability index	see eq. 3
SMG	same	SMOG readability index	see eq. 4
READ	same	Readability	mean(*FOG, FLE, SMOG*)
NEG	same	Negative words	Word count, Loughran & McDonald, 2011
POS	same	Positive words	Word count, Loughran & McDonald, 2011
BIAS	same	Bias	absolute(*POS-NEG*)
COS	same	Similarity of two texts	Cosine between two vectors of word counts
CSRQ	none	Textual score	Combination of *TOP, READ, BIAS, COS, LEN*
REV	none	Increase in revenues	Datastream
ROA	none	Return on assets	Datastream
MV	logarithm	Market value	Datastream
BM	none	Book to market	Datastream
BA	logarithm	Bidask	yearly average of the daily differences between the downloaded from Datastream ask and bid prices divided by their midpoint, Leuz & Verrecchia, 2000
VLT	logarithm	Earnings volatility	Datastream
VOL	logarithm	Trading volume	Datastream
ESG	none	CSR ratings	Asset 4

Notes: *when continuous textual indexes are converted to ranks on a scale from 1 to 10, we indicate it by adding "_10" symbol to the variable denomination

APPENDIX 5

Evidence on value - relevance of CSR disclosures

Author, years	Independent	Dependent	Country	Results
Clarkson et al., 2013	EPA TRI database	Cost of equity	USA	Positive association between TRI and cost of capital, none for the voluntary environmental disclosures
Egginton and McBrayer, 2019	Bloomberg CSR Disclosure score	Bid-ask spread	Score coverage	Better score associated with narrower spreads, results are more pronounced for firms without analyst following
Mittelbach-Hörmanseder et al., 2020	topic-specific disclosures measures	Share price	STOXX Europe-600	Positive relationship under voluntary regime, negative under mandatory
Verbeeten et al., 2016	Availability of CSR report, number of words	Share price	Germany	Marginally positive association of CSR report and share price
Dhaliwal et al., 2011	issuance of stand-alone CSR report	Cost of equity	USA	Negative association
Cho et al., 2013	KLD strength and concerns	Bid-ask spread	KLD coverage	Better bid-ask spreads for both strength and concerns
Matsumura et al., 2014	carbon emissions disclosure	Market value	S&P 500 firms	Negative association
Cormier et al., 2011	Disclosure index on environmental reporting	Market to book	Germany, France, Canada	Positive effect only in Germany
Plumlee et al., 2015	environmental disclosure index	Cost of equity	USA	Voluntary environmental disclosures are associated with cost of equity
Cahan et al., 2016	Proprietary data on CSR reporting	Tobin's Q	International	Positive association, moderated by country characteristics
Reverte, 2016	Proprietary data on CSR reporting	Share price	Spain	Positive association
Barth et al., 2017	Integrated reporting quality EY scores	among others: Bid-ask spreads, Tobin's Q, Cost of capital	South Africa	Positive association except for cost of capital and analysts' error
Ioannou and Serafeim, 2017	Bloomberg CSR Disclosure score	Tobin's Q	China, Denmark, Malaysia, and South Africa	Increases in sustainability disclosure driven by the regulation are value-relevant