ANALYST MOTHER TONGUE LANGUAGE NEGATIVITY AND FORECAST OPTIMISM

ZIHUA LIU

PhD

The Hong Kong Polytechnic University

2023

The Hong Kong Polytechnic University

School of Accounting and Finance

Analyst Mother Tongue Language Negativity and Forecast Optimism

Zihua LIU

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

May 2023

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

____ (Signed)

_____________Zihua LIU_______________ (Name of student)

Analyst Mother Tongue Language Negativity and Forecast Optimism

Abstract

I examine whether and how the level of negativity of US financial analysts' mother tongue language affects their earnings forecast optimism. By collecting negative emotional words (expressing death, diseases and violence) from 25 different languages, I construct a novel measure of language negativity at the country-level covering 47 countries around the world, capturing a country's general tendency to use negative narratives in citizen's daily life. I find that financial analysts with their mother tongue language characterized by a high level of negativity tend to issue less optimistic earnings forecasts. Additional result suggests that the effect of language negativity on analyst forecast optimism tends to be stronger (1) during financial crisis period; (2) for firms with loss, a high level of earnings volatility, and analyst with limited attention; (3) for younger analysts and analysts working for a smaller brokerage firm. Additional results suggest that higher levels of language negativity may dissuade analysts from making excessively optimistic forecasts, ultimately result in a decrease in analyst overall forecast errors. Overall, the finding of this study supports the conjecture that the level of narrative negativity across languages can have a significant impact on capital market participants' behavior. Thus, it sheds light on how culturally inherited emotion can affect analyst forecast optimism.

Keywords: equity analysts; forecast optimism; negativity; culture; emotion

JEL Classification: *F30,* G30, G40

Acknowledgements

First and foremost, I want to express my deepest thanks to my chief supervisor, Prof. Jingran ZHAO, for her steadfast guidance and support throughout my doctoral studies and research. I am extremely grateful for her thorough attention and care in every aspect of my research and job searching. Dr. ZHAO has consistently been attentive and offered invaluable assistance during each stage of my doctoral dissertation. This has ultimately allowed me to complete my thesis.

I also want to express my heartfelt appreciation to my former supervisor, Prof. Albert TSANG, who also closely followed my research and also provided guidance on my Ph.D. thesis. I am also grateful to Professor TSANG for generously sharing his academic expertise and research skills with me. Prof. TSANG has a knack for clarifying complex academic concepts with simple, daily life examples, inspiring and stimulating my thinking. This has not only ignited my academic interests but also laid a solid foundation for my long-term academic career.

I am also grateful to my former supervisor Prof. Agnes CHENG, who, despite her busy schedule, regularly organizes internal discussion seminars, invites world-renowned scholars, and genuinely cares for the well-being of her doctoral students by organizing team-building activities. This has allowed me to enjoy both the cutting-edge academic frontier and heartfelt human care throughout my doctoral journey, making the entire process truly enjoyable. Additionally, I want to give special thanks to the Hong Kong Polytechnic University for providing an academic platform that has allowed me to connect with like-minded scholars and classmates, creating a mutually beneficial learning environment.

Lastly, I want to express my sincere gratitude to my family for their unconditional support and love. I am especially thankful to my parents for constantly encouraging me to be bold and enterprising and for backing up my personal life. My heartfelt thanks also go to my loving wife, Iris, and my son, Eason, for their patience and unwavering support throughout my Ph.D. studies. Without their sacrifices, this accomplishment would not have been possible.

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

An extensive research documents that language has a significant effect on the economic behavior of capital market participants and/or economic entities. For instance, previous studies find that a language's future-time reference (FTR) structure (i.e., to what extent a language grammatically separates the future and the present) explains individual speakers' saving rates, health behaviors, and retirement assets (Chen 2013). Following Chen (2013), subsequent studies document a significant association between FTR and earnings management (Kim et al., 2017), corporate future orientation (Liang et al., 2018), management earnings forecasts (Guan et al., 2022), and tax avoidance (Na and Yan 2022). Another strand of literature suggests that emotion can have an influential role in affecting economic decisions and outcomes in financial markets (e.g., Hirshleifer and Shumway 2003; Goetzmann et al., 2015; Cortés et al., 2016; Momtaz 2020; Chen et al., 2022; Wang et al., 2022).

Despite the importance of languages and emotion in explaining corporate behaviors and outcomes as documented by prior studies, to my knowledge, no study examines whether and how languages would have an effect on capital market participants' emotion and ultimately influence their behaviors. This is surprising, as evidence lends a strong support to the significant link between languages and emotions (e.g., Lindquist et al., 2006; Barrett et al., 2007; Lieberman et al., 2007; Gendron et al., 2012; Lindquist and Gendron 2013; Gendron et al., 2014; Lindquist et al., 2014; Lindquist et al., 2015). Thus, in this study, I fill this void, by examining the link between the level of negativity associated with the mother tongue language[1] of financial analysts and their

[1] In accordance with Dodds et al.'s (2015) research, I define language negativity as a linguistic attribute characterized by the comparatively lower average tone of frequently used words within the specific language. Language negativity is a neutral and objective tool that has been used in scientific research to measure the tone pattern of language. It is important to note that this measure is not intended to imply any form of discrimination or prejudice towards a particular language or culture. Rather, it is a tool that can help researchers better understand the emotional content of language and its potential impact on decision-making.

earnings forecast optimism (i.e., an outcome variable likely to be affected by individual's emotion). I focus on analyst EPS forecasts because it often used by investors to make investment decisions and by managers to inform corporate decision-making. In contrast to other analyst outputs, such as stock recommendations, EPS forecasts are continuous variables with a wide range of variations, making it possible to precisely quantify the impact of language negativity on forecast optimism.

In this paper, I propose that the general tendency to use negative narratives in the mother tongue language of a financial analyst can have an influential effect in shaping the emotion of financial analysts. Specifically, I predict that a greater level of language negativity in the mother tongue language of a financial analyst is negatively associated with the level of analyst earnings forecast optimism.

To test this prediction, I start with collecting the most emotionally negative words commonly perceived by native speakers of all languages examined by Dodds et al. (2015). These words[2] could be grouped into three major categories: death, disease, and violence. [3] I then conduct translation using Google translation and identify the synonym for each of the emotionally negative words selected for each of the 25 languages (covering 47 countries) examined in my study, with the help of native speakers. As a result, I construct a novel language negativity measure for a large number of languages in my study. Specifically, based on the synonym collected, I define narrative negativity at the language level. I define language negativity as the total number of emotionally negative words describing the most negative events (i.e., death, disease, and violence) divided by

[2]These most emotionally negative words including *Cancer, Death, Massacre, Suicide, Drug Abuse, Rape* and *AIDS*.
[3] Although by estimating the emotional content of 100,000 words spread across 25 corpora in 10 languages, Dodds et al. (2015) provides a ranking of negativity by 10 languages, given the limited number of languages examined, I do not rely on Dodds et al. (2015)'s finding in my study. Instead, after I construct the data of language negativity for 25 languages, I validate my data using the language negativity ranking provided by Dodds et al. (2015).

the total word counts of the largest dictionary[4] of each language. This language negativity measure indicates that while Arabic and Chinese tend to have the most pronounced level of language negativity, Finnish and Portuguese appear to have the lowest level of language negativity.

To further validate the quality of the language negativity measure created by me, I employ two strategies. First, I validate my measurement using the average tone of all types of news in a country. To the extent that a high level of language negativity is correlated with the level of emotion of the speakers, I posit that people speaking such language would exhibit a greater level of emotional contents in their daily life which can be captured by the word choices of media. The data is from the GDELT Project.[5] By regressing the average tone of all types of news in a country provided by the GDELT project on the language negativity measure created by us, I find a significant positive association between them supporting the validity of my measure in measuring emotional content of a language. Second, I conduct a similar validation test by regressing the tone score of each language provided by Dodds et al. (2015) on my language negativity measure. [6] Again, I find a significant positive association between them indicating support to the validity of my measure in capturing the emotional content of each language.

I next explore the implication of language negativity on financial analysts' earnings optimism. I focus on financial analyst working for US brokerage firms because relative to a cross country setting, a single country setting reduces the concern that any significant finding may be resulted

[4] Wikipedia lists the main dictionaries for the different languages and corresponding total number of words. The specific link is: https://en.wikipedia.org/wiki/List_of_dictionaries_by_number_of_words . If a language has more than one dictionary, I pick the dictionary with largest number of words. My dictionary word count statistics was obtained as of 10 December 2021, the latest updated word count statistics may slightly higher than my data.

[5] The GDELT project is supported by Google Jigsaw and the project monitors the tones/sentiments of all articles on broadcast, print, and the web news disseminated in over 100 languages in countries around the world (https://www.gdeltproject.org/).

[6] Although it is possible to examine the research question by directly employing data from Dodds et al. (2015)'s tone score measure, Dodds et al (2015).'s measure is only available for 10 languages while my measure covers 25 languages. In additional test, using Dodds et al. (2015) as an additional test in examining the association between language negativity and analyst forecast optimism, I find results consistent with my finding.

by the omitted but correlated country-level institutions. To empirical test the conjecture that language negativity of the mother tongue language of financial analysts will affect their earnings forecast optimism, I first infer the analysts' countries of ancestry/mother tongue language by matching each analyst's surname with the most common surnames of each country around the world (Pan et al., 2017). Accordingly, I classify a financial analyst with a common foreign surname as an analyst with mother tongue language from that foreign country.

Using 956,105 earnings forecasts associated with 10,401 firms from 1994 to 2021 made by 4,302 financial analysts, I first find that relative to financial analysts with mother tongue language characterized with a low level of language negativity, those with mother tongue languages characterized with higher level of language negativity tend to exhibit a lower level of earnings forecast optimism. This result holds after controlling for various analyst and firm characteristics documented by prior studies with the potential to affect analyst earnings forecasts along with firm- and year-fixed effects. This result also holds after I take into consideration the potential influence of six country-level cultural factors suggested by Hofstede which captures a country's power distance, individualism, uncertainty avoidance, masculinity, long-term orientation, and indulgence value (Hofstede 1980 a & b; Hofstede and Minkov 2010) on analyst optimism. The finding also indicates a substantial economic magnitude. For instance, analysts whose native languages are characterized by a high level of negativity on average tend to exhibit earnings forecast optimism levels approximately 4.4% lower than those of analysts whose native languages are less negative.

Through additional cross-sectional tests, I further find the effect of language negativity on analyst forecast optimism tends to be stronger (1) during financial crisis period; (2) for firms with loss, a high level of earnings volatility and for analyst with limited attention, and (3) for younger analysts and analysts working for a smaller brokerage firm. Taken together, these findings suggest that analysts' resource and experience, analyst limited attention and firms' information uncertainties (alternatively the level of difficulty in making analysts forecasts) as well, play an important moderating role in the effect of culturally inherited emotion[7] in affecting analyst forecast optimism. This finding is also consistent with the argument that emotion/mood likely exhibit stronger influence in the judgement process when the economic agents face a greater level of forecast difficulty which is proxied by uncertainty, limited attention and incomplete information access (Forgas 1995, 2008; de Vries et al., 2008).

I also perform a series of additional robustness tests to further reinform my conclusion. These tests include, for example, excluding all financial analysts who are classified as US-domestic analysts, excluding all foreign firms listed in the U.S., excluding countries with more than one official/dominant language (i.e., Canada, Switzerland, and Belgium), controlling for country level variables such as Economic Policy Uncertainty and/or a comprehensive set of country-level variables identified by Isidro et al., (2020),[8] using weight-least square regression estimation instead of ordinary least square regression, and defining language negativity measure differently and employing alternative analyst forecast optimism variable (Hong and Kubik 2003).

This paper contributes to the literature in few major ways. First, to my knowledge the paper is among the first in the accounting and finance literature which quantifies language negativity across a large number of languages worldwide and subsequently apply such metric to examine the

[7] The culturally inherited emotion I define, akin to Hofstede's six key dimensions of culture (power distance, individualism, uncertainty avoidance, masculinity, long-term orientation, and indulgence value), is a cultural attribute possessing a "sticky" characteristic and being capable of intergenerational transmission (Guiso et al., 2006).

[8] Isidro et al. (2020) survey a large number of studies published in the last two decades and identify a total of 72 country-level variables that measure differences in economic, cultural, institutional, and societal development across countries. Specifically, by performing factor analysis on all these institutional features, they find that four core country-level institutional factors collectively explain a substantial amount of the observed variation in financial reporting quality across countries. Specifically, these four country-level institutional factors are likely to capture the four distinct country-level factors — (1) *legal system*, (2) *creditor and/or investor rights*, (3) *political process*, and (4) *societal closeness*—that separately explain the heterogeneity in corporate voluntary disclosure practices across countries.

effect of culturally inherited emotion on analyst forecast optimism. Although a well-established literature examines the role of language or culture in affecting a wide range of economic activities, [9] no study has examined whether and how language would have an effect on capital market participants' emotion which in turn affect their decision-makings. This is surprising as evidence suggesting that analysts who play a valuable role in improving market efficiency (Healy and Palepu 2001) tend to have diverse cultural background (Merkley et al., 2020) which may play a significant role in their forecast properties. I fill this gap by introducing a new emotion-related cultural variable (i.e., language negativity) into the literature and examine the effect of mother tongue language negativity of financial analysts on their earnings forecasts. Unlike the temporary state emotion(i.e., weather-induced mood) that studied in the prior literature (i.e., Hirshleifer and Shumway 2003; Goetzmann et al., 2015; Cortés et al., 2016), my constructed language-induced mood, which may be inherited culturally, could potentially influence the emotional experiences of language users in their daily lives, and may even be transmitted to future generations through intergenerational communication (i.e., Chen, 2013; Lindquist et al. 2015; Liu, 2016; Cao et al., 2022b). My study shed light on how persistent and long-lasting emotions can affect agents' economic decision-making.

Second, the study contributes to studies examining factors affecting analyst earnings forecast in general and forecast optimism in particular. A large number of studies find that analysts are systematically optimistic in their forecasts and identify various determinants (e.g., Francis and Philbrick 1993; Kang et al., 1994; Dugar and Nathan 1995; Rajan and Servaes 1997; Das et al.,

[9] For the role of language in financial markets, see, for example Chen (2013), Kim et al. (2017), Liang et al. (2018), Guan et al. (2022), Na and Yan (2022); for the role of culture in financial decisions see, for instance, Stulz and Williamson (2003), House et al. (2004), Hope et al. (2008), Han et al. (2010), Shao et al. (2010), Dhaliwal et al. (2012), Ioannou, and Serafeim (2012), Bilinski et al. (2013), Li et al. (2013); Shao et al. (2013), Zheng et al. (2013), Kanagaretnam et al. (2014), Alesina and Giuliano (2015), Liu (2016), Brochet et al. (2019).

1998; Lin and McNichols 1998; Chaney et al., 1999; Easterwood and Nutt 1999; Darrough and Russell 2002; Hong and Kubik 2003; Cowen et al., 2006; Libby et al., 2008). The paper finding adds to this literature by presenting evidence suggesting the important role which language and/or culturally inherited emotion can play in affecting analyst forecast optimism. Moreover, my study responds to the call of Ramnath et al., (2008, page 68) by exploring cultural factors with the potential to create cross-country differences in the properties of analyst forecasts and sheds understanding of the impact of culture on accounting information and capital market activities.

Finally, this study adds to the studies examining cross-country factors influencing analyst earnings forecast properties. While previous studies tend to focus more on the effect of formal institutions including for example, the mandatory adoption of International Financial Reporting Standards, the level of stakeholder protection, the enforcement of accounting standards, country-level media competition, legal institution and product market competition in affecting the properties of analyst forecasts (e.g., Barniv et al., 2005; DeFond and Hung 2007; Bae et al., 2008; Haw et al., 2010; Dhaliwal et al., 2012; Haw et al., 2015; Cao et al., 2022a), this study presents evidence suggesting that informal institution (specifically, culturally inherited emotion resulted by language negativity) also affects analyst forecast properties.

The rest of the paper is structured as follows. Section 2 reviews relevant literature and develops the hypotheses. In Section 3, I discuss the sample, variable definition and research methodology. Section 4 presents the main empirical results examining whether and how financial analysts' mother tongue language negativity affects their earnings forecast optimism. I also examine whether such an effect would be moderated by analyst and firm characteristics. In Section 5 I discuss additional robustness test results. Finally, Section 5 concludes the paper.

Chapter 2. Literature Review and Hypothesis Development

2.1. Determinants and Analyst Earning Forecast Optimism

In examining the determinants of cross-sectional differences in analysts' forecast optimism, previous studies find that analysts' forecast optimistic bias is associated with analysts' incentive to obtain access to private information, the predictability of earnings, international diversification, the profitability of firms, firm size, analyst following and analyst affiliation and trading incentive (e.g., Das et al., 1998; Lim 2001; Brown 2001a; Duru and Reeb 2002; Cowen et al., 2006; Mola and Guidolin 2009). Prior studies also document significant heterogeneity in analysts' earnings forecast properties across countries suggesting that cultural factors may play a role in explaining level of analysts' forecast optimism across countries (e.g., Barniv et al., 2005; Bae et al., 2008; Dhaliwal et al., 2012; Bilinski et al., 2013; Cao et al., 2022b; Pursiainen 2022; Tsang et al., 2022).

2.2. Cultures and Economic Decision-making

A growing line of research suggests that one's culture would have significant impacts on economic decision-making and outcomes (Guiso et al., 2006; Giannetti and Yafeh 2012; Ahern et al., 2015; Liu 2016). The literature suggests that inherited cultural heritage can have a lasting and persistent effect on individuals (Brochet et al., 2019). Consistent with this view, studies suggest that culture which is an important part of the tradition that is transmittable from generation to generation (Guiso et al., 2006) can play a crucial role in analyst forecasts. For example, Bhagwat and Liu (2020) find that cultural trust affects how analysts process information from outside sources. Merkley et al. (2020) find that greater level of cultural diversity among analysts is associated with higher quality of consensus forecasts.

2.3 Culturally Inherited Emotion and Analyst Earning Forecast Optimism

An extensive psychological literature suggests that language represents an important element in shaping human emotions (Barrett et al., 2007; Gendron et al., 2012; Lindquist and Gendron 2013; Lindquist et al., 2015).[10] For instance, according to the *Conceptual Act Theory (CAT)* by Lindquist et al. (2015), emotions are not innate but are constructed by individuals through the integration of affective and conceptual processes. In this framework, language plays a crucial role in assisting humans to acquire concept knowledge, subsequently enabling humans to make sense of emotional perceptions and experiences. Moreover, Chen (2013) shows that the human language system could create attention and precision effects for its users. The study reveals that embedding time markers in language could enhance users' attention on time (linguistic-attention effect) and results in more accurate beliefs regarding the timing of future rewards (linguistic-precision effect). In light of this, I posit that when a language's commonly used words consist of more negative terms (e.g., death, violence, and disease), it draws the attention of language users to these negative events (linguistic-attention effect) and bring about vivid perception of these adverse events to language users (linguistic-precision effect). Through a series of psychological processes mentioned by Lindquist et al. (2015), the attention and vivid perception of the negative side elicited by language ultimately lead language users to a relatively downbeat emotional status. Compared to short-lived emotions like those influenced by the weather (i.e., Hirshleifer and Shumway 2003; Goetzmann et al., 2015; Cortés et al., 2016), language-induced emotions tend to have a more persistent presence in daily life and may be conveyed to future generations through intergenerational communication (i.e., Chen, 2013; Lindquist et al., 2015; Liu, 2016; Cao et al., 2022b). Hence, I label the language-induced emotion as a culturally inherited emotion.

Consistent with the notion that concept knowledge supported by language plays a constitutive

[10] Lending support to this view, research documents that impairing people's access to the meaning of emotion words (e.g., disgust, anger, fear) impairs their ability to perceive emotions on faces subsequently (Lindquist et al. 2006; 2014; Gendron et al. 2012).

role in emotions, evidence from neuroscience also points to a crucial connection between language and emotion. Lieberman et al. (2007) discover that using emotion words to label emotional facial expressions helps reduce activity in the brain's uncertainty-related regions (e.g., amygdala), indicating that language assists participants in making sense of ambiguous emotions. Evidence from cross-cultural emotional research lends further support to the conclusion that language plays a constitutive role in emotion. Gendron et al. (2014) reveal that speakers of different languages use distinct perceptual cues to differentiate emotion categories, such as anger, disgust, fear, happiness, sadness, and neutrality. In line with these studies, based on most commonly used words across 10 languages (English, Spanish, French, German, Brazilian Portuguese, Korean, Chinese (simplified), Russian, Indonesian, and Arabic) collected from an array of sources including books, news outlets, social media, the web, television and movie subtitles, and music lyrics, Dodds et al. (2015) reveals a significant interlanguage variation in the emotional spectrum of languages.[11]

In addition, existing studies present evidence suggesting that the bad (good) emotion of managers can lead pessimism (optimism) corporate decisions (Hirshleifer and Shumway 2003; Goetzmann et al., 2015; Cortés et al., 2016; Momtaz 2020; Chen et al., 2022; Wang et al., 2022). This can be attributed to the mood congruency effect, which suggests that people in negative moods are more likely to notice and focus on negative information, while those in positive moods are more inclined to recognize and concentrate on positive information. As a result, when individuals are experiencing negative moods, they tend to exhibit more unfavorable evaluations across various aspects, including future prospects, life satisfaction, past experiences, and even interpersonal relationships. (e.g., Isen et al. 978; Forgas and Bower 1987). More importantly, given

[11] Their research finds that among the 10 languages examined, Chinese presents the highest level of negativity whereas Spanish tends to exhibit the lowest level of negativity (see, “It’s official: Chinese is the saddest language”, available at https://www.thatsmags.com/china/post/8771/its-official-chinese-is-the-saddest-language).

the well-documented empirical regularity about analysts' forecast optimism in countries around the world (Bradshaw et al., 2019), and the robust evidence on cross-country differences in analyst forecast properties (e.g., Barniv et al., 2005; Bae et al., 2008; Dhaliwal et al., 2012; Bilinski et al., 2013; Cao et al., 2022 a & b; Pursiainen 2022; Tsang et al., 2022), it follows from these two literatures that analysts' earnings forecast optimism are likely to vary across cross-country cultural differences in general, and culturally inherited emotion associated with language in particular.

As such, in my study, I argue that the level of negativity associated with the mother tongue language of financial analysts can have a significant impact on analyst average emotion status and further affect their earnings forecast optimism. Based on the discussion above, I formulate my main hypothesis in the following:

> **H1:** *Analysts' earnings forecast optimism is negatively associated with the level of negativity of the mother tongue language of the analysts.*

Next, I explore for contexts in which the language negativity effect on analysts' forecast optimism is expected to have cross-sectional variations. Studies in psychology argue that mood tends to exhibit a greater level of influence in the judgement process of economic agents in setting characterized with a high level of ambiguity, uncertainty and incomplete information (Forgas 1995, 2008; de Vries et al., 2008). Similarly, other studies (e.g., Clore et al., 1994; Cortés et al., 2016) show a stronger effect of mood on economic decisions when such decisions require more subjective judgment and discretion from the decision-makers. Following this view, I expect language negativity to have a stronger effect in influencing analysts' forecast optimism during financial crisis period, when analyst face limited attention, and when firms have a high level of volatile earnings or experience loss. Presumably, in these situations, making accurate analyst earnings forecasts is inherently more difficult given the greater level of ambiguity and uncertainty associated with firms' future performance and analyst limited attention to do rigorous analytical thinking. Following the discussion above, I propose the second hypothesis of my study in the following:

> **H2:** *The effect of language negativity on analysts' earnings forecast optimism is more pronounced when firms' forecast difficulty is high.*

Finally, a growing literature examines how analyst forecast can be affected by observable analyst characteristic (Hanlon et al., 2022). Prior studies on analyst forecast accuracy find that analyst forecast can be affected by various analyst specific characteristics or attributes, including for instance, analyst ability, resources and experience (e.g., Mikhail et al., 1997; Clement 1999; Jacob et al., 1999; Brown 2001b; Drake and Myers 2011; Lehmer et al., 2022). Lending support to this view, Hong and Kubik (2003) show that analysts' forecast optimism facilitates more favorable analyst career outcomes especially for less experienced analysts. Similarly, Dong et al. (2021) find that analyst experience has a negative association with analyst forecast optimism.

Turning to the size of brokerage, research suggests that larger brokers are more likely to provide analysts with superior resources in terms of information access and training etc. in their forecast activities (Lim 2001; Mohanram and Sunder 2006). Consistent with this view, Drake and Myers (2011) find that analysts working for larger brokerage houses tend to have less optimistic forecasts. This leads to my prediction that language negativity to have a stronger effect in influencing analysts' forecast optimism for analysts with fewer experience or work for smaller brokerage. Therefore, I state my third hypothesis as the following:

> **H3:** *The effect of language negativity on analysts' earnings forecast optimism is more pronounced for analysts with fewer experience or work for smaller brokerage.*

Chapter 3. Samples and Empirical Design

In this section, I describe how I construct my measure of language negativity and the validation strategies I have employed. I also describe the firm-year-analyst level data construction procedure and present descriptive statistics for the main variables.

3.1 Language Negativity Measurement and Validation

Dodds et al. (2015) reveals the presence of language negativity heterogeneity across languages and countries/regions. Language negativity is a linguistic attribute characterized by a relatively lower average tone of frequently used words within a specific language, it reflects the propensity of language users to convey sorrowful and negative emotions. In their study, Dodds et al. (2015) collected 24 corpora spanning books, news outlets, social media, the web, television and movie subtitles, and music lyrics in 10 languages, and sorted out the 10,000 most frequently used words in each language. They then identify a large number of native speakers to rate each word's emotional level on a nine-point scale, with 1 denoting the most negative or saddest, 5 representing neutrality, and 9 indicating the most positive or happiest.

Following their work, I develop a method to expand language negativity measurement from 10 languages to 25 languages (covering 47 countries). I start my language negativity construction by gathering the most frequently perceived emotionally negative words across the 10 languages analyzed by Dodds et al. (2015). These seven most emotionally negative words could be categorized into three groups: death, disease, and violence. I primarily rely on language-specific dictionaries to find synonyms for each of these most emotionally negative words in each language. In an effort to capture all words describing death, disease, and violence in various languages, I also utilize Google searches to find additional relevant synonyms and slang terms. Finally, to control for the potential influence of languages with large number of vocabularies, I normalize the total number of negativity words in each language by dividing the total word counts of that language's largest dictionary. Consequently, this methodology offers a broader scope of language negativity measurement (expanding the language negativity measurement from 10 languages to 25 languages), allowing for a more comprehensive analyses across a larger number of languages in countries around the world.

To validate my measurement of language negativity, I implement two strategies. Firstly, I compare my language negativity measurement with the country average news tone calculated from the Global Database of Events, Language, and Tone (GDELT) project. The GDELT dataset extracts location and tone information from various mediums such as broadcast, print, and web sources in over 100 languages via NLP techniques (Leetaru and Schrodt 2013; Manacorda and Tesei 2016; Campante and Yanagizawa 2018). I utilize the "Historical Backfile" subset of GDELT, which covers events from January 1, 1979 to March 31, 2013. For each news event, the data includes the exact date of event occurrence, the precise location (in terms of latitude and longitude of the centroid), and the calculated news tone. Events without location are excluded from my analysis. I average all parsed news tone from various countries and then compare it to my language negativity measurement. Panel A of Appendix D presents evidence of a significant positive correlation between my language negativity measurement and the country's average news tone negativity.[12] Secondly, I also compare the 10 language tones provided by Dodds et al. (2015) with my language negativity measurement. I compute the average tone scores of the 24 corpora by language provided by Dodds et al. (2015) and then regress the 10 language tone score on my language negativity

[12] GDELT project assigns a positive tone value to each news article, with a higher value indicating a more positive tone. To facilitate comparison, I transform the country average news tone to country average news negativity tone using the formula: $1 - {(countryTone - minTone)}\big/{(maxTone - minTone)}$ where *countryTone* equals to country average news tone. *minTone /maxTone* equals to the minimum/maximum country average news tone for matched 45 countries.

measure.[13] In Panel B of Appendix D, I show that the coefficient on language negativity is positive (0.798) and significant at the 1% level, indicating a positive correlation between the two measurements.[14] The two tests provide evidence that my language negativity measurement indeed captures the heterogeneity among users of different languages in expressing sorrowful and negative emotions.

Fig 1. and Fig 2. present language negativity score by language and by countries/regions, respectively. Fig 1. shows that Arabic and Chinese exhibit the most pronounced levels of language negativity, while Finnish and Portuguese have the lowest levels of language negativity. Additionally, English displays a relatively lower level of language negativity. Fig 2. displays the language negativity scores for matched 47 countries that speaks the 25 languages. Yemen and Egypt have the highest language negativity scores due to Arabic is their primarily spoken language, followed by Mainland China, Taiwan (SAR), and Hong Kong (SAR). Finland has the lowest language negativity score due to its use of Finnish. Angola and Brazil, which use Portuguese, have the second-lowest language negativity scores.

3.2 Sample Construction

According to previous research, a person's surname, which is passed down from their parents, can serve as a hereditary indicator of their ancestry, even if their family migrated to the United States many generations ago (Jobling, 2001; Hanks, 2003; Goldstein and Stecklov, 2016). Liu (2016) provides the first evidence in a corporate setting showing that surname-based methods can infer firm managers' ancestry country. Liu (2016) finds that the corruption culture of the manager's ancestry country, inferred from the manager's surname, is positively associated with various forms

[13] Indonesian language is excluded from my validation tests because of lack of analysts from that country, so my regression analysis involves 9 languages from Dodds et al. (2015) only.

[14] To facilitate comparison, I utilized the same algorithm employed in GDELT validation test (footnote 12) to transform Dodds et al. (2015)'s language tone to language negativity tone.

of firm misconduct. Following that, a series of literature provide evidence for the effectiveness of surname-based methods in inferring an individual's ancestry country (e.g., Brochet et al., 2019; Pan et al., 2020). Similar to previous studies, in my study, I match the analyst's surname with their ancestry's country mainly using the Forebears database, a name database that contains over 27 million surnames from 195 countries and lists up to 200 of the most common surnames from each country. (Jung et al., 2019; Bradley et al., 2020; Merkley et al., 2020; Cao et al., 2022b).

To create the firm-year-analyst level data for my empirical analyses, I begin with the complete U.S. Compustat sample spanning from 1994 to 2021. I narrow down my firm-year data by excluding those with missing stock prices in the CRSP dataset and then merge it with the U.S. detailed analyst forecast data from Thomson Reuters' Institutional Brokers' Estimate System (I/B/E/S) database. To address the forecast horizon issue, I focus on one-year-ahead annual earnings forecasts record of U.S.-listed firms between 1994 and 2021 for analysts who work for U.S. brokerage firms. [15] I then use each country's most common surname tables, collected from the Forebears database or Ancestry.com, to infer the analyst's ancestral country from their surname (Pan et al., 2017; Jung et al., 2019; Pacelli, 2019). I designate an analyst to a specific ancestral country if their surname is listed among that country's 200 most common surnames. When a surname appears on multiple countries' most common surname list, I assign the analyst to the country with a higher incidence of individuals bearing the name. When I cannot determine the country of origin of a surname, I expand my surname search using Ancestry.com. After that, I match my country language negativity score with the ancestral country of each financial analyst. My final sample consisting of 956,105 earnings forecasts (i.e., the firm-year-analyst unit of analysis) issued by 4,302 analysts and associated with 10,431 distinct firms during the sample

[15] I exclude any earnings forecasts issued after a firm's actual earnings announcement date, as these forecasts are likely caused by data errors.

period of 1994 to 2021.

3.3 Regression Models

The regression model used in my baseline analysis to examine the relationship between analyst ancestral country language negativity and their earnings forecast optimism is as follows:

$$\begin{aligned} Forecast\_Optimism_{ijt} &= \alpha_0 + \alpha_1 \times Negativity_c + \theta \times X_{ijt-1} + Firm\ Fixed\ Effect \\ &+ Year\ Fixed\ Effect + \epsilon_{ijt} (1) \end{aligned}$$

where i, j, t and c index analyst, firm, and time and analyst ancestry country, respectively. $X_{ijt-1}$is a vector of firm and analyst level control variables for firm/analyst in the most recent year. The dependent variable analyst forecast optimism, *Forecast_Optimism,* equals to the difference between analyst i' s forecasted earnings per share (EPS) for firm j for year t and the actual EPS of firm j for year t scaled by firm j' s stock price on the day prior to the earnings forecast (×100) (e.g. Jackson 2005; Dong et al., 2021). The variable of interest language negativity (*Negativity*) defined as the total number of the most negative words collected by my method, adjusted by the total word count of the largest dictionary for that language. I expect a significant negative coefficient on *Negativity,* suggesting that analyst from country characterized with a high level of language negativity tend to issue less optimism EPS forecast.

Following prior literature, I first control various firm-level characteristics including firm size (*Size*) (Lys and Soo, 1995; Lim, 2001), financial leverage ratio (*Leverage*), book to market Ratio (*BM*) (Atiase, 1985), whether firm experienced loss before analyst forecast (*LOSS*), R&D intensity (*RDIntensity*), return on assets (*ROA*) (Dhaliwal et al., 2011), and firm's most recent five-year earnings volatility (*EarnVol*). I use firm size (*SIZE*) as a proxy for various factors that may influence analysts' incentives to cover a firm and the characteristics of their forecasts, such as overall information availability and investors' attention (Gu and Wu 2003). Studies also shows that analysts have greater motivation to issue optimism forecasts for smaller firms as a means of promoting communication with management and acquiring firm private information (Lim, 2001). High-growth firms may offer more information to aid analysts in making informed earnings forecasts regarding their growth potential (García-Meca and Sánchez-Ballesta, 2011). Hence, I use book-to-market ratio (*BM*) as a proxy for growth opportunities to account for this effect. Research by Gu and Wang (2005) suggests that analysts' forecast biases are more significant for companies with innovative technologies. In line with this, I add R&D intensity (*RDIntensity*) as control variables. Brown (2001a) suggests that estimating losses becomes more difficult due to managerial incentives such as big baths, which can decrease the accuracy of analysts' forecasts. To control for this effect, I include a binary indicator variable (*LOSS*).

I also include a battery of analyst-level characteristics as control variables, which include whether the forecasted firm headquartered country equals to analyst's ancestral country (*Proximity*) (Du et al., 2017), the distance of the U.S. capital city to the analyst's ancestral country capital city (*Distance*), analyst's forecasting ability proxied by previous year average forecast accuracy (*LagAcc*), forecast horizon (*Horizon*) measured as the number of days between the analyst's forecast date and the earnings announcement date, the number of analysts following the forecasted firm (*AnalystFollowing*), analyst firm-specific experience (*FirmExp*), and analyst general experience (*AnalysExp*). Finally, I also include firm and year fixed effects in my analysis to account for time-invariant heterogeneity and other firm-specific invariant characteristics that may have an impact on both an analyst's forecast optimism and analyst ancestry country language negativity.

3.4 Descriptive Statistics

Table 1 presents the summary statistics for the main variables. Panel A reports the sample mean and standard deviation of forecast optimism to be 0.599 and 4.764, respectively, which is consistent with prior research showing that sell-side analysts tend to issue overly optimistic earnings forecasts. Panel B of Table 1 shows mean results of forecast optimism and language negativity by country. The panel (*ID=10*) indicates that I have matched 435 analysts from China, who together issued 49,559 EPS forecasts covering 1,284 US-listed firms. This is comparable to Du et al.'s (2017) manual verification of Chinese ethnic analysts, where they confirmed 333 Chinese ethnic analysts within the sample periods of 1990-2010. Panel B also shows that for identified 1,595 (37.1% of all analysts) US surname analysts together issued 392,270 (41.0% of all forecast) EPS forecasts.

I plot the country mean language negativity and forecast optimism on the panel A and Panel B of Fig 3, respectively. From Fig 3., I could observe that the language negativity measurement covers most countries/regions around the world. Besides, the country language negativity value seems to exhibit a negative correlation with country average forecast optimism. For instance, China is characterized with high language negativity level (dark blue) while present a low average forecast optimism value (light blue). On the other hand, The United States and Canada exhibit relatively low levels of language negativity (light blue), yet analysts in the two countries are more willing to issue optimistic forecasts (dark blue) compared to analysts from other countries. In Fig. 4, where I divide language negativity into four quartiles and calculate the corresponding average analyst forecast optimism, shows that as the level of language negativity increases, the average analyst optimism monotonically decreases (dropped from 0.646 to 0.429). The univariate negative correlation between language negativity and analyst forecast optimism is further confirmed by the correlation analysis in Table 2. Table 2 reports that the pairwise correlation between language negativity and forecast optimism is -0.016, significant at the 1% level.

Chapter 4. Empirical Results

4.1. Main Empirical Results

Table 3 reports the results of estimating equation (1) using two specifications: one with the control variables and another without the control variables. In both specifications, the coefficient on *Negativity* is consistently significant and negative. These results indicate that financial analysts with their mother tongue language characterized by a high level of negativity tend to issue less optimistic earnings forecast. These results support H1 that the level of negativity of financial analysts' mother tongue language can negatively affect their earnings forecast optimism. This finding is consistent with my argument that the general tendency to use negative narratives in the mother tongue language of a financial analyst can have an influential effect in shaping the emotion of financial analysts, thereby affecting financial analysts' earnings forecast optimism.

The observed effect is also economically significant. For example, when I exclude firm and analyst level control variables in Column (1), the estimated coefficient on *Negativity* is negative and statistically significant at the 1% level. As expected, the coefficient on *Negativity* becomes smaller when all the firm and analyst level control variables are added to the regression in Columns (2), but it remains highly significant. These results reveal that when financial analysts with their mother tongue language characterized by a high level of negativity, the earnings forecast issued by them tend to be less optimistic. The observed effects are also economically significant. For example, in Column (2), all of the control variables are included, and the coefficient on *Negativity* is -0.153, which implies that an increase of one standard deviation in language negativity is, on average, associated with a 4.4% (0.153×0.174/0.599) decrease in U.S. analyst EPS forecast optimism. In contrast, one standard deviation increase in analyst forecast ability proxied by analyst previous year forecast accuracy is associated with 24.1% decrease in analyst forecast optimism. To summarize, language negativity's impact on forecast optimism accounts for nearly one-fifth of the influence of analyst forecasting ability on forecast optimism.

As the analysts with U.S. surname contributes to 41.1% of my final sample, I exclude all observations associated with analysts with U.S. surname. The results are displayed in Columns (3) to (4). The coefficient estimate of *Negativity* is negative and significant. In addition, given the analysts associated with countries with multiple official language, it is difficult to identify which language play a more important role in shaping the emotion of financial analysts. Thus, I exclude analysts associated with countries with multiple official languages.[16] The results are displayed in Columns (5) to (6). The coefficient on *Negativity* becomes more negative and highly significant. Taken together, these results are consistent with the argument that financial analysts with their mother tongue language characterized by a high level of negativity are more likely to issue less optimistic earnings forecast as culturally inherited negative emotion should impede their forecast optimism. In terms of the control variables, most of the coefficients have the expected signs. For example, *Analyst Forecast Optimism* is positively associated with, *Loss* and *EarVol* and negatively associated with firm size (*Size*) and analyst experience and ability(*AnalysExp and LagAcc)*(e.g., Lim, 2001; Jackson, 2005; Wong and Zhang, 2014; Brown et al., 2022; Zhao et al., 2022).

<Insert Tables 3 Here>

4.2. Robustness Checks

4.2.1. Alternative Language Negativity Measures

In the benchmark analysis, I use the sum of three subcategories of language negativity scores

---

[16] Countries with multiple official language is defined as countries where less than 80% of the population speaks its most widely spoken language. These countries *are Angola, Belgium, Canada, India, Israel, Luxembourg*, and *Switzerland*.

as a proxy for the general tendency to use negative narratives in the mother tongue language of a financial analyst. [17]To explore whether the results are consistent using alternative construction of language negativity, I further consider three alternative language negativity measures. First, language negativity may vary systematically between three subcategories. To consider this variation, I normalize the three-language negativity subcategory score between 0 and 1 and then average the score, the weighted language negativity could avoid certain subcategory's score dominating the aggregated score. Second, given the mixed effect stemming from financial analysts associated with countries with multiple official languages, I construct the population-weighted average of language negativity scores for all primary languages spoken in the country to address this concern. Third, to address the dimensionality problem, I perform first principal component of three language negativity subcategory score of the country to identify patterns in the variables and explore whether the observed patterns can proxy the level of negativity of financial analysts' mother tongue language. Then I rerun the regressions using the alternative language negativity measures. As shown in Table 4, the coefficients on *Negativity Weight1, Negativity Weight2 and Negativity PCA* are all negative and significant from all models, suggesting that the results don't depend on the use of alternative language negativity measures.

<Insert Tables 4 Here>

4.2.2. Language Negativity Categories

Given that the language negativity measure is composed of three categories, a natural follow-up question is whether specific components of language negativity play a more important role in affecting analysts' earnings forecast optimism. To address this question, I consider the three language negativity dimensions. Table 5 reports the results using the score for each of these three

[17] The three subcategories of language negativity include the following: *Death, Disease*, and *Violence.*

categories as the explanatory variable. I find that except for specification (6), the coefficients on these three subcategories all remain significantly negative, indicating that the results are robust to the each of these three subcategories. The coefficient on *disease category* is bigger and more significant than other two categories, suggesting that the language negativity's negative effect on analysts' earnings forecast optimism potentially results from the aspects of disease.

<Insert Tables 5 Here>

4.3. Cross-sectional Tests

In this section, to shed light on the potential channels through which language negativity decrease analysts' earnings forecast optimism, I conduct several cross-sectional tests.

4.3.1. Moderating Effect of Financial Crisis

It is well documented that the emotion/mood can play a more important role in shaping economic agents' judgement process behavior when facing a greater level of ambiguity, uncertainty and incomplete information, such as financial crisis (Forgas 1995, 2008; de Vries et al., 2008). This evidence implies that it is inherently more difficult for financial analysts to make accurate analyst earnings forecast given the greater level of ambiguity and uncertainty associated with firms' future performance during financial crisis, even though such information is readily available on firms' websites and in their financial reports as before. Therefore, I predict that the effect of language negativity on financial analysts' earnings forecast optimism tends to be stronger during financial crisis. I examine this argument in the following analysis.

To test this conjecture, I generate two indicator variables that measure market condition associated with firms' future performance. Specifically, *Crisis* is an indicator variable that coded 1 if the year belongs to 2000-2002 and 2008, and 0 otherwise; *Boom* is an indicator variable coded 1 if the year belongs to 1997-1999 and 2007, and 0 otherwise. The first proxy indicates the period of financial crisis. The second proxy indicates the opposite market condition. Table 6 presents the results of the cross-sectional tests based on market condition. According to the results in Columns (1) and (3) of Table 6, I find that the coefficients on *Negativity* × *Crisis* is negative and significant. The coefficients on *Negativity* × *Boom* in Columns (2) and (4) is negative but insignificant. These findings are consistent with my conjecture that the effect of culturally inherited emotion on analyst forecast optimism is stronger during financial crisis period as it is inherently more difficult for financial analysts to make accurate analyst earnings forecast during financial crisis.

<Insert Tables 6 Here>

4.3.2. Moderating Effect of Loss/Earnings Volatility/Limited Attention

Prior studies have documented that there is a stronger effect of mood on economic decisions when such decisions require more subjective judgment and discretion from the decision-makers (Clore et al., 1994; Cortés et al., 2016). In line with the view, I predict that effect of language negativity on analysts' earnings forecast optimism becomes stronger when firms have a high level of earnings volatility or experience loss. Similarly, when analysts face limited attention, the impact of the mother tongue language of analysts may play a more important role. As under these situations, making accurate earnings forecast becomes more difficult for analysts given a high level of ambiguity, uncertainty, incomplete information and limited attention, which increase analyst forecast difficulty and require more subjective judgment and discretion.

To test this prediction, I construct three variables to proxy firm's future earnings uncertain and analyst limited attention. Specifically, *Loss* is an indicator variable coded 1 if firm's net income is smaller than zero, and 0 otherwise. *EarnVol* equals to the standard deviation of EPS over the last five years. *LtdAttn* equals to the total number of companies that the analyst followed in that year. Then, I interact *Loss*, *EarnVol,* and *LtdAttn* with *Negativity*, respectively. Table 7 presents the results of the cross-sectional tests based on firm's future earnings uncertain and analyst limited attention. Consistent with my prediction, I find that the coefficients on *Negativity* × *Loss*, *Negativity* × *EarnVol,* and *Negativity* × *LtdAttn* are all negative and significant. This finding is consistent with my conjecture that the effect of language negativity on analysts' earnings forecast optimism tends to be more pronounced when firms exhibit a high level of earnings volatility or analyst face greater level of limited attention since the high level of future earnings uncertain and analyst limited attention makes it more difficult for analysts to make accurate earnings forecast.

<Insert Tables 7 Here>

4.3.3. Moderating Effect of Broker Size/Analyst Experience

In order to further explore the relevance of analysts' characteristics in making accurate earnings forecast in explaining the effect of language negativity on analyst's earnings forecast optimism, I attempt to examine the moderating effect of analysts' resource and experience in the association between the language negativity and analyst's earnings forecast optimism. Previous studies find that analyst forecast accuracy can be affected by various analyst specific characteristics or attributes, including for instance, analyst ability, resources and experience (e.g., Mikhail et al., 1997; Clement 1999; Jacob et al., 1999; Brown 2001b; Drake and Myers 2011; Lehmer et al., 2022). Following this view, Hong and Kubik (2003) show that analysts' forecast optimism facilitates more favorable analyst career outcomes especially for less experienced analysts. Similarly, Dong et al. (2021) find that analyst experience has a negative association with analyst forecast optimism. Turning to the size of brokerage, research suggests that larger brokers are more likely to provide analysts with superior resources in terms of information access and training etc. in their forecast activities (Lim 2001; Mohanram and Sunder 2006).

Therefore, I propose that the effect of language negativity on analysts' earnings forecast optimism is more pronounced for analysts with fewer experience or work for smaller brokerage. To test this hypothesis, I use analysts working experience (*YoungAnalys*) and the broker size (*SmallBroker*) to capture the different analysts' characteristics. Specifically, *YoungAnalys* equals the reverse 10 deciles value of the analyst's working years. *SmallBroker* equals the reverse 10 decile value of broker size, measured as the broker's total number of analysts at a given year. Consistent with my prediction, in Columns (1) to (4) of Table 8, the coefficients on the interaction terms with the two measures of analysts' characteristics are both significantly negative, suggesting that the effect of language negativity on analysts' earnings forecast optimism is more pronounced among younger analysts and analysts working for a smaller brokerage firm as their limited working resource and experience makes it difficult for those to make accurate earnings forecasts.

<Insert Tables 8 Here>

4.4. Additional Analyses

Our findings indicate that the level of negativity of analysts' mother tongue language is negatively associated with their earnings forecast optimism. I attribute this finding to the level of narrative negativity across languages can have a significant impact on financial analysts' behavior due to the culturally inherited emotion, which helps to explain analysts' earnings forecast optimism. In this section, I conduct several additional analyses to illuminate the underlying channels through which language negativity decrease the likelihood of financial analysts' earnings forecast optimism.

Table 9 examines the economic implications of language negativity. Several studies indicate that negative moods often stimulate individuals to engage in detailed analytical activities, while positive moods are linked to less critical information processing methods (Schwarz 1990; Sinclair and Mark 1995; Petty et al. 2020). So that analyst in bad mood may make more accuracy decision. Hence, in table 9, I test whether the language negativity deterrence effect on forecast optimism through emotion mechanism could translate into lower forecast errors of financial analysts. Specifically, I replaced forecast optimism with analyst forecast error, which defined as the absolute difference between analyst forecasted EPS and actual EPS scaled by the stock price on the day prior to the earnings forecast date ×100. The regression results in Table 9 indicates that the coefficients on *Negativity* remain consistently negative and statistically significant across all regression specifications, lending supportive evidence that the language negativity may act as a deterrent to analysts' tendency to issue optimistic forecasts and finally reduce analysts' overall forecast errors.

<Insert Tables 9 Here>

Table 10 examines the robustness of my results. First, to investigate whether the results are consistent using alternative construction of analysts' earnings forecast optimism, I further consider the alternative analysts' earnings forecast optimism measures. Panel A of Table 10 reports the results. Following prior studies (e.g., Cowen et al., 2006), I use the analysts' relative optimism (*Relative_Optimism*)[18], which compared the optimism of a given analyst's earnings forecast with those of all analysts who made forecasts for the same firm-year within a similar forecast horizon to retest my results. The coefficients on *Negativity* remain negative and statistically significant across the alternative analysts' earnings forecast optimism measures, which is consistent with the main findings.

Second, in the benchmark analysis, I focus on the earnings forecast of all analysts and I find that there is a negative association between language negativity and analysts' earnings forecast

[18] $Relative\_Optimism_{ijt} = {(F_{ijt} - Mean(F_{jt}))}\big/{SD(F_{jt})}$ where F_{ijt} is analyst i's forecast of firm j's earnings for year t. $Mean(F_{jt})$/ $SD(F_{jt})$ is the average/ standard deviation of forecast for all analysts who made forecasts for firm j's earnings for year t within the same forecast horizon (e.g., 90/180/360 days before earning announcement), respectively.

optimism. Given that my sample is composed of multiple earnings forecasts issued by the same analysts for a particular firm in a given year, to further ensure the robustness of the findings, I only keep the initial forecast issued by each analyst for firm i in year t. Panel B of Table 10 reports the results. Consistent with the main findings, I find that the estimated coefficients on *Negativity* is consistently negative and significant.

Third, I posit that whether my findings are driven by the external economic policy uncertainty. To rule out the possible alternative explanations and further strengthen my inferences, I expand my benchmark models with additional country level variable, such as *Economic Policy Uncertainty (EPU)*, following previous study (Baker et al., 2016). Panel C of Table 10 displays the results of the regressions with *EPU*. I find that the coefficient on *Negativity* remains significantly negative, indicating that the results are robust to the inclusion of *EPU*.

Forth, to further rule out the alternative explanation that the association between the language negativity and analysts' earnings forecast optimism is mainly triggered by country-level characteristics across countries, I include an additional comprehensive set of country-level variables identified by Isidro et al. (2020), capturing financial analysts' ancestry countries' legal system, creditor and/or investor rights, political process, and societal closeness. Adding these country-level characteristics allows me to consider any unobserved country-specific, time-invariant characteristics, such as formal and/or informal institutions, that may correlate with analysts' earnings forecast optimism. Then, I rerun the regressions with the comprehensive set of country-level variables. Panel D of Table 10 shows the results of the regressions and the finding is unchanged.

Finally, thus far, I have established the overall effect of language negativity on analysts' earnings forecast optimism. To further check the robustness of my results and explore whether the results are consistent when using different models, I employ the weight-least square regression estimation (WLS) instead of ordinary least square regression (OLS) to reset my results and I report the results in Panel E of Table 10. The coefficients on *Negativity* are all negative and statistically significant across all specifications, thereby lending support to my inference that the level of narrative negativity across languages, through culturally inherited emotion, influence analysts' forecast properties. In Panel F, I exclude all foreign firms listed in the US and the coefficient on *Negativity* consistently present significant negative sign.

<Insert Tables 10 Here>

However, I acknowledge that alternative arguments may still exist outside of the robustness check that was performed above. For example, it is possible that the personality traits of analysts' parents are influenced by negativity of their mother tongue, and it is these traits that discourage analysts from issuing optimistic earnings forecasts, rather than solely language-induced emotions. Besides, the selection of analysts by brokerages and the selective coverage by analysts may bias my estimation of language negativity on analyst forecast optimism. It is plausible that brokerages choose analysts based on their linguistic background, while analysts may also selectively decide to follow familiar firms (i.e., firms located in the analyst's ancestry country). Nevertheless, prior research indicates that U.S. analysts exhibit greater forecasting accuracy when predicting EPS for firms domiciled in their ancestry countries without presenting increased optimism for ancestry country firms' earnings prospects (Du et al., 2017). Consequently, the selective coverage by analysts might lead to an underestimation of the influence of language negativity on forecast optimism. These considerations reinforce our primary hypothesis that analysts exhibiting greater language negativity tend to issue less optimistic earnings per share (EPS) forecasts.

Chapter 5. Conclusion

How analysts make their earnings forecast is crucial to the understanding of the efficiency of capital markets. Extending prior studies examining the cross-country formal institutions in affecting analyst forecast properties, in this study, I posit that culturally inherited emotion can affect analyst forecast optimism as well. To test this conjecture, I construct a novel language negativity measure for a large number of languages as prior literature suggests that language plays a crucial role in affecting human emotion.

I define language negativity as the total number of emotionally negative words describing the most negative events (i.e., death, disease, and violence) divided by the total word counts of the main dictionary of each language. Based on these emotionally negative words collected from 25 different languages examined in my study, I measure level of language negativity at the country-level for 47 countries around the world. My primary result indicates that relative to financial analysts with mother tongue language characterized by a lower level of narrative negativity, financial analysts whose mother tongue language is characterized by a high level of negativity tend to issue less optimistic earnings forecasts. In additional cross-sectional tests, I further find that the effect of language negativity on analyst forecast optimism is stronger when analyst forecast difficulty is high. More specifically, the negative association is more pronounced during, financial crisis period, for firms with loss and a high level of earnings volatility, for analyst faced limited attention and for analysts who are less experienced and who works for a smaller brokerage firm. The findings also indicate that language negativity may discourage analysts from issuing overly optimistic forecasts, ultimately leading to a reduction in their overall forecast errors.

Taken together, my study provides evidence supporting the conjecture that the level of narrative negativity across languages can have a significant impact on analyst forecast properties.

I also innovate beyond prior literature by introducing a new emotional dimension of culture into the literature on cross-cultural studies. Thus, my results should be useful to both academic and capital market participants in better understanding cross-country heterogeneity in analyst forecast properties.

References

Ahern, K. R., Daminelli, D., and Fracassi, C. 2015. Lost in translation? The effect of cultural values on mergers around the world. *Journal of Financial Economics* 117(1): 165-189.

Alesina, A., and Giuliano. 2015. Culture and institutions. *Journal of Economic Literature* 53(4): 898-944.

Atiase, R. K. 1985. Predisclosure information, firm capitalization, and security price behavior around earnings announcements. *Journal of Accounting Research*, 23(1), 21-36.

Bae, K. H., Tan, H., and Welker, M. 2008. International GAAP differences: The impact on foreign analysts. *The Accounting Review* 83: 593-628.

Baker, S. R., Bloom, N., and Davis, S. J. 2016. Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, 131(4), 1593-1636

Barniv, R., Myring, M., and Thomas, W. 2005. The association between the legal and financial reporting environments and forecast performance of individual analysts. *Contemporary Accounting Research* 22: 727-758.

Barrett, L. F., Lindquist, K. A., and Gendron, M. 2007. Language as context for the perception of emotion. *Trends in Cognitive Sciences* 11(8): 327-332.

Bhagwat, V., and Liu, X. 2020. The role of trust in information processing: Evidence from security analysts. *The Accounting Review* 95(3): 59-83.

Bilinski, P., Lyssimachou, D., and Walker, M. 2013. Target price accuracy: International evidence. *The Accounting Review* 88(3): 825-851.

Bradshaw, M. T., Huang, A. G., and Tan, H. 2019. The effects of analyst-country institutions on biased research: Evidence from target prices. *Journal of Accounting Research* 57(1): 85-120.

Bradley, D., S. Gokkaya, and X. Liu. 2020. Ties that bind: The value of professional connections to sell-side analysts. *Management Science* 66 (9): 4118-4151.

Brochet, F., Miller, G. S., Naranjo, P., and Yu, G. 2019. Managers' cultural background and disclosure attributes. *The Accounting Review* 94(3): 57-86.

Brown, L. D. 2001a. A temporal analysis of earnings surprises: Profits versus losses. *Journal of Accounting Research* 39: 221-242.

Brown, L. D. 2001b. How important is past analyst forecast accuracy? *Financial Analysts Journal* 57: 44-49.

Brown, A. B., Lin, G., and Zhou, A. 2022. Analysts' forecast optimism: The effects of managers' incentives on analysts' forecasts. *Journal of Behavioral and Experimental Finance*, 35, 100708

Campante, F., and Yanagizawa-Drott, D. 2018. Long-range growth: economic development in the global network of air links. *The Quarterly Journal of Economics*, 133(3), 1395-1458.

Cao, Y., Keskek, S., Myers, L. A., and Tsang, A. 2022a. The effect of media competition on analyst forecast properties: Cross-country evidence. *Journal of Accounting, Auditing, and Finance* 37(2): 378-406.

Cao, Y., Hao, R., and Yang, Y. G. 2022b. National Culture and Analysts' Forecasting Behavior. *Available at SSRN* 3713715.

Chaney, P., Hogan, C., and Jeter, D. 1999. The effect of reporting restructuring charges on analysts' forecast revisions and errors. *Journal of Accounting and Economics* 27: 261-284.

Chen, M. K. 2013. The effect of language on economic behavior: Evidence from saving rates, health behaviours, and retirement assets. *American Economic Review* 103(2): 690-731.

Chen, C., Chen, Y., Pittman, J. A., Podolski, E. J., and Veeraraghavan, M. 2022. Emotions and managerial judgement: Evidence from sunshine exposure. *The Accounting Review* 97(3): 179-203.

Clement, M. 1999. Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? *Journal of Accounting and Economics* 27: 285-303.

Clore, G., Schwarz, N., and Conway, M. 1994. Affective causes and consequences of social information processing. In: Wyer Jr.,,R.S., Srull, T.K. (Eds.), Handbook of Social Cognition, vol. 1. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 323–418 second ed.

Cortés, K., Duchin, R., and Sosyura, D. 2016. Clouded judgement: The role of sentiment in credit origination. *Journal of Financial Economics* 121: 392-413.

Cowen, A., Groysberg, B., and Healy, P. 2006. Which types of analyst firms are more optimistic? *Journal of Accounting and Economics* 41: 119-146.
Darrough, M., and Russell, T. 2002. A positive model of earnings forecasts: Top down versus bottom up. *Journal of Business* 75: 127-152.
Das, S., Levine, C., and Sivaramakrishnan, K. 1998. Earnings predictability and bias in analysts' earnings forecasts. *The Accounting Review* 73: 277-294.
DeFond, M, and Hung, M. 2007. Investor protection and analysts' cash flow forecasts around the world. *Review of Accounting Studies* 12: 377-419.
De Vries, M., Holland, R. W., and Witteman, C. L. M. 2008. Fitting decisions: Mood and intuitive versus deliberative decision strategies. *Cognition and Emotion* 22(5): 931-943.
Dhaliwal, D. S., Li, O. Z., Tsang, A., and Yang, Y. G. 2011. Voluntary nonfinancial disclosure and the cost of equity capital: The initiation of corporate social responsibility reporting. *The Accounting Review*, 86(1), 59-100.
Dhaliwal, D. S., Radhakrishnan, S., Tsang, A., and Yang, Y. G. 2012. Nonfinancial disclosure and analyst forecast accuracy: International evidence on corporate social responsibility disclosure. *The Accounting Review* 87(3): 723-759.
Dodds, P. S., Clark, E. M., Desu, S., Frank, M. R., Reagan, A. J., Williams, J. R., Mitchell, L., Harris, K. D., Kloumann, I. M., Bagrow, J. P., Megerdoomian, K., McMahon, M. T., Tivnan, B. F., and Danforth, C. M. 2015. Human language reveals a universal positivity bias. *Psychological and Cognitive Sciences* 112(8): 2389-2394.
Dong, R., Fisman, R., Wang, Y., and Xu, N. 2021. Air pollution, affect, and forecasting bias: Evidence from Chinese financial analysts. *Journal of Financial Economics* 139: 971-984.
Drake, M. S., and Myers, L. A. 2011. Analysts' accrual-related over-optimism: Do analyst characteristics play a role? *Review of Accounting Studies* 16: 59-88.
Du, Q., Yu, F., and Yu, X. 2017. Cultural proximity and the processing of financial information. *Journal of Financial and Quantitative Analysis*, 52(6), 2703-2726.
Dugar, A., and Nathan, S. 1995. The effect of investment banking relationships on financial analysts' forecasts and investment recommendations. *Contemporary Accounting Research* 11: 131-160.
Duru, A., and Reeb, D. M. 2002. International diversification and analysts' forecast accuracy and bias. *The Accounting Review* 77(2): 415-433.
Easterwood, J., and Nutt, S. 1999. Inefficiency in analysts' earnings forecasts: Systematic misreaction or systematic optimism? *Journal of Finance* 54: 1777-1797.
Forgas, J. P., and Bower, G. H. 1987. Mood effects on person-perception judgments. *Journal of personality and social psychology*, 53(1), 53.
Forgas, J. P. 1995. Mood and judgment: The affect infusion model (AIM). *Psychological Bulletin* 117(1): 39-66.
Forgas, J. P. 2008. Affect and cognition. *Perspectives on Psychological Science* 3(2): 94-101.
Francis, J., and Philbrick, D. 1993. Analysts' decisions as products of a multi-task environment. *Journal of Accounting Research* 31: 216-230.
García-Meca, E., and Sánchez-Ballesta, J. P. 2011. Ownership structure and forecast accuracy in Spain. *Journal of International Accounting, Auditing and Taxation*, 20(2), 73-82.
Gendron, M., Lindquist, K. A., Barsalou, L., and Barrett, L. F. 2012. Emotion words shape emotion percepts. *Emotion* 12(2): 314-325.
Gendron, M., Roberson, D., van der Vyver, J M., and Barrett, L. F. 2014. Perceptions of emotion from facial expressions are not culturally universal: Evidence from a remote culture. *Emotion* 14: 251-262.
Giannetti, M., and Yafeh, Y. 2012. Do cultural differences between contracting parties matter? Evidence from syndicated bank loans. *Management Science* 58: 365-383.
Goetzmann, W., Kim, D., Kumar, A., and Wang, Q. 2015. Weather-induced mood, institutional investors, and stock returns. *Review of Financial Studies* 28: 73-111.
Goldstein, J. R., and Stecklov, G., 2016. From Patrick to John F.: Ethnic names and occupational success in the last era of mass migration. *American Sociological Review*, 81(1): 85-106.

Gu, F., and Wang, W. 2005. Intangible assets, information complexity, and analysts' earnings forecasts. *Journal of Business Finance & Accounting*, 32(9-10), 1673-1702.

Gu, Z., and Wu, J. S. 2003. Earnings skewness and analyst forecast bias. *Journal of Accounting and Economics*, 35(1), 5-29.

Guan, Y., Wang, Z., Wong, M. H. F., and Xin, X. 2022. Language and management forecasts around the world. *Contemporary Accounting Research* 39(1): 50-86.

Guiso, L., Sapienza, P., and Zingales, L. 2006. Does culture affect economic outcomes? *Journal of Economic Perspectives* 20(2): 23-48.

Han, S., Kang, T., Salter, S., and Yoo, Y. K. 2010. A cross-country study on the effects of national culture on earnings management. *Journal of International Business Studies* 41: 121-141.

Hanks, P., 2003. Dictionary of American Family Names. *Oxford Univeristy Press*, Oxford, UK.

Hanlon, M., Yeung, K., and Zuo, L. 2022. Behavioral economics of accounting: A review of archival research on individual decision makers. *Contemporary Accounting Research* 39(2): 1150-1214.

Haw, I. M., Ho, S. S. M., Hu, B., and Wu, W. 2010. Analysts' forecast properties, concentrated ownership and legal institutions. *Journal of Accounting, Auditing and Finance* 25(2): 235-259.

Haw, I. M., Hu, B., and Lee, J. 2015. Product market competition and analyst forecasting activity: International evidence. *Journal of Banking & Finance* 56: 48-60.

Healy, P. M., and Palepu, K. G. 2001. Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics* 31: 405-440.

Hirshleifer, D., and Shumway, T. 2003. Good day sunshine: Stock returns and the weather. *Journal of Finance* 58: 1009-1032.

Hofstede, G. 1980a. Culture and organizations. *International studies of management & organization* 10(4): 15-41.

Hofstede, G. 1980b. *Culture's Consequences: International differences in work related values*. Beverly Hills, CA: Sage Publications.

Hofstede, G., and Minkov, M. 2010. Long-versus short-term orientation: new perspectives. *Asia Pacific Business Review* 16(4): 493-504.

Hong, H., and Kubik, J. 2003. Analyzing the analysts: Career concerns and biased earnings forecasts. *Journal of Finance* 58: 313-351.

Hope, O., Kang, T., Thomas, W., and Yoo ,Y. K. 2008. Culture and auditor choice: A test of the secrecy hypothesis. *Journal of Accounting and Public Policy* 27: 357-373.

House, R. J., Hanges, P. J., Javidan, M., Dorfman, P. W., and Gupta, V. 2004. Culture, leadership and organizations: The GLOBE study of 62 societies. Sage Publications.

Ioannou, I., and Serafeim, G. 2012. What drives corporate social performance? The role of nation-level institutions. *Journal of International Business Studies* 43: 834-864.

Isen, A. M., Shalker, T. E., Clark, M., and Karp, L. 1978. Affect, accessibility of material in memory, and behavior: A cognitive loop?. *Journal of personality and social psychology*, 36(1), 1.

Isidro, H., Nanda, D., and Wysocki, P. D. 2020. On the relation between financial reporting and country attributes: Research challenges and opportunities. *The Accounting Review* 95(3): 279-314.

Jackson, A. R. 2005. Trade generation, reputation, and sell-side analysts. *The Journal of Finance*, 60(2), 673-717.

Jacob, J., Lys, T., and Neale, M. 1999. Expertise in forecasting performance of security analysts. *Journal of Accounting and Economics* 28: 51-82.

Jobling, M. A., 2001. In the name of the father: Surnames and genetics. *Trends in Genetics*, 17(6): 353-357.

Jung, J. H., Kumar, A., Lim, S. S., and Yoo, C. Y. 2019. An analyst by any other surname: Surname favorability and market reaction to analyst forecasts. *Journal of Accounting and Economics*, 67(2-3), 306-335.

Kanagaretnam, K., Lim, C. Y., and Lobo, G. J. 2014. Influence of national culture on accounting conservation and risk-taking in the banking industry. *The Accounting Review* 89(3): 1115-1149.

Kang, S., O'Brien, J., and Sivaramakrishnan, K. 1994. Analysts' interim earnings forecasts: Evidence on

the forecasting process. *Journal of Accounting Research* 32: 103-112.
Kim, J., Kim, Y., and Zhou, J. 2017. Languages and earnings management. *Journal of Accounting and Economics* 63: 288-306.
Lehmer, T., Lourie, B., and Shanthikumar, D. 2022. Brokerage trading volume and analysts' earnings forecasts: A conflict of interest? *Review of Accounting Studies* 27: 441-476.
Leetaru, Kalev, and Philip A. Schrodt, "GDELT: Global Data on Events, Location and Tone, 1979–2012, *International Studies Association Annual Conference*, 2013.
Li, K., Griffin, D., Yue, H., and Zhao, L. 2013. How does culture influence corporate risk-taking? *Journal of Corporate Finance* 23: 1-22.
Liang, H., Marquis, C., Renneboog, L., Sun, S. L. 2018. Future-time framing: The effect of language on corporate future orientation. *Organization Science* 29(3): 1093-1111.
Lim, T. 2001. Rationality and analysts' forecast bias. *Journal of Finance*, 56(1), 369-385.
Libby, R., Hutton, J. E., Tan, H. T., and Seybert, N. 2008. Relationship incentives and the optimistic/pessimistic pattern in analysts' forecasts. *Journal of Accounting Research* 46: 173-198.
Lieberman, M. D., Eisenberger, N. I., Crockett, M. J., Tom, S. M., Pfeifer, J. H., and Way, B. M. 2007. Putting feelings into words. *Psychological Science* 18(5): 421-428.
Lin, H., and McNichols, M. 1998. Underwriting relationship, analyst' earnings forecasts and investment recommendations. *Journal of Accounting and Economics* 25: 101-127.
Lindquist, K. A., Barrett, L. F., Bliss-Moreau, E., and Russell, J. A. 2006. Language and the perception of emotion. *Emotion* 6: 125-138.
Lindquist, K. A., and Gendron, M. 2013. What's in a word: Language constructs emotion perception. *Emotion Review* 5(1): 66-71.
Lindquist, K. A., Gendron, M., Barrett, L. F., and Dickerson, B. C. 2014. Emotion, but not affect perception, is impaired with semantic memory loss. *Emotion* 14: 375-387.
Lindquist, K. A., MacCormack, J. K., and Shablack, H. 2015. The role of language in emotion: Predictions from psychological constructionism. *Frontiers in Psychology* 6: 1-17.
Liu, X. 2016. Corruption culture and corporate misconduct. *Journal of Financial Economics* 122(2): 307-327.
Lys, T., and Soo, L. G. 1995. Analysts' forecast precision as a response to competition. *Journal of Accounting, Auditing & Finance*, 10(4), 751-765.
Manacorda, Marco, and Andrea Tesei 2016, Liberation Technology: Mobile Phones and Political Mobilization in Africa, *CEP Discussion Paper* No. 1419.
Merkley, K. J., Michaely, R., and Pacelli, J. 2020. Cultural diversity on Wall Street: Evidence from sell-side analysts' forecasts. *Journal of Accounting and Economics* 70(1): 1-24.
Mikhail, M. B., Walther, B. R., and Willis, R. H. 1997. Do security analysts improve their performance with experience? *Journal of Accounting Research* 35: 131-157.
Mohanram, P. S., and Sunder, S. V. 2006. How has regulation FD affected the operations of financial analysts? *Contemporary Accounting Research* 23: 491-525.
Mola, S., and Guidolin, M. 2009. Affiliated mutual funds and analyst optimism. *Journal of Financial Economics* 93: 108-137.
Momtaz, P. P. 2020. CEO emotions and firm valuation in initial coin offerings: An artificial emotional intelligence approach. *Strategic Management Journal* 42: 558-578.
Na, K., and Yan, W. 2022. Languages and corporate tax avoidance. *Review of Accounting Studies* 27: 148-184.
Pacelli, J., 2019. Corporate culture and analyst catering. *Journal of Accounting and Economics*, 67(1): 120-143.
Pan, Y., Siegel, S., and Wang, T. Y. 2017. Corporate risk culture. *Journal of Financial and Quantitative Analysis*, 52(6), 2327-2367.
Pan, Y., Siegel, S., and Yue Wang, T. (2020). The cultural origin of CEOs' attitudes toward uncertainty: Evidence from corporate acquisitions. The Review of Financial Studies, 33(7), 2977-3030.
Petty, R. E., Gleicher, F., and Baker, S. M. 2020. Multiple roles for affect in persuasion. *In Emotion and*

social judgments (pp. 181-200). Garland Science.
Pursiainen, V., 2022. Cultural biases in equity analysis. *The Journal of Finance* 77(1): 163-211.
Schwarz, N., 1990, Feelings as information. Informational and motivational functions of affective states, in R. Sorrentino, and E.T.Higgins, eds.: *Handbook of Motivation and Cognition* (Guilford Press, NewYork).
Shao, L., Kwok, C., and Guedhami, O. 2010. National culture and dividend policy. *Journal of International Business Studies* 41: 1391-1414.
Shao, L., Kwok, C., and Zhang, R. 2013. National culture and corporate investment. *Journal of International Business Studies* 44: 745-763.
Sinclair, R. C., and Mark, M. M. 1995. The effects of mood state on judgmental accuracy: Processing strategy as a mechanism. *Cognition & Emotion*, 9(5), 417-438.
Stulz, R. M., and Williamson, R. 2003. Culture, openness, and finance. *Journal of Financial Economics* 70: 313-349.
Tsang, A., Wang, K. T., Wu, Y., and Lee, J. 2022. Nonfinancial corporate social responsibility reporting and firm value: International evidence on the role of financial analysts. *European Accounting Review*, forthcoming.
Ramnath, S., Rock, S., and Shane, P. 2008. The financial analyst forecasting literature: A taxonomy with suggestions for future research. *International Journal of Forecasting* 24: 34-75.
Rajan, R., and Servaes, H. 1997. Analyst following of initial public offerings. *Journal of Finance* 52: 507-529.
Wang, L., Lin, Y., Jiang, W., Yang, H., and Zhao, H. 2022. Does CEO emotion matters? CEO affectivity and corporate social responsibility. *Strategic Management Journal,* forthcoming.
Wong, M. F., and Zhang, X. F. (2014). CEO optimism and analyst forecast bias. *Journal of Accounting, Auditing & Finance*, 29(3),
Zheng, X., Ghoul, S. E., Guedhami, O., and Kwok, C. 2013. Collectivism and corruption in bank lending. *Journal of International Business Studies* 44: 363-390.
Zhao, C., Li, Y., Govindaraj, S., and Zhong, Z. K. (2022). CDS trading and analyst optimism. *The British Accounting Review*, 54(4), 101109.

Appendix A Variable Definition

Variable	Definition	Datasource
Dependent Variable		
Forecast_Optimism	(analyst EPS forecast - actual EPS) / stock price on the day prior to the earnings forecast date ×100.	IBES
Independent Variable		
DeathRatio	The ratio of total number of death related words divided by total words of country largest dictionary (multiply by 1000).	Manual, Forebears
DiseaseRatio	The ratio of total number of disease related words divided by total words of country largest dictionary (multiply by 1000).	Manual, Forebears
ViolenceRatio	The ratio of total number of violence related words divided by total words of country largest dictionary (multiply by 1000).	Manual, Forebears
Negativity	Equals to the sum of three subcategories of language negativity scores divided by the total word counts of the language largest dictionary (DeathRatio, DiseaseRatio, and ViolenceRatio). More specifically, Negativity = *DeathRatio +DiseaseRatio+ViolenceRatio*.	Manual, Forebears
Negativity Weight1	Equals to the average score of three standardized language negativity subcategories (*DeathRatio, DiseaseRatio,* and *ViolenceRatio*). By normalizing the three-language negativity subcategory score between 0 and 1 and then averaging the score, the weighted language negativity could avoid certain subcategory's score dominating the aggregated score.	Manual, Forebears
Negativity Weight2	Equals to the population-weighted average of language negativity scores for all primary languages spoken in the country.	Manual, Forebears
Negativity PCA	First principal component of three language negativity subcategory score of the country.	Manual, Forebears
Firm-Level Control Variables		
Size	The natural log of one plus total asset of the firm in previous year.	Compustat
Leverage	The sum of short-term debt and long-term debt scaled by total assets.	Compustat
BM	The ratio of book equity to market equity for a firm, measured at the most recent December preceding the forecast date.	Compustat, CRSP
LOSS	An indicator variable, with the value 1 if net income is smaller than zero, and 0 otherwise.	Compustat
RDIntensity	The ratio of total research and development expenses to total sales for a given year.	Compustat
Ret	The average monthly return over the last 12 months.	CRSP
ROA	The ratio of income before extraordinary items to total assets.	Compustat
EarnVol	The standard deviation of EPS over the last five years.	Compustat
Analyst-Level Control Variables		
Proximity	Equals to one if analyst originated country equals to forecasted firm headquartered country, and zero otherwise.	Forebears, Compustat
Distance	The logarithm of number of kilometers between US capital city to analyst's ancestry country capital city.	CEPII
LagAcc	Analyst previous year average EPS forecast accuracy, measured as the negative of the average absolute difference between actual and forecasted EPS scaled by stock price on the prior day of earnings forecast ×100.	IBES
Horizon	The logarithm of one plus the number of days between the forecast issue date and the earnings announcement date.	IBES
AnalystFollowing	The natural logarithm of one plus the number of analysts following a firm in the preceding year.	IBES
FirmExp	The natural logarithm of one plus the number of years an analyst has issued one-year-ahead earnings forecasts for a firm.	IBES
AnalysExp	The natural logarithm of one plus the number of years an analyst has appeared in I/B/E/S.	IBES

Appendix B Language Negativity Measure

Panel A Summary Statistics

Variable	**N**	**Mean**	**STD**	**25%**	**50%**	**75%**
Death Ratio	47	0.197	0.142	0.100	0.139	0.215
Disease Ratio	47	0.025	0.017	0.012	0.018	0.033
Violence Ratio	47	0.040	0.029	0.019	0.026	0.056
Negativity	47	0.261	0.174	0.131	0.188	0.376
Negativity Weight1	47	0.306	0.265	0.092	0.176	0.572
Negativity Weight2	47	0.251	0.177	0.131	0.170	0.376
Negativity PCA	47	0.000	1.555	-1.255	-0.756	1.557

Panel B Pearson Correlation

ID	Variables	1	2	3	4	5	6
1	*Death Ratio*						
2	*Disease Ratio*	**0.520**					
3	*Violence Ratio*	**0.668**	**0.918**				
4	*Negativity*	**0.980**	**0.675**	**0.801**			
5	*Negativity Weight1*	**0.796**	**0.918**	**0.970**	**0.901**		
6	*Negativity Weight2*	**0.967**	**0.669**	**0.788**	**0.987**	**0.889**	
7	*Negativity PCA*	**0.794**	**0.920**	**0.970**	**0.900**	**0.990**	**0.888**

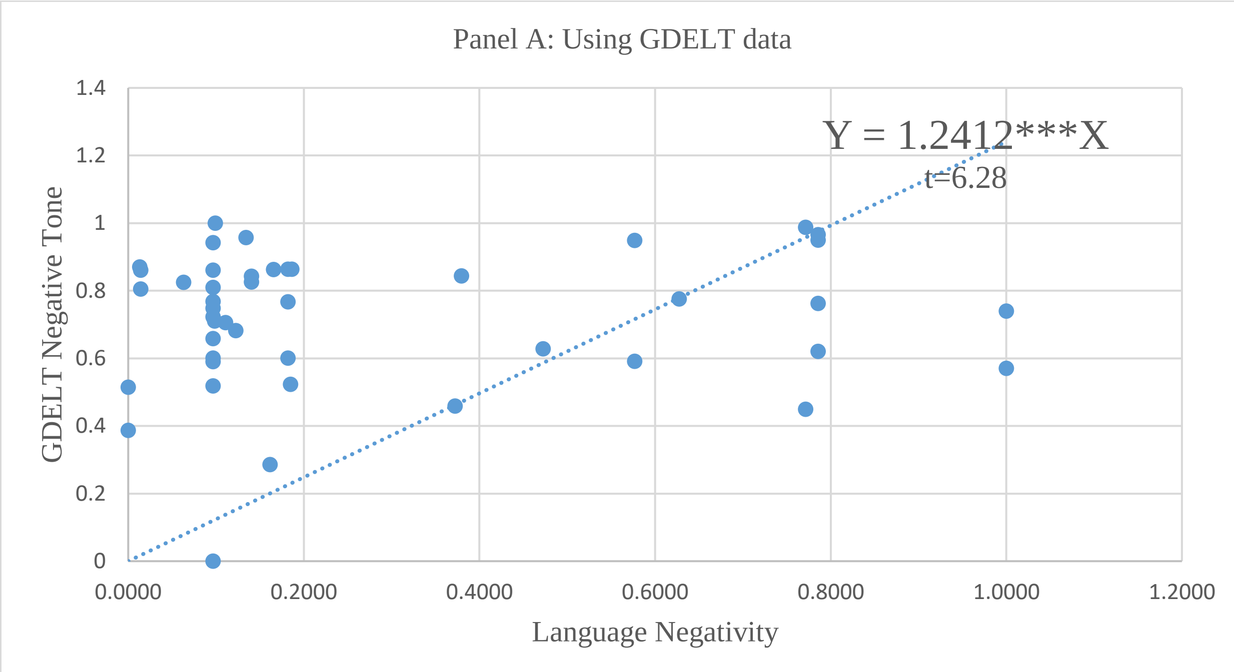
Note: This appendix presents the summary statistics and correlation for my language negativity measures. The top three rows (*Death/Disease/Violence Ratio*) are the three language negativity subcategories. The bottom four row are the aggregate language negativity scores calculated using different aggregation methods. Detailed calculation methods are included in Appendix A.

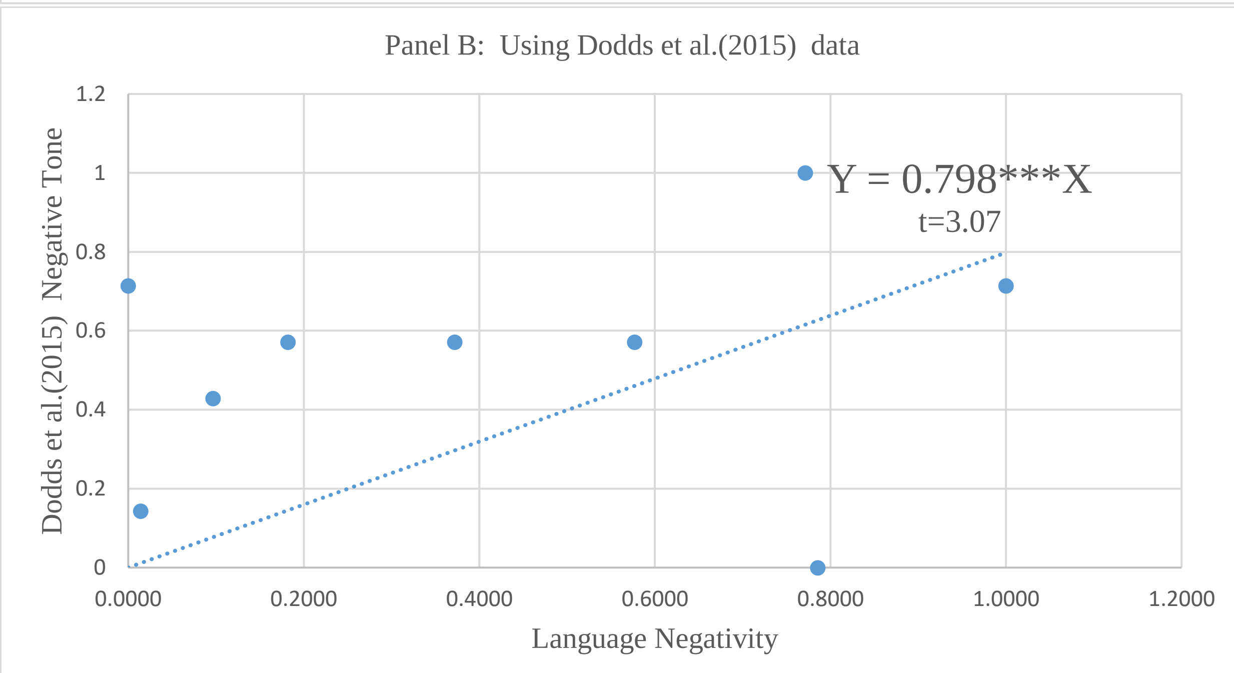
Appendix C Country/Region and Language

ID	Country/Region	Widely Spoken Language	Primary Languages
1	Angola	Portuguese	Portuguese (71.1%)
2	Argentina	Spanish	Spanish
3	Australia	English	English
4	Austria	German	German
5	Bahamas	English	English
6	Belgium	Dutch	Dutch (59%)
7	Belize	English	English (82.9%)
8	Brazil	Portuguese	Portuguese
9	Canada	English	English (54%); French (19%)
10	China	Chinese	Chinese
11	Denmark	Danish	Danmark
12	Egypt	Arabic	Arabic
13	El Salvador	Spanish	Spanish
14	Finland	Finnish	Finnish (88.3%)
15	France	French	French
16	Germany	German	German
17	Greece	Greek	Greek
18	Guinea	French	French
19	Hong Kong (SAR)	Chinese	Chinese (90.9%)
20	Hungary	Hungarian	Hungarian
21	India	Hindi	Hindi (79.80%)
22	Ireland	English	English
23	Israel	Hebrew	Hebrew (49%); Russian (15%);English(2%)
24	Italy	Italian	Italian
25	Jamaica	English	English
26	Lesotho	Sesotho	Sesotho
27	Liberia	English	English
28	Luxembourg	Luxembourgish	Luxembourgish (52%); German (3.2%)
29	Mexico	Spanish	Spanish
30	Netherlands	Dutch	Dutch
31	North Korea	Korean	Korean
32	Norway	Norwegian	Norwegian
33	Poland	Polish	Polish
34	Romania	Romanian	Romanian (91.55%)
35	Russia	Russian	Russian
36	Singapore	English	English (80%); Malay (17%);Tamil (4%)
37	South Korea	Korean	Korean
38	Spain	Spanish	Spanish
39	Sweden	Swedish	Swedish
40	Switzerland	German	German (64%)
41	Taiwan (SAR)	Chinese	Chinese
42	Turkey	Turkish	Turkish
43	United Kingdom	English	English
44	United States	English	English
45	Uruguay	Spanish	Spanish
46	Vietnam	Vietnamese	Vietnamese
47	Yemen	Arabic	Arabic

Note: This table presents the main and primary language spoken by the populations of 47 countries/regions in my sample.

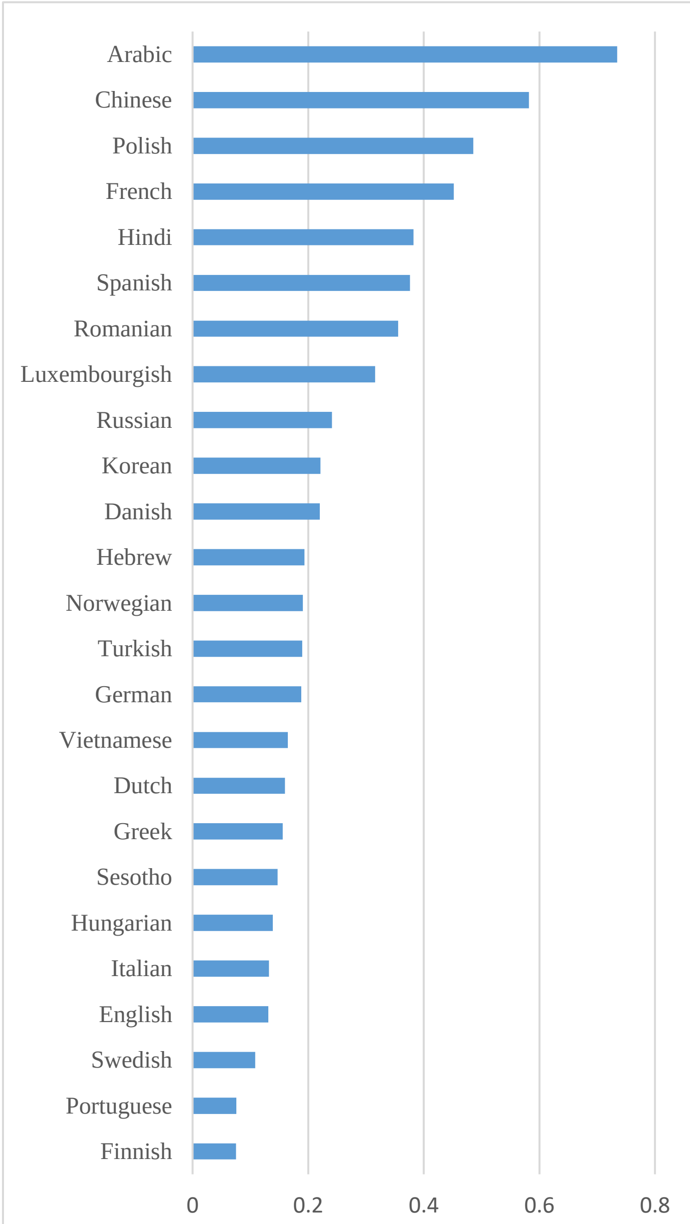
Appendix D: Language Negativity Measure's Validation Test





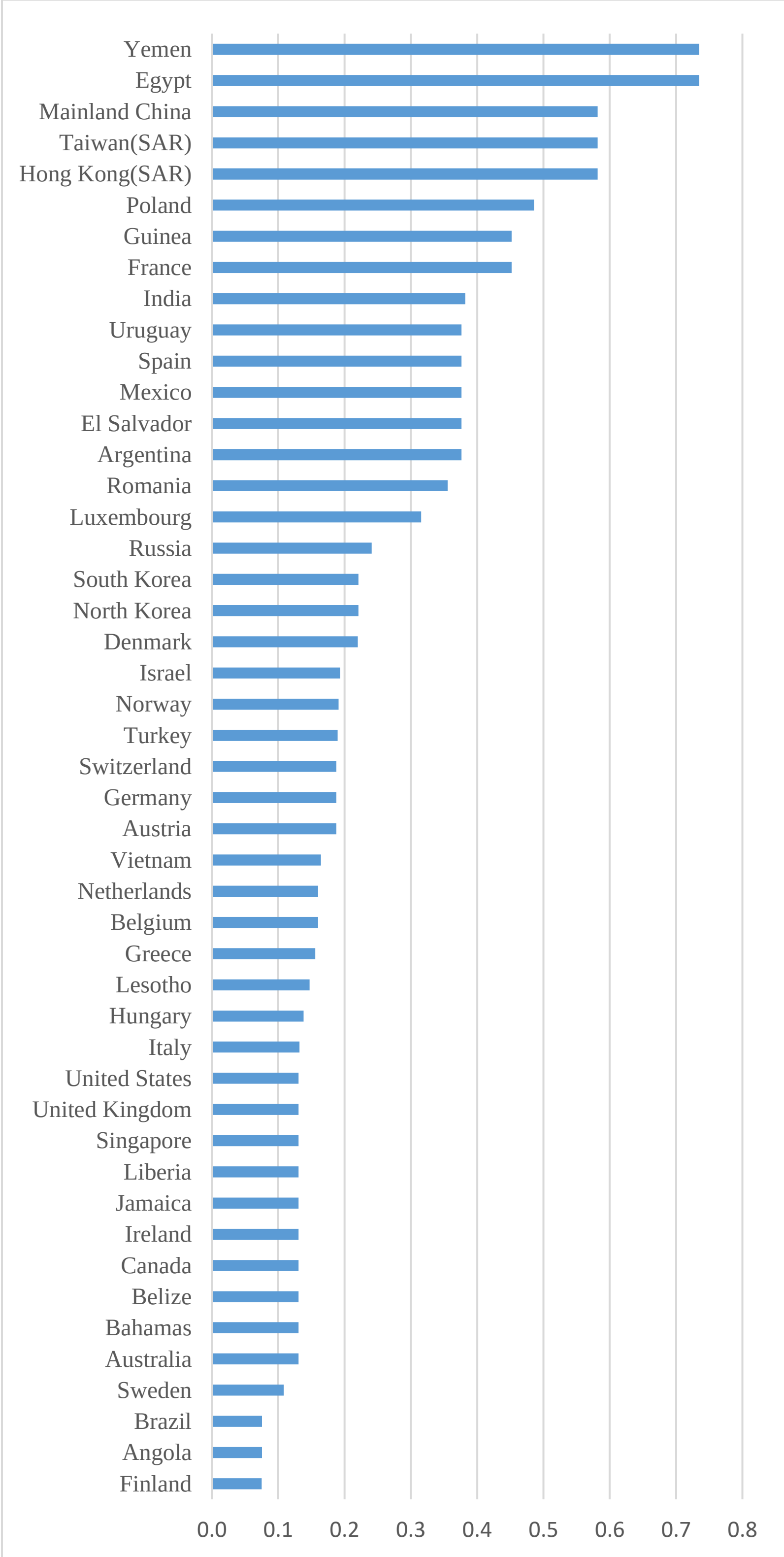
Note: This appendix plots the relationship between language negativity measure constructed by my study and the data from GDELT news negative tone and Dodds et al. (2015) language negative tone. Supported by Google, the GDELT corpus collected world's web news from nearly all countries with over 100 languages. The Google algorithm automatically identifies the news corpus with location and emotion. In Panel A, I averaged all the news tones from the 45 countries matched in my sample and compared it with my country's language negativity measurement. Dodds et al. (2015) calculated the average tone of 10 languages using 100,000 words scored manually by 5 million individual humans. In panel B, I plot the matched 9 languages tone (Indonesian language is excluded due to mismatching.) with my country language negativity measurement. To facilitate comparison, I transform all three measurements to negative tone and scale them between 0 and 1. The higher the value, the more negative the language tone.

Figure 1 Language Negativity Score by Language



Notes: This figure present the language negativity score by language.

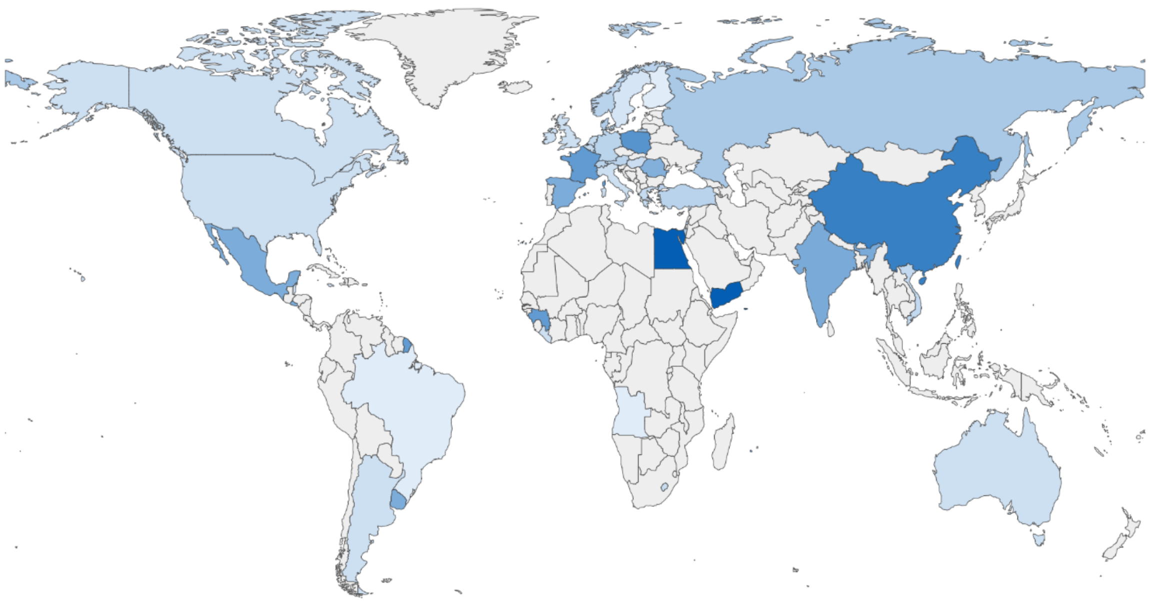
Figure 2 Language Negativity Score by Country/Region



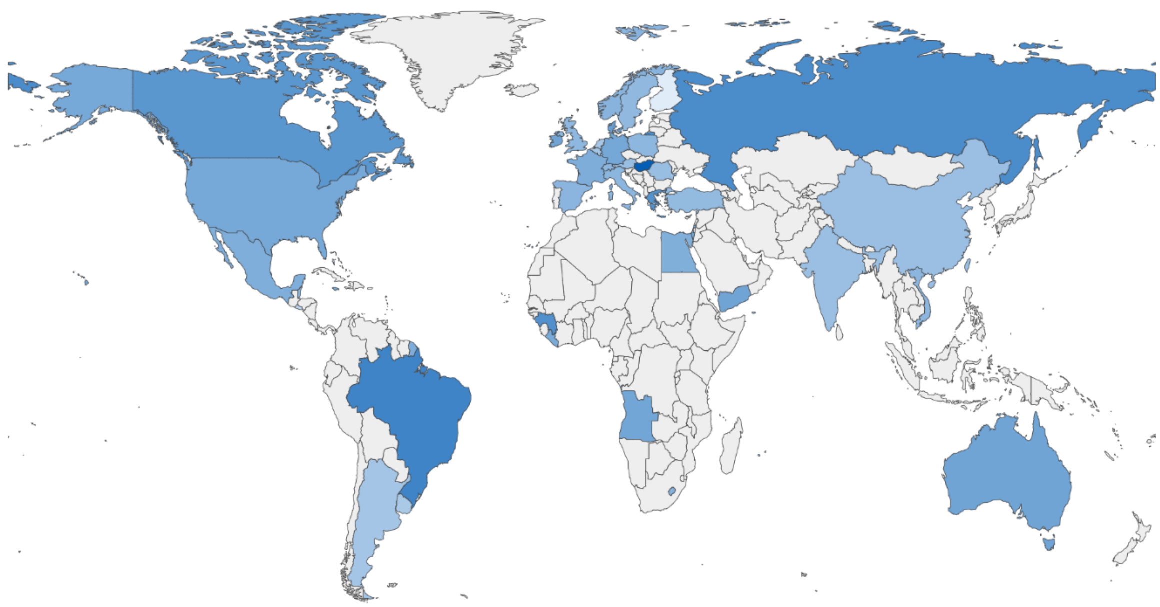
Notes: This figure present the language negativity score by country/region.

Figure 3 Language Negativity and Analyst Forecast Optimism by Country/Region

Panel A: Language Negativity by Country/Region

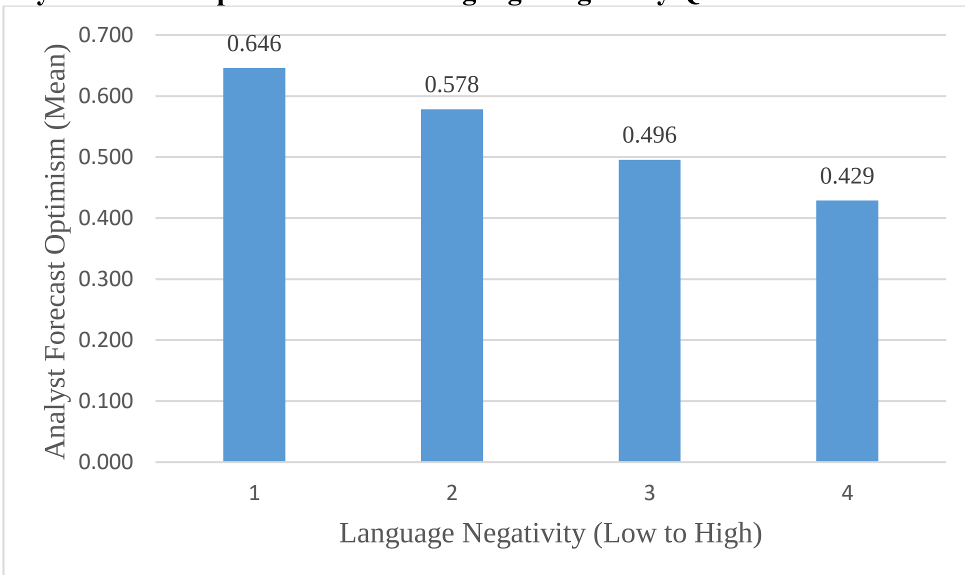


Panel B: Analyst Forecast Optimism by Country/Region



Notes: This figure shows the distribution of language negativity and analyst forecast optimism across various countries and regions. The color scale indicates the magnitude of the values, with darker shades representing higher values.

Figure 4 Analyst Forecast Optimism across Language Negativity Quartiles



Notes: This figure shows the mean analyst forecast optimism across four quartiles based on language negativity score, where quartile 1 contains the observations with the lowest language negativity score and quartile 4 contains the observations with the highest.

Table 1 Descriptive Statistics

Panel A- full sample

Variable	No. of Forecasts	Mean	STD	25%	50%	75%
Forecast_Optimism	956,105	0.599	4.764	-0.432	-0.030	0.573
Negativity	47	0.261	0.174	0.131	0.188	0.376
Negativity Weight1	47	0.306	0.265	0.092	0.176	0.572
Negativity Weight2	47	0.251	0.177	0.131	0.170	0.376
Negativity PCA	47	0.000	1.555	-1.255	-0.756	1.557
Size	956,105	7.665	2.371	6.290	7.763	9.213
Leverage	956,105	0.550	0.264	0.363	0.553	0.735
BM	956,105	0.495	0.414	0.206	0.404	0.681
LOSS	956,105	0.107	0.309	0.000	0.000	0.000
RDIntensity	956,105	0.149	0.707	0.000	0.000	0.051
Ret	956,105	0.113	0.506	-0.177	0.047	0.323
ROA	956,105	0.096	0.169	0.029	0.114	0.184
EarnVol	956,105	1.337	4.437	0.175	0.409	0.913
Proximity	956,105	0.408	0.491	0.000	0.000	1.000
Distance	956,105	8.100	0.956	7.058	8.603	8.872
LagAcc	956,105	-1.616	1.827	-2.114	-1.055	-0.475
Horizon	956,105	5.143	0.731	4.727	5.313	5.684
AnalystFollowing	956,105	2.378	0.816	1.946	2.485	2.996
FirmExp	956,105	1.237	0.849	0.693	1.099	1.792
AnalysExp	956,105	2.281	0.858	1.792	2.485	2.944

Panel B- by country

ID	Country	No. of Forecasts	No. of Firms	No. of Analyst	Forecast Optimism	Negativity
1	Angola	1,042	95	3	0.686	0.076
2	Argentina	1,959	149	6	0.280	0.376
3	Australia	1,918	223	13	0.710	0.131
4	Austria	10,678	505	32	0.377	0.188
5	Bahamas	4,876	245	29	0.117	0.131
6	Belgium	838	98	5	0.279	0.160
7	Belize	1,946	55	7	0.471	0.131
8	Brazil	6,087	452	42	1.111	0.076
9	Canada	11,889	567	58	0.877	0.131
10	China	49,559	1,284	435	0.357	0.582
11	Denmark	7,598	170	47	0.764	0.220
12	Egypt	2,871	60	14	0.569	0.735
13	El Salvador	2,336	2	4	0.109	0.376
14	Finland	1,652	18	6	-0.210	0.075
15	France	15,365	212	64	0.588	0.452
16	Germany	70,883	1,315	271	0.662	0.188
17	Greece	3,286	27	9	0.925	0.156
18	Guinea	3,470	24	5	0.960	0.452
19	Hong Kong	9,585	69	76	0.342	0.582
20	Hungary	1,341	6	6	1.585	0.139
21	India	32,202	344	191	0.363	0.382
22	Ireland	95,091	891	335	0.709	0.131
23	Israel	55,651	420	181	0.671	0.193
24	Italy	19,342	128	83	0.480	0.132
25	Jamaica	14,693	97	53	0.605	0.131
26	Lesotho	938	5	4	0.526	0.147
27	Liberia	1,772	15	12	0.634	0.131
28	Luxembourg	4,582	44	16	0.319	0.316
29	Mexico	13,657	135	95	0.579	0.376
30	Netherlands	2,760	10	19	0.274	0.160
31	North Korea	1,580	9	11	0.131	0.221
32	Norway	5,480	24	30	0.511	0.191
33	Poland	7,783	46	31	0.469	0.486
34	Romania	2,307	13	16	0.381	0.356
35	Russia	1,058	2	12	1.007	0.241
36	Singapore	2,452	19	16	1.373	0.131
37	South Korea	8,048	44	94	0.359	0.221
38	Spain	3,069	18	10	0.487	0.376
39	Sweden	3,420	32	27	0.438	0.108
40	Switzerland	15,593	467	31	0.473	0.188
41	Taiwan	1,129	4	20	0.372	0.582
42	Turkey	5,748	34	14	0.418	0.190
43	UK	38,963	586	168	0.457	0.131
44	US	392,270	1,433	1,595	0.647	0.131
45	Uruguay	856	7	2	0.239	0.376
46	Vietnam	15,574	34	99	0.476	0.165
47	Yemen	908	4	5	0.707	0.735
	Overall	*956,105*	*10,441*	*4,302*	*0.599*	*0.198*

Panel C- by year

Year	No. of Forecasts	No. of Firms	No. of Analyst	Forecast Optimism	Negativity
1994	21,579	2,867	649	0.769	0.170
1995	22,778	2,991	655	0.839	0.173
1996	23,729	3,340	728	0.764	0.176
1997	25,356	3,539	811	0.720	0.183
1998	28,364	3,482	908	0.772	0.182
1999	26,381	3,298	953	0.771	0.181
2000	25,577	3,116	1,012	0.948	0.179
2001	28,183	2,775	1,016	1.515	0.184
2002	26,523	2,658	932	0.962	0.186
2003	27,630	2,554	914	0.261	0.187
2004	32,355	2,715	916	0.313	0.194
2005	33,941	2,864	959	0.234	0.197
2006	35,549	2,952	991	0.441	0.200
2007	37,403	3,024	992	0.968	0.201
2008	40,343	2,794	955	1.115	0.203
2009	39,104	2,574	899	0.591	0.204
2010	41,171	2,589	974	0.106	0.209
2011	42,545	2,573	1,003	0.470	0.207
2012	43,245	2,597	974	0.579	0.205
2013	42,001	2,612	925	0.443	0.207
2014	42,206	2,677	922	0.395	0.205
2015	42,209	2,643	883	0.355	0.205
2016	39,598	2,552	853	0.315	0.203
2017	37,840	2,499	813	0.048	0.205
2018	37,530	2,518	778	0.297	0.211
2019	36,981	2,536	761	0.704	0.210
2020	41,978	2,545	768	0.056	0.210
2021	34,006	2,609	822	-0.456	0.212
Overall	*956,105*	*10,441*	*4,302*	*0.599*	*0.198*

Panel D - by Industry

ID	Industry	No. of Forecasts	No. of Firms	No. of Analyst	Forecast Optimism	Negativity
1	Mining/Construction	18,925	190	412	1.104	0.193
2	Food	20,384	171	290	0.309	0.217
3	Textiles/Print/Publish	27,154	336	629	0.834	0.194
4	Chemicals	21,454	203	370	0.479	0.186
5	Pharmaceuticals	65,065	1,097	620	0.028	0.251
6	Extractive	101,696	372	444	0.479	0.210
7	Manf: Rubber/Glass/Etc.	10,014	144	388	0.721	0.203
8	Manf: Metal	14,093	206	457	0.818	0.182
9	Manf: Machinery	27,593	272	542	0.403	0.190
10	Manf: Electrical Equipment	21,446	379	780	1.147	0.202
11	Manf: Transport Equipment	15,322	177	363	0.975	0.171
12	Manf: Instruments	39,711	623	788	0.446	0.191
13	Manf: Misc.	4,075	78	202	1.332	0.181
14	Computers	144,728	1,871	1,605	0.463	0.216
15	Transportation	65,718	541	751	1.106	0.170
16	Utilities	21,809	258	312	0.277	0.163
17	Retail: Wholesale	18,861	320	770	0.524	0.178
18	Retail: Misc	69,219	499	722	0.588	0.193
19	Retail: Restaurant	14,955	149	207	0.409	0.197
20	Financial	133,979	1,219	756	0.690	0.189
21	Insurance/Real Estate	4,831	131	305	0.681	0.193
22	Services	61,576	1,027	1,296	0.763	0.190
23	Others	33,497	3,614	1,163	0.724	0.178
	Overall	*956,105*	*10,441*	*4,302*	*0.599*	*0.198*

Note: This table presents the summary statistics for the main variables used in the regression analyses. All continuous variables are winsorized at the 1st and 99th percentiles. All the variables are defined in Appendix A. Panel A, B, C, and D, present the sample distributions by full sample, country, year, and industry respectively.

Table 2 Pairwise Correlation Table

	Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	*Forecast_Optimism*																
2	*Negativity*	**-0.016**															
3	*Size*	**-0.055**	**0.013**														
4	*Leverage*	**0.022**	**-0.025**	**0.504**													
5	*BM*	**0.078**	**-0.027**	**0.269**	**0.107**												
6	*Loss*	**0.034**	**0.058**	**-0.273**	**-0.105**	**-0.088**											
7	*RDIntensity*	**0.017**	**-0.075**	**0.190**	**0.125**	**0.119**	**-0.454**										
8	*Ret*	**-0.194**	**0.005**	**-0.006**	**0.036**	**0.059**	**0.003**	**0.007**									
9	*ROA*	**-0.035**	**-0.039**	**0.230**	**-0.019**	**-0.052**	**-0.712**	**0.496**	**0.007**								
10	*EarnVol*	**0.082**	0.001	**0.033**	**0.051**	**0.126**	**0.121**	**-0.076**	**-0.064**	**-0.122**							
11	*Proximity*	**0.009**	**-0.410**	**-0.033**	**-0.009**	**0.015**	**-0.014**	**0.028**	**-0.009**	**0.014**	0.002						
12	*Distance*	**-0.011**	**0.538**	**0.027**	0.001	**-0.020**	**0.031**	**-0.046**	**0.009**	**-0.027**	**-0.003**	**-0.894**					
13	*LagAcc*	**-0.071**	**-0.005**	**0.029**	**-0.043**	**-0.186**	**-0.194**	**0.099**	**-0.072**	**0.174**	**-0.119**	**-0.005**	-0.001				
14	*Horizon*	**0.051**	**0.005**	**-0.028**	**-0.016**	**-0.011**	**0.028**	**-0.026**	-0.001	**-0.020**	**0.004**	-0.001	0.002	**-0.013**			
15	*AnalystFollowing*	**-0.092**	**0.027**	**0.602**	**0.093**	**0.007**	**-0.178**	**0.095**	**0.006**	**0.255**	**0.012**	**-0.044**	**0.037**	**0.089**	**-0.012**		
16	*FirmExp*	**-0.023**	**-0.019**	**0.286**	**0.128**	**0.104**	**-0.139**	**0.086**	**-0.005**	**0.106**	**0.020**	**-0.010**	**0.004**	**-0.035**	**0.029**	**0.290**	
17	*AnalysExp*	**-0.011**	**-0.044**	**0.140**	**0.112**	**0.056**	**-0.060**	**0.055**	0.002	**0.031**	0.001	**0.005**	**-0.008**	**-0.077**	**0.022**	**0.066**	**0.469**

Note: This table reports Pearson's correlation matrix among the key variables used in the regression analyses. All continuous variables are winsorized at the 1st and 99th percentiles. Bold indicates that the correlation is significant at the 1% level or below. All these variables are defined in Appendix A.

Table 3 Language Negativity and Analyst Forecast Optimism

Dep. Var.	***Analyst Forecast Optimism***					
	All Analysts		Excluding Analysts with U.S. Surname		Excluding Analysts Associated with Countries with Multiple Official Languages	
	(1)	(2)	(3)	(4)	(5)	(6)
Negativity	**-0.199*** **	**-0.153** **	**-0.176** **	**-0.160** **	**-0.252*** **	**-0.200** **
	(3.07)	**(2.04)**	**(2.30)**	**(2.03)**	**(2.85)**	**(2.11)**
Size		-0.061*		-0.040		-0.071**
		(1.83)		(1.09)		(2.02)
Leverage		0.782***		0.750***		0.883***
		(5.22)		(4.66)		(5.42)
BM		1.278***		1.374***		1.309***
		(9.85)		(10.51)		(9.42)
LOSS		0.275**		0.150		0.320**
		(1.99)		(0.96)		(2.17)
RDIntensity		0.112**		0.113**		0.099*
		(2.28)		(2.12)		(1.70)
Ret		-1.699***		-1.675***		-1.745***
		(28.65)		(24.93)		(27.81)
ROA		0.531*		0.559*		0.619**
		(1.89)		(1.91)		(2.00)
EarnVol		0.035***		0.026*		0.037***
		(2.59)		(1.83)		(2.60)
Proximity		-0.001		-0.222		0.088*
		(0.03)		(0.99)		(1.90)
Distance		-0.008		-0.012		0.037
		(0.37)		(0.52)		(1.37)
LagAcc		-0.079***		-0.058***		-0.088***
		(6.56)		(4.11)		(6.77)
Horizon		0.326***		0.308***		0.345***
		(17.92)		(14.01)		(17.50)
AnalystFollowing		-0.039		-0.005		-0.044
		(1.05)		(0.12)		(1.11)
FirmExp		0.051***		0.035**		0.063***
		(3.74)		(2.03)		(4.19)
AnalysExp		-0.028**		-0.017		-0.035***
		(2.51)		(1.14)		(2.89)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Hofstede's Culture Index	Yes	Yes	Yes	Yes	Yes	Yes
N	956,105	956,105	563,835	563,835	834,308	834,308
Overall. R-sq	0.202	0.231	0.211	0.239	0.206	0.240

Note: This table presents the regression results for the effect of language negativity on analyst forecast optimism. Columns (1) and (2) integrate all countries analysts in the analysis. Columns (3) and (4) exclude analysts from the US. Columns (5) and (6) omit analysts from countries with multiple official language, defined as countries where less than 80% of the population speaks its most widely spoken language. These countries are *Angola, Belgium, Canada, India, Israel, Luxembourg,* and *Switzerland*. The dependent variable, *Forecast_Optimism*, is calculated as (analyst EPS forecast - actual EPS) / stock price on the day prior to the earnings forecast date ×100. *Negativity* is the number of language negativity-related words divided by the total number of words in the language's largest dictionary (See Appendix). *Hofstede's Culture Index* is to control for the Hofstede's six national cultural dimensions, which capture a country's power distance, individualism, uncertainty avoidance, masculinity, long-term orientation, and indulgence value, respectively. All other variables are defined in Appendix A. All continuous variables are winsorized at 1% and 99% to mitigate outliers. ***, **, * indicate statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. The t values reported in parentheses are based on two-tailed tests. Standard errors are clustered by firm.

Table 4 Language Negativity and Analyst Forecast Optimism - Alternative Negativity Measures

Dep. Var.	***Analyst Forecast Optimism***					
	All Analysts			Excluding Analysts with U.S. Surname		
	(1)	(2)	(3)	(4)	(5)	(6)
Negativity Weight1	**-0.117****			**-0.095***		
	(2.23)			**(1.75)**		
Negativity Weight2		**-0.153****			**-0.175****	
		(2.08)			**(2.26)**	
Negativity PCA			**-0.020****			**-0.016***
			(2.23)			**(1.75)**
Size	-0.049	-0.049	-0.049	-0.044	-0.042	-0.044
	(1.49)	(1.49)	(1.49)	(1.22)	(1.14)	(1.22)
Leverage	0.781***	0.782***	0.781***	0.703***	0.695***	0.703***
	(5.21)	(5.22)	(5.21)	(4.39)	(4.35)	(4.39)
BM	1.278***	1.279***	1.278***	1.325***	1.315***	1.325***
	(9.84)	(9.85)	(9.84)	(9.89)	(9.74)	(9.89)
LOSS	0.275**	0.275**	0.275**	0.157	0.168	0.157
	(1.99)	(1.99)	(1.99)	(1.00)	(1.07)	(1.00)
RDIntensity	0.112**	0.112**	0.112**	0.115**	0.112**	0.115**
	(2.28)	(2.28)	(2.28)	(2.11)	(2.04)	(2.11)
Ret	-1.699***	-1.699***	-1.699***	-1.660***	-1.662***	-1.660***
	(28.65)	(28.65)	(28.65)	(24.91)	(24.69)	(24.91)
ROA	0.531*	0.531*	0.531*	0.502*	0.488*	0.502*
	(1.89)	(1.89)	(1.89)	(1.72)	(1.66)	(1.72)
EarnVol	0.035***	0.035***	0.035***	0.030**	0.031**	0.030**
	(2.59)	(2.59)	(2.59)	(2.10)	(2.13)	(2.10)
Proximity	-0.012	-0.004	-0.012	-0.231	-0.267	-0.231
	(0.32)	(0.12)	(0.32)	(1.03)	(1.20)	(1.03)
Distance	-0.013	-0.010	-0.013	-0.022	0.039	-0.022
	(0.65)	(0.52)	(0.65)	(1.03)	(1.46)	(1.02)
LagAcc	-0.079***	-0.079***	-0.079***	-0.057***	-0.055***	-0.057***
	(6.55)	(6.56)	(6.55)	(4.08)	(3.87)	(4.08)
Horizon	0.326***	0.326***	0.326***	0.307***	0.303***	0.307***
	(17.92)	(17.92)	(17.92)	(14.13)	(14.04)	(14.13)
AnalystFollowing	-0.039	-0.039	-0.039	-0.007	-0.012	-0.007
	(1.05)	(1.05)	(1.05)	(0.17)	(0.28)	(0.17)
FirmExp	0.051***	0.051***	0.051***	0.039**	0.038**	0.039**
	(3.74)	(3.74)	(3.74)	(2.29)	(2.20)	(2.29)
AnalysExp	-0.028**	-0.028**	-0.028**	-0.018	-0.015	-0.018
	(2.53)	(2.52)	(2.53)	(1.24)	(0.97)	(1.24)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Hofstede's Culture Index	Yes	Yes	Yes	Yes	Yes	Yes
N	956,105	956,105	956,105	563,835	563,835	563,835
Overall. R-sq	0.231	0.231	0.231	0.240	0.239	0.240

Note: This table presents the regression results for the effect of multiple alternative language negativity on analyst forecast optimism. The dependent variable, *Negativity* is the number of language negativity-related words divided by the total number of words in the language's largest dictionary. *Negativity Weight1* equals to the average score of three standardized negativity subcategories (*DeathRatio, DiseaseRatio, ViolenceRatio*). *Negativity Weight2* equals to the equally-weighted negativity score of all language the country speak. *Negativity PCA* is the first principal component of three negativity subcategory score of the country. All other variables are defined in Appendix A. All continuous variables are winsorized at 1% and 99% to mitigate outliers. ***, **, * indicate statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. The t values reported in parentheses are based on two-tailed tests. Standard errors are clustered by firm.

Table 5 Language Negativity (Measured by Each Subcategory) and Analyst Forecast Optimism

Dep. Var.	***Analyst Forecast Optimism***					
	All Analysts			Excluding Analysts with U.S. Surname		
	(1)	(2)	(3)	(4)	(5)	(6)
DeathRatio	**-0.168***			**-0.211****		
	(1.87)			**(2.21)**		
DiseaseRatio		**-1.926****			**-1.455***	
		(2.34)			**(1.69)**	
ViolenceRatio			**-1.018***			**-0.727**
			(1.92)			**(1.31)**
Size	-0.061*	-0.061*	-0.061*	-0.042	-0.042	-0.042
	(1.83)	(1.83)	(1.83)	(1.14)	(1.14)	(1.14)
Leverage	0.782***	0.781***	0.781***	0.695***	0.694***	0.694***
	(5.22)	(5.21)	(5.21)	(4.35)	(4.34)	(4.35)
BM	1.279***	1.278***	1.278***	1.315***	1.315***	1.315***
	(9.85)	(9.84)	(9.85)	(9.74)	(9.74)	(9.74)
LOSS	0.275**	0.275**	0.275**	0.169	0.168	0.168
	(1.99)	(1.99)	(1.99)	(1.07)	(1.07)	(1.07)
RDIntensity	0.112**	0.112**	0.112**	0.112**	0.112**	0.112**
	(2.28)	(2.28)	(2.28)	(2.04)	(2.04)	(2.04)
Ret	-1.699***	-1.699***	-1.699***	-1.662***	-1.662***	-1.662***
	(28.65)	(28.65)	(28.65)	(24.69)	(24.69)	(24.69)
ROA	0.531*	0.531*	0.531*	0.488*	0.488*	0.488*
	(1.89)	(1.89)	(1.89)	(1.66)	(1.66)	(1.66)
EarnVol	0.035***	0.035***	0.035***	0.031**	0.031**	0.031**
	(2.59)	(2.59)	(2.59)	(2.13)	(2.13)	(2.13)
Proximity	0.001	-0.022	-0.018	-0.266	-0.277	-0.276
	(0.03)	(0.59)	(0.48)	(1.20)	(1.24)	(1.24)
Distance	-0.007	-0.020	-0.019	0.047*	0.016	0.018
	(0.35)	(1.01)	(0.96)	(1.69)	(0.61)	(0.69)
LagAcc	-0.079***	-0.079***	-0.079***	-0.055***	-0.055***	-0.056***
	(6.56)	(6.55)	(6.56)	(3.87)	(3.88)	(3.89)
Horizon	0.326***	0.326***	0.326***	0.303***	0.303***	0.303***
	(17.92)	(17.92)	(17.92)	(14.04)	(14.04)	(14.04)
AnalystFollowing	-0.039	-0.039	-0.039	-0.012	-0.012	-0.012
	(1.05)	(1.05)	(1.05)	(0.28)	(0.29)	(0.29)
FirmExp	0.051***	0.051***	0.051***	0.037**	0.038**	0.038**
	(3.74)	(3.74)	(3.75)	(2.19)	(2.20)	(2.20)
AnalysExp	-0.028**	-0.028**	-0.028**	-0.014	-0.014	-0.014
	(2.50)	(2.52)	(2.53)	(0.95)	(0.94)	(0.94)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Hofstede's Culture Index	Yes	Yes	Yes	Yes	Yes	Yes
N	956,105	956,105	956,105	563,835	563,835	563,835
Overall. R-sq	0.231	0.231	0.231	0.239	0.239	0.239

Note: This table presents the regression results for the effect of three language negativity sub-component on analyst forecast optimism. The dependent variable, *Death /Disease/Violence* ratio are language-related words divided by the total number of words in the language's largest dictionary (See Online Appendix). All other variables are defined in Appendix A. All continuous variables are winsorized at 1% and 99% to mitigate outliers. ***, **, * indicate statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. The t values reported in parentheses are based on two-tailed tests. Standard errors are clustered by firm.

Table 6 Language Negativity and Analyst Forecast Optimism - Cross-sectional Tests on Financial Crisis/Market Condition

Dep. Var.	***Analyst Forecast Optimism***			
	All Analysts		Excluding Analysts with U.S. Surname	
	(1)	(2)	(3)	(4)
Negativity×Crisis	**-1.146*** **		**-1.529*** **	
	(4.00)		**(4.64)**	
Negativity×Boom		**-0.051**		**-0.236**
		(0.26)		**(1.11)**
Negativity	-0.030	-0.199	0.043	0.086
	(0.39)	(1.03)	(0.54)	(0.42)
Size	-0.050	-0.049	-0.044	-0.044
	(1.50)	(1.49)	(1.23)	(1.22)
Leverage	0.779***	0.782***	0.699***	0.704***
	(5.20)	(5.22)	(4.37)	(4.40)
BM	1.276***	1.278***	1.322***	1.326***
	(9.83)	(9.84)	(9.88)	(9.89)
LOSS	0.274**	0.275**	0.155	0.157
	(1.98)	(1.99)	(0.99)	(1.00)
RDIntensity	0.111**	0.112**	0.114**	0.115**
	(2.27)	(2.28)	(2.10)	(2.11)
Ret	-1.700***	-1.699***	-1.661***	-1.660***
	(28.66)	(28.65)	(24.92)	(24.91)
ROA	0.528*	0.531*	0.498*	0.502*
	(1.88)	(1.89)	(1.70)	(1.72)
EarnVol	0.035***	0.035***	0.030**	0.030**
	(2.59)	(2.59)	(2.10)	(2.10)
Proximity	-0.004	-0.001	-0.225	-0.226
	(0.10)	(0.04)	(1.00)	(1.01)
Distance	-0.008	-0.008	-0.018	-0.017
	(0.41)	(0.37)	(0.83)	(0.76)
LagAcc	-0.079***	-0.079***	-0.057***	-0.057***
	(6.55)	(6.56)	(4.09)	(4.08)
Horizon	0.326***	0.326***	0.308***	0.307***
	(17.93)	(17.92)	(14.14)	(14.13)
AnalystFollowing	-0.040	-0.039	-0.008	-0.007
	(1.06)	(1.05)	(0.20)	(0.17)
FirmExp	0.051***	0.051***	0.039**	0.039**
	(3.75)	(3.74)	(2.28)	(2.29)
AnalysExp	-0.027**	-0.028**	-0.018	-0.018
	(2.45)	(2.51)	(1.18)	(1.23)
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Hofstede's Culture Index	Yes	Yes	Yes	Yes
N	956,105	956,105	563,835	563,835
Overall. R-sq	0.232	0.231	0.240	0.240

Note: This table presents the moderating effect of market condition on the relation between language negativity and analyst forecast optimism. *Boom* and *Crisis* are derived from the dot-com bubble and the 2008 financial crisis. Boom equals one if the year equals 1997-1999 and 2007, and Crisis set to one if the year equals 2000-2002 and 2008. All other variables are defined in Appendix A. All continuous variables are winsorized at 1% and 99% to mitigate outliers. ***, **, * indicate statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. The t values reported in parentheses are based on two-tailed tests. Standard errors are clustered by firm.

Table 7 Language Negativity and Analyst Forecast Optimism - Cross-sectional Tests on Loss/Earnings Volatility/ Limited Attention

Dep. Var.	*Analyst Forecast Optimism*					
		All Analysts		Excluding Analysts with U.S. Surname		
	(1)	(2)	(3)	(4)	(5)	(6)
Negativity×Loss	**-0.648****			**-0.650***		
	(2.28)			**(1.95)**		
Negativity×EarnVol		**-2.292****			**-2.422****	
		(2.06)			**(2.17)**	
Negativity×LtdAttn			**-0.018*****			**-0.014***
			(2.71)			**(1.77)**
LtdAttn			0.002			0.002
			(1.18)			(0.66)
Negativity	-0.064	-0.026	0.166	-0.045	-0.003	0.194
	(0.85)	(0.29)	(1.16)	(0.57)	(0.03)	(1.17)
Size	-0.050	-0.052	-0.049	-0.044	-0.047	-0.047
	(1.49)	(1.56)	(1.48)	(1.22)	(1.31)	(1.28)
Leverage	0.785***	0.775***	0.783***	0.708***	0.693***	0.744***
	(5.24)	(5.17)	(5.22)	(4.43)	(4.32)	(4.64)
BM	1.279***	1.277***	1.279***	1.326***	1.324***	1.334***
	(9.85)	(9.83)	(9.85)	(9.89)	(9.88)	(9.91)
Loss	0.408***	0.287**	0.275**	0.324*	0.173	0.188
	(2.63)	(2.08)	(1.99)	(1.72)	(1.10)	(1.19)
RDIntensity	0.110**	0.112**	0.112**	0.112**	0.114**	0.092*
	(2.24)	(2.27)	(2.28)	(2.06)	(2.10)	(1.69)
Ret	-1.699***	-1.698***	-1.699***	-1.660***	-1.659***	-1.662***
	(28.66)	(28.64)	(28.65)	(24.91)	(24.89)	(25.06)
ROA	0.532*	0.525*	0.532*	0.501*	0.495*	0.584**
	(1.90)	(1.87)	(1.90)	(1.71)	(1.69)	(1.99)
EarnVol	0.035***	0.036***	0.035***	0.030**	0.031**	0.030**
	(2.58)	(2.64)	(2.60)	(2.08)	(2.16)	(2.07)
Proximity	-0.001	-0.001	-0.007	-0.231	-0.221	-0.201
	(0.02)	(0.02)	(0.18)	(1.03)	(1.00)	(0.73)
Distance	-0.007	-0.007	-0.011	-0.016	-0.016	-0.015
	(0.32)	(0.33)	(0.53)	(0.73)	(0.74)	(0.68)
LagAcc	-0.079***	-0.079***	-0.079***	-0.056***	-0.057***	-0.058***
	(6.53)	(6.58)	(6.56)	(4.04)	(4.10)	(4.12)
Horizon	0.326***	0.326***	0.326***	0.307***	0.308***	0.310***
	(17.93)	(17.93)	(17.94)	(14.13)	(14.14)	(14.13)
AnalystFollowing	-0.039	-0.035	-0.039	-0.007	-0.002	-0.003
	(1.04)	(0.94)	(1.04)	(0.16)	(0.04)	(0.06)
FirmExp	0.051***	0.050***	0.051***	0.039**	0.038**	0.038**
	(3.76)	(3.70)	(3.75)	(2.30)	(2.23)	(2.20)
AnalysExp	-0.028**	-0.028**	-0.023**	-0.018	-0.018	-0.009
	(2.53)	(2.51)	(2.00)	(1.22)	(1.20)	(0.53)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Hofstede's Culture Index	Yes	Yes	Yes	Yes	Yes	Yes
N	956,105	956,105	956,105	563,835	563,835	563,835
Overall. R-sq	0.231	0.231	0.231	0.240	0.240	0.242

Note: This table presents the moderating effect of firm characteristics on the relation between language negativity and analyst forecast optimism. *Loss* equals to one if most recent year net income is smaller than zero, and 0 otherwise. *EarnVol* equals to the standard deviation of EPS over the last five years. *LtdAttn* equals to the total number of companies that the analyst followed in that year. All other variables are defined in Appendix A. All continuous variables are winsorized at 1% and 99% to mitigate outliers. ***, **, * indicate statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. The t values reported in parentheses are based on two-tailed tests. Standard errors are clustered by firm.

Table 8 Language Negativity and Analyst Forecast Optimism - Cross-sectional Tests on Broker Size/Analyst Experience

Dep. Var.	*Analyst Forecast Optimism*			
	All Analysts		Excluding Analysts with U.S. Surname	
	(1)	(2)	(3)	(4)
Negativity×YoungAnalys	**-0.060*****		**-0.066*****	
	(2.78)		**(2.75)**	
Negativity×SmallBroker		**-0.043****		**-0.061*****
		(2.18)		**(2.66)**
YoungAnalys	0.037***		0.028**	
	(3.59)		(2.23)	
SmallBroker		0.008		0.014**
		(1.58)		(2.06)
Negativity	0.114	0.037	0.208	0.140
	(0.81)	(0.32)	(1.36)	(1.11)
Size	-0.044	-0.050	-0.044	-0.045
	(1.29)	(1.50)	(1.22)	(1.26)
Leverage	0.790***	0.782***	0.702***	0.704***
	(5.23)	(5.22)	(4.39)	(4.40)
BM	1.291***	1.278***	1.324***	1.326***
	(9.77)	(9.84)	(9.88)	(9.90)
LOSS	0.280**	0.276**	0.156	0.158
	(2.02)	(1.99)	(1.00)	(1.01)
RDIntensity	0.116**	0.112**	0.115**	0.115**
	(2.34)	(2.28)	(2.11)	(2.11)
Ret	-1.704***	-1.699***	-1.660***	-1.660***
	(28.26)	(28.65)	(24.91)	(24.92)
ROA	0.515*	0.530*	0.500*	0.500*
	(1.81)	(1.89)	(1.71)	(1.71)
EarnVol	0.033**	0.035***	0.030**	0.030**
	(2.48)	(2.59)	(2.10)	(2.10)
Proximity	-0.018	-0.005	-0.246	-0.212
	(0.47)	(0.13)	(1.09)	(0.97)
Distance	-0.010	-0.010	-0.020	-0.019
	(0.47)	(0.46)	(0.88)	(0.88)
LagAcc	-0.082***	-0.079***	-0.058***	-0.057***
	(6.60)	(6.55)	(4.08)	(4.08)
Horizon	0.330***	0.326***	0.308***	0.308***
	(17.83)	(17.94)	(14.13)	(14.15)
AnalystFollowing	-0.047	-0.039	-0.007	-0.006
	(1.25)	(1.04)	(0.16)	(0.15)
FirmExp	0.053***	0.051***	0.039**	0.039**
	(3.81)	(3.73)	(2.32)	(2.27)
AnalysExp	-0.107***	-0.029***	-0.055	-0.020
	(3.28)	(2.58)	(1.35)	(1.32)
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Hofstede's Culture Index	Yes	Yes	Yes	Yes
N	956,105	956,105	563,835	563,835
Overall. R-sq	0.231	0.231	0.240	0.240

Note: This table presents the moderating effect of analyst/broker characteristic on the relation between language negativity and analyst forecast optimism. *YoungAnalys* equals the reverse 10 deciles value of the analyst's working years. The year the analyst began working is set to the first time he appears in the IBES database. *SmallBroker* equals the reverse 10 decile value of broker size, measured as the broker's total number of analysts at a given year. All other variables are defined in Appendix A. All continuous variables are winsorized at 1% and 99% to mitigate outliers. ***, **, * indicate statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. The t values reported in parentheses are based on two-tailed tests. Standard errors are clustered by firm.

Table 9 Economic Implication of Language Negativity

Dep. Var.	***Analyst Forecast Error***			
	All Analysts		Excluding Analysts with U.S. Surname	
	(1)	(2)	(3)	(4)
Negativity	**-0.184*****	**-0.125***	**-0.176****	**-0.148***
	(3.01)	**(1.81)**	**(2.03)**	**(1.68)**
Size		-0.225***		-0.220***
		(6.92)		(6.43)
Leverage		2.045***		1.970***
		(12.98)		(11.80)
BM		1.755***		1.840***
		(16.67)		(15.57)
LOSS		0.694***		0.574***
		(5.76)		(4.28)
RDIntensity		0.110***		0.151***
		(3.17)		(3.48)
Ret		-0.898***		-0.889***
		(19.22)		(16.62)
ROA		-0.792***		-0.773***
		(3.10)		(2.80)
EarnVol		0.055***		0.055***
		(4.34)		(4.01)
Proximity		-0.003		-0.168
		(0.08)		(0.87)
Distance		-0.005		0.002
		(0.23)		(0.09)
LagAcc		-0.137***		-0.128***
		(11.59)		(8.85)
Horizon		-0.096***		-0.089***
		(4.91)		(3.67)
AnalystFollowing		-0.007		0.030
		(0.20)		(0.82)
FirmExp		0.058***		0.052***
		(4.44)		(2.96)
AnalysExp		-0.026***		-0.023*
		(2.62)		(1.66)
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Hofstede's Culture Index	Yes	Yes	Yes	Yes
N	956,105	956,105	563,835	563,835
Overall. R-sq	0.260	0.287	0.272	0.297

Note: This table presents the effect of language negativity on analyst forecast error. The dependent variable, *Analyst Forecast Error,* equals to the absolute difference between analyst forecasted EPS and actual EPS scaled by the stock price on the day prior to the earnings forecast date ×100. *Negativity* is the number of language negativity-related words divided by the total number of words in the language's largest dictionary (See Appendix). *Hofstede's Culture Index* is to control for the Hofstede's six national cultural dimensions, which capture a country's power distance, individualism, uncertainty avoidance, masculinity, long-term orientation, and indulgence value, respectively. All other variables are defined in Appendix A. All continuous variables are winsorized at 1% and 99% to mitigate outliers. ***, **, * indicate statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. The t values reported in parentheses are based on two-tailed tests. Standard errors are clustered by firm.

Table 10 Robustness Test

Dep. Var.	***Analyst Forecast Optimism***			
	All Analysts		Excluding Analysts with U.S. Surname	
	(1)	(2)	(3)	(4)
Panel A: Alternative Optimism Measurement - *Relative_Optimism*				
Negativity	**-0.039*** **	**-0.027** **	**-0.043*** **	**-0.031** **
	(3.13)	**(2.32)**	**(2.93)**	**(2.16)**
Controls	NO	Yes	NO	Yes
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Hofstede's Culture Index	Yes	Yes	Yes	Yes
N	956,105	956,105	563,835	563,835
Overall. R-sq	0.007	0.008	0.012	0.013
Panel B: Keep Only the First Analyst Forecast for Each Firm-Year				
Negativity	**-0.232*** **	**-0.209** **	**-0.250*** **	**-0.183* **
	(2.80)	**(2.13)**	**(2.61)**	**(1.82)**
Controls	NO	Yes	NO	Yes
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Hofstede's Culture Index	Yes	Yes	Yes	Yes
N	244,992	244,992	144,125	144,125
Overall. R-sq	0.205	0.252	0.207	0.253
Panel C: Alternative Explanation - Economic Policy Uncertainty (EPU)				
Negativity	**-0.266*** **	**-0.344*** **	**-0.178* **	**-0.304** **
	(3.62)	**(2.81)**	**(1.95)**	**(2.27)**
EPU	-0.064	-0.004	-0.128	-0.129
	(0.32)	(0.02)	(0.61)	(0.58)
Controls	NO	Yes	NO	Yes
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Hofstede's Culture Index	Yes	Yes	Yes	Yes
N	956,105	956,105	563,835	563,835
Overall. R-sq	0.205	0.239	0.216	0.255
Panel D: Additional Country Variable Controls (Isidro, Nanda, and Wysocki,2020)				
Negativity	**-0.477*** **	**-0.354** **	**-0.439** **	**-0.322* **
	(2.75)	**(1.98)**	**(2.27)**	**(1.70)**
Controls	NO	Yes	NO	Yes
Four Country Attributes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Hofstede's Culture Index	Yes	Yes	Yes	Yes
N	956,105	956,105	563,835	563,835
Overall. R-sq	0.206	0.236	0.220	0.250
Panel E: WLS Regression				
Negativity	**-0.267*** **	**-0.233** **	**-0.333** **	**-0.228* **
	(3.14)	**(2.07)**	**(2.35)**	**(1.66)**
Controls	NO	Yes	NO	Yes
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Hofstede's Culture Index	Yes	Yes	Yes	Yes
N	956,105	956,105	563,835	563,835
Overall. R-sq	0.231	0.260	0.246	0.278

Panel F: Exclude all foreign firms listed in the US				
Negativity	**-0.206******	**-0.150****	**-0.173****	**-0.159****
	(3.15)	**(1.98)**	**(2.24)**	**(1.99)**
Controls	NO	Yes	NO	Yes
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Hofstede's Culture Index	Yes	Yes	Yes	Yes
N	941,422	941,422	553,410	553,410
Overall. R-sq	0.200	0.230	0.210	0.238

Note: This table presents the robustness check on the relation between language negativity and analyst forecast optimism. Panel A replaces the dependent variable as *Relative_Optimism* (Cowen et al., 2006), which compared the optimism of a given analyst's forecast with those of all analysts who made forecasts for the same company and time period within a comparable forecast horizon. Panel B retains only the initial forecast made by each analyst for firm i on fiscal year t. Panel C controlled analyst ancestry country's *Economic Policy Uncertainty*(EPU) (Baker et al., 2016). Panel D controlled four comprehensive country factors derived from 72 country characteristic by Isidro, Nanda, and Wysocki (2020). The four factors are likely to capture a country's legal system, creditor and/or investor rights, political process, and societal characteristic. Panel E changed the regression specification to WLS regression. Panel F keep only listed firm headquartered in the United States.