

THE TONE CHANGE IN EARNINGS CONFERENCE CALLS AS ESG
DISCLOSURES BECOME MANDATORY



ZÜLAL SEZİCİ

BOĞAZİÇİ UNIVERSITY

2025

THE TONE CHANGE IN EARNINGS CONFERENCE CALLS AS ESG
DISCLOSURES BECOME MANDATORY

Thesis submitted to the
Institute for Graduate Studies in Social Sciences
in partial fulfillment of the requirement for the degree of

Master of Arts
in
International Trade Management

by
Zülal Sezici

BOĞAZİÇİ UNIVERSITY

2025

The Tone Change in Earnings Conference Calls
as ESG Disclosures Become Mandatory

The thesis of Zülal Sezici
has been approved by:

Assoc. Prof. Gözde Erhan
(Thesis Advisor)

Assoc. Prof. Ali oşkun
(Thesis Co-Advisor)

Assist. Prof. Mehtap Işık

Prof. Neslihan Yılmaz

Assist. Prof. Erdiñ Akyıldırım
(External Member)

May 2025

DECLARATION OF ORIGINALITY

I, Zülal Sezici, certify that

- I am the sole author of this thesis and that I have fully acknowledged and documented in my thesis all sources of ideas and words, including digital resources, which have been produced or published by another person or institution;
- this thesis contains no material that has been submitted or accepted for a degree or diploma in any other educational institution;
- this is a true copy of the thesis approved by my advisor and thesis committee at Boğaziçi University, including final revisions required by them.

Signature.....

Date.....

ABSTRACT

The Tone Change in Earnings Conference Calls as ESG Disclosures Become Mandatory

This thesis investigates whether mandatory Environmental, Social, and Governance (ESG) disclosure regulations influence the tone of corporate communication during earnings conference calls (ECCs). Using a large global panel of transcripts from 16,327 firms across 95 countries (2005–2024), the study employs advanced natural language processing tools to quantify tone. Specifically, FinBERT-Tone is used to detect sentence-level sentiment (positive, neutral, negative), while FinBERT-ESG and ESG-BERT models classify ESG-related content across environmental, social, and governance dimensions. For each ECC, the most optimistic ESG-related sentence is extracted to construct ESG_Tone_max, E_Tone_max, S_Tone_max, and G_Tone_max indicators. A Difference-in-Differences (DiD) approach compares tone changes in firms from countries with ESG mandates (treatment group) to those without (control group), while controlling for firm, year, and sector-year fixed effects.

Results indicate a statistically significant increase in social tone (S_Tone_max) following ESG mandates, whereas environmental and governance tones show no robust change. Overall ESG tone improves slightly but is sensitive to model specification. Firms with stronger ESG performance consistently use more positive ESG language. This study bridges NLP and causal inference to show that regulation affects not only what firms disclose but also how they speak about ESG, enriching the literature on corporate communication and sustainability.

ÖZET

Zorunlu ESG Açıklamalarıyla Birlikte Kazanç Duyuru Toplantılarında Ton Değişimi

Bu tez, çevresel, sosyal ve yönetim (ESG) konularında zorunlu açıklama düzenlemelerinin, şirketlerin kazanç açıklama toplantılarındaki (ECC) iletişim tonunu etkileyip etkilemediğini incelemektedir. 95 ülkedeki 16.327 firmadan alınan büyük bir küresel transkript panelinin kullanıldığı çalışmada (2005–2024), tonlamayı ölçmek için gelişmiş doğal dil işleme araçları kullanılmıştır: FinBERT-Tone modeli cümle düzeyinde duyguyu (olumlu, nötr, olumsuz) belirlerken, FinBERT-ESG ve ESG-BERT modelleri çevresel, sosyal ve yönetim içeriklerini sınıflandırmaktadır. Her ECC metninden ESG ile ilgili en olumlu cümle seçilerek ESG_Tone_max, E_Tone_max, S_Tone_max ve G_Tone_max göstergeleri oluşturulmuştur. Farklılaştırılmış-etkiler (Difference-in-Differences) yöntemiyle, ESG düzenlemesi getiren ülkelerdeki firmalar (tedavi grubu) ile bu tür zorunluluk bulunmayan ülkelerdeki firmalar (kontrol grubu) karşılaştırılmıştır. Analizlerde firma, yıl ve sektör-yıl sabit etkileri kontrol edilmiştir. Sonuçlar, sosyal tonun (S_Tone_max) ESG düzenlemeleri sonrasında anlamlı biçimde arttığını, çevresel ve yönetim tonlarında ise belirgin bir değişim olmadığını göstermektedir. Genel ESG tonu (ESG_Tone_max) hafifçe artmakta, ancak bu etki model duyarlılığına bağlıdır. ESG performansı yüksek firmalar daha olumlu bir dil kullanmaktadır. Bu çalışma, NLP ve nedensel çıkarımı birleştirerek, düzenlemelerin yalnızca şirketlerin neleri ifşa ettiğini değil, aynı zamanda ESG hakkında nasıl konuştuklarını da etkilediğini göstererek kurumsal iletişim ve sürdürülebilirlik literatürünü zenginleştirmektedir.

ACKNOWLEDGMENTS

This thesis has been a journey shaped not only by academic effort but also by countless moments of kindness, patience, laughter, and quiet resilience. I would like to extend my heartfelt gratitude to those who stood by me along this journey of life. First and foremost, I would like to thank my advisors, Dr. Gözde Erhan and Dr. Ali Coşkun, for their invaluable guidance, unwavering support, and generous mentorship. I am also deeply grateful to Dr. Erdinç Akyıldırım, whose thoughtful insights enriched this study and encouraged me to think beyond the frame. I am especially thankful to Melih, to whom I could always turn for technical questions and practical advice throughout this process.

To my dear colleagues and companions — Cansu, Melih, Fatih, and Moşe — thank you for walking this path with me. Your presence made the road as beautiful as the destination. To Merve, Mehmet, İslam, Gülşah, and the entire Department of International Trade — thank you for every kind of support and solidarity.

To my soul-sisters Rüya and Bori — for always being there, no matter the distance. And to Elif — our true friendship began in a Chinese restaurant in Ljubljana, where a shared meal opened the door to a friendship that never let go.

To Asiye and Ayşe — thank you for your delicate souls and unwavering support. To Mehtap — for always showing up even before I asked. Your presence made all the difference.

To our amazing MA team — İnci, Bika, Fatma, and Nehir — we wouldn't have succeeded without believing in one another. To Selin, Muteber, and Meryem — your emotional support was a quiet force that lifted me when I needed it most.

To Tansu, Esra, Merve, and little Duru — thank you for brightening my days with joyful nostalgia and affection. To all the members of the Çelik (Dupljak) and Sezici families — thank you for always being there for me and for all the memories we share.

To my grandfathers — Yusuf Çelik (Dupljak) and Yunus Sezici — two strong migrant men. I miss you both. Your strength still guides me. To my brother Efe — you will always be the golden-haired little boy I adore. To my mother Fatma — your strength, love, and unrealized artistic dreams inspire me endlessly. I wish I could have given you a life full of recognition and joy — not just for you, but for all women. To my father — thank you for teaching me how to deal with boys if one messed with me in primary school. To my aunts — I send my deepest and most heartfelt thanks. Your presence has always felt like home.

To my dear friend Şüheda — for your tender spirit and quiet strength. To Melike and Sevilcan — thank you for being my home away from home. To Emre, Gizem, Aleyna and Hande thank you for the friendship that stretched from Avcılar Campus to Berlin to Eskişehir, and for every shared moment along the way.

Finally, to the art of cinema and photography — especially Scandinavian cinema; to Balkan songs; to Russian novels and Greek melodies; to Mediterranean food; to green and to blue. To Yabancı Damat, to the Oslo Trilogy; to Rome, Vienna, Oslo, and Bodrum. To the sea — and to the sea again. And most of all — to you.

I dedicate this thesis to my beautiful sisters, *Ayşenur Halil* and *İkbal Uzunur*, whose lives were stolen too soon. This work is also for all women silenced before their time — may their stories be heard, and their light never fade.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION	1
CHAPTER 2: LITERATURE REVIEW	3
2.1 From voluntary to mandatory esg disclosure	3
2.2 Earnings conference calls (eccs) as a vehicle for ESG dialogue.....	5
2.3 Measuring tone: from dictionary counts to domain-specific BERT models	7
CHAPTER 3: PREVIOUS STUDIES, RESEARCH DESIGN & HYPOTHESES	11
3.1 Previous empirical findings mandatory ESG disclosure and firm behaviour .	11
3.2 Theoretical framework	13
3.3 Research design.....	17
3.4 Research question and hypotheses	23
CHAPTER 4: DATA & METHODOLOGY	26
4.1 Data	26
4.2 Methodology	32
CHAPTER 5: ANALYSIS AND EMPIRICAL RESULTS	37
5.1 Descriptive statistics.....	37
5.2 Correlation matrix	38
5.3 Parallel trends verification.....	39
5.4 Main regression results.....	43
5.5 Placebo tests	50
CHAPTER 6: CONCLUSION & DISCUSSION	54
CHAPTER 7: LIMITATIONS & FURTHER STUDIES.....	56
REFERENCES.....	58

LIST OF TABLES

Table 1. Variable Definitions and Sources	31
Table 2. Descriptive Statistics of Full Sample	37
Table 3. Correlation Matrix of Key Variables	39
Table 4. Main DiD Results: Impact on Overall ESG Tone	45
Table 5. Main DiD Results: Impact on Environmental Tone	47
Table 6. Main DiD Results: Impact on Social Tone	47
Table 7. Main DiD Results: Impact on Governance Tone.....	48
Table 8. Placebo Test Results: Impact on Overall ESG Tone with Alternative Intervention Years.....	52
Table 9. Placebo Test Results: Impact on Environmental Tone with Alternative Intervention Years.....	52
Table 10. Placebo Test Results: Impact on Social Tone with Alternative Intervention Years.....	53
Table 11. Placebo Test Results: Impact on Governance Tone with Alternative Intervention Years.....	53

CHAPTER 1

INTRODUCTION

Environmental, Social and Governance (ESG) issues, which have become an important part of corporate strategies, are rapidly gaining importance in line with the increasing regulatory pressures and stakeholder demands on a global scale. Earnings Conference Calls (ECC), which can be communicated directly with investors and analysts, stand out as one of the main communication channels through which companies convey not only their financial performance but also their strategic and operational developments in the ESG field (Huang et al., 2022).

Regulations such as the Non-Financial Reporting Directive (NFRD) and later the Corporate Sustainability Reporting Directive (CSRD) implemented in the European Union have significantly transformed corporate transparency and accountability. These regulations force companies to report their ESG performance in an open, systematic and reliable manner. As a result, companies operating in countries where such mandatory ESG reporting regimes are in effect are under pressure to convey their sustainability performance in a more positive and strategic tone. This situation reveals the need for new research on how the content and tone of public disclosures change (Amel-Zadeh & Serafeim, 2018; Kaplan & Ramanna, 2021).

This study aims to analyze how mandatory ESG regulations implemented at the country level affect the tone of the language used in ECCs. In this study, which uses advanced natural language processing (NLP) techniques such as FinBERT and ESG-BERT models, the general ESG tone and the changes in environmental, social

and governance dimensions in the pre- and post-regulation periods are examined separately. The study draws its theoretical basis from signaling, impression management and legitimacy theories. In line with these theories, it is assumed that mandatory ESG reporting leads companies to adopt a more positive and strategic communication language (Eccles et al., 2014; Matsumoto et al., 2011).

The empirical analysis was conducted with ECC data from 16,327 companies from 95 countries covering the period 2005–2024. The Difference-in-Differences (DiD) model was used as the analysis method. With this method, companies in countries that made ESG reporting mandatory were compared with companies in countries where there was no such obligation, and the tone differences before and after the regulation were evaluated. In addition, the focus was on tone changes separately in environmental, social and governance dimensions.

This study makes significant contributions to the literature on understanding the effects of ESG regulations on companies' communication strategies. By analyzing the tone change in ECCs, it reveals how companies respond to regulatory frameworks and provides significant findings for policy makers, investors and managers.

CHAPTER 2

LITERATURE REVIEW

2.1 From voluntary to mandatory ESG disclosure

Environmental, Social, and Governance (ESG) disclosures initially emerged within the framework of voluntary stakeholder-focused standards that companies adopted voluntarily. Standards such as the Global Reporting Initiative (GRI), established in the late 1990s, and the Sustainability Accounting Standards Board (SASB), established in 2011, have guided companies to report their sustainability performance transparently.

These voluntary efforts have gained importance with the rise of ethical investments and the strengthening of corporate social responsibility movements, and many companies have begun to report on ESG issues under stakeholder pressure and reputation management motivation.

Over time, regulatory authorities have realized that ESG data plays a critical role in evaluating company performance and have taken steps to make reporting mandatory. The European Union's Non-Financial Reporting Directive (NFRD, 2014) and the subsequent Corporate Sustainability Reporting Directive (CSRD, 2022) are milestones in this area. As of 2018, the NFRD required large-scale European companies to publish reports containing environmental and social information, covering approximately 6,000 companies and establishing a minimum standard for ESG disclosures.

The CSRD expanded the scope even further, including tens of thousands of companies across the EU in the scope of sustainability reporting, requiring

independent auditing of reporting, and introducing detailed ESG standards.

Similarly, some other countries and regulatory bodies have joined this trend and have begun to make voluntary frameworks, such as the TCFD (Task Force on Climate-related Financial Disclosures), mandatory. For example, the United Kingdom has made TCFD-compliant climate reporting mandatory for large companies as of 2021.

This evolutionary process indicates that sustainability reporting has moved from a niche practice to a globally expected corporate standard. The global implementation of mandatory reporting regulations has enabled companies to disclose a wider range of information and comply with stakeholder expectations. However, differences in implementation timing and inconsistencies in monitoring mechanisms across countries continue to present significant challenges (Eccles & Serafeim, 2013). For example, while the EU has taken a leading role in this regard, a comprehensive mandatory ESG reporting standard has not yet been enacted in the US; regulatory guidance has been left to industry initiatives.

As a result, some companies report under strict legal requirements, while others still rely on a voluntary basis. This creates problems in data comparability for investors and companies operating internationally. When historical development is examined, it is seen that there are different approaches to the design of rules: Some regulations include flexible mechanisms such as “comply or explain”, while others demand strict compliance with detailed metrics and sanctions.

Early findings in the literature on the effects of mandatory ESG disclosures generally indicate positive results. Mandatory reporting increases the transparency of companies and increases the reliability of disclosed ESG information. In a comprehensive study, Ioannou and Serafeim (2017) showed that mandatory

sustainability reporting encourages more responsible management practices within the firm, significantly increases the level of ESG disclosure, and increases the likelihood that companies will submit this information to third-party verification. In terms of market reactions, the same study indicates that companies that improve their ESG performance after the regulation and report this improvement have increased their market value. Similarly, in their analysis of global data sets, Christensen, Hall, and Leuz (2021) found that ESG disclosure requirements increase stock liquidity at the firm level and improve the information content of price formation. In other words, investors can make more effective pricing by reflecting the additional information provided by mandatory reporting into prices.

In summary, the transition from a voluntary basis to a mandatory system has transformed ESG reporting in terms of both scope and reliability, and the general consensus in the literature is that appropriately designed mandatory disclosures create value for both stakeholders and firms.

2.2 Earnings conference calls (ECCs) as a vehicle for ESG dialogue

Earnings Conference Calls (ECCs) are important communication platforms that take place between company senior management and financial analysts, and where quarterly financial results are usually discussed. In ECCs, management presents quarterly performance and future expectations, while analysts ask questions about company performance.

The place of ESG dialogue in ECCs has increased significantly in recent years. While historically these discussions focused on financial performance and operational issues, analysts began to ask ESG-related questions in ECC sessions towards the end of the 2010s, as investors' interest in sustainability and climate risks

increased. It has been observed that ESG themes such as climate change, carbon emissions, and diversity policies have rapidly increased on the ECC agenda, especially after 2019 (Dzieliński et al., 2022; Sautner et al., 2023). For example, Dzieliński et al.'s (2022) study documented a dramatic increase in the frequency of climate-related words used in ECCs after the 2015 Paris Agreement and increasing regulatory pressures around 2018. It was noted that in ECCs in many sectors during this period, analysts directly asked management questions such as “Does your company have a carbon neutrality target?” or “How do you assess the impact of climate regulations on your financials?” Similarly, Sautner et al. (2023) found a significant increase in the number of words related to ESG in ECC transcripts of large-capital companies in the last few years and interpreted this as a reflection of investor pressure and regulatory signals.

In short, ECCs are evolving from platforms where only financial results are discussed to a comprehensive communication channel where companies also share their ESG performance and approach. The increasing importance of ESG issues in ECCs means that the tone of managers in these areas has also gained importance. Investors and analysts interpret the tone of language in the answers given by managers to ESG questions as well as the content of their answers. For example, if a manager responds defensively (e.g., by changing the subject or using vague expressions) to an analyst asking a concerned question about climate risk, this can create a negative perception in the market.

Indeed, Francis (2008) showed that the narrative tone in the public statements of senior managers can shape investor reactions; a positive and proactive tone can build trust, while a defensive or evasive tone can create doubt and uncertainty.

Similarly, Alta'any (2020) stated that a positive narrative style in ECCs can have a positive impact on stock returns in the short term, but an overly optimistic or unrealistic discourse can lead to erosion of confidence in the later period.

According to these findings, examining the ESG dialogue in ECCs in terms of both content and tone has become critical to understanding the impact of companies' sustainability communication on the market.

2.3 Measuring tone: From dictionary counts to domain - specific BERT models

Sentiment analysis in financial texts has been carried out using traditional methods for a long time. Initially, dictionary approaches based on the number of positive and negative words in documents such as company statements or ECC transcripts were widely used. One of the most well-known methods in the field of finance is the Loughran–McDonald word lists, developed specifically for financial texts. This dictionary aims to reflect the tone in financial reports more accurately by eliminating some words that create errors in general-purpose sentiment analysis (for example, the word “debt” is considered negative in a general sentiment lexicon, but neutral in a financial context) (Loughran & McDonald, 2011).

Early studies on tone measurement in ECCs also used such dictionary-based counting methods. For example, the ratio of the number of positive words to the number of negative words in a ECC transcript was interpreted as an indicator of the level of optimism. Using such an approach, Druz et al. (2020) found that the increase in negativity in managers' language was associated with future stock returns, and that the market initially failed to fully price these signals after periods when managers used a pessimistic tone.

However, dictionary approaches are limited in capturing the meaning of words in context. Especially in new topics such as ESG, which may have different meanings depending on the company context, simple positive/negative word lists cannot capture the nuances. For example, while the frequency of words such as “carbon” or “emission” does not indicate a positive or negative situation on its own, the expression “decrease in carbon emissions” should be evaluated together with its context. For this reason, dictionary-based methods have been inadequate in tone analyses in the ESG context (Druz, 2020).

Today, with the development of artificial intelligence and deep learning techniques for language processing, much more advanced methods are used to detect tone and emotion in texts. Transformer-based models, in particular, have brought revolutionary innovations. Standing out among these models, BERT (Bidirectional Encoder Representations from Transformers) offers a bidirectional learning architecture that enables contextual understanding of language. Since the BERT model can determine the meaning of a word in a sentence by looking at both the previous and following words, it is extremely successful in capturing the nuances of linguistic tone in the text.

Recent advances in Natural Language Processing (NLP) have enabled researchers to extract more precise and context-sensitive insights from financial and sustainability-related texts. In this thesis, three distinct transformer-based models were employed to capture both tone and ESG-specific content in Earnings Conference Calls (ECCs).

FinBERT-Tone, introduced by Huang et al. (2023), is a BERT-based model trained on financial texts to classify sentences into positive, negative, or neutral tone.

Unlike traditional lexicon-based methods, this model can capture tone nuances by considering the full sentence context. The same authors also developed FinBERT-ESG, which classifies text into Environmental (E), Social (S), Governance (G), or None categories.

To address these limitations, this study also used a more specialized set of models developed by Schimanski et al. (2024), known as ESG-BERT. These include EnvironmentalBERT, SocialBERT, and GovernanceBERT, each fine-tuned separately using ESG-specific datasets with over 13 million sentences. According to the authors, this domain-specific training significantly improves classification accuracy compared to more general models such as FinBERT-ESG.

By combining FinBERT-Tone with both FinBERT-ESG and ESG-BERT classifiers, this thesis conducts a dual-layer analysis: it measures the sentiment of corporate communication and identifies the presence, density, and distribution of ESG-related content. This integrated NLP approach enables a more detailed and robust examination of how firms address Environmental, Social, and Governance topics under varying regulatory frameworks.

Transformer-based models are especially useful for capturing tone in complex topics like ESG, where meanings vary across firms. These models can interpret sentences with mixed sentiments and detect nuances such as idioms or implicit meanings that dictionary-based methods often miss. As a result, tools like FinBERT and ESG-BERT have become more common in recent studies analyzing ESG tone in earnings calls.

Yet, there is still a significant gap in the literature when it comes to integrating state-of-the-art natural language processing (NLP) tools with causal

inference methods. While transformer-based models such as FinBERT and ESG-BERT have enhanced the precision of textual analysis in finance and sustainability research, they are rarely combined with rigorous econometric techniques. Few studies employ advanced language models alongside Difference-in-Differences (DiD) designs to evaluate the impact of regulatory changes on firm behavior. This thesis addresses that methodological gap by embedding NLP-based tone and content measures into a DiD framework. Through this integrated approach, the study offers a novel contribution by demonstrating how mandatory ESG disclosure affects not only what firms say, but also how they say it. The findings contribute to a deeper understanding of the role of regulation in shaping corporate narrative strategies, thereby bridging the gap between textual analysis, financial communication, and policy evaluation.

CHAPTER 3

PREVIOUS STUDIES, RESEARCH DESIGN & HYPOTHESES

3.1 Previous empirical findings mandatory ESG disclosure and firm behaviour

There has recently been an increasing number of studies in the literature on the effects of mandatory ESG disclosures on firm behavior and the market. Early findings suggest that mandatory sustainability reporting creates greater transparency and accountability in internal processes and positively impacts capital markets. For example, multi-country studies have shown that mandatory ESG reporting significantly increases the level of ESG disclosure, increases the likelihood that companies will have their data independently verified in this area, and even encourages firms to voluntarily adopt international reporting guidelines (such as the GRI) (Ioannou & Serafeim, 2017).

When looking at the financial impacts of mandatory reporting regulations, positive market outcomes are also generally found. A global analysis by Christensen et al. (2021) found that liquidity and price discovery accelerated in the stocks of firms subject to ESG disclosure requirements following the implementation of these requirements. Similarly, Krueger et al. (2023) studies indicate that ESG reporting requirements around the world reduce uncertainty in firm valuations and lead to a decrease in information asymmetry.

In short, the literature is dominated by findings that mandatory ESG disclosures generally create a confidence-building effect on investors, encourage companies to improve their sustainability performance in the long term, and contribute to healthier price formation in the markets.

3.1.1 Earnings - conference - call (ECC) tone research

The Earnings Conference Call (ECC) literature has also provided rich findings on the importance of tone of language in recent years. The tone of language used by senior executives in ECCs has measurable effects on investor reactions and market dynamics. For example, studies have shown that the optimistic or pessimistic expression of executives during ECCs affects the movements in stock prices after the meeting.

Francis (2008) established a relationship between the attitude of the language used by senior management when providing public information and investor behavior, and showed that the statements of executives who speak in a positive tone generally receive a positive market reaction, whereas an uncertain or defensive tone can create distrust in the market and lead to a negative reaction. It has been found that ECC tone can also affect information asymmetry; managements that communicate with more open and positive language can fill in the gaps in analysts' and investors' knowledge about the company to some extent, thus reducing expectation divergences. Alta'any (2020) examined ECC records and emphasized that even when companies' financial performance is good, when management deliberately adopts a more cautious tone (for example, as a requirement for expectation management), market reactions remain more measured and tone information is taken into account by investors.

In addition, studies such as Druz et al. (2020) found that the negative tone increase in management's ECC is not fully priced in by the markets immediately, but this tone signals future risks and can therefore predict future returns. These results suggest that ECC tone information has become a critical component of financial

analysis and valuation. Empirical studies on ECC dialogues specific to ESG issues are just beginning to emerge. Dzieliński et al. (2022) and Sautner et al. (2023) are important studies documenting the increasing trend of ESG dialogue in ECCs. Michał Dzieliński and colleagues examined how statements about climate change and environmental sustainability in ECCs increased over time and under what conditions this increase occurred. Their findings show that managers began to address climate issues more, especially after 2015, and that this trend accelerated by 2019-2020.

Furthermore, Dzieliński et al. found that firms that emphasized climate issues more in their ECCs tended to take concrete steps to reduce carbon emissions in the subsequent period. This examines the relationship between whether companies “talk about climate” and “act on climate,” suggesting that there are situations where words turn into action. On the other hand, Sautner et al. (2023) reported in their study that questions from investors and analysts about ESG issues have increased significantly in recent years, and that there are differences between sectors in the answers given by managers to these questions.

3.2 Theoretical framework

Multiple theoretical perspectives guide us in examining ESG disclosures and managers’ communication tone. Three main theoretical frameworks stand out in this study: Signaling Theory, Impression Management Theory, and Legitimacy Theory.

Signaling Theory was initially developed to explain the ways in which parties send reliable signals to the other party under asymmetric information conditions in economic markets (Spence, 1973). In the institutional context, this theory suggests

that corporate management can use various indicators to convey its private information to external stakeholders. In financial communication, tone of language can be one of these indicators. When managers want to give a positive signal about their company's future performance or the real situation in sustainability issues, they can send a message to the market using more optimistic language. Companies with strong ESG performance and who really care about sustainability may want to signal this strong stance by adopting a positive and constructive tone on platforms such as ECC. For example, following the implementation of mandatory ESG disclosures, the CEO of a firm recognized as a leader in ESG may adopt a positive tone during an ECC, highlighting the company's sustainability initiatives to signal its strong performance to investors. According to Signalling Theory, such a positive tone can only be used convincingly by firms that truly perform well; otherwise, the "empty signal" loses its credibility over time. Therefore, a positive tone is expected to be a signalling tool that indicates sustainability commitment (Eccles, Newquist, & Schatz, 2014).

In the context of our study, if the signalling effect is dominant, firms with high ESG performance are expected to use a more positive tone in their ESG statements after mandatory ESG regulations, thus sending distinctive signals to the market.

Impression Management Theory suggests that managers can act strategically in their communications to influence the perception of external stakeholders (Merkl-Davies & Brennan, 2007). According to this theory, corporate executives may use the flexibility of language to portray their performance as better than it actually is or to reduce the impact of negative news. To example, it has been documented in the

financial reporting literature that executives gloss over failures with vague expressions or emphasize successes with exaggerated language as a common impression management tactic.

ECCs are also environments where impression management is staged live. To soften the negative assessments that may arise, especially after mandatory ESG disclosures, management may use more cautious or rosy language in ECCs. For example, if the company's carbon emissions are high as a result of mandatory reporting, management may address this issue with strategic optimism in the ECC and try to turn the perception into a positive one by emphasizing, "Although our emissions are above the industry average, we have set aggressive reduction targets and are making significant investments." From an Impression Management perspective, companies with relatively weak ESG performance are likely to engage in damage control through more verbal manipulation and positive framing in ECC speeches after mandatory ESG data is disclosed. This is a phenomenon that can perhaps be detected in tone analysis as a mismatch between performance reality and tone. Therefore, the impression management motivation creates an expectation that some companies will try to influence investor perception by using overly optimistic or overly defensive language on ESG issues after mandatory disclosure.

Legitimacy Theory emphasizes that the long-term survival of companies depends on their social acceptance and legitimacy (Suchman, 1995; Deegan, 2002). According to this theory, companies try to establish corporate legitimacy by aligning their activities and communications with the norms, values, and expectations of the society they are in.

ESG issues have become an important part of social expectations today. Responsible stances of companies on issues such as climate change, social justice, and labor rights are not only a regulatory requirement, but also an expectation in the eyes of the public. Mandatory ESG disclosures are, in fact, the reflection of these social expectations on the corporate world through regulation. From a Legitimacy Theory perspective, companies feel the need to communicate more about ESG issues, not only because it is required by law or regulation, but also to be perceived as a legitimate and acceptable institution. This situation is also reflected in ECCs: By talking about sustainability and responsibility issues and voicing their efforts in these areas, managers give the message to society and investors that “we are also complying with the norms, doing our part”. Especially large and visible companies may take care to defend or emphasize their ESG performance in ECCs to protect their corporate citizenship image.

In terms of tone, the search for legitimacy can be manifested by managers using a balanced and reputation-protective language on ESG issues. It can be seen that they adopt a moderate tone to take a legitimate position in the eyes of stakeholders, neither over-bragging (which would damage credibility) nor ignoring the issue (which could also draw a reaction).

These three theories guide the formation of the hypotheses of our study. While Signalling Theory supports the expectation that firms with strong ESG performance will exhibit a more positive ESG tone in ECCs after mandatory reporting, Impression Management Theory suggests that those with poorer ESG performance may use similarly positive language (perhaps independently of actual performance) to improve their image. Legitimacy Theory suggests that, in general,

all firms will feel the need to speak up and express themselves more about ESG due to societal pressures and norms, and that this may affect tone. When these theoretical frameworks are considered together, it is thought that there will be a significant change in the time and tone devoted to ESG issues in ECCs after mandatory ESG disclosures come into effect; however, the direction and nature of this change may differ depending on firms' intrinsic motivations and performance levels. Our study aims to empirically test how ESG-related themes are reflected in the tone and content of earnings conference calls, in line with theoretical expectations.

3.3 Research design

Treatment definition: Country - level start year of mandatory ESG reporting (e.g., NFRD/CSRD in the EU; analogous rules elsewhere).

Control group: Firms headquartered in countries without a mandate (or pre - adoption years).

The main estimation equation is specified as follows:

$$\text{Tone}_{it} = \beta_1(\text{Treated}_i \times \text{Post}_t) + X_{it}\gamma + \alpha_i + \delta_t + (\alpha \times \delta)_{st} + \varepsilon_{it} \quad (1)$$

- Tone_{it} : ESG tone score of firm i in year t , measured using BERT-based sentiment models.
- Treated_i : Dummy variable equal to 1 if firm i is headquartered in a country with mandatory ESG regulation.
- Post_t : Dummy variable equal to 1 for years after the regulation takes effect.
- X_{it} : Vector of control variables: ROA, ROE, Leverage, log(Firm Size), ESG Score.
- α_i : Firm fixed effects.
- δ_t : Year fixed effects.
- $(\alpha \times \delta)_{st}$: Sector-year fixed effects.
- ε_{it} : Error term.

Difference – in – Differences specification:

Outcome variables: BERT - based tone scores: ESG_Tone_max, E_Tone_max, S_Tone_max, G_Tone_max.

Validation: Parallel - trends diagnostic; placebo years 2010, 2016, and a random placebo.

In this study, a quasi-experimental design (Difference-in-Differences) was used to examine how mandatory ESG reporting regulations affect the tone and frequency of ESG-related content in earnings conference calls.

In our study, the treatment group consists of companies in countries that implemented mandatory ESG reporting regulations during the study period. For example, since European Union countries are subject to such a requirement under the NFRD/CSRD, they were included in the treatment group.

The control group consists of companies in countries with similar characteristics but without mandatory ESG reporting during the same time period. In this way, the regulatory shock experienced by firms in the treatment group is compared with the situation in which the control group was not exposed, and whether the observed change is really due to the regulation is analysed.

In simple terms, the difference between “the change in ECC tone of firms in countries that introduced mandatory reporting after this rule” and “the tone change of firms not subject to this rule in the same period” is examined. If the regulation has an effect, the change in tone in the treatment group is expected to be statistically significantly different (i.e. more positive or negative) compared to the control group. The basic DiD model of the study is as follows:

$$\text{Tone}_{it} = \beta_1 \cdot (\text{Treated}_i \times \text{Post}_t) + X_{it} \cdot \gamma + \alpha_i + \delta_t + (\alpha \times \delta)_{\{s\}t} + \varepsilon_{it}$$

The main independent variable in this model is the interaction term $\text{Treated} \times \text{Post}$, which captures the effect of mandatory ESG disclosure policies.

In this equation, Tone_{it} is the tone indicator of firm i measured in ECC in period t . Treated_i is a fixed indicator indicating whether firm i is in the treatment group (i.e., under a country/regulation with mandatory ESG regulation). Post_t is a time indicator indicating observations after the period when the regulation came into effect. The product of this two ($\text{Treated} \times \text{Post}$) term is the key component of the DiD design and is multiplied by the coefficient β_1 , which measures the additional effect of being in the treatment group after the adjustment.

In the model, the vector X_{it} contains other control variables that may have an impact on tone, and their coefficients are denoted by γ . α_i represents firm fixed effects, and δ_t represents year (time) fixed effects. The additional term $(\alpha \times \delta)_{\{s\}t}$

is included in the model and represents industry-year fixed effects – that is, fixed effects added to control for common factors specific to a particular year in a particular industry. ε_{it} is the error term and includes other effects that cannot be included in the model. The scope and variables of the model can be summarized as follows:

Treated_i: Takes the value 1 for firms located in jurisdictions with mandatory ESG reporting rules and 0 for others. This way, it is determined which companies are affected by the regulation.

Post_t indicator: It takes the value of 1 for the year after the relevant mandatory reporting rule came into force, and 0 for the period before. Thus, the policy effect is captured in the time dimension.

Dependent variables (Tone indicators): More than one tone metric was used in the study. ESG_Tone_max represents the highest (most optimistic) value of tone measured in the ESG-related sections of the ECC text. Similarly, E_Tone_max, S_Tone_max, and G_Tone_max reflect the highest level of language tone in the environmental, social and governance dimensions, respectively. These metrics are based on sentiment scores obtained by applying an advanced NLP model (e.g., FinBERT/ESG-BERT) to the relevant sections in the text. The most positive expression score (or tone) detected for the relevant category in each text was used as the metric. This quantifies the extent to which the company used positive language on ESG issues in ECC during that period.

Fixed effects

- Firm fixed effects (α_i): Controls for firm-specific characteristics that do not change. This leaves fixed differences, such as some firms

having a generally more optimistic or pessimistic communication style, out of the equation.

- Year fixed effects (δ_t): Controls for macroeconomic conditions or general year-based trends that may affect all firms simultaneously. For example, the year of a global crisis may affect all ECC tones; the year fixed effect captures and extracts this.
- Sector-year fixed effects ($(\alpha \times \delta)_{s t}$): Common shocks or trends that specific sectors may experience in certain years (for example, regulations in the energy sector in a certain year) may also have an impact on tone. Sector-year fixed effects reflect the time-varying effects of sectoral differences in the model by including a fixed effect for each sector $\times$ year combination.

Control variables (X_{it}): Various financial and performance indicators are included as control variables in the model to isolate other influences on tone.

The main control variables are:

- ROA (Return on Assets): Indicates how efficiently a firm uses its assets. The communication tone of profitable firms may differ due to their financial well-being.
- Leverage ratio: Reflects the firm's level of financial risk (debt/equity ratio). Management of heavily leveraged firms may use cautious language about the future.
- Firm size (Log): An indicator of the firm's size, measured as the logarithm of its total assets or market value. Larger firms are

generally under greater public scrutiny, so they carefully adjust their communication tone.

- ESG score (ESG_Score): An index or score that numerically summarizes a firm's environmental, social, and governance performance. Managers of firms with higher ESG scores may want to highlight their achievements in this area (more positive tone), while those with lower scores may be cautious or defensive to avoid criticism.

These control variables help account for firm-level characteristics such as financial performance, leverage, size, and ESG maturity. By including these controls, the model aims to reduce potential confounding and isolate the effect of ESG disclosure mandates on communication tone.

To ensure internal validity, the study applies a Difference-in-Differences (DiD) approach. This method compares changes in tone scores between firms located in countries with mandatory ESG disclosure regulations (treatment group) and those without such regulations (control group), before and after the implementation period.

A key assumption of this approach is that the treatment and control groups would have followed parallel trends in the absence of the regulatory change. To evaluate the plausibility of this assumption, the research design includes visual inspection of pre-treatment trends and formal checks using interaction terms between treatment status and time.

In addition, placebo analyses are built into the design framework to test whether spurious effects could emerge when treatment timing is artificially shifted to periods before the actual implementation. The placebo setup includes multiple pre-

regulation years and country-level random assignments. This helps validate the causal identification strategy without relying solely on observed treatment effects.

Although the study includes visual plots that track tone dynamics over time, no formal event-study regression is presented in this section. These visualizations serve primarily to support the assessment of pre-treatment comparability and the temporal structure of the design.

By combining fixed-effects panel regressions with firm-year observations and rich text-based tone metrics derived from NLP models, the empirical framework is structured to capture tone shifts associated with regulatory exposure while controlling for unobserved heterogeneity.

3.4 Research question and hypotheses

This study investigates the impact of mandatory ESG reporting regulations on the language tone of corporate executives in their earnings conference calls (ECCs).

Although there is increasing interest in the literature on how ESG disclosures affect firm behaviour, investor perceptions, and market performance, the direct impact of mandatory regulations on managerial communication tone has not yet been sufficiently examined. The following hypotheses, which are formulated within the framework of Signalling Theory, Impression Management Theory, and Legitimacy Theory, aim to fill this gap. The hypotheses aim to test the direction of changes in ESG tone in ECCs following the regulations and how these changes differ according to firm characteristics.

Overarching Research Question:

Does the implementation of country-level ESG mandates influence the tone of corporate ESG disclosures?

Main Hypotheses:

- H1 (Overall ESG Tone): The implementation of country-level ESG mandates leads to an increase in the overall ESG tone (ESG_Tone) of corporate disclosures for firms operating in those countries, compared to firms in countries without such mandates (or relative to the pre-mandate period, controlling for common trends).
 - Null Hypothesis (H0): The implementation of country-level ESG mandates has no significant effect on the overall ESG tone of corporate disclosures.
- H2 (Environmental Tone): The implementation of country-level ESG mandates leads to an increase in the Environmental Tone (E_Tone) of corporate disclosures for firms operating in those countries, compared to those without such mandates.
 - Null Hypothesis (H0): The implementation of country-level ESG mandates has no significant effect on the Environmental tone of corporate disclosures.
- H3 (Social Tone): The implementation of country-level ESG mandates leads to an increase in the Social Tone (S_Tone) of corporate disclosures for firms operating in those countries, compared to those without such mandates.
 - Null Hypothesis (H0): The implementation of country-level ESG mandates has no significant effect on the Social tone of corporate disclosures.
- H4 (Governance Tone): The implementation of country-level ESG mandates leads to an increase in the Governance tone (G_Tone) of corporate

disclosures for firms operating in those countries, compared to those without such mandates.

- Null Hypothesis (H0): The implementation of country-level ESG mandates has no significant effect on the Governance tone of corporate disclosures.

3.4.1 Implicit assumptions/supporting arguments tested

- Parallel Trends Assumption: The code includes a plot to visually inspect this.

While not a formal hypothesis, the validity of the DiD results relies on the assumption that, in the absence of the treatment, the average change in ESG tone for the treated group would have been the same as the average change for the control group.

- Placebo Test: The placebo test implicitly tests the null hypothesis that there is no effect associated with a fake implementation date (placebo_year). A non-significant result in the placebo test supports the idea that the main findings are specific to the actual treatment timing and not due to pre-existing differential trends.

This analysis is set up to test whether the coefficient on $I(\text{Treated} * \text{Post})$ is statistically significantly positive for each of the tone variables, after accounting for controls and fixed effects.

CHAPTER 4

DATA & METHODOLOGY

4.1 Data

4.1.1 Data sources

Earnings Conference Call (ECC) Transcripts

The primary text data for this study consists of transcripts that took place between 2005 and 2024. These meetings belong to 16,327 public companies operating in 95 countries worldwide and were obtained from the S&P Global Capital IQ platform. Standard and Poor's Capital IQ Pro database provides data on "banks, private equity firms, insurance companies, broker dealers and financial markets worldwide." Its scope is defined as "ownership information, market analysis, filings, financial statements, mergers and acquisitions' data, and sectoral news." (European University Institute, 2025)

ECCs reflect the strategic and financial communication that takes place between company management and financial analysts each quarter. These transcripts were analysed at the sentence level using transformer-based models developed in the field of Natural Language Processing (NLP) in recent years.

In this thesis, pre-trained transformer-based language models tailored to finance and ESG domains were used to analyse the tone and ESG content of firms' earnings conference calls (ECCs). Unlike dictionary-based methods, these models rely on the BERT architecture, which processes bidirectional context to classify sentences based on semantic meaning (Devlin et al., 2019).

First, the FinBERT-Tone model (Huang et al., 2022), trained on financial texts such as analyst reports, ECCs, and SEC filings, was used to classify each sentence as

positive, negative, or neutral. These sentence-level results were aggregated to construct a firm-level Tone Score.

To measure tone specifically within ESG-related content, FinBERT was also applied to sentences tagged under environmental (E), social (S), and governance (G) themes. For each ESG dimension, the highest tone score in a given year was retained, creating the variables:

- Max E Tone
- Max S Tone
- Max G Tone

Additionally, FinBERT-ESG (Huang et al., 2022) was employed to identify ESG-relevant content in each ECC through multi-label sentence classification into E, S, and G categories. As a complementary tool, the ESG-BERT models developed by HKUST (Schimanski et al., 2023) were used to enhance ESG content detection, especially in sustainability narratives. Each sentence was thus evaluated both in terms of sentiment (via FinBERT-Tone) and ESG content (via FinBERT-ESG and ESG-BERT). Based on this, the following core variables were constructed:

- Tone Score: Net sentiment across all sentences
- Max E/S/G Tone: Maximum tone for each ESG component
- ESG Score: Proportion of ESG-related sentences in each ECC

All models were used in their pre-trained form without fine-tuning, ensuring generalizability across firms and years. This approach bridges NLP methods with ESG and accounting research, offering a scalable way to quantify tone and ESG communication in corporate disclosures.

Financial and Control Variables

The firm-level financial data used in this study were obtained from various data sources, corresponding to the companies' headquarters countries. Financial information for companies headquartered in the United States (US) was obtained directly from the Compustat North America database. The Compustat Global database obtained financial data for companies operating in countries other than the US. Compustat is a database provided by Standard & Poor's that provides comprehensive data on firms' financial performance measures (Standard & Poor's, 2024).

This approach ensured that each company obtained the most appropriate data source for the country it was affiliated with and increased the accuracy and consistency of financial variables.

In this direction, the control variables used in the analyses (Total Assets, Company Size, Leverage, ROA) were created by taking them from the database to which the companies were affiliated, the financial indicators were matched with the relevant ECC data on a yearly basis and used as independent variables in the regression models.

The main financial and control variables used are:

- Total Assets: The total reported assets of the company. Used as raw data, no transformation was applied.
- Firm Size: Calculated by taking the natural logarithm of total assets:

$$\text{Firm Size} = \ln(\text{Total Assets})$$

- Leverage Ratio: Calculated by the ratio of total liabilities to total assets:

$$\text{Leverage} = \frac{\text{Total Liabilities}}{\text{Total Assets}}$$

- Profitability Ratio: ROA (Return on Assets) was used.
- Sector Classification: Used to control for sector fixed effects.

In addition, to reduce the effect of outliers in continuous variables, the data was Winsorized at the 1% and 99% percentiles.

All financial variables were synchronized based on the year in which the relevant ECC was conducted, and this synchronization was maintained in the analyses. All financial and market data used in the study were accessed via the Wharton Research Data Services (WRDS) platform. WRDS provides high-quality firm-level data that ensures data consistency over time and enables accurate data matching via ISIN codes and firm identifiers (e.g. gvkey).

4.1.2 Identification of countries with mandatory ESG disclosure

The separation of countries into “treatment” and “control” groups in the study is based on the international mandatory ESG disclosure regulations dataset compiled by Krueger, Sautner, Tang, and Zhong (2024). This dataset is accessible through the Open Science Framework (OSF) and documents when and within which frameworks (full compliance regimes or “comply or explain” models) countries introduced environmental, social, and governance (ESG) reporting obligations (Krueger, Sautner, Tang & Zhong, 2024).

For EU member states, the obligations introduced under the Non-Financial Reporting Directive (NFRD) and subsequently the Corporate Sustainability Reporting Directive (CSRD) were included in the treatment group. These regulations require large-scale firms to submit comprehensive, standardized ESG reporting (Krueger et al., 2024).

In contrast, countries that did not have binding national ESG disclosure requirements during the specified period, such as the United States, were defined as the control group; these firms voluntarily disclose ESG information (Krueger et al., 2024). The gradual and variable implementation of ESG reporting requirements in different countries provides our study with a quasi-empirical DiD (differences in differences) opportunity.

Each firm's "treatment" status was coded based on the ESG obligation status of the country where the accounting center was located in that year. When the country phased in the mandatory disclosure regime for all Environmental, Social, and Governance dimensions, the treatment indicator is activated from that year onwards (Krueger et al., 2024).

The integration of this regulatory timeline with firm-level ECC data enables the DiD estimation strategy, which is as follows: It allows us to examine the causal effects of mandatory ESG disclosures on corporate communication tone and ESG content by controlling for sector-year and firm fixed effects.

To measure the timing of ESG regulation, we compiled country-level ESG mandate start years from publicly available databases and regulatory records. Each firm-year observation is categorized as:

Treated = 1 if ESG disclosure was mandatory in that firm's country at any point;

Post = 1 if the year is equal to or greater than the country's mandate start year.

Winsorization is applied at the 1st and 99th percentiles for continuous variables (e.g., ROA, Leverage) to mitigate the influence of outliers. ECCs or firm observations without total asset information were dropped due to the importance of firm size as a control.

Table 1. Variable Definitions and Sources

Variable	Definition / construction	Source
<i>Dependent and treatment variables</i>		
ESG_Tone_max	Maximum positive sentiment across <i>all</i> ESG-related sentences in an ECC transcript (FinBERT-Tone).	FactSet Call-Street; FinBERT-Tone
E_Tone_max	Maximum positive sentiment in environmental sentences.	EnvironmentalBERT
S_Tone_max	Maximum positive sentiment in social sentences.	SocialBERT
G_Tone_max	Maximum positive sentiment in governance sentences.	GovernanceBERT
ESG_Score_W	Share of ESG-tagged sentences in total sentences (winsorised). Mapped from R variable.	FinBERT-ESG (Assumed, update if different)
Treated _{<i>i</i>}	1 if firm <i>i</i> is headquartered in a jurisdiction with a mandatory ESG mandate; 0 otherwise.	Krueger2024 (Assumed, for mandate timing); Constructed
Post _{<i>t</i>}	1 for years $\geq$ mandate start year for that country; 0 otherwise.	Constructed
Treated $\times$ Post	Difference-in-Differences interaction term.	Constructed
<i>Financial control variables</i>		
Log(Total Assets)	Natural logarithm of Total Assets. Represents 'Firmsize' in R code.	Compustat
Leverage	(Long-term debt + current liabilities) $\div$ Total assets. 'Leverage _W ' is the winsorised version.	Compustat
ROA	Net income $\div$ Total assets. 'ROA _W ' is the winsorised version.	Compustat

Notes: This table provides definitions for key variables based on the provided image, adapted to match variables used in the R analysis (e.g., winsorized controls, log transformation). Tone variables were derived from models cited in the source image (e.g., FinBERT-Tone Huang2023, Environmental/Social/GovernanceBERT models Schimanski2024). Country-level mandate years follow Krueger2024. Firm-level accounting data are from S&P Compustat. Winsorized variables (_W suffix) are clipped at the 1st and 99th percentiles.

4.2 Methodology

4.2.1 Preliminary analysis and data integrity

Python-Based Variable Preparation Process

All variables in our thesis analysis were created in the Python environment. In the first step, HuggingFace Transformers (Wolf et al., 2020), PyTorch (Paszke et al., 2019), and pandas (McKinney, 2010), required for Natural Language Processing (NLP) and data manipulation, were installed. These libraries provided the calling of BERT-based models, GPU/CPU acceleration, and tabular data processing, respectively.

On the hardware side, large language models were effectively run using NVIDIA A100 GPUs offered at Google Colab. Thus, FinBERT - tone, FinBERT - ESG, and ESG - BERT were executed with GPU acceleration.

The three models are used for:

FinBERT - tone to determine the sentiment (positive, negative, neutral) of financial sentences. FinBERT - ESG and ESG - BERT to label sentences as Environmental (E), Social (S) or Governance (G).

Earnings Conference Call (ECC) texts were parsed at the sentence level by loading tokenizer objects belonging to each model. All PDFs were converted to CSV format; each sentence was put on a separate line in the SENTENCE column. The models were applied to each line to produce (i) sentiment, (ii) E/S/G labels. The intersection of FinBERT - ESG and ESG - BERT was taken for a more robust

$$\mathbf{Tone}_{it} = \frac{\text{Number of positive sentences}_{it} - \text{Number of negative sentences}_{it}}{\text{Total number of sentences}_{it}}$$

classification. Only sentences for which both models gave the same label were included in the final classification.

For each ECC i in year t , the Tone variable captures the precise sentiment in the language used by managers and analysts. Positive values indicate an optimistic tone, and negative values indicate a more pessimistic tone.

$$\mathbf{ESG\ Score}_{it} = \frac{\text{Number of sentences labelled } E, S \text{ or } G_{it}}{\text{Total number of sentences}_{it}}$$

The ESG Score indicates the weight of sustainability-themed conversation in a call. For example, a value of 0.15 means that 15% of all sentences are devoted to environmental, social, or governance issues.

$$\mathbf{E_Tone}_{ft} = \max_{i \in \mathcal{C}_{ft}} \left\{ \frac{E_i^+ - E_i^-}{\text{Total sentences}_i} \right\}, \quad \text{analogously for } \mathbf{S_Tone}_{ft}, \mathbf{G_Tone}_{ft}.$$

Notation.

- f = firm, t = year, i = individual call in the set $\mathcal{C}_{ft}$.
- E_i^+ and E_i^- denote, respectively, the number of environment-related sentences classified as positive and negative in call i .
- The same logic applies to the social (S) and governance (G) dimensions.

As a result of these steps, differences in both sentiment intensity and ESG content coverage were measured quantitatively at the firm-year level, and a data set ready for Difference-in-Differences (DiD) analyses was obtained.

4.2.2 Descriptive statistics and visual analysis

Descriptive statistics were used to measure the accuracy of the difference-in-difference (DiD) study. Before running regressions, this first stage was done to guarantee data integrity and to find any possible problems, including missing values, outliers, or data entry mistakes. Furthermore, verifying whether the variables followed normal distribution was important since non-normal distributions—especially in skewed financial ratios—would call for suitable transformations.

Second, essential measurements between the treated and control groups were compared using analysis of sample properties. This comparison allows one to evaluate whether the two groups were balanced in size, industry, and ESG ratings and gives first ideas on possible variations between them.

Finally, we investigated the validity of the assumption of parallel trends. If Tone Score trends had diverged before the ESG mandate, the DiD findings might have been skewed. Consequently, a pre-treatment trend analysis was performed before starting the DiD investigation.

4.2.3 Difference-in-Differences (DiD)

In this study, the Difference-in-Differences (DiD) method was applied to measure the impact of mandatory ESG disclosure regulations on the communication tone in companies' earnings disclosure meetings (ECC).

The DiD approach allows isolating the policy effect by comparing the tone changes of firms operating in countries with ESG regulations (treatment group) and firms not subject to such regulations (control group) in the pre- and post-implementation periods. This method allows causal effect estimation by controlling

for firm characteristics that do not change over time and common shocks affecting all firms (Angrist & Pischke, 2009).

The basic regression model is expressed as follows:

$$Tone_{it} = \beta_0 + \beta_1(Treated_i \times Post_t) + \beta_2 X_{it} + \alpha_i + \gamma_{st} + \epsilon_{it}$$

Where:

- $Tone_{it}$: Communication tone measure (ESG_Tone_max, E_Tone_max, S_Tone_max, G_Tone_max).
- $Treated_i \times Post_t$: Interaction term for treatment effect.
- X_{it} : Control variables (ROA, ROE, Leverage, ESG Score, log(Firm Size)).
- α_i : Firm fixed effects.
- γ_{st} : Sector-year fixed effects.
- ϵ_{it} : Error term (clustered at the firm level).

4.2.4 Placebo tests

Using placebo tests to test the reliability of Difference-in-Differences (DiD) estimates and eliminating potentially misleading statistical significance results (e.g., due to serial correlation) is an important step that increases the strength of causal inference. In particular, Bertrand, Duflo, and Mullainathan (2004) have shown that standard DiD approaches can produce significant results even if the intervention has no real effect, highlighting the need for such robustness checks.

A common placebo testing approach uses pre-intervention periods to test the plausibility of the common trend assumption. In practice, one of the pre-treatment periods is treated as the reference (e.g., T=0) and the next (also pre-treatment) period is treated as the pseudo-treatment period (e.g., T=1). Since the treatment has not yet been applied in any period, a statistically significant pseudo-treatment effect found in

this period indicates that the common trend (or stability of deviation) assumption has been violated (Huber, 2023, p. 200). This approach is closely related to the possible trend differences and deviation stability issues among treatment groups, which are also addressed by Rambachan and Roth (2020); these studies offer solutions to such problems in model-based DiD inference.

In this framework, three separate placebo DiD models were estimated to strengthen the causal inferences of the study:

- Year 2010 Test: It was selected as a pre-policy reference point, thus testing the assumption of a common trend between the treatment and control groups in the periods before the actual implementation began.
- Year 2016 Test: It was determined as a fake intervention year to control for the potential confounding effect of the Paris Climate Agreement (UNFCCC, 2015) on company communications at the end of 2015.
- Random Year Test: A year other than the actual mandatory ESG implementation years was randomly selected for each country to test whether the effect found in the main analysis was specific to a specific year or was due to general time trends.

CHAPTER 5

ANALYSIS AND EMPIRICAL RESULTS

5.1 Descriptive statistics

Table 2 reports descriptive statistics for the main variables used in the analysis. The dataset includes up to 109,400 firm-year observations, although some variables (e.g., financial ratios) have fewer observations due to missing data.

Table 2. Descriptive Statistics

Statistic	N	Mean	St. Dev.	Min	Median	Max
ESG Tone (Max)	109,338	0.006	0.009	-0.030	0.003	0.335
E Tone (Max)	109,338	0.003	0.007	-0.030	0.000	0.333
S Tone (Max)	109,338	0.003	0.005	-0.029	0.002	0.200
G Tone (Max)	109,338	0.0001	0.001	-0.009	0.000	0.100
ROA (Wins.)	44,726	-0.013	0.229	-1.387	0.034	0.312
Leverage (Wins.)	45,247	0.283	0.236	0.000	0.249	1.215
Log(Total Assets)	94,335	14.400	2.214	3.738	14.408	22.592
ESG Score (Wins.)	109,400	0.025	0.026	0.000	0.017	0.138
Treated Group	109,400	0.252	0.434	0	0	1
Post Period (for Treated)	109,400	0.242	0.428	0	0	1

Notes: Firm_Size is Log(Total Assets). Winsorized variables are clipped at 1% and 99%. N denotes the number of non-missing observations for each variable. ROE is excluded from the analysis.

The average maximum ESG tone score (ESG_Tone_max) is 0.006, with a standard deviation of 0.009, indicating that ESG-related sentences in earnings calls generally reflect mild sentiment. The Environmental (E_Tone_max) and Social (S_Tone_max) tone indicators both have a mean of 0.003, with standard deviations of 0.007 and 0.005, respectively. The Governance tone score (G_Tone_max) has the lowest average value, 0.0001, and a standard deviation of 0.001.

The winsorized ESG Score (ESG_Score_W), calculated as the proportion of ESG-tagged sentences in an ECC transcript, has a mean of 0.025. This implies that, on average, 2.5% of sentences in earnings calls are related to ESG topics.

The average firm size (Firm_Size), measured as the logarithm of total assets, is 14.40 with a standard deviation of 2.21. The sample also shows financial diversity: the winsorized Return on Assets (ROA_W) has a mean of -0.013 , and the Leverage ratio (Leverage_W) averages 0.283.

Furthermore, 25.2% of the observations are from the treatment group (i.e., firms from countries with mandatory ESG disclosure), and 24.2% fall into the post-treatment period.

These summary statistics, presented in Table 2, describe the key characteristics of the dataset used in the Difference-in-Differences (DiD) analysis.

5.2 Correlation matrix

Table 3 shows the correlations between the main variables used in the regression analysis. This matrix helps us see how strongly the variables are related and whether multicollinearity may be a concern.

As expected, ESG_Tone_max is strongly correlated with E_Tone_max ($r = 0.887$) and also with S_Tone_max ($r = 0.595$). The correlation with G_Tone_max is lower ($r = 0.181$), but still positive. This makes sense, since overall ESG tone includes signals from the E, S, and G components. ESG_Tone_max is also highly correlated with ESG_Score_W ($r = 0.752$), showing that firms with higher ESG-related content tend to use more positive ESG language.

The financial control variables, such as ROA_W and Leverage_W, have weak correlations with tone variables, mostly between -0.10 and +0.07. This means firm performance and capital structure are not strongly related to tone, which helps reduce the risk of omitted variable bias.

Importantly, no correlations exceed 0.9, so multicollinearity does not appear to be a problem in our main DiD models.

Table 3. Correlation Matrix of Key Variables

	ESG Tone (Max)	E Tone (Max)	S Tone (Max)	G Tone (Max)	ROA (W)	Leverage (W)	Log(Total Assets)	ESG Score (W)
ESG Tone (Max)	1.000							
E Tone (Max)	0.887	1.000						
S Tone (Max)	0.595	0.167	1.000					
G Tone (Max)	0.181	0.064	0.087	1.000				
ROA (Wins.)	0.020	0.020	0.009	0.000	1.000			
Leverage (Wins.)	0.033	0.027	0.025	-0.002	-0.099	1.000		
Log(Total Assets)	0.099	0.106	0.024	0.033	0.068	0.081	1.000	
ESG Score (Wins.)	0.752	0.678	0.434	0.116	0.043	0.049	0.124	1.000

Notes: Pairwise correlations based on observations with non-missing values for both variables involved. Values rounded to three decimal places.

5.3 Parallel trends verification

The parallel trend assumption is a fundamental condition for the validity of the difference-in-differences (DiD) analysis strategy. This assumption predicts that firms in the treatment and control groups would have followed similar trends in outcome variables such as communication tone if no intervention (e.g. mandatory ESG disclosure obligation) had taken place. Demonstrating that this assumption holds strengthens the conclusion that the estimated DiD coefficient truly reflects the policy effect.

As a first verification step, a visual inspection of temporal trends was performed. Figures 1 through 4 illustrate the average tone trends for the treatment

and control groups over time for the main outcome variable (ESG_Tone_max) and its subcomponents (E_Tone_max, S_Tone_max, and G_Tone_max). In each figure, the dashed line represents the earliest policy implementation year.

- Figure 1 (ESG_Tone_max) shows that the average tone scores for both groups remained close and parallel until the late 2010s. From 2018 onwards, there is a visible increase, with the treatment group rising more sharply starting in 2019.
- Figure 2 (E_Tone_max) presents a similar trend, with the treatment group diverging more clearly from the control group after 2019.
- Figure 3 (S_Tone_max) indicates milder changes, but still shows a post-2018 increase more pronounced in the treatment group.
- Figure 4 (G_Tone_max) reflects relatively stable trends, with only modest separation after 2020.

These visual trends suggest that, prior to regulation, the groups followed similar patterns, supporting the parallel trends assumption.

Secondly, to further assess the validity of this assumption, placebo tests were conducted focusing on pre-treatment periods. In these tests, the treatment variable was artificially shifted to earlier years (2010 and 2016), and the corresponding DiD regressions were estimated using the full model with firm and sector-year fixed effects.

Placebo Year 2010

The regression assuming a policy implementation in 2010 produced a coefficient of +0.0007 for the interaction term $\text{Treated} \times \text{Post}_{2010}$, which was statistically insignificant ($p = 0.161$). This result supports the parallel trends assumption, indicating no early divergence in tone scores between the two groups.

Placebo Year 2016

In the placebo test using 2016, the interaction term remained at +0.0007 but was statistically significant ($p = 0.005$). This suggests a mild divergence prior to formal regulation. However, this result is not unexpected, as the Paris Agreement was signed in 2016, which may have encouraged earlier ESG discourse, especially in environmentally proactive regions. Therefore, the 2016 result likely reflects anticipation effects or voluntary ESG actions, not a violation of the DiD assumption.

In conclusion, while the 2016 placebo finding should be noted, the overall results from visual trends and placebo regressions provide reasonable support for the validity of the identification strategy used in the main DiD analysis.

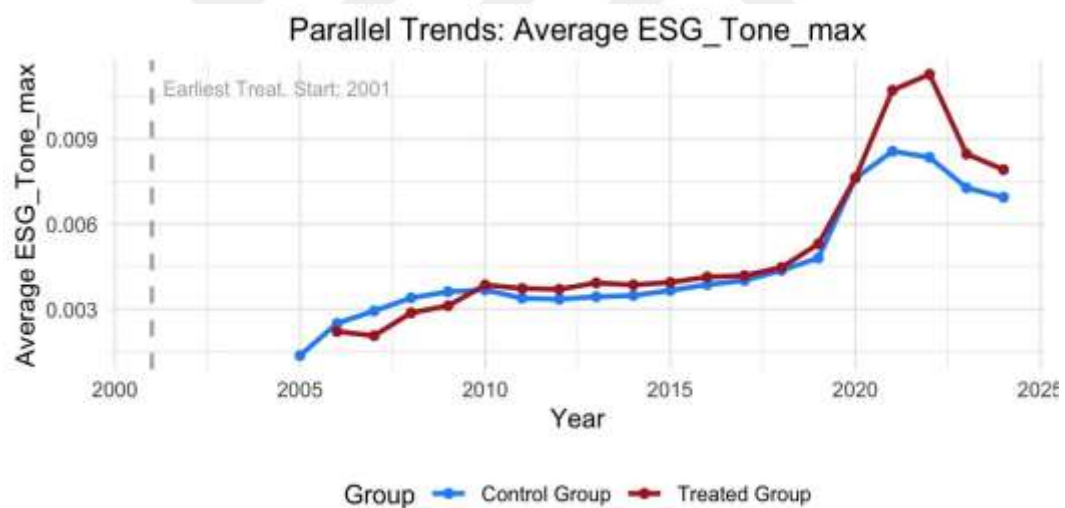


Figure 1. Average ESG_Tone_max

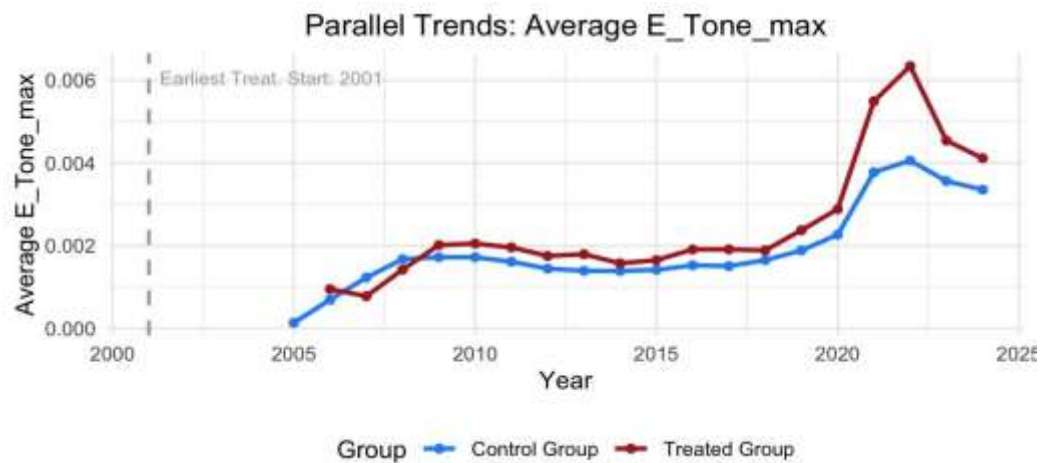


Figure 2. Average E_Tone_max

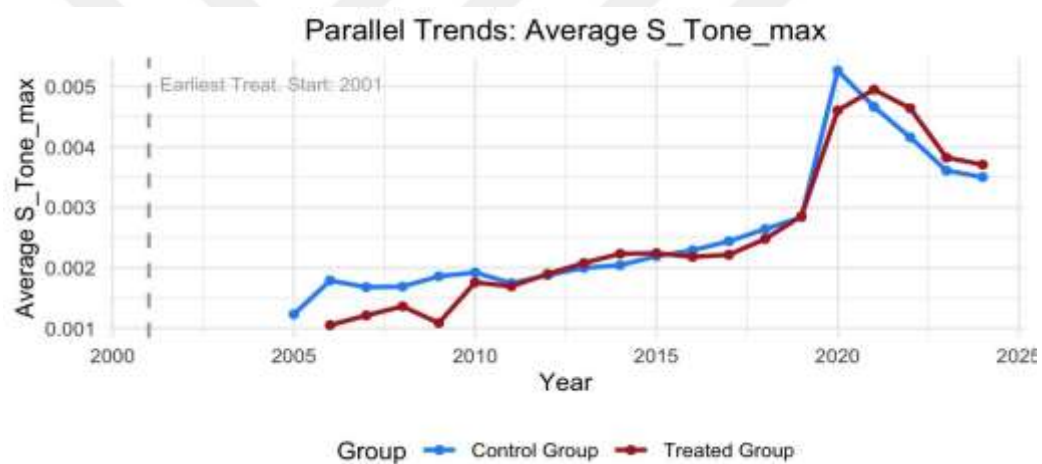


Figure 3. Average S_Tone_max

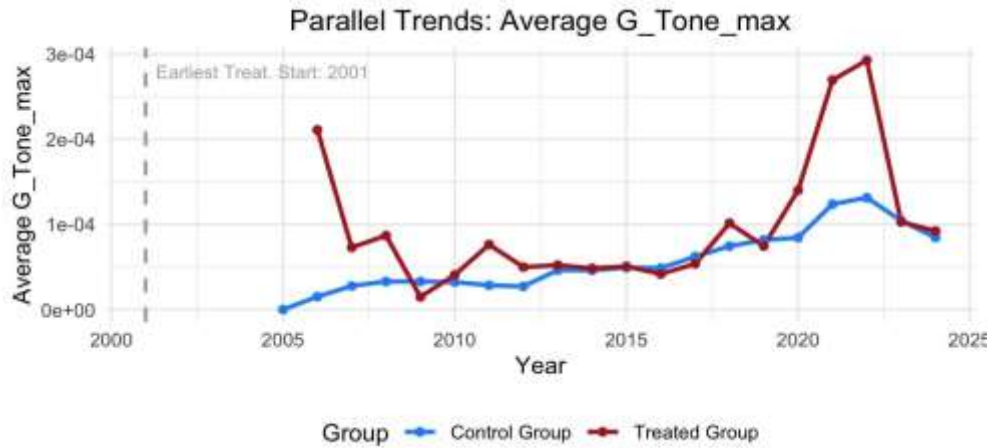


Figure 4. Average G_Tone_max

5.4 Main regression results

Tables 4, 5, and 5 present the main findings of the Difference-in-Differences (DiD) estimation models used to assess the impact of mandatory ESG disclosure policies on the tone of corporate communication during earnings conference calls. The analyses focus on four key dependent variables: maximum overall ESG tone (ESG_Tone_max), Environmental tone (E_Tone_max), Social tone (S_Tone_max), and Governance tone (G_Tone_max), each examined under six different fixed-effects specifications.

The baseline DiD regression model estimated for each outcome variable is specified as follows:

$$Tone_{it} = \beta_0 + \beta_1(Treated_i \times Post_t) + \beta_2 X_{it} + \alpha_i + \gamma_{st} + \epsilon_{it}$$

Where:

- $Tone_{it}$: Communication tone measure (ESG_Tone_max, E_Tone_max, S_Tone_max, G_Tone_max).
- $Treated_i \times Post_t$: Interaction term for treatment effect.
- X_{it} : Control variables (ROA, ROE, Leverage, ESG Score, log(Firm Size)).
- α_i : Firm fixed effects.
- γ_{st} : Sector-year fixed effects.
- ϵ_{it} : Error term (clustered at the firm level).

The key coefficient of interest is the interaction term $Treated \times Post$ (DiD), which estimates the average differential change in tone for treated firms after the policy implementation, compared to control firms.

Individual E, S, and G Tones

- General ESG Tone (ESG_Tone_max)

Table 4 shows that the $Treated \times Post$ coefficient is positive across all model specifications (Columns 1 to 6). In the simpler models without firm-level fixed effects (Cols 1, 3, and 5), the coefficient is consistently +0.0007 and statistically significant at the 1% level. When firm fixed effects are included (Col 2), the coefficient slightly increases to +0.0008 ($p < 0.05$).

However, in more robust models that also control for time and industry trends (Cols 4 and 6), the coefficient drops to +0.0003 and becomes statistically insignificant ($p = 0.33$ and 0.35). This means the positive effect is not robust when we account for firm-specific and sectoral factors that vary over time.

Even in the significant models, the effect size is modest. Considering that the average ESG tone is about 0.003–0.004 before the policy (see Figure 1), the observed increase represents a 17–23% change. While small, this is still meaningful in the context of corporate communication.

Table 4. Main DiD Results: Impact on Overall ESG Tone (ESG_Tone_max)

FE Spec:	(1) None	(2) F	(3) Y	(4) F+Y	(5) SY	(6) F+SY
Treated × Post (DiD)	0.0007*** (0.0001)	0.0008** (0.0003)	0.0007*** (0.0001)	0.0003 (0.0003)	0.0007*** (0.0001)	0.0003 (0.0003)
ROA (Wins.)	-0.0007*** (0.0002)	0.0001 (0.0002)	-0.0007*** (0.0002)	0.0000 (0.0002)	-0.0007*** (0.0002)	-0.0000 (0.0002)
Leverage (Wins.)	-0.0003 (0.0002)	0.0002 (0.0003)	-0.0003 (0.0002)	-0.0002 (0.0003)	-0.0004** (0.0002)	-0.0003 (0.0003)
Log(Total Assets)	0.0000* (0.0000)	0.0004*** (0.0001)	0.0001** (0.0000)	-0.0002** (0.0001)	0.0001** (0.0000)	-0.0001 (0.0001)
ESG Score (Wins.)	0.2759*** (0.0053)	0.2566*** (0.0046)	0.2721*** (0.0055)	0.2422*** (0.0048)	0.2717*** (0.0054)	0.2392*** (0.0048)
Observations	43,146	43,146	43,146	43,146	42,418	42,418
Adj. R2	0.567	0.734	0.571	0.738	0.571	0.739
Firm FE	No	Yes	No	Yes	No	Yes
Year FE	No	No	Yes	Yes	No	No
Sector x Year FE	No	No	No	No	Yes	Yes

Notes: Standard errors clustered by Firm_ID in parentheses. FE Spec Abbreviations: F=Firm, Y=Year, F+Y=Firm+Year, SY=Sector x Year, F+SY=Firm+Sector x Year. Significance levels: * p<0.1, ** p<0.05, *** p<0.01. Controls are winsorized at 1%/99%. Firm_Size is Log(Total Assets). ROE is excluded from controls.

- Environmental Tone (E_Tone_max)

As shown in Table 5 (Columns 7 to 12), the coefficient on Treated × Post is positive and significant in models without firm fixed effects (Cols 7, 9, and 11), ranging from +0.0003 to +0.0005 ($p < 0.05$). But once firm fixed effects are added (Cols 8, 10, and 12), the results are no longer statistically significant.

This suggests that the observed changes in environmental tone may be due to firm-specific characteristics rather than the policy itself. In short, mandatory ESG

disclosure does not appear to have a strong or consistent effect on how firms communicate environmental issues.

- Social Tone (S_Tone_max)

Table 6 (Columns 13 to 18) shows the clearest results. The Treated $\times$ Post coefficient is positive and statistically significant across all specifications. It ranges from +0.0002 to +0.0009 and remains significant even in the strictest model (Col 18), where the coefficient is +0.0003 ($p = 0.041$).

This indicates that the policy had a small but consistent positive impact on how firms talk about social issues. Compared to other tone dimensions, the effect here is the most robust and reliable.

- Governance Tone (G_Tone_max)

In Table 7 (Columns 19 to 24), the policy's effect on governance tone is weak and inconsistent. The Treated $\times$ Post coefficient is only significant in models without firm fixed effects (Cols 19, 21, and 23), with values around +0.0001 ($p < 0.05$). Once firm fixed effects are added (Cols 20, 22, and 24), the effect becomes statistically insignificant and close to zero.

Therefore, we find no strong evidence that the policy changed how firms communicated about governance.

Table 5. Main DiD Results: Impact on Environmental Tone (E_Tone_max)

FE Spec:	(1) None	(2) F	(3) Y	(4) F+Y	(5) SY	(6) F+SY
Treated × Post (DiD)	0.0003** (0.0001)	-0.0000 (0.0002)	0.0004*** (0.0001)	-0.0001 (0.0002)	0.0005*** (0.0001)	-0.0000 (0.0002)
ROA (Wins.)	-0.0005** (0.0002)	-0.0000 (0.0002)	-0.0004** (0.0002)	-0.0001 (0.0002)	-0.0005** (0.0002)	-0.0000 (0.0002)
Leverage (Wins.)	-0.0003* (0.0002)	-0.0003 (0.0002)	-0.0002 (0.0002)	-0.0003 (0.0002)	-0.0004* (0.0002)	-0.0003 (0.0002)
Log(Total Assets)	0.0001*** (0.0000)	-0.0000 (0.0001)	0.0001*** (0.0000)	-0.0003*** (0.0001)	0.0001** (0.0000)	-0.0001 (0.0001)
ESG Score (Wins.)	0.2000*** (0.0052)	0.1570*** (0.0043)	0.2009*** (0.0054)	0.1524*** (0.0046)	0.1958*** (0.0054)	0.1486*** (0.0045)
Observations	43,146	43,146	43,146	43,146	42,418	42,418
Adj. R2	0.460	0.703	0.465	0.707	0.476	0.712
Firm FE	No	Yes	No	Yes	No	Yes
Year FE	No	No	Yes	Yes	No	No
Sector x Year FE	No	No	No	No	Yes	Yes

Notes: Standard errors clustered by Firm_ID in parentheses. FE Spec Abbreviations: F=Firm, Y=Year, F+Y=Firm+Year, SY=Sector x Year, F+SY=Firm+Sector x Year. Significance levels: * p<0.1, ** p<0.05, *** p<0.01. Controls are winsorized at 1%/99%. Firm_Size is Log(Total Assets). ROE is excluded from controls.

Table 6. Main DiD Results: Impact on Social Tone (S_Tone_max)

FE Spec:	(1) None	(2) F	(3) Y	(4) F+Y	(5) SY	(6) F+SY
Treated × Post (DiD)	0.0004*** (0.0001)	0.0009*** (0.0001)	0.0002** (0.0001)	0.0003** (0.0002)	0.0002* (0.0001)	0.0003** (0.0002)
ROA (Wins.)	-0.0002 (0.0001)	0.0002 (0.0001)	-0.0002* (0.0001)	0.0001 (0.0001)	-0.0002 (0.0001)	0.0001 (0.0001)
Leverage (Wins.)	0.0001 (0.0001)	0.0006*** (0.0002)	0.0000 (0.0001)	0.0001 (0.0002)	-0.0000 (0.0001)	0.0000 (0.0002)
Log(Total Assets)	-0.0001*** (0.0000)	0.0004*** (0.0001)	-0.0000* (0.0000)	0.0001 (0.0001)	0.0000 (0.0000)	0.0000 (0.0001)
ESG Score (Wins.)	0.0723*** (0.0027)	0.0939*** (0.0026)	0.0678*** (0.0027)	0.0839*** (0.0026)	0.0723*** (0.0027)	0.0848*** (0.0027)
Observations	43,146	43,146	43,146	43,146	42,418	42,418
Adj. R2	0.190	0.554	0.202	0.564	0.239	0.567
Firm FE	No	Yes	No	Yes	No	Yes
Year FE	No	No	Yes	Yes	No	No
Sector x Year FE	No	No	No	No	Yes	Yes

Notes: Standard errors clustered by Firm_ID in parentheses. FE Spec Abbreviations: F=Firm, Y=Year, F+Y=Firm+Year, SY=Sector x Year, F+SY=Firm+Sector x Year. Significance levels: * p<0.1, ** p<0.05, *** p<0.01. Controls are winsorized at 1%/99%. Firm_Size is Log(Total Assets). ROE is excluded from controls.

Table 7. Main DiD Results: Impact on Governance Tone (G_Tone_max)

FE Spec:	(1) None	(2) F	(3) Y	(4) F+Y	(5) SY	(6) F+SY
Treated × Post (DiD)	0.0000** (0.0000)	-0.0000 (0.0000)	0.0000** (0.0000)	-0.0000 (0.0000)	0.0000** (0.0000)	-0.0000 (0.0000)
ROA (Wins.)	-0.0000 (0.0000)	-0.0001 (0.0000)	-0.0000 (0.0000)	-0.0001 (0.0000)	-0.0000 (0.0000)	-0.0001 (0.0000)
Leverage (Wins.)	-0.0000 (0.0000)	-0.0001 (0.0001)	-0.0000* (0.0000)	-0.0001 (0.0001)	-0.0000 (0.0000)	-0.0001 (0.0001)
Log(Total Assets)	0.0000*** (0.0000)	-0.0000 (0.0000)	0.0000*** (0.0000)	-0.0000 (0.0000)	0.0000*** (0.0000)	-0.0000 (0.0000)
ESG Score (Wins.)	0.0037*** (0.0004)	0.0057*** (0.0007)	0.0034*** (0.0004)	0.0059*** (0.0008)	0.0035*** (0.0004)	0.0058*** (0.0008)
Observations	43,146	43,146	43,146	43,146	42,418	42,418
Adj. R2	0.014	0.292	0.016	0.292	0.014	0.296
Firm FE	No	Yes	No	Yes	No	Yes
Year FE	No	No	Yes	Yes	No	No
Sector x Year FE	No	No	No	No	Yes	Yes

Notes: Standard errors clustered by Firm_ID in parentheses. FE Spec Abbreviations: F=Firm, Y=Year, F+Y=Firm+Year, SY=Sector x Year, F+SY=Firm+Sector x Year. Significance levels: * p<0.1, ** p<0.05, *** p<0.01. Controls are winsorized at 1%/99%. Firm_Size is Log(Total Assets). ROE is excluded from controls.

Control Variables

- ESG Performance (ESG_Score_W)

This stands out as the most consistent and influential factor across all models and tone categories. The coefficient is always positive and highly significant ($p < 0.01$), confirming that firms with stronger ESG performance tend to use more optimistic language in their communications. This aligns with signalling theory: firms that perform better want to highlight their strengths.

- Other financial controls — ROA_W, Leverage_W, and Firm_Size — generally show weak or inconsistent results. Their significance and direction vary across models and tone types. This suggests that ESG performance matters more than traditional financial indicators when it comes to tone.

Model Explanatory Power (R^2)

The inclusion of firm fixed effects significantly improves the explanatory power of the models. In particular, the adjusted R^2 values exceed 0.70 for ESG_Tone_max and E_Tone_max in the most comprehensive specifications (see Table 4 and Table 5). This suggests that a substantial portion of the variation in tone levels is driven by stable, unobserved firm-specific characteristics, such as corporate culture or managerial preferences.

In contrast, the within R^2 — which captures variation within firms over time — remains relatively low. This implies that although policy effects and control variables account for some of the tone variation over time, much of the explanatory power is attributed to differences across firms rather than changes within firms. The findings highlight the persistent role of firm identity in shaping ESG communication practices.

Overall, the regression results show that mandatory ESG disclosure policies led to a small but statistically significant increase in Social tone (see Table 6), while effects on ESG_Tone_max and E_Tone_max are not robust under full fixed effects. No consistent effect is found for G_Tone_max. Across all models, ESG_Score_W remains the strongest and most consistent predictor of communication tone.

5.5 Placebo tests

In this study, various placebo tests were conducted to verify the reliability of the main findings and the underlying causal relationship. These tests help determine whether the observed effects are genuinely due to the mandatory ESG reporting policy or potentially driven by random factors.

Fake Intervention Year (2010)

As detailed in Section 2, a placebo test was conducted using 2010, a year well before the actual policy implementation. The estimated interaction coefficient ($\text{Treated} \times \text{Post}_{2010}$) is +0.0007 and statistically insignificant ($p = 0.156$), as reported in Table 8, Column (1). This finding supports the parallel trends assumption and confirms that the main results are not driven by pre-existing differences between treatment and control groups.

Random Year Test

Another placebo test was carried out by randomly assigning a pre-policy year as the treatment period for each firm. The interaction coefficients in Table 8, Column (3), and Table 9, Column (3), were statistically insignificant for both ESG_Tone_max and E_Tone_max . This supports the conclusion that the main findings are not driven by coincidental timing or random firm-level shocks.

Pseudo-Intervention Year (2016)

Table 8, Column (2), and Table 9, Column (2), present results using 2016 as a pseudo-intervention year. The coefficient for ESG_Tone_max is +0.0007 ($p < 0.05$), and for E_Tone_max it is +0.0006 ($p < 0.05$), indicating statistically significant tone differences even before the policy came into effect. These results may reflect early

tone improvements in anticipation of upcoming regulations or responses to international developments.

This result suggests that firms in the treatment group may have already begun improving their ESG communication tone even before the formal introduction of the mandatory disclosure policy. It may also indicate that tone trends between the treatment and control groups started to diverge slightly prior to policy implementation. Although this implies that the baseline estimates might be slightly overstated, the magnitude of the effect is small and does not undermine the overall validity or robustness of the main results.

The Impact of the Paris Agreement

The small but significant improvement in tone observed in 2016 may not be attributable solely to regulatory expectations. Global developments, particularly the Paris Climate Agreement signed in December 2015, likely played an important role. The Paris Agreement raised global awareness and urgency regarding environmental sustainability issues, encouraging firms to strengthen their ESG-related communication.

Supporting this view, a recent study by Fan, Shen, Wang, and Zhou (2025) found that firms operating in carbon-intensive sectors in China significantly increased their carbon information disclosure after the Paris Agreement. Moreover, these firms reduced speculative (i.e., symbolic or showy) disclosure behaviours, with stronger emphasis on substantive reporting.

Table 8. Placebo Test Results: Impact on Overall ESG Tone (E_Tone_max) with Alternative Intervention Years

	(1)	(2)	(3)
Placebo Specification:	Placebo Year 2010	Placebo Year 2016	Placebo Random Year
FE Specification:	F+SY	F+SY	F+SY
Treated x Placebo Post (Placebo DiD Term)	0.0007 (0.0005)	0.0007** (0.0002)	-0.0005* (0.0003)
ROA (Wins.)	-0.0000 (0.0002)	-0.0000 (0.0002)	-0.0000 (0.0002)
Leverage (Wins.)	-0.0003 (0.0003)	-0.0003 (0.0003)	-0.0003 (0.0003)
Log(Total Assets)	-0.0001 (0.0001)	-0.0001 (0.0001)	-0.0001 (0.0001)
ESG Score (Wins.)	0.2393*** (0.0048)	0.2392*** (0.0048)	0.2393*** (0.0048)
Observations	42,418	42,418	42,418
Adj. R2	0.739	0.740	0.739
Firm FE	Yes	Yes	Yes
Sector x Year FE	Yes	Yes	Yes

Notes: Dependent variable is ESG_Tone_max. All models include Firm (F) and Sector x Year (SY) fixed effects. Standard errors clustered by Firm_ID in parentheses. Significance levels: * p<0.1, ** p<0.05, *** p<0.01. Controls are winsorized at 1%/99%. Firm_Size is Log(Total Assets). ROE is excluded from controls. Column (1) uses 2010 as the fake intervention year. Column (2) uses 2016. Column (3) uses a randomly assigned pre-treatment year per country.

Table 9. Placebo Test Results: Impact on Environmental Tone (E_Tone_max) with Alternative Intervention Years

	(1)	(2)	(3)
Placebo Specification:	Placebo Year 2010	Placebo Year 2016	Placebo Random Year
FE Specification:	F+SY	F+SY	F+SY
Treated x Placebo Post (Placebo DiD Term)	0.0004 (0.0003)	0.0006** (0.0002)	-0.0001 (0.0003)
ROA (Wins.)	-0.0000 (0.0002)	-0.0000 (0.0002)	-0.0000 (0.0002)
Leverage (Wins.)	-0.0003 (0.0002)	-0.0003 (0.0002)	-0.0003 (0.0002)
Log(Total Assets)	-0.0001 (0.0001)	-0.0001 (0.0001)	-0.0001 (0.0001)
ESG Score (Wins.)	0.1486*** (0.0045)	0.1485*** (0.0045)	0.1486*** (0.0045)
Observations	42,418	42,418	42,418
Adj. R2	0.712	0.712	0.712
Firm FE	Yes	Yes	Yes
Sector x Year FE	Yes	Yes	Yes

Notes: Dependent variable is E_Tone_max. All models include Firm (F) and Sector x Year (SY) fixed effects. Standard errors clustered by Firm_ID in parentheses. Significance levels: * p<0.1, ** p<0.05, *** p<0.01. Controls are winsorized at 1%/99%. Firm_Size is Log(Total Assets). ROE is excluded from controls. Column (1) uses 2010 as the fake intervention year. Column (2) uses 2016. Column (3) uses a randomly assigned pre-treatment year per country.

Table 10. Placebo Test Results: Impact on Social Tone (S_Tone_max) with Alternative Intervention Years

	(1)	(2)	(3)
Placebo Specification:	Placebo Year 2010	Placebo Year 2016	Placebo Random Year
FE Specification:	F+SY	F+SY	F+SY
Treated x Placebo Post (Placebo DiD Term)	0.0001 (0.0002)	0.0002 (0.0001)	-0.0002 (0.0002)
ROA (Wins.)	0.0001 (0.0001)	0.0001 (0.0001)	0.0001 (0.0001)
Leverage (Wins.)	0.0000 (0.0002)	0.0000 (0.0002)	0.0000 (0.0002)
Log(Total Assets)	0.0000 (0.0001)	0.0000 (0.0001)	0.0000 (0.0001)
ESG Score (Wins.)	0.0848*** (0.0027)	0.0848*** (0.0027)	0.0848*** (0.0027)
Observations	42,418	42,418	42,418
Adj. R2	0.567	0.567	0.567
Firm FE	Yes	Yes	Yes
Sector x Year FE	Yes	Yes	Yes

Notes: Dependent variable is S_Tone_max. All models include Firm (F) and Sector x Year (SY) fixed effects. Standard errors clustered by Firm_ID in parentheses. Significance levels: * p<0.1, ** p<0.05, *** p<0.01. Controls are winsorized at 1%/99%. Firm_Size is Log(Total Assets). ROE is excluded from controls. Column (1) uses 2010 as the fake intervention year. Column (2) uses 2016. Column (3) uses a randomly assigned pre-treatment year per country.

Table 11. Placebo Test Results: Impact on Governance Tone (G_Tone_max) with Alternative Intervention Years

	(1)	(2)	(3)
Placebo Specification:	Placebo Year 2010	Placebo Year 2016	Placebo Random Year
FE Specification:	F+SY	F+SY	F+SY
Treated x Placebo Post (Placebo DiD Term)	-0.0000 (0.0000)	-0.0000 (0.0000)	-0.0000 (0.0000)
ROA (Wins.)	-0.0001 (0.0000)	-0.0001 (0.0000)	-0.0001 (0.0000)
Leverage (Wins.)	-0.0001 (0.0001)	-0.0001 (0.0001)	-0.0001 (0.0001)
Log(Total Assets)	-0.0000 (0.0000)	-0.0000 (0.0000)	-0.0000 (0.0000)
ESG Score (Wins.)	0.0058*** (0.0008)	0.0058*** (0.0008)	0.0058*** (0.0008)
Observations	42,418	42,418	42,418
Adj. R2	0.296	0.296	0.296
Firm FE	Yes	Yes	Yes
Sector x Year FE	Yes	Yes	Yes

Notes: Dependent variable is G_Tone_max. All models include Firm (F) and Sector x Year (SY) fixed effects. Standard errors clustered by Firm_ID in parentheses. Significance levels: * p<0.1, ** p<0.05, *** p<0.01. Controls are winsorized at 1%/99%. Firm_Size is Log(Total Assets). ROE is excluded from controls. Column (1) uses 2010 as the fake intervention year. Column (2) uses 2016. Column (3) uses a randomly assigned pre-treatment year per country.

CHAPTER 6

CONCLUSION & DISCUSSION

This thesis examined how mandatory Environmental, Social, and Governance (ESG) disclosure rules influence the tone of firms' communication in their earnings conference calls (ECCs). The main aim was to contribute to the academic literature and offer practical insights for policymakers. Based on Difference-in-Differences (DiD) estimations, a significant increase was found in the Social tone (S_Tone) after ESG mandates, suggesting that firms use more positive language when discussing social topics after such regulations are introduced.

However, no consistent or statistically significant effect was found on the overall ESG tone (ESG_Tone_max), Environmental tone (E_Tone), or Governance tone (G_Tone) once firm and sector-by-year fixed effects were added. This shows that mandatory ESG policies do not have a uniform effect on all ESG components. The positive change in social tone might be driven by firms' need to respond to rising public and stakeholder expectations on social issues.

Additionally, the ESG Score was positively related to tone scores in most models. This indicates that firms with higher ESG performance also use more positive language in their communication (Ioannou & Serafeim, 2015). In contrast, financial control variables like ROA, and leverage showed weak or inconsistent effects, suggesting that tone is more strongly influenced by ESG factors than by short-term financial indicators.

This research also made a methodological contribution by using advanced natural language processing (NLP) models. The FinBERT-Tone model classified

each sentence in ECCs as positive, negative, or neutral (Huang, Wang, & Yang, 2022). In addition, the FinBERT-ESG model was used to identify whether each sentence referred to Environmental, Social, Governance topics, or none. This model was fine-tuned on manually annotated ESG-related sentences from company filings, enabling it to capture domain-specific ESG signals with high precision (Huang, Wang, & Yang, 2022). Furthermore, ESG-BERT models—EnvironmentalBERT, SocialBERT, and GovernanceBERT—developed by Schimanski et al. (2023), were employed to further categorize ESG-related themes in firm communication. This ensemble of models allowed for a more nuanced and topic-specific tone analysis, providing deeper insights into how firms adapt their communication in response to ESG disclosure mandates.

In conclusion, this thesis shows that mandatory ESG rules can shape how companies talk about ESG issues, especially social ones. But the effects depend on firm-level characteristics and sectoral dynamics. This suggests that firms respond to both external regulatory pressures and internal factors when shaping their communication tone.

CHAPTER 7

LIMITATIONS & FURTHER STUDIES

This study provides important findings on how the tone of earnings conference calls (ECCs) changes after ESG disclosure rules become mandatory. However, there are several limitations.

First, the analysis looks at the full text of ECCs without separating different parts, such as the management's prepared speech and the Q&A session with analysts. These sections serve different purposes. The Q&A often reflects more natural and direct answers, while prepared parts are more controlled. Because we combine both, we may miss tone differences between them. Future studies could analyse these parts separately to understand how managers use tone in different settings.

Second, this study focuses only on written text. It does not include how the manager's voice sounds—such as pitch, volume, or pauses—which can also affect how listeners understand the message (Mayew & Venkatachalam, 2012). Using speech data in the future could help improve the analysis.

Moreover, the study uses a Difference-in-Differences design to compare before and after the ESG rules. This helps find causal effects. But it is still a short-term analysis. Future work can use longer time periods to see if the tone changes are temporary or long-lasting.

Additionally, we do not explore differences across sectors. ESG reporting often depends on the industry. For example, the energy and finance sectors face more pressure. Future studies can compare tone changes between sectors (Berg et al., 2022).

Finally, while this study shows that tone changes after ESG rules, it does not look at how this affects investors or the public. Future studies could connect tone scores with outcomes such as stock returns, news coverage, or firm reputation (Eccles & Serafeim, 2013; Tetlock, 2007).

REFERENCES

- Amel-Zadeh, A. & Serafeim, G. (2018). Why and how investors use ESG information: Evidence from a global survey. *Financial Analysts Journal*, 74(3), 87–103. <https://doi.org/10.2469/faj.v74.n3.2>
- Angrist, J. D. & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.
- Alta'any, M., Kayed, S., Meqbel, R., & Albitar, K. (2024). Speaking success: Managerial tone in earnings conference calls and financial performance. *Corporate Governance*, 25(2), 311–336. <https://doi.org/10.1108/CG-09-2023-0381>
- Berg, F., Koelbel, J. F., Pavlova, A., & Rigobon, R. (2022). Aggregate confusion: The divergence of ESG ratings. *Review of Finance*, 26(6), 1315–1344. <https://doi.org/10.1093/rof/rfac033>
- Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How Much Should We Trust Differences-in-Differences Estimates? *Quarterly Journal of Economics*, 119(1), 249–275. <https://doi.org/10.1162/003355304772839588>
- Christensen, H. B., Hail, L., & Leuz, C. (2021). Mandatory sustainability disclosure and stock liquidity: Evidence from the EU Directive. *Review of Accounting Studies*, 26(3), 686–717.
- Deegan, C. (2002). Introduction: The legitimising effect of social and environmental disclosures. *Accounting, Auditing & Accountability Journal*, 15(3), 282–311.
- Devlin, J. D., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *NAACL 2019 Proceedings*, 4171–4186. <https://doi.org/10.18653/v1/N19-1423>
- Druz, M., Petzev, I., Wagner, A. F., & Zeckhauser, R. J. (2020). When managers change their tone, analysts and investors change their tune. *Financial Analysts Journal*, 76(2), 1–18.
- Dzieliński, M., Eugster, F., Sjöström, E., & Wagner, A. F. (2022). Climate talk in corporate earnings calls (*MISUM Working Paper 2022/6*). https://swopec.hhs.se/hamisu/papers/hamisu2022_006.1.pdf
- Eccles, R. G., Ioannou, I. & Serafeim, G. (2014). The impact of corporate sustainability on organizational processes and performance. *Management Science*, 60(11), 2835–2857.
- Eccles, R. G. & Serafeim, G. (2013). The performance frontier: Innovating for a sustainable strategy. *Harvard Business Review*, 91(5), 50–60.

- Fan, X., Shen, X., Wang, D., & Zhou, C. (2025). The Paris Agreement and firms' carbon disclosure: Honesty or catering? *International Review of Economics and Finance*, 99. <https://doi.org/10.1016/j.iref.2025.103988>
- Huang, A. H., Wang, H., & Yang, Y. (2022). FinBERT: A Large Language Model for Extracting Information from Financial Text. *Contemporary Accounting Research*. <https://doi.org/10.1111/1911-3846.12832>
- Huang, W., Yang, Q., Wilson, C., & Li, B. (2023). FinBERT-Tone: A transformer-based model for financial sentiment analysis. *Journal of Financial Data Science*, 5(4), 10–30.
- Huber, M. (2023). *Causal Analysis, Impact Evaluation, and Causal Machine Learning with Applications in R*. MIT Press. <https://doi.org/10.7551/mitpress/12337.001.0001>
- Ioannou, I. & Serafeim, G. (2017). The consequences of mandatory corporate sustainability reporting. *HBS Working Paper 11 -100*. <https://doi.org/10.2139/ssrn.1799709>
- Kaplan, R. S. & Ramanna, K. (2021). Accounting for climate change. *Harvard Business Review*, 99(6), 120–131.
- Krueger, P., Sautner, Z., Tang, D. Y., & Zhong, R. (2024). The effects of mandatory ESG disclosure. *Journal of Accounting Research*, 62(5), 1795–1847. <https://doi.org/10.1111/1475-679X.12548>
- Loughran, T. & McDonald, B. (2011). When is a liability not a liability? *The Journal of Finance*, 66(1), 35–65.
- Mayew, W. J. & Venkatachalam, M. (2012). The power of voice. *The Journal of Finance*, 67(1), 1–43.
- Matsumoto, D., Pronk, M. & Roelofsen, E. (2011). What makes conference calls useful? *The Accounting Review*, 86(4), 1383–1414.
- Paszke, A., et al. (2019). PyTorch: An imperative style, high-performance deep learning library. *NeurIPS*, 32, 8024–8035.
- Rambachan, A. & Roth, J. (2020). Modeling dynamic treatment effects. *Journal of Econometrics*, 218(1), 104–133.
- Sautner, Z., van Lent, L., Vilkov, G., & Zhang, R. (2023). Firm-level climate change exposure. *The Journal of Finance*, 78(3), 1449–1498.
- Schimanski, S., Müller, K., & Schmidt, F. (2024). ESG-BERT: Domain-specific models for sustainability. *Finance Research Letters*, 60, 103636.

- Schimanski, T., Reding, A., et al. (2023). Bridging the Gap in ESG Measurement. *Finance Research Letters*. <https://doi.org/10.1016/j.frl.2024.104979>
- Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, 87(3), 355–374.
- Suchman, M. C. (1995). Managing legitimacy. *Academy of Management Review*, 20(3), 571–610.
- Tetlock, P. C. (2007). Giving content to investor sentiment. *The Journal of Finance*, 62(3), 1139–1168.
- Wolf, T., Debut, L., Sanh, V., et al. (2020). Transformers: State-of-the-art NLP. *EMNLP Demonstrations*, 38–45. <https://doi.org/10.18653/v1/2020.emnlp-demos.6>