PhD THESIS DECLARATION

I, the undersigned

FAMILY NAME	Bogachek
NAME	Olga Stanislavovna
Student ID no.	1824275

Thesis title:

Essays on the Economics of Corporate Regulatory Disclosures

PhD in	Economics and Finance
Cycle	30
Student's Advisor	Miles Gietzmann
Calendar year of thesis defence	2019

DECLARE

under my responsibility:

1) that, according to Italian Republic Presidential Decree no. 445, 28th December 2000, mendacious declarations, falsifying records and the use of false records are punishable under the Italian penal code and related special laws. Should any of the above prove true, all benefits included in this declaration and those of the temporary "embargo" are automatically forfeited from the beginning;

2) that the University has the obligation, according to art. 6, par. 11, Ministerial Decree no. 224, 30th April 1999, to keep a copy of the thesis on deposit at the "Biblioteche Nazionali Centrali" (Italian National Libraries) in Rome and Florence, where consultation will be permitted, unless there is a temporary "embargo" protecting the rights of external bodies and the industrial/commercial exploitation of the thesis;

3) that the Bocconi Library will file the thesis in its "Archivio Istituzionale ad Accesso Aperto" (Institutional Registry) which permits online consultation of the complete text (except in cases of temporary "embargo");

4) that, in order to file the thesis at the Bocconi Library, the University requires that the thesis be submitted online by the student in unalterable format to Società NORMADEC (acting on behalf of the University), and that NORMADEC will indicate in each footnote the following information:

- PhD thesis (*write thesis title*) **Essays on the Economics of Corporate Regulatory Disclosures** ;
- by (S*tudent's family name and name)* **Bogachek Olga Stanislavovna**;
- defended at Università Commerciale "Luigi Bocconi" – Milano in the year (year of defence) **2019** ;
- the thesis is protected by the regulations governing copyright (Italian law no. 633, 22nd April 1941 and subsequent modifications). The exception is the right of Università Commerciale "Luigi Bocconi" to reproduce the same, quoting the source, for research and teaching purposes;

- **<u>only</u> when a separate "Embargo" Request has been undersigned**: the thesis is subject to "embargo" for (indicate duration of the "embargo") ……… months;

5) that the copy of the thesis submitted online to Normadec is identical to the copies handed in/sent to the members of the Thesis Board and to any other paper or digital copy deposited at the University offices, and, as a consequence, the University is absolved from any responsibility regarding errors, inaccuracy or omissions in the contents of the thesis;

6) that the contents and organization of the thesis is an original work carried out by the undersigned and does not in any way compromise the rights of third parties (Italian law, no. 633, 22nd April 1941 and subsequent integrations and modifications), including those regarding security of personal details; therefore the University is in any case absolved from any responsibility whatsoever, civil, administrative or penal, and shall be exempt from any requests or claims from third parties;

7) **choose 7a <u>or</u> 7b**:

<u>7a) that the thesis is not subject to "embargo", i.e. that it is not the result of work included in the regulations governing industrial property; it was not written as part of a project financed by public or private bodies with</u>

restrictions on the diffusion of the results; is not subject to patent or protection registrations.

or

7b) that the PhD thesis is subject to "embargo" as per the separate undersigned "PhD Thesis Temporary ″Embargo″ Request".

Date **_17/04/2019___**

Olga Bogachek

Abstract

This PhD thesis is composed of three chapters on corporate disclosures and their interaction with regulation.

The first chapter examines the timing of the decision by management to disclose details of internal control weaknesses. It presents a model of a disclosure game where managers are requested to communicate their ex-ante beliefs on the reliability of a firm's reporting system. However, they have certain incentives to delay disclosing their negative beliefs in the hope that ex-post, an unreliable reporting system indicated by a restatement will not be detected. A theoretical model is presented to explain why such a disclosure gamble may be an equilibrium strategy. This study also tests for empirical evidence of the existence of the delaying equilibrium using the data on disclosures of material weaknesses (MW) and financial restatements (FSR) of U.S. public companies in the 2005-2014 period. It is shown how specific underlying characteristics of firms determine whether a delaying strategy is likely to be followed.

The second chapter examines the content of corporate Codes of Conduct and predicts that the degree of detailed risk focus in such documents is driven by an entity’s international exposure to the risk of corruption. This study uses the Latent Dirichlet Allocation model on the Codes of Conduct of S&P 500 firms. We construct a quantitative index which measures the CEO’s anti-fraud emphasis by automatically detecting and extracting fraud-related topics. We test and find a positive association between the CEO’s anti-fraud emphasis and the risk (business corruption) exposure of the firm’s foreign subsidiaries.

Finally, the third chapter examines the incremental value relevance of annual report management commentary beyond financial statement information. We find that absolute tone and change in tone are positively correlated with the equity price, while abnormal tone is largely discounted by investors. We also find that the equity price is negatively associated with the average number of words per sentence of the MD&A section of 10-K, suggesting that investors appreciate readable and succinct information.

Acknowledgements

I am indebted to my advisor Professor Miles Gietzmann whose continuous support, guidance and assistance were indispensable for the successful completion of this dissertation. I am extremely grateful for his scientific advice, mentorship, time and patience during my PhD years.

I thank my dissertation committee members Professor Mara Cameran and Professor Angela Pettinicchio, Professor Francesco Grossetti, my program director Professor Nicola Pavoni, Professor Hannes Wagner and all my colleagues and instructors at Bocconi for professional and constructive guidance of my scientific work. Furthermore, I would like to thank Professor Mark Clatworthy and Professor Steven Young, who have kindly accepted to be the reviewers of this work. This dissertation benefitted from useful comments received from the participants and discussants of academic conferences, presentations and workshops that I was able to attend thanks to the generous support and fellowship of the PhD program at the Università Commerciale Luigi Bocconi.

Finally, this accomplishment would not be possible without the unconditional love, reassurance and support I received from my family: my mother, husband and son Oleg who followed me to Italy to pursue my dreams. I dedicate this thesis to the memory of my father, who was my inspiration to pursue a doctoral degree. “You raise me up so I can stand on mountains”.

Milan, April 17, 2019 Olga Bogachek

Keeping Quiet Until a Restatement: Are Disclosures of Material Weaknesses Delayed?

Olga Bogachek[a] and Miles Gietzmann[a]

[a]Bocconi University. Via Sarfatti 25, 20136 - Milan, Italy

ARTICLE HISTORY

Compiled April 17, 2019

ABSTRACT

This paper presents a model of a disclosure game where managers are requested to communicate their *ex-ante* beliefs on the reliability of a firm's reporting system. However, they have certain incentives to delay disclosing their negative beliefs in the hope that *ex-post*, an unreliable reporting system indicated by a restatement will not be detected. A theoretical model is presented to explain why such a disclosure gamble may be an equilibrium strategy. This study then tests for empirical evidence of the existence of the *delaying* equilibrium using the data on disclosures of material weaknesses (MW) and financial restatements (FSR) of U.S. public companies in the 2005-2014 period. It is shown how specific underlying characteristics of firms determine whether a delaying strategy is likely to be followed.

KEYWORDS

Material weakness, financial restatement, audit game

CONTACT Olga Bogachek. Email: olga.bogachek@phd.unibocconi.it

1. Introduction

One of the primary objectives of the Sarbanes-Oxley Act (SOX, 2002) was to require firms to proactively report on the underlying reliability of their financial reporting systems. In particular, firms are required to test for and report any material weaknesses (MWs)[1] in their reporting systems in order for remedial action to be taken in a timely manner before negative reporting events (such as restatements) arise (Aobdia, Choudhary, & Sadka, 2016). In practice, a trend of firms reporting fewer material weaknesses has been observed, superficially implying that their financial reporting systems are more reliable. However, at the same time, this has been accompanied by an increased frequency of subsequent financial restatements[2] (FSRs), which seems to suggest that many MWs remain undisclosed (Franzel, 2015).

In an attempt to explain this apparent contradiction, this research demonstrates how, under reasonable conditions, an equilibrium exists, within which firms choose to delay the disclosure of MWs in the hope that subsequently, somehow the firm's unreliable reporting system is not confirmed by the auditor. Variations in firm characteristics and penalties are introduced in order to understand how and when a separating equilibrium exists. Using empirical data, we show how recent experience is consistent with the existence of this stylized delaying equilibrium.

We develop a model of a disclosure game, which builds upon the model of Gietzmann and Sen (2002). The players - a firm's owner ("she") and a manager ("he") - have misaligned preferences; that is, the owner only wants the firm to continue if it is expected to be profitable, while the manager wants the firm to always continue and get compensated. The manager's job is to issue an intermediary report attesting to the state of the firm, which is validated by the owner in the final period. This setting creates an incentive for the manager to choose a disclosure strategy that allows him to delay the disclosure of negative news concerning the firm, which improves the odds of satisfying his preferences.[3] While stylized and solved using pure strategies, the model

[1] A material weakness is "a deficiency, or a combination of deficiencies, in internal control over financial reporting, such that there is a reasonable possibility that a material misstatement of the company's annual or interim financial statements will not be prevented or detected on a timely basis" (PCAOB, 2004); disclosure required by Section 404 of the Sarbanes-Oxley Act of 2002 (SOX).

[2] Restatement is defined as the revising of previously issued financial statements to reflect the correction of an error (Financial Accounting Standards Board, 2005).

[3] Our theoretical model may be applicable in other situations where an intermediary report is issued before the

captures essential features of misaligned preferences that could lead to a separating equilibrium (with some managerial types choosing the delaying strategy).

In our model, managers are requested to make two types of disclosures. First, they disclose realizations of **financial restatements**, denoted as $\hat{F}$. Observation of $\hat{F}$, together with other available evidence (*prior information*) is used by managers to formulate their *ex-ante* beliefs on the reliability of the firm's reporting system. If the reliability falls below a certain threshold W ($\tilde{R} < W$), the firm has an **material weakness** in the reporting system, which is the second requested disclosure.[4] In our game, the threshold simplistically takes two values $\hat{W} > \hat{W}_{other}$, where $\hat{W}$ captures the reliability required by regulators (and involves a costly effort) and $\hat{W}_{other}$ captures a substandard level of reliability. That is, the probability of the manager disclosing an MW depends on which threshold W is used:

$$p(MW|\hat{W}) > p(MW|\hat{W}_{other}) \tag{1}$$

Ex-post, the manager's report is verified by the auditor, who may penalize the manager for not disclosing an underlying MW.

The timing of the disclosures is critical in our analysis. Upon observing $\hat{F}$ the manager faces two choices - adopt threshold $\hat{W}$ and disclose both FSR and MW simultaneously in year t (FSR-and-MW) or adopt $\hat{W}_{other}$ and only disclose FSR in year t in the hope that the MW won't be identified by the auditor in the current year, effectively disclosing MW in $t+1$[5] (FSR-then-MW).[6] Thus, in our analysis, adopting $\hat{W}_{other}$ is equivalent to delaying the MW disclosure to the next year.

To summarize, in our model, the manager first observes and discloses $\hat{F}$, chooses a threshold $W \in [\hat{W},\hat{W}_{other}]$, forms ex-ante beliefs $\tilde{R}$, and discloses an MW when $\tilde{R} < W$, which is ex-post confirmed by the auditor. Therefore, choosing a lower and

final realization of the true state of the company is confirmed. We limit our attention to FSR/MW disclosure settings, which allows us to specifically model the manager behavior using a two-step approach (detection and identification of the root-cause; discussed in Section 3) and later empirically test the model predictions in Section 5.

[4] In other words, W is the manager's internal threshold of the reliability of the firm's reporting system, below which he believes an MW exists.

[5] We assume that the unreliable reporting system is not concealed for long and in year $t+1$, either the manager will disclose it or the auditor will reveal it.

[6] In Sec. 5, we discuss all practically possible disclosure patterns and how they fit within our analyses.

less costly threshold $\hat{W}_{other}$ (which we denote as a *Delaying* strategy) reduces the current costs of the manager as well as the likelihood of him disclosing an MW, but increases the chance of subsequent (restatement-only) penalty by the verifier.

We show that, depending on preferences and payoffs, two equilibria can be identified:

(1) **Full Compliance** equilibrium: All types choose costly threshold $\hat{W}$ and report an MW whenever $\tilde{R} < \hat{W}$

(2) **Partial Compliance** equilibrium: Some types choose non-costly threshold $\hat{W}_{other}$ and only report an MW if $\tilde{R} < \hat{W}_{other}$, facing the probability of a penalty for delayed disclosure

Anticipating subsequent arguments, we provide some informal intuition for why a manager may choose to delay the disclosure of an MW. First, MW disclosures are penalized more severely than restatements, creating a pressure to avoid the former (Kinney, William, & Shepardson, 2011; Kinney Jr, Martin, & Shepardson, 2013; Rice & Weber, 2012; Rice, Weber, & Wu, 2014). Consequently, if the efforts associated with disclosing MWs are too high as compared to the incentives for timely disclosure, firms may be tempted to gamble and choose the delaying strategy. Franzel (2015) notes, "[t]he vast majority of companies restating prior year financial statements received a 'clean' audit opinion on ICFR[7] in the year in which the restatement was announced." Second, delaying a disclosure of an MW does not preclude the manager from remediating the underlying deficiency in the reporting system. The delay allows him to fix the error in a timely manner and obtain a "clean" internal controls certification or, if not remediated, replace it with an adverse opinion later on. Franzel (2015) notes, "[a]nother interesting data point is the number of clean ICFR audit opinions that are subsequently withdrawn and replaced with adverse opinions on ICFR," indicating a strategic choice of non-disclosure.

The following is a critical issue. On the one hand, a MW in a firm's financial reporting system suggests that the firm's financial statements are unreliable. Managers may want to withhold this bad news to avoid short-term disclosure implications, but potentially face litigation risks of untimely disclosure. On the other hand, firms that

[7]Internal Control over Financial Reporting (ICFR)

choose to disclose a MW bear these short-term disclosure costs in hopes to gain long-term reputation for their transparency and compliance. The issue regarding how to balance these competing effects arises here.

Our analysis makes several contributions to the extant literature. First, we develop a theoretical model and determine equilibria conditions under which different types of firms choose to adopt either a Timely or a Delaying strategy. We show that a separating equilibrium (with some types choosing the Delaying strategy) is possible when a certain manager's concern over current effort costs outweighs potential penalties, causing him to delay the disclosure of an MW. Second, we test our model in a practical setting of MWs / financial restatements disclosures using existing empirical evidence. We shed new light on the factors associated with the timeliness of these disclosures and offer a new perspective to show that managers may benefit strategically by delaying the disclosure of MWs. Our contributions may be of practical interest as we provide insights for practitioners and regulators seeking to establish penalties for untimely disclosures.

This paper is organized as follows. Section 2 provides a literature review. Section 3 describes the modeling of MWs and restatements. Section 4 describes the model of the disclosure game and derives equilibria conditions. Empirical results, including the research question, regression model, sample construction, and robustness tests, are discussed in Section 5. Section 6 concludes the paper. All proofs are given in the Appendix.

2. Literature Review and SOX Background

This paper forms a part of the growing literature that studies voluntary disclosure decisions, as discussed in Grossman (1981) and Milgrom (1981). Dranove and Jin (2010) provide an extensive review of the theoretical and empirical literature on quality disclosures and certification. Healy and Palepu (2001) review empirical literature on voluntary disclosures. Dobler (2008) reviews discretionary disclosure and cheap talk models in the context of risk reporting. Pae (2005) presents a two-signal theoretical model, and R. Dye (2013); R. A. Dye (2017) offers a recent example of a voluntary disclosure model that includes a duty to disclose, liability of non-disclosure and disclosure materiality threshold as well as fact-finders to verify the truthfulness of the disclosure. This literature studies situations where an agent's summary disclosure is mandatory but there is only limited verifiability of the agent complying with all the detailed requirements. We adapt the approach described in Gietzmann and Sen (2002) to model a disclosure game that demonstrates how strategic behavior of the manager is determined.

Our theoretical model is related to analytical audit literature - a subset of literature on inspection games,[8] pioneered by Dresher (1961). Inspection games have been widely applied in auditing; see Avenhaus, Von Stengel, and Zamir (2002); Trockel (2013) for an extensive survey and applications.

We contribute to the literature on managerial aversion to disclose bad news (Kothari, Li, & Short, 2009; Kothari, Shu, & Wysocki, 2009). There are conflicting forces affecting the timeliness of managements unfavorable disclosures. On one hand, litigation risks are likely to motivate the manager to disclose bad news early (Baginski, Hassell, & Kimbrough, 2002; Kasznik & Lev, 1995). Quick release of info also reduces information asymmetry and potentially lowers the firm's cost of capital (Healy & Palepu, 2001; Verrecchia, 2001). On the other hand, managements career or compensation concerns might motivate managers to delay the release of bad news (Kothari, Shu, & Wysocki, 2009; Nagar, Nanda, & Wysocki, 2003). Managers may also be concerned if the disclosures reveal proprietary information to competitors (Verrecchia, 2001) or time good/bad news to maximize the value of their option grants (Aboody & Kasznik, 2000; Yermack, 1997). Kothari, Li, and Short (2009) show that on average, the management delays the release of bad news to investors. Hermalin and Weisbach (2007) conclude that optimal disclosure is less than fully transparent, especially with respect to bad news. Overall, provided arguments and evidence suggest that given conflicting incentives, the timeliness of good and bad news is difficult to predict (Kothari, Li, & Short, 2009).

Our empirical work is closely related to contemporary research on the implications

[8]An inspection game can be described as a decision-making model involving a vertical conflict between an inspection authority (in our case, the owner) and an agent (the manager). The manager has a contractual commitment to comply with disclosure requirements, while the owner has to execute monitoring measures in order to guarantee that the manager conforms with the rules.

of the Sarbanes-Oxley Act (SOX) (SOX, 2002) on market behavior and firm characteristics, which we review below.

The SOX Act requires a firm's management to formally assess and report on the effectiveness of its internal control over financial reporting (ICFR), which we refer to as the reliability of the financial reporting system. It also requires a firm's independent auditor to issue an "attestation" that provides an independent reason to rely on the management's assertion of the effectiveness of the firm's ICFR. Importantly, "management is not permitted to conclude that the registrants internal control over financial reporting is effective if there are one or more material weaknesses"(SOX, 2002). MW is the most severe kind of an internal control deficiency (as opposed to deficiencies or significant deficiencies, defined in PCAOB (2004), paragraphs A3 and A11). It is reported to the company's audit committee and, if the company receives an adverse ICFR opinion, publicly disclosed. The two sections of the SOX law that have particular implications for firms are Sections 302 and 404.[9] This study focuses on disclosures of MWs related to SOX Section 404.

Any communication pertaining to an MW is costly. Firms disclosing MWs experience negative stock reaction, less coverage from financial analysts, negative impact on the cost of equity, decrease in chief financial officer compensation, higher management and auditor turnover, and higher investor-related costs (Bardos & Mishra, 2014; Bedard, 2006; Beneish, Billings, & Hodder, 2008; Burks, 2011; Clinton, Pinello, & Skaife, 2014; Ge & McVay, 2005; Gordon & Wilford, 2012; Hoitash, Hoitash, & Johnstone, 2012; Kryzanowski & Zhang, 2013; Palmrose, Richardson, & Scholz, 2004). Audit fees increase and remain significantly higher, even after the remediation of a MW, as compared to a sample of companies that never report an adverse SOX 404 opinion (Hoag & Hollingsworth, 2011). Prior studies by Ashbaugh-Skaife, Collins, Kinney Jr, and LaFond (2008); Ashbaugh-Skaife, Collins, Lafond, et al. (2009); Chan, Farrell, and Lee (2008); Donelson, Ege, and McInnis (2016); J. T. Doyle, Ge, and McVay (2007); Hammersley, Myers, and Shakespeare (2008); Hoitash, Hoitash, and Bedard (2008) document an association between MWs and lower reporting quality, such as

[9]Section 302 concerns the quarterly ICFR disclosure assessment, while Section 404 is an annual assessment of the effectiveness of ICFR. Section 404 also requires accelerated filers to obtain an audit opinion attesting the result of management's assessment of ICFR. For a definition of accelerated filer, please refer to SEC (2005).

accrual quality, fraud, and higher cost of capital. Rice et al. (2014) also find that the likelihood of class action lawsuits, management and auditor turnover, and Securities and Exchange Commission (SEC) sanctions also increases when a control weakness has been reported before a restatement, which is consistent with the hypothesis that the disclosure of control weaknesses makes it difficult for the management to plausibly claim later on that they were unaware of the underlying conditions that led to a restatement.

We note that while some FSRs stem from underlying control weaknesses, many financial restatements are innocuous and, therefore, not indicative of existing MW. Contrarily, they may, in fact, re-establish investors' confidence in reported numbers (Hranaiova & Byers, 2007). This means that penalties for restatements in order to deter non-disclosure of a MW cannot be too high. In accordance with Rice et al. (2014), we study restatements that were made because previously filed financial reports were deemed unreliable (therefore, focusing on non-reliance restatements as opposed to less pervasive revisions or adjustments[10]). As defined in SEC (2004), these occur if the company or its auditors conclude that the company's previously issued financial statements "[...] no longer should be relied upon because of an error in such financial statements." The rule also mandates firms to disclose such restatements on the Form 8-K, specifically Item 4.02 (Scholz, 2014; SEC, 2004). Similar to MWs, voluntary financial restatements disclosures (FSR) also have a negative effect on the stock price, but the reaction is not uniform (Agrawal & Chadha, 2005; Callen, Livnat, & Segal, 2005; Chen, Cheng, & Lo, 2013; Hennes, Leone, & Miller, 2008).

While disclosures of FSRs and MWs are governed by different processes within the company, the two are closely related.

> [...] "*The underlying assumption is that material financial reporting errors stem from material weaknesses in internal controls. Specifically, the intention [of SOX404] was to modify problems in internal controls before they potentially lead to material errors and disclose the existence of problematic internal control in advance of material errors Accordingly, evidence that ICFR auditing is working as intended would suggest a greater prevalence of material weaknesses relative to material errors [and] would result in ma-*

[10]For types of restatements see: www.auditanalytics.com/blog/different-types-of-financial-statement-error-corrections.

terial weakness disclosure that frequently precede the discovery of material errors. " [...] (Aobdia et al., 2016)

However, empirical evidence demonstrates the opposite. Academic research as well as the SEC and the Public Company Accounting Oversight Board (PCAOB),[11] have raised concerns regarding a misaligning trend between disclosures of financial restatements and MWs. MWs have been consistently lagging behind in number (Baumann, 2010; Plumlee & Yohn, 2010). Firms fail to report MWs in a timely manner, and MW disclosures infrequently precede a future material misstatement (Rice & Weber, 2012; Rice et al., 2014). While the decline in the number of MWs may be viewed as a positive sign of overall better ICFR, the SEC has suggested that the decrease in reporting control weaknesses in recent years "could be due to material weaknesses not being identified or reported" as opposed to improvements in underlying controls (Whitehouse, 2010a, 2010b). In 2014, of 3874 issuers with an ICFR opinion, 6 percent received an adverse ICFR opinion while 11.3 percent announced a financial restatement (Franzel, 2015). Furthermore, the PCAOB reported that 35 percent of the internal control audits performed by the U.S. Big 4 firms that it reviewed for 2014 were deficient (Franzel, 2015). Overall, the literature agrees that the linkage between MW and FSR disclosures is an important, but under-researched topic (Francis, 2011; Hogan, Lambert, & Schmidt, 2013; Nagy, 2010). Most recently, Wans, He, and Sarath (2014) study pricing effects of the joint disclosure of a MW and restatement, and find that such a dual disclosure is viewed more adversely by the market as compared by a disclosure of just a MW or just a restatement.

Extant literature has explored the underlying reasons for the different trends between the two disclosures. First, internal audit procedures may be insufficient and *unable to uncover that control weaknesses exist.* Second, determining the severity of internal control deficiencies is difficult, subjective, and involves considerable expertise and judgment. Thus, auditors may find it difficult to assess an internal control deficiency as an MW in the absence of an existing error (Hoitash et al., 2008; Kinney Jr

[11]The Public Company Accounting Oversight Board (PCAOB) was created to oversee the audits of public companies and other issuers in order to protect the interests of investors and further the public interest in the preparation of informative, accurate, and independent audit reports (Sarbanes-Oxley Act of 2002, Public law 107-204, Section 101).

et al., 2013). The management may be generally aware that a control deficiency exists but *fail to classify it as an MW* and, subsequently, not disclose it. Most recently, Aobdia et al. (2016) find that auditors' inability to properly identify relevant internal controls is a contributing reason for why MWs are not discovered and disclosed prior to material error restatements.[12] Third, the pressure to not issue an adverse opinion arises from the fact that timely disclosures are not rewarded by the market (Kasznik & Lev, 1995). Consistent with the negativity associated with MW disclosures discussed above, managers *prefer to avoid MW disclosures* and may attempt to deter auditors from disclosing control weaknesses (Kinney et al., 2011; Kinney Jr et al., 2013; Rice & Weber, 2012; Rice et al., 2014; Turner & Weirich, 2006).

All three disclosure determinants discussed above - error detection, MW identification, and willingness to disclose - have been incorporated in our theoretical disclosure game, which we proceed to describe in the following sections.

3. Material Weakness vs. Financial Restatement and Delaying vs. Timely

3.1. MW vs. FSR

In this subsection, we describe the two types of disclosures a manager is requested to make in our game and the relationship between them. At this point, we assume that the manager is not acting strategically and making the disclosure in a timely manner.

Previously, we summarized the current notion in the literature indicating that the manager needs to know whether a material weakness exists in a firm's financial reporting system before he discloses it to the investors. The lack of his knowledge may be attributed to two factors: inability to detect an error or inability to evaluate and understand the error's magnitude/pervasiveness. We, therefore, make a distinction between the manager's ability to (a) "observe" negative events and collect evidence ("detect" financial restatements) and (b) evaluate collected evidence to understand the "root-cause" of the observed error (identify material weaknesses). We will discuss this distinction both in mathematical and economic terms below.

---

[12] "It is surprisingly rare to see management identify a material weakness in the absence of a material misstatement. This could be either because the deficiencies are not being identified in the first instance or otherwise because the severity of deficiencies is not being evaluated appropriately." (Croteau, 2013)

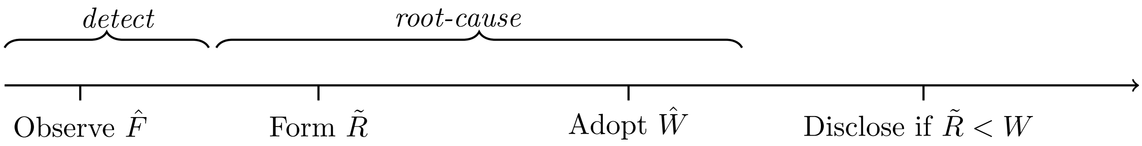


Figure 1. Progression from FSR to MW

Managers are requested to disclose realizations of **financial restatements**, denoted as $\hat{F}$. Observation of $\hat{F}$ triggers the manager to evaluate whether the disclosure of a **material weakness** is necessary. In order to do so, the manager is requested to form his *ex-ante* beliefs about the reliability of the firm's reporting system, which we denote as $\tilde{R}$. These beliefs are based upon observing $\hat{F}$ as well as other available evidence (*prior information*) (Corless, 1972; Crosby, 1981; Mock, Wright, Washington, & Krishnamoorthy, 1990).[13] As not all FSRs indicate an existing MW, observing $\hat{F}$ does not imply that an MW exists:

$$P(\hat{F}) > P(MW) \text{ and } P(MW|\hat{F}) \neq 1 \tag{2}$$

The manager then proceeds to adopt a reporting threshold W, such that if the reliability $\tilde{R}$ is below this threshold[14], the firm has an MW, e.g.:

$$\text{Disclose MW} \iff \tilde{R} < W \tag{3}$$

Figure 1 summarizes the progression, from observing $\hat{F}$ to disclosing an MW.
The economic interpretation of this distinction is as follows. The manager is hired to collect evidence ("detect") and evaluate the state of the firm's financial reporting system (identify the "root-cause") in order to prepare a decision-relevant report (disclose or not disclose an MW). We assume that his evaluation is triggered[15] by the observation of actual financial results ($\hat{F}$). He then proceeds to assess other financial and non-financial signals of the firm's performance (deficiencies, excursions, strength of

[13]The process of how his ex-ante beliefs are obtained has been left out of the scope of this paper; however, we mention that the concepts of the reliability theory (Bodnar, 1975; Cushing, 1974; Mohamed, Qureshi, & Behnezhad, 2005; Srinidhi & Vasarhelyi, 1989; Stratton, 1994) can be applied to model $\tilde{R}$ as a probability distribution over an observable signal.
[14]In practice, establishing such a threshold is a demanding task.
[15]In practice, MWs may exist in the absence of FSR.

procedures, results of past audits, anti-fraud and ethics trainings, employee hiring and evaluation processes, etc.) in order to evaluate the deficient process and the scope of the potential error.[16] Therefore, for disclosing an MW, the manager needs to perform a two-step process - *detection* followed by identification of *root-cause*.

This distinction allows us to separate the disclosure of an FSR, a "low-hanging fruit" disclosure that does not involve severe consequences to the manager, from a more comprehensive and potentially punishing disclosure of an MW, which requires the manager to perform more work. We also separate the incentives of the manager (acting as an internal auditor) and external auditors: The manager may get a reputation bonus if he finds the error before the auditor is paid a high fee. This practice has been noted in existing literature: When issues are identified internally or management concedes to a deficiency (as opposed to denying it), auditors assess the significance of the deficiency lower (Wolfe, Mauldin, & Diaz, 2009). It also allows us to model a varying degree of effort invested in obtaining the information ("*root-cause*" will involve a costly effort). To further support the distinction between *detection* and identification of *root-cause*, Aobdia et al. (2016) find that situations where the management is generally aware that a control deficiency exists, but misjudges its severity and fails to classify it as an MW, are not uncommon.

3.2. Delaying vs. Timely

In this subsection, we will allow the manager to act strategically and choose the disclosure strategy as well as the amount of effort that he wishes to invest into disclosing.

Consider a simplified scenario wherein the manager faces a dichotomous choice between two possible thresholds, $\hat{W} > \hat{W}_{other}$. $\hat{W}$ captures the reliability required by regulators (and involves a costly effort from the manager), and $\hat{W}_{other}$ captures a substandard level of reliability. In this case, the probability of disclosing an MW depends

[16]As insiders, managers are well-equipped to estimate $\tilde{R}$ based on observable factors. To demonstrate the differences between MW and non-MW firms based on several financial predictors, we perform the following exercise. Using actual MW data, we predict the probability of a firm having an MW using a set of covariates based on the research of Rice and Weber (2012). In Figure A2, we plot the kernel densities of the predicted probabilities for two subsamples - the firms that disclosed an MW in the year it occurred (corresponds to density plot f_1), and firms that never had an MW (f_2). We also fit a two-parameter beta distribution (Corless, 1972; Crosby, 1981) to a distribution of predicted probabilities for both subsamples (by using *betafit* in Stata (Buis, Cox, & Jenkins, 2012), which performs the fit by maximum likelihood (Gupta & Gupta, 2000).

on which threshold W is used by the manager. A manager adapting a substandard threshold $\hat{W}_{other}$ is less likely to disclose an MW (see Equation 1).[17] Thus, we denote the choice $\hat{W}_{other}$[18] as a "delaying" disclosure strategy. The intuition behind the terminology is as follows. The disclosure of the FSR has already been made. The manager may be withholding the disclosure of an MW hoping to gain the additional time required to remediate the issue before the next audit. Alternatively, if the efforts associated with disclosing MWs are too high, relative to the incentives present for timely disclosure, he may be tempted to gamble and wait in the hope that an unreliable reporting system will not be timely detected by the external auditor.[19] Conversely, the choice of $\hat{W}$ is denoted as a "timely" disclosure strategy.

In our game, the choice of W also determines the amount of effort the manager invested in MW identification. We allow the "delaying" manager (choice of $\hat{W}_{other}$) to voluntarily choose to not "dig deeper" and obtain more information, such that his disclosure remains at the level of the observable signal $\hat{F}$ without additional costly effort associated with $\hat{W}$. Thus, our model is compliant with the usual assumption of truthful disclosure, allowing for a varying degree of effort invested in obtaining the information.[20] Based on practices in existing literature, we assume that if the manager does not learn information (e.g., does not observe $\hat{F}$), he necessarily makes no disclosure.

[17] Assume that $\hat{W}_{other} < \tilde{R} < \hat{W}$. A manager who uses a reliability threshold $\hat{W}$ discloses an MW. A manager whose adopted threshold is $\hat{W}_{other}$, upon formation of the same beliefs, does not disclose an MW. In this setting, a manager's disclosure of an MW is equivalent to the choice of W.

[18] Formation of the wrong set of beliefs and adoption of $\hat{W}_{other}$ may occur for various reasons - insufficient testing, failure to classify the severity of the observable signal, inability to identify relevant internal controls, or strategic choice of non-disclosure.

[19] We choose to call it *delaying* and not *concealing* as, in reality, the issue will either manifest itself in later periods or the auditor will unveil it in subsequent audits. Thus, complete concealment is unlikely. To support this, we note evidence by Franzel (2015) that there exist situations when clean ICFR audit opinions are subsequently withdrawn and replaced with adverse opinions. Additionally, we observe instances where the existence of a material weakness in prior periods is acknowledged by the management, but at the time of the disclosure the issue has already been remediated, and thus no material weakness disclosure is required. This is in line with evidence by Graham, Harvey, and Rajgopal (2005) that some CFOs claim that they delay bad news disclosures in the hope that they may never have to release the bad news if the firm's status improves before the required information release.

[20] Existing literature assumes that the manager only has two options - disclosing or withholding the information - meaning, that he can neither "partially" disclose the information nor distort the information he receives before disclosing it (R. Dye, 2013).

Step 0	Step 1	Step 2
Nature chooses state of firm Nature chooses type of manager	Manager chooses his strategy (Delay/Timely) Manager adopt W, forms beliefs $\tilde{R}$ and reports the state to owner	Owner employs auditor to verify manager's report Payoffs are realized

Figure 2. The Timeline

4. The Disclosure Game

4.1. The Model

In this section, we present the model of a one-period[21] disclosure game in order to demonstrate how strategic behavior of the manager is determined. We adapt the dynamic auditor rotation model given by Gietzmann and Sen (2002) to our setting.

There are two risk-neutral players involved in a project - the owner (principal), who considers the firm as an investment project, and the manager (agent), who oversees operations on a day-to-day basis. There are three steps (step 0, 1, and 2) that, along with the choices of the players, are outlined in Figure 2. Detailed actions of all players are described in the following subsections and graphically presented in Figure A1.

4.1.1. The Firm and the Owner

The inherent reliability of the firm's financial reporting system can be classified as either good, g, or bad, b. At the beginning of the game, the owner assumes that the unconditional probability that the firm is good is $p(g)$. If the state of the firm's reporting system is good (bad), its next year net present value is V_g (V_b).

Economically, a good firm has a robust and effective ICFR system and thus, is less likely to receive an adverse audit opinion at the end of the year. Conversely, the ICFR of a bad firm is ineffective, so there exists a strictly positive probability that the firm will suffer financial distress (caused by failures in financial reporting or an adverse audit opinion). We assume that $V_g > V_b$.

In order to generate a potential demand for a manager to attest to the state of the firm's financial reporting system, we assume that if the state is bad (before it suffers financial distress), the owner can immediately liquidate the firm for a value Z, such that $V_g > Z > V_b$; i.e., the owner would prefer to keep the good firm and liquidate a

[21] Details of a two-period game can be obtained from the authors.

bad one. However, without the manager, he continues with the firm, because:

$$\mathrm{p(g)V_g + (1 - p(g))V_b > Z} \tag{4}$$

4.1.2. The Manager

The manager is hired by the owner to obtain private information concerning the state of the firm and disclose it to her at the end of step 1. He does so by issuing a report. We introduce a *tilde* notation to differentiate between the manager's report on the state of the firm and the intrinsic states, g and b. Therefore, $\tilde{g}$ ($\tilde{b}$) indicates that the manager reports that the firm is good (bad).

Given the stochastic nature of the firm, one of the main responsibilities of the manager is to act as an "internal auditor", setting up the necessary controls and procedures for ensuring that the firm expects to continue being profitable. However, the manager's truthfulness may be compromised since his desire to maximize expected profits may induce him to "not find" evidence that places the firm in a negative light. While there may be other reasons for the managerial choice of disclosure, we simplistically assume that the manager's desire to disclose in a timely or delayed manner is determined by his expected compensation.

As discussed in Section 3.1, the manager learns about the bad state in a two-step process, as described in Figure 1. His "detection" effectiveness is denoted as d and "root-cause" effectiveness is denoted as r. With probability r, his root-cause efforts are successful, so he discloses the findings in a bad state report $\tilde{b}$ to the owner at the end of step 1. However, there exists a probability $1 - r$ that, despite "*detecting*" (observing $\hat{F}$), he is unable to identify the root-cause of the issue and still issues a report $\tilde{g}$. In the case of probability $1 - d$, he doesn't observe the signal and issues a good report $\tilde{g}$.

We introduce the possibility of different strategic behaviors by different managers in two ways. Our model includes two (unobservable) types of managers. The H type makes a firm-specific investment in internal controls, denoted by I,[22] unobservable

[22]Examples of I include technological investment in firm's system of internal controls, training for employees, risk assessments, and other procedures to reasonably assure that financial errors are prevented and detected timely to maintain a "good" state of the firm.

by the owner, resulting in the chance of observing $\hat{F}$ being d_H. Conversely, type L manager is unable to improve the system of internal controls. Thus, his detection probability remains d_L. Similarly, the root-cause abilities of type H (L) manager is r_H (r_L). This investment is productive in the sense that it increases the probability of a manager internally detecting a bad state, with:

$$1 > \mathrm{d_H} > \mathrm{d_L} > 0,\ 1 > \mathrm{r_H} > \mathrm{r_L} > 0 \tag{5}$$

There is a continuum of managers uniformly distributed on the interval $[0,1]$. The proportion of type H managers in the economy is denoted as $m \in (0,1)$.

Second, we allow both types of managers to act strategically and choose the reliability disclosure threshold $W \in [\hat{W},\hat{W}_{other}]$, which is equivalent to adopting a Timely (if $\hat{W}$ is chosen) or Delaying ($\hat{W}_{other}$) strategy. The choice of Timely strategy requires the manager to bear a firm-specific cost of disclosure effort,[23] denoted by $e > 0$; unobservable by the owner, fixed and equal for both types of managers. We note that the cost of this effort is different from a negative market reaction of publicly disclosing an MW and involves additional work on behalf of the manager to set up a timely communication channel with the owner and, potentially, the verifier. The choice of Delaying strategy carries no effort and thus, the costs of delaying are normalized to zero. To distinguish between different strategic choices of the two types of managers, we denote that type H (L) manager chooses a Delaying strategy with probability c_H (c_L) and, in general, the Delaying probability is denoted as $c_m \in [c_H, c_L]$.

A type H manager learns the signal $\hat{F}$ with probability d_H and finds the root-cause of the issue with effectiveness r_H, in which case, he issues a bad report $\tilde{b}$. If he observes $\hat{F}$, but is unable to identify the root-cause, he issues a good report $\tilde{g}$ and discloses FSR. If he is unable to observe $\hat{F}$, he reports $\tilde{g}$.

A type L manager's behavior is different in the following sense. He observes $\hat{F}$ with probability d_L but does not proceed to find the root-cause of the issue. He always issues a good report $\tilde{g}$ and discloses FSR if he observes $\hat{F}$.

[23]Economically, e involves setting up a disclosure channel between the management and shareholders, time investment in alignment with external auditors, additional testing and communication efforts, etc.

4.1.3. The Auditor

At step 2, the owner hires an independent external auditor ("fact-finder") to verify the managerial report $\tilde{g}$, who, at step 2, learns the true state of the firm with probability $\omega > 0$. Analogous to the notation in Section 4.1.2, we introduce a double *tilde* notation for the auditor's report.[24] Parameter ω is the effectiveness of the audit, where the audit is deemed effective if it was able to reveal the bad state, conditional on it being bad (report $\tilde{\tilde{b}}$). We assume that since the audit technology is imperfect, audits do not perfectly reveal the firm's underlying state, so $\omega < 1$. We also assume that external auditors always report what they see and do not strategically misrepresent their findings: If the auditor detects a bad state, the firm will end in the period. Conversely, with probability $(1-\omega)$ that the external auditor is unable to identify a bad state (conditional on it being bad), he issues a clean report $\tilde{\tilde{g}}$ and the firm continues. If the manager reports a bad state, the external auditor confirms it with probability 1 (there is no need to verify a bad report and thus, $p(\tilde{\tilde{g}}|\tilde{b}) = 0$ and $p(\tilde{\tilde{b}}|\tilde{b}) = 1$). If the external auditor issues a bad report $\tilde{\tilde{b}}$ before the manager, the firm is considered as having failed and the owner gets V_b.

As in R. Dye (2013), we assume that both the manager and external auditor may only make one type of error - they may fail to detect a bad state (Type I error, false positive) - but cannot err to claim the firm is bad when it is not (Type II error, false negative). Naturally, an auditor will be unable to issue an adverse conclusion on the firm's ICFR when it is, in fact, effective.[25]

4.1.4. The Fees and Payoffs

The manager is paid a compensation $\mathcal{F}$ at the end of step 2 after the owner verifies his report with the external auditor. He is incentivized to learn more: In case of disclosure of the observed signal $\hat{F}$, he gets an additional bonus $a > 0$; if he learns and reports the bad state before the external auditor, his bonus is A ($A > a$). The economic interpretation of these incentives is the following. If the owner is able to obtain a report of the true state of the firm by incentivizing the manager to work, she can

---

[24]Therefore, $\tilde{\tilde{g}}$ indicates that the auditor issued a report that the firm is good

[25]There are several institutional barriers for the auditors to trivially or strategically misrepresent their findings, such as legal liabilities and requirements to maintain a defensible paperwork trail.

save on the audit fees. We can interpret a and A as the financial resources that the owner can potentially reinvest in growth opportunities that positively contribute to the overall manager's performance-dependent compensation package. Additionally, A and a can be viewed as the reputation fees earned by the manager for exerting effort, which does not affect the payoff of the owner. We assume that $\mathcal{F}$ is sufficiently large for covering the manager's operating costs so that both types of managers can participate.

If the manager reports a good state $\tilde{g}$, but the external auditor reveals that the state is, in fact, b, the manager is penalized and required to pay a damage compensation α to the owner (e.g., to compensate the latter for the additional audit fees associated with substantive audits required to confirm a bad state).[26]

The goal of the manager is to maximize his compensation at the end of step 2 by choosing a disclosure behavior as the solution to the following optimization problem:

$$\max_{c_m=[c_H,c_L]\in[0,1]} \quad \mathrm{E}(\mathcal{Y}_{\mathrm{M}}(\mathrm{m},\mathrm{c_m},\sigma)) \tag{6}$$

where $\mathcal{Y}_M(\cdot) = f(\mathcal{F}, I, e, A, a, \alpha)$ is the manager's payoff function at the beginning of the period, m refers to the manager's type, c_m is the strategy of the manager denoting the probability with which the manager decides to delay the information, and $\sigma \in [g, b]$ indicates the state of the firm.

The owner's payoffs $\mathcal{Y}_O$ when hiring a manager are as follows:[27]

$$\mathcal{Y}_O = \begin{cases} V_g - \mathcal{F} \text{ if the manager's and the auditor's reports are } \tilde{g} \text{ and } \tilde{\tilde{g}} \\ V_b - \mathcal{F} \text{ if the manager's and the auditor's reports are } \tilde{g} \text{ and } \tilde{\tilde{b}} \\ Z - \mathcal{F} \text{ if the manager's report is } \tilde{b} \end{cases}$$

4.1.5. Game Summary

We can summarize the flow of the one-period disclosure game in the following manner. At step 0, before the game starts, nature chooses the state of the firm $\sigma \in [g, b]$, with respective prior probabilities of $p(g)$ and $1 - p(g)$. Nature also chooses the type

[26] In this model, since the manager is responsible for the day-to-day operations of the firm, he is liable for not setting up appropriate ICFR and thus, is guilty of negligence.

[27] The owner's payoffs without hiring a manager are defined as V_g if the firm's state is g and V_b if state is b.

of manager (H or L),[28] whereby $m \in (0,1)$ is the proportion of type H managers capable of investing an amount I in internal controls. The manager is hired and paid a compensation $\mathcal{F}$, provided that he produces decision-relevant information.

The manager chooses his disclosure strategy (Timely or Delaying) and proceeds to detect the error and possibly, the root-cause of the state of the firm, issuing an attestation report in step 1. At step 2, his report is validated by the external auditor (with effectiveness ω), such that the owner obtains the final report at the end of step 2 and the payoffs are realized.

The sequence of strategic moves and payoffs of the manager in a one-period game are summarized in Figure A1. Note that nodes HIGH and HIGH' (as well as LOW and LOW') are informationally equivalent. We restrict our analysis to pure strategies, so the admissible disclosure strategy choices are Delay or Timely.

4.2. Analysis of the Equilibrium

4.2.1. Owner's demand for manager's information

We assume that the owner's decision on whether or not to hire a manager is based upon the evaluation whether he produces decision-relevant information. This information should have a strictly positive value to the owner. The extent to which the manager produces such information leads the owner to revise her beliefs concerning the state of the firm.

If the state is b, the probability that the manager will not detect it is given as:

$$\begin{gathered}\nu \equiv m(1 - k_H(1 - c_H)) + (1 - m)(1 - k_L(1 - c_L)) \\ k_i = d_i \cdot r_i,\ i \in [H, L],\ 1 > k_H > k_L > 0\end{gathered} \tag{7}$$

The unconditional probability of the manager producing a good report $\tilde{g}$ is:

$$p(\tilde{g}) = p(g) + (1 - p(g))\nu \tag{8}$$

[28]Since managerial type is unobservable by the owner, it is considered as though nature assigns the type to the firm.

The revised belief of the owner that the state of the firm is g after having observed a good report $\tilde{g}$ is the following:

$$\rho(g|\tilde{g}) = \frac{p(g)}{(p(g) + (1 - p(g))\nu} \tag{9}$$

Theorem 4.1. *The manager's report has value to the owner when the following inequality is satisfied:*

$$p(\tilde{g})(\rho(g|\tilde{g})V_g + (1 - \rho(g|\tilde{g}))V_b) + (1 - p(\tilde{g}))Z - F > p(g)V_g + (1 - p(g))V_b \tag{10}$$

Proof: See Appendix.

4.2.2. Equilibrium Conditions

To identify equilibrium relationship in this game, we use the concept of the Bayesian-Nash sequential equilibrium as defined by Kreps and Wilson (1982), which we adapt to our situation. Please refer to Definition B.1 in the Appendix.

Since (by Assumption 5) $d_H > d_L$ and $r_H > r_L$, a separating equilibrium is only possible when type H manager discloses timely, but type L manager delays the disclosure. Following the notation adopted in Gietzmann and Sen (2002), we denote this equilibrium as a **Partial Compliance (PC)** equilibrium. A pooling equilibrium in which both types do not delay, is denoted as a **Full Compliance (FC)** equilibrium. Another pooling equilibrium, in which both types delay the information, is denoted as a **No Compliance (NC)** equilibrium.

Theorem 4.2. *There exist an FC equilibrium and a PC equilibrium of the single-period disclosure game presented above. An NC equilibrium does not exist.*
An FC equilibrium exists if:

$$e \leq (1 - p(g))k_L\tilde{A}, \ where \ \tilde{A} = A - a + \omega(\alpha + a) \tag{11}$$

A PC equilibrium exists if (note $\Delta d = d_H - d_L,\ \Delta k = k_H - k_L$):

$$(1-p(g))k_L\tilde{A} \leq e \leq (1-p(g))k_H\tilde{A} \tag{12}$$

$$I \geq (1-p(g))[\tilde{A}\Delta k + a(1-\omega)\Delta d] \tag{13}$$

Proof: See Appendix.

4.2.3. Analysis of the Equilibrium

To induce the manager of type H to disclose, the incentive to detect and identify root-cause, denoted as $\tilde{A}$ in Condition 11, should be strictly positive, that is:

$$A > a(1-\omega) - \alpha\omega \tag{14}$$

We also note that for an FC equilibrium to exist, the liability for untimely disclosure needs to be higher than for a PC equilibrium. A PC equilibrium will follow if $\tilde{A}$ is chosen to satisfy:

$$\frac{e}{(1-p(g))k_H} < \tilde{A} < \frac{e}{(1-p(g))k_L} \tag{15}$$

We interpret $\tilde{A}$ as a ω-weighted liability level.[29] If damage costs α are sufficiently large, or the auditor is sufficiently effective, it is easy to satisfy Condition 11 to induce both types to work. However, low damage costs or an ineffective auditor (consider an extreme case of $\alpha = 0$ or $\omega = 0$) make it unattractive for the manager to exert effort, despite his potential future bonuses for timely disclosure. Depending on type L abilities, $\tilde{A}$ that is sufficient enough to induce type H to work, may be insufficient for type L, so a **PC** equilibrium will result. However, if the level of liability α is very high, the manager will always exert effort regardless of his type, and an **FC** equilibrium will follow.

[29] A disclosing manager always gets $A - a$, and, if the auditor is effective, he will also avoid paying the penalty α.

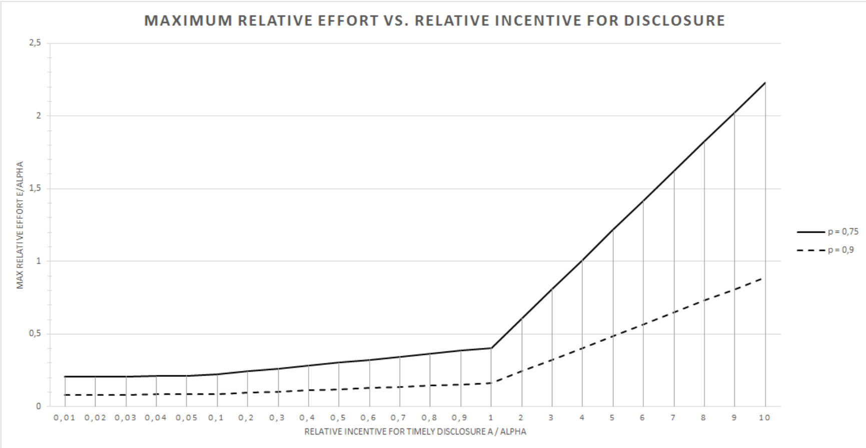


Figure 3. Maximum Relative Effort vs. Relative Incentive for Timely Disclosure
Graphical illustration of maximum relative effort $\frac{e}{\alpha}$ as a function of relative incentive for timely disclosure $\frac{A}{\alpha}$ for $d = r = 0.9$ and $p = 0.75$ and $p = 0.9$.

In Condition 11, the expected value of the liability $(1-p)\tilde{A}$ is reduced by k_L, so the reduction of the liability compensates for the extra effort. In equilibrium, the minimal liability level $\tilde{A}$ is determined by Condition 11 or 12, and for a given effort e it depends on operational parameters ω and $p(g)$ and managerial abilities d and r.

To illustrate the relative maximum effort (as compared to potential penalties), we divide both parts of Equation 11 by α. Assuming $a = 0$ (no incentive for the disclosure of FSR) and $\omega = 1$ (highly effective external auditor), the maximum relative effort that type L manager is willing in invest in MW discovery is given by the following:

$$\frac{e}{\alpha} = (1-p) \cdot d_L \cdot r_L \cdot (1 + \frac{A}{\alpha}) \tag{16}$$

Figure 3 plots $\frac{e}{\alpha}$ as a function of $\frac{A}{\alpha}$ for $d_L = r_L = 0.9$ and two different values of p. When $\frac{A}{\alpha}$ is close to 0, the maximum effort e represents only a small fraction of the potential penalty α ($\frac{e}{\alpha} = 0.2$ for $p = 0.75$). The effort is financially equivalent to the penalty ($\frac{e}{\alpha} = 1$) when $\frac{A}{\alpha} = 4$ for $p = 0.75$. Therefore, to induce an FC equilibrium, policymakers need to balance setting a relatively high incentive for timely disclosure $\frac{A}{\alpha}$ and, at the same time, lowering firm's disclosure efforts e.

5. Empirical Results

5.1. Research Question and Hypothesis

A firm that reports an MW sends a signal to market participants that the reliability of its internal financial reports is low (Feng, Li, & McVay, 2009). This signal will likely lead investors to demand a higher expected return from the firm's expected cash flows (Ashbaugh-Skaife et al., 2009; Hoag & Hollingsworth, 2011). Intuitively, the manager may try to avoid the consequences by attempting to push the disclosure of the MW to a later period if possible.[30]

Building on the theoretical model outlined in Section 4, we empirically test whether the firm that chooses to delay its MW announcement (e.g., discloses FSR before its first MW) are characterized by a different level of disclosure effort and/or detection abilities. Unfortunately, none of these characteristics are directly observable. We focus on the following financial characteristics that help us infer effort and detection abilities of a firm. First, we hypothesize that higher detection ability is associated with higher audit scrutiny, and thus we expect that a timely disclosing firm is likely to be associated with higher level of audit fees and auditor rotations. We also believe that high-risk financial accounts related to revenue recognition and tax transactions, as well as issues related to fraud, warrant higher management attention and auditors effort; thus we expect that a timely disclosing firm is likely to be associated with higher likelihood of errors identified in the areas of revenue, tax or fraud. A firm exhibiting poorer financial results or in distress is more likely to be scrutinized for an underlying material weakness, therefore, a timely disclosing firm is likely to be associated with a higher probability of financial distress and lower financial performance indicators.

Our primary reference point in the selection of above characteristics is research conducted by Rice and Weber (2012), which identifies various characteristics (financial distress, auditor effort, size, fees, external capital needs, auditor characteristics, etc.) as the main driving factors of disclosing an MW as well as the research of Rice et al. (2014), which studies how disclosing an MW before an FSR is associated with a greater likelihood of litigation, such as class action lawsuits and turnover, against the

[30]We relate this to the Delaying strategy below.

$$
\begin{array}{lccc}
 & \textbf{Non-MW} & \textbf{Delayed} & \textbf{Timely} \\
 & \begin{array}{cc} Y_{t-1} & Y_t \end{array} & \begin{array}{cc} Y_{t-1} & Y_t \end{array} & \begin{array}{cc} Y_{t-1} & Y_t \end{array} \\
\begin{array}{l} \text{FSR disclosure} \\ \text{MW disclosure} \end{array} & \begin{bmatrix} any & any \\ no & no \end{bmatrix} & \begin{bmatrix} FSR & any \\ no & MW \end{bmatrix} & \begin{bmatrix} no & any \\ no & MW \end{bmatrix}
\end{array}
$$

Figure 4. Possible MW-FSR Disclosure Patterns

The figure depicts different MW-FSR disclosure combinations possible around the first MW. Non-MW firms have various patterns of FSR disclosures, but never disclose a MW. "Delaying" firms have an FSR in the year preceding to the year of first MW disclosure. "Timely" firms will have no FSR in the year preceding the first MW. Y stands for Year.

firm. The former is also related to research by Ashbaugh-Skaife, Collins, and Kinney (2007); J. Doyle, Ge, and McVay (2007); J. T. Doyle et al. (2007), which shows that profitability and complexity are associated with the probability of disclosing an MW. Consequently, by controlling for the inherent differences between MW firms, we should be able to separate the incremental effect of the strategic choice of disclosure on firm performance. Additionally, Hennes et al. (2008) show that restatements involving fraud are more penalized by the market.

In our analysis, we distinguish between different patterns of MW-FSR disclosure. In particular, we identify firms whose first MW disclosure was preceded by an FSR. We call this firm "delaying." All other MW disclosing patterns, where the first MW is not preceded by FSR, are considered to be "timely" (e.g., if a firm discloses the first MW before or together with FSR or discloses MW-only without an associated FSR). Figure 4 provides an illustration of different MW/FSR disclosure combinations. Importantly, we focus on the first instance of disclosing an MW in order to identify "delaying" firms. In case of repeat MWs, the negative consequences have been already established and there is no incentive to delay the bad news concerning the unreliability of the financial reporting system.

The theory discussed in Section 4 suggests that a type L manager may be induced towards a Delaying strategy equilibrium if he is more concerned with current costs than potential penalties. Since prior empirical research has shown that MW-disclosing firms are inherently different from non-MW firms-disclosing, we hypothesize that, conditional on disclosing a MW, firms that choose a Timely strategy are likely to exhibit different financial characteristics as compared to Delaying or all other firms:

Table 1. Sample Breakdown By Disclosure Pattern

Pattern	N of obs	No of firms	Ave Obs./Firm	% of Sample
MW delaying	955	121	7.89	2%
MW timely	28,481	5,564	5,12	53%
FSR only	7,744	975	7.94	14%
Clean	16,534	2,717	6.09	31%
Total Sample	**53,714**	**9.377**	**5.73**	100%

Hypothesis 1. The financial characteristics of a "delaying" firm are different from the respective characteristics of "timely" firms in all years ("delaying" $\neq$ "timely").

Hypothesis 2. The financial characteristics of a "timely" firm are different from the respective characteristics of control firms in all years ("timely" $\neq$ control).[31]

5.2. Sample Selection

We obtain our sample from the Audit Analytics database on financial restatements and SOX 404 opinions, extracting all SOX 404 reports between 2005 (the year after the first effective SOX 404 year) and 2014 (118,363 firm-year observations), and matching them with non-reliance restatement information from Audit Analytics and financial data from Compustat.

We eliminate duplicate observations (same CIK-year combination) as well as observations missing fiscal year or SIC code, ones without a definitive opinion on internal controls, and missing Compustat data to construct the variables in our model or to obtain uninterrupted years between FSR and MW disclosures (e.g., when a panel data of financial information is missing one year between FSR and MW disclosures). This resulted in a usable sample of 53,714 observations (9,377 unique firms), which includes 28,481 observations with "timely" firms, 7,744 observations with FSR-only disclosing firms, 955 observations with "delaying," and 16,534 observations of firms without MW or FSR disclosures. Table C1 summarizes the sample selection process. Table 1 reports the breakdown of the sample by different possible disclosure patterns, as discussed in Section 5.1.

All empirical analyses discussed in this paper have been carried out in Stata (R).

[31]While Hyp. 2 has been extensively studied in prior literature, it is important in our research setting. By empirically demonstrating the difference between "timely" vs. "delaying" and "timely" vs. "control," we highlight the fundamental reason behind some managers adopting a *delaying* strategy

5.3. Research Design

5.3.1. Construction of the Variables

We use various financial and audit characteristics as dependent variables Y. We include the following financial measures:

- Financial performance (Return on Assets (ROA))
- Tobin's Q (TQ)[32]
- Profitability ($Profitability$)
- A decile of Altman Z score (Altman et al., 2000)($AZtile$).

Our audit-related measures include:

- Audit fees (log of $audit\ fees$ scaled by $total\ assets$)
- Audit rotation ($audit\ rotation$)
- An indicator variable of the deficiency account ($core$)[33]
- An indicator variable of a fraud-related deficiency ($fraud$)[34]

Our independent variables are defined in the following manner. D_i is an indicator variable of a firm exhibiting a "delayed" disclosure pattern; T_i is an indicator variable of a firm disclosing a "timely" pattern. By construction, D_i (T_i) is coded as 1 for all firm-year observations of a "delaying" ("timely") firm i.

Since SOX 404 MW attestations occur at the end of the firm's fiscal year (the effectiveness of ICFR is certified as of the last day of the fiscal year), we consider the year of the MW disclosure and years before as pre-treatment years and the years after the disclosure as post-treatment. Thus, $Post_{it}$ is as a dummy variable equal to 1 if the firm-year observation corresponds to a firm i in the year t after disclosure of a first MW and zero otherwise.

Based on prior research, we include control variables, such as $marketcap$, $foreign$, $bookmarket$, $salegrowth$, $tangibility$, $big4$, $leverage$, $dividendpayer$ and $finslack$, in our regression model in Equation 17.[35] All variables have been described in Table C2.

[32] While Tobin's Q is often used as a proxy for operating performance, Dybvig and Warachka (2010) find that higher Tobin's Q may be indicative of an underinvestment.

[33] An indicator variable is equal to 1 if the MW or FSR disclosed by firm i in year t is related to Revenue or Tax accounts.

[34] An indicator variable is equal to 1 if the MW or FSR disclosed by firm i in year t is fraud-related.

[35] The data on compensation is available only for a small fraction of firms in our sample. We will explore the

All continuous control variables are winsorized at the 0.01 level.

5.3.2. Regression Model

We apply the following panel regression model to examine how the firm's financial characteristics affect the choice of a Timely or Delayed MW disclosure:

$$Y_{it} = \alpha + \beta \cdot X_{it} + \gamma \cdot C_{it} + \theta_1 \cdot industry_i + \theta_2 \cdot year_t + \epsilon_{it} \tag{17}$$

where Y is one of the dependent variables; X_{it} is a vector of disclosure dummies, such that $X_{it}^T = [D_i, T_i, Post_{it}, T_i \cdot Post_{it}]$; C_{it} are financial controls; and $industry_i$ and $year_t$ are industry (two-digit SIC code) and year dummies respectively. All variables have been described in Section 5.3.1. All regressions were performed with cluster-robust standard errors in order to avoid overstating t-statistics due to multiple observations of the same firm within the dataset (Gow, Ormazabal, & Taylor, 2010; Petersen, 2009). Results of the model in Equation17 are reported in Table C6.

We are aware of the endogeneity concerns in our study. Its main source - simultaneity - exists in Equation 17 as $E(\epsilon_{it}|D_i, C_{it}) \neq 0$ and $E(\epsilon_{it}|T_i, C_{it}) \neq 0$. Simultaneity arises in our model in the following manner. If, as prior research suggests, a firm's disclosure of an MW affects it financial performance, the reverse may also be true, as the probability of having ineffective ICFR and thus reporting an MW may also be affected by current and past performance. In this case, MW disclosures and performance are simultaneously determined and both OLS and fixed-effects estimates of Equation 17 will be biased. To alleviate these concerns, we include two lags of dependent variables in Equation17 (Angrist & Pischke, 2008; Wintoki, Linck, & Netter, 2012), effectively testing the following model:

$$Y_{it} = \alpha + \delta_1 \cdot L.Y_{it} + \delta_2 \cdot L2.Y_{it} + \beta \cdot X_{it} + \gamma \cdot C_{it} + \theta_1 \cdot ind_i + \theta_2 \cdot year_t + \epsilon_{it} \tag{18}$$

where L is a lag operator. Results of the model in Equation 18 are given in Table C7.

Dynamic panel regression estimation does not solve all endogeneity problems. How-

effect of compensation on strategic disclosure choices in future studies.

ever, in the absence of a natural experiment, we believe that, as compared to using a traditional OLS, our inferences are improved.

5.4. Results

In this section, we present the main empirical results of the paper, including descriptive statistics, correlation matrix, model estimation, and robustness tests. All tables can be found in Appendix C.

5.4.1. Descriptive Statistics

The summary statistics for our sample are given in Table C3. An average firm in our sample is not profitable, has an *ROA* of -0.24, with Tobin's Q of 4.3, has a debt to total assets ratio (*leverage*) of 0.2, and cash to total assets ratio (*finslack*) of 0.15. 57% of the firms are audited by Big4 audit firms, 60% have foreign subsidiaries, and 46% pay dividends. 4.2% of deficiencies are related to Revenue or Tax accounts (*core*) and only a small fraction (0.7%) have fraud-related deficiencies. The average firm age in our sample is 12 years. 53% of firm-year observations in our sample are associated with "timely" disclosing firms, 1.8% with "delaying" firms. Figure C1 visually illustrates our sample composition by industry, year, and size of the firms.

Table C4 provides the pairwise correlation coefficients between each covariate and denotes the significance level of 5% or better with a *. Correlation of dependent and control variables is the source of endogeneity in our model, as discussed in Section 5.3.2.

We begin our analysis with the visual inspection of financial performance trends of firms with different disclosure patterns. The left graph in Figure A3 shows the difference in financial characteristics between non-MW, "delaying," and "timely" firms across all years available in our dataset. Interestingly, "delaying" firms exhibit characteristics in between those of non-MW firms and "timely" ones. For example, an average ROA of a "delaying" firm is slightly less that that of a non-MW firm but higher than the average ROA of a "timely" one. The right graph in Figure A3 demonstrates the difference in financial characteristics between "delaying" and "timely" firms around the first MW year. The latter exhibit weaker ROA and profitability and higher external financing needs and the differences decrease during or after the disclosure of MW.

5.4.2. Model Estimates

Table C5 displays the univariate test results of a two-group mean comparison test between various samples. Panel A reports and compares the variables' mean value for the control sample (non-MW) with the MW sample. The results indicate that the observations in the MW sample tend to be smaller and financially weaker (as characterized by lower ROA, $profitability$, and $marketcap$), which is in line with existing research results (Ashbaugh-Skaife et al., 2007; J. Doyle et al., 2007; Gordon & Wilford, 2012). MW firms exhibit higher sales growth and audit fees. They are less likely to pay out dividends, have operations in foreign countries, or be audited by a Big4 auditor.

Panel B compares the variables' mean value for the sample of "timely" firms with the "delaying" sample. These results indicate that "delaying" firms exhibit better financial performance, have lower audit fees, are bigger, are more likely to pay out dividends or be audited by Big4 auditors, and have less sales growth pressure. Almost all the variables exhibit a statistically significant difference in means, except *leverage* and *tangibility*. The above results foreshadow our main empirical results discussed below, providing weak preliminary support to Hypothesis 1. The results indicate that a "delaying" firm is financially different from its MW peers and closer with regard to its financial characteristics to a non-MW firm.

Table C6 reports the empirical results of the model formalized in Equation 17 and provides evidence to support Hypothesis 1 ("delaying" $\neq$ "timely") and Hypothesis 2 ("timely" $\neq$ control).

The coefficients associated with a "delaying" firm D are statistically significant only for a subset of dependent variables. In particular, we find no differences between "delaying" and control firms regarding financial characteristics such as ROA, profitability, and Tobin's Q. "Delaying" firms are associated with a higher likelihood of a deficiency being in a core account or associated with fraud. Naturally, we compare them with firms that never disclosed a deficiency.

Conversely, the coefficient of "timely" firm T is statistically significant across the dependent variables. Such firms have lower ROA and profitability, higher external financing needs, higher Tobin's Q, and exhibit a slightly higher decile of an Altman Z score than their control peers. Furthermore, their audit fees are also higher as compared to their non-MW peers. Interestingly, the results indicate that timely firms are associated with a higher probability that the auditor has rotated in the prior year, suggesting that a new auditor may be playing a contributing role in timely identification of an MW. Finally, the coefficients on $core$ and $fraud$ are positive and statistically significant, but the magnitude of the coefficients is less than that for "delaying" firms. This indicates that although "timely" firms are associated with pervasive issues in revenue-recognition, tax- and fraud-related accounts as compared to non-MW firms, this association is much more pronounced for "delaying" firms.

The fact that the coefficient of D lacks statistical significance supports our visual observations that a "delayed" firm resembles a non-MW firm with regard to its financial characteristics. The coefficients of $Post$ and $T \cdot Post$ demonstrate overall weak significance for financial performance variables, indicating that after the disclosure of MWs, the firms are not different from their control counterparts. Such firms are associated with slightly better ROA and lower probability of deficiencies in their $core$ accounts. The analysis of post-disclosure behavior of the firm is beyond the scope of the current study, as firms may be improving their internal processes as part of the remediation efforts following an MW disclosure (Felix & Wilford, 2019).

In Table C7, we report the results of a dynamic panel regression, as described in Equation 18, with lags of dependent variables. The results are consistent with the static model described above. We note that the coefficients associated with lagged dependent variables are statistically significant, indicating that a firm's previous performance plays an important role in determining its future financial characteristics. Overall, we confirm all previously discussed inferences: "Timely" disclosing firms exhibit weaker financial performance as measured by ROA and profitability; have higher Tobin's Q and higher decile of Altman Z score; and are more likely to have a new auditor. "Delaying" firms are associated with higher probability of a deficiency being related to revenue or tax accounts.

Overall, our findings support existing research that suggests that MW-disclosing firms are, indeed, inherently different from other firms and indicate that MW firms have lower performance, possibly because of inefficient operations (Cheng, Dhaliwal, & Zhang, 2013; Ge & McVay, 2005). The positive coefficient associated with the audit fees suggests that the auditor's efforts increase after the disclosure of an MW, which is in line with Hoag and Hollingsworth (2011). Finally, our results suggest that the firm choosing a "delaying" disclosure pattern is associated with better financial results (ROA and profitability) and lower probability of having a new auditor as compared to its "timely" disclosing peers. A possible explanation for this could be that the compensation contract for the managers of "delaying" firms places higher focus on current fees (which is likely to be contingent on firm's financial results) as opposed to potential future penalties. Complacency of a non-rotating auditor may render additional ICFR investments unattractive as these resources may be diverted into more profitable venues to enjoy better short-term financial results. We are aware of possible endogeneity issues and thus, establish association results without inferring the direction of causality.

5.4.3. Robustness Tests

In this section, we conduct a series of sensitivity tests for our primary results.

Our first robustness test is performed using a subsample of MW firms only when non-MW firms are excluded. The regression specification is identical to Equation 17, with covariates D and $Post$ being excluded for collinearity reasons. Excluding non-MW observations certainly introduces a selection bias into the interpretation of the results. Table C8 presents the results of this regression. Overall, we confirm our main results, demonstrating that "timely" firms exhibit significantly different characteristics than their control counterparts and are associated with weaker financial characteristics; higher Tobin's Q, higher audit fees; and lower probability of a deficiency related to revenue, tax, or fraud as compared to "delaying" firms.

In our second robustness test, we perform propensity score matching using the nearest neighbor methodology (we employ *psmatch2* command in Stata (Leuven &

Sianesi, 2018) in order to calculate propensity scores and run an OLS regression using Equation 17 to estimate the coefficient on D and $D \cdot Post$. The matching is performed on the firm's size (using *logat*), age, and one-digit industry code to satisfy the balancing property of the propensity score. Covariate T, $Post$, and $T \cdot Post$ are omitted due to collinearity. Results are presented in Table C9. Similar to our main results, we find that "delaying" firms exhibit lower audit fees (which increase after the disclosure of MW) and are more susceptible to having issues related to revenue or tax accounts than their timely disclosing peers.

Overall, our main findings in Tables C6 and C7 show consistent evidence to support both Hypotheses 1 and 2.

6. Conclusion

In this paper, we present an economic model of a disclosure game, where managers are requested to communicate their *ex-ante* beliefs regarding the reliability of the firm's reporting system along with realizations of financial restatements. However, they have incentives to delay full disclosure in the hope that an *ex-post* auditor verification does not reveal that the management ignored bad news concerning the unreliability of the financial statements. Our model demonstrates how and when such a disclosure gamble may become an equilibrium strategy.

We empirically test the model using the data on disclosures of material weaknesses and financial restatements of U.S. public companies in the 2005-2014 period. We examine the relationship between a firm's financial characteristics and various MW-FSR disclosure patterns and find that reporting a MW (material weakness) in a timely manner is associated with the firm's lower financial performance and higher probability of new auditor's appointment. As suggested by the theory, MW-delaying firms are statistically different from their timely-disclosing peers. Thus, a separating equilibrium may exist, which explains the empirical evidence of seemingly low relative frequency of reported MWs as opposed to financial restatements.

References

Aboody, D., & Kasznik, R. (2000). CEO stock option awards and the timing of corporate voluntary disclosures. *Journal of Accounting and Economics*, *29*(1), 73–100.

Agrawal, A., & Chadha, S. (2005). Corporate governance and accounting scandals. *The Journal of Law and Economics*, *48*(2), 371–406.

Altman, E. I., et al. (2000). Predicting financial distress of companies: Revisiting the Z-score and ZETA models. *Stern School of Business, New York University*, 9–12.

Angrist, J. D., & Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion.* Princeton University Press.

Aobdia, D., Choudhary, P., & Sadka, G. (2016). Do auditors correctly identify and assess internal control deficiencies? Evidence from the PCAOB data. *Georgetown McDonough School of Business Research Paper. Available at SSRN: https://ssrn.com/abstract=2838896*.

Ashbaugh-Skaife, H., Collins, D. W., & Kinney, W. R. (2007). The discovery and reporting of internal control deficiencies prior to SOX-mandated audits. *Journal of Accounting and Economics*, *44*(1), 166–192.

Ashbaugh-Skaife, H., Collins, D. W., Kinney Jr, W. R., & LaFond, R. (2008). The effect of SOX internal control deficiencies and their remediation on accrual quality. *The Accounting Review*, *83*(1), 217–250.

Ashbaugh-Skaife, H., Collins, D. W., Lafond, R., et al. (2009). The effect of SOX internal control deficiencies on firm risk and cost of equity. *Journal of Accounting Research*, *47*(1), 1–43.

Avenhaus, R., Von Stengel, B., & Zamir, S. (2002). *Handbook of game theory with economic applications.* Elsevier.

Baginski, S. P., Hassell, J. M., & Kimbrough, M. D. (2002). The effect of legal environment on voluntary disclosure: Evidence from management earnings forecasts issued in US and Canadian markets. *The Accounting Review*, *77*(1), 25–50.

Bardos, K. S., & Mishra, D. (2014). Financial restatements, litigation and implied cost of equity. *Applied Financial Economics*, *24*(1), 51–71.

Baumann, M. (2010). *Speech presented at the 2010 AICPA national conference on current SEC and PCAOB developments, Washington, DC (07.12. 10).*

Bedard, J. (2006). Sarbanes Oxley internal control requirements and earnings quality.

Beneish, M. D., Billings, M. B., & Hodder, L. D. (2008). Internal control weaknesses and information uncertainty. *The Accounting Review*, *83*(3), 665–703.

Bodnar, G. (1975). Reliability modeling of internal control systems. *The Accounting Review*, *50*(4), 747–757.

Buis, M., Cox, N., & Jenkins, S. (2012). *Betafit: Stata module to fit a two-parameter beta distribution.* Retrieved from `https://EconPapers.repec.org/RePEc:boc:bocode:s435303`

Burks, J. J. (2011). Are investors confused by restatements after Sarbanes-Oxley? *The Accounting Review*, *86*(2), 507–539.

Callen, J. L., Livnat, J., & Segal, D. (2005). Accounting restatements: Are they always bad news for investors? *NYU Working Paper No. JOSHUA LIVNAT-04. Available at SSRN: https://ssrn.com/abstract=1280738*.

Chan, K. C., Farrell, B., & Lee, P. (2008). Earnings management of firms reporting material internal control weaknesses under Section 404 of the Sarbanes-Oxley Act. *Auditing: A Journal of Practice & Theory*, *27*(2), 161–179.

Chen, X., Cheng, Q., & Lo, A. K. (2013). Is the decline in the information content of earnings following restatements short-lived? *The Accounting Review*, *89*(1), 177–207.

Cheng, M., Dhaliwal, D., & Zhang, Y. (2013). Does investment efficiency improve after the disclosure of material weaknesses in internal control over financial reporting? *Journal of Accounting and Economics*, *56*(1), 1–18.

Clinton, S. B., Pinello, A. S., & Skaife, H. A. (2014). The implications of ineffective internal control and SOX 404 reporting for financial analysts. *Journal of Accounting and Public Policy*, *33*(4), 303–327.

Corless, J. C. (1972). Assessing prior distributions for applying bayesian statistics in auditing. *The Accounting Review*, *47*(3), 556–566.

Crosby, M. A. (1981). Bayesian statistics in auditing: A comparison of probability elicitation techniques. *Accounting Review*, 355–365.

Croteau, B. (2013). Remarks before the 2013 AICPA national conference on current SEC and PCAOB developments. Audit policy and current auditing and internal control matters. *Speech available at https://www. sec. gov/News/Speech/Detail/Speech/1370540472057*.

Cushing, B. E. (1974). A mathematical approach to the analysis and design of internal control systems. *The Accounting Review*, *49*(1), 24–41.

Dobler, M. (2008). Incentives for risk reporting a discretionary disclosure and cheap talk approach. *The International Journal of Accounting*, *43*(2), 184–206.

Donelson, D. C., Ege, M. S., & McInnis, J. M. (2016). Internal control weaknesses and financial reporting fraud. *Auditing: A Journal of Practice & Theory*, *36*(3), 45–69.

Doyle, J., Ge, W., & McVay, S. (2007). Determinants of weaknesses in internal control over financial reporting. *Journal of Accounting and Economics*, *44*(1), 193–223.

Doyle, J. T., Ge, W., & McVay, S. (2007). Accruals quality and internal control over financial reporting. *The Accounting Review*, *82*(5), 1141–1170.

Dranove, D., & Jin, G. Z. (2010). Quality disclosure and certification: Theory and practice. *Journal of Economic Literature*, *48*(4), 935–963.

Dresher, M. (1961). *Games of strategy: theory and applications* (Tech. Rep.). RAND Corp Santa Monica CA.

Dybvig, P. H., & Warachka, M. (2010). Tobins Q does not measure performance: Theory, empirics, and alternative measures. *Unpublished Working paper, Washington University, Saint Louis, United Sates*.

Dye, R. (2013). Voluntary disclosure and the duty to disclose. *Working Paper, Northwestern University*.

Dye, R. A. (2017). Optimal disclosure decisions when there are penalties for nondisclosure. *The RAND Journal of Economics*, *48*(3), 704–732.

Felix, R., & Wilford, A. (2019). Does it pay to remediate? an analysis of the internal and external benefits of remediation. *Accounting and Business Research*, *49*(2), 181–205.

Feng, M., Li, C., & McVay, S. (2009). Internal control and management guidance. *Journal of Accounting and Economics*, *48*(2-3), 190–209.

Financial Accounting Standards Board, F. (2005). *Statement of financial accounting standards No. 154*.

Francis, J. R. (2011). A framework for understanding and researching audit quality. *Auditing: A Journal of Practice & Theory*, *30*(2), 125–152.

Franzel, J. (2015). Current Issues, Trends, and Open Questions in Audits of Internal Control over Financial Reporting. In *American Accounting Association Annual Meeting. Chicago, IL*.

Ge, W., & McVay, S. (2005). The disclosure of material weaknesses in internal control after the Sarbanes-Oxley Act. *Accounting Horizons*, *19*(3), 137–158.

Gietzmann, M. B., & Sen, P. K. (2002). Improving auditor independence through selective mandatory rotation. *International Journal of Auditing*, *6*(2), 183–210.

Gordon, L. A., & Wilford, A. L. (2012). An analysis of multiple consecutive years of material weaknesses in internal control. *The Accounting Review*, *87*(6), 2027–2060.

Gow, I. D., Ormazabal, G., & Taylor, D. J. (2010). Correcting for cross-sectional and time-series dependence in accounting research. *The Accounting Review*, *85*(2), 483–512.

Graham, J. R., Harvey, C. R., & Rajgopal, S. (2005). The economic implications of corporate financial reporting. *Journal of accounting and economics*, *40*(1-3), 3–73.

Grossman, S. J. (1981). The informational role of warranties and private disclosure about product quality. *The Journal of Law and Economics*, *24*(3), 461–483.

Gupta, P. L., & Gupta, R. C. (2000). The monotonicity of the reliability measures of the beta distribution. *Applied Mathematics Letters*, *13*(5), 5–9.

Hammersley, J. S., Myers, L. A., & Shakespeare, C. (2008). Market reactions to the disclosure of internal control weaknesses and to the characteristics of those weaknesses under Section 302 of the Sarbanes Oxley Act of 2002. *Review of Accounting Studies*, *13*(1), 141–165.

Healy, P. M., & 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*(1), 405–440.

Hennes, K. M., Leone, A. J., & Miller, B. P. (2008). The importance of distinguishing errors from irregularities in restatement research: The case of restatements and ceo/cfo turnover. *The Accounting Review*, *83*(6), 1487–1519.

Hermalin, B. E., & Weisbach, M. S. (2007). *Transparency and corporate governance* (Tech. Rep.). National Bureau of Economic Research.

Hoag, M. L., & Hollingsworth, C. W. (2011). An intertemporal analysis of audit fees and Section 404 material weaknesses. *Auditing: A Journal of Practice & Theory*, *30*(2), 173–200.

Hogan, C., Lambert, T., & Schmidt, J. (2013). Do management internal control certifications increase the likelihood of restatement-related litigation? *Available at SSRN: https://ssrn.com/abstract=2169553*.

Hoitash, R., Hoitash, U., & Bedard, J. C. (2008). Internal control quality and audit pricing under the Sarbanes-Oxley Act. *Auditing: A Journal of Practice & Theory*, *27*(1), 105–126.

Hoitash, R., Hoitash, U., & Johnstone, K. M. (2012). Internal control material weaknesses and CFO compensation. *Contemporary Accounting Research*, *29*(3), 768–803.

Hranaiova, J., & Byers, S. (2007). Changes in market responses to financial statement restatement announcements in the Sarbanes-Oxley era. *Available at SSRN: https://ssrn.com/abstract=1319354*.

Kasznik, R., & Lev, B. (1995). To warn or not to warn: Management disclosures in the face of an earnings surprise. *The Accounting Review*, 113–134.

Kinney, J., William, R., & Shepardson, M. L. (2011). Do control effectiveness disclosures require SOX 404 (b) internal control audits? A natural experiment with small US public companies. *Journal of Accounting Research*, *49*(2), 413–448.

Kinney Jr, W. R., Martin, R. D., & Shepardson, M. L. (2013). Reflections on a decade of SOX 404 (b) audit production and alternatives. *Accounting Horizons*, *27*(4), 799–813.

Kothari, S. P., Li, X., & Short, J. E. (2009). The effect of disclosures by management, analysts, and business press on cost of capital, return volatility, and analyst forecasts: A study using content analysis. *The Accounting Review*, *84*(5), 1639–1670.

Kothari, S. P., Shu, S., & Wysocki, P. D. (2009). Do managers withhold bad news? *Journal of Accounting Research*, *47*(1), 241–276.

Kreps, D. M., & Wilson, R. (1982). Sequential equilibria. *Econometrica: Journal of the Econometric Society*, 863–894.

Kryzanowski, L., & Zhang, Y. (2013). Financial restatements and Sarbanes-Oxley: Impact on Canadian firm governance and management turnover. *Journal of Corporate Finance*, *21*, 87–105.

Leuven, E., & Sianesi, B. (2018). Psmatch2: Stata module to perform full mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing.

Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics*, 380–391.

Mock, T. J., Wright, A., Washington, M., & Krishnamoorthy, G. (1990). Auditor's probabilistic reasoning in a multi-stage risk assessment task. *School of Business Discussion Papers*, 1.

Mohamed, A.-A. M., Qureshi, M. A., & Behnezhad, A. R. (2005). Reliability evaluation and design of aics: A survey of models and experiments. *Review of Accounting and Finance*, *4*(2), 59–85.

Nagar, V., Nanda, D., & Wysocki, P. (2003). Discretionary disclosure and stock-based incentives. *Journal of Accounting and Economics*, *34*(1-3), 283–309.

Nagy, A. L. (2010). Section 404 compliance and financial reporting quality. *Accounting Horizons*, *24*(3), 441–454.

Pae, S. (2005). Selective disclosures in the presence of uncertainty about information endowment. *Journal of Accounting and Economics*, *39*(3), 383–409.

Palmrose, Z.-V., Richardson, V. J., & Scholz, S. (2004). Determinants of market reactions to restatement announcements. *Journal of accounting and economics*, *37*(1), 59–89.

PCAOB. (2004). An audit of internal control over financial reporting performed in conjunction with an audit of financial statements, Auditing Standard 5.

Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *The Review of Financial Studies*, *22*(1), 435–480.

Plumlee, M., & Yohn, T. L. (2010). An analysis of the underlying causes attributed to restatements. *Accounting Horizons*, *24*(1), 41–64.

Rice, S. C., & Weber, D. P. (2012). How effective is internal control reporting under SOX 404? Determinants of the (non-) disclosure of existing material weaknesses. *Journal of Accounting Research*, *50*(3), 811–843.

Rice, S. C., Weber, D. P., & Wu, B. (2014). Does SOX 404 have teeth? Consequences of the failure to report existing internal control weaknesses. *The Accounting Review*, *90*(3), 1169–1200.

Scholz, S. (2014). Financial restatement trends in the United States: 2003–2012. *Center for Audit Quality*.

SEC. (2004). Additional form 8-K disclosure requirements and acceleration of filing date.

SEC. (2005). Revisions to accelerated filer definition and accelerated deadlines for filing periodic reports. *Washington, DC*.

SOX. (2002). *The Sarbanes-Oxley Act of 2002. Public Law 107-204.* GPO Washington, DC.

Srinidhi, B. N., & Vasarhelyi, M. (1989). Adaptation and use of reliability concepts in internal control evaluation. *Advances in Accounting: Supplement I (July)*, 141–158.

Stratton, W. O. (1994). Reliability theory applied to accounting internal control: a field study. *IIE transactions*, *26*(6), 44–55.

Trockel, J. (2013). Changing bonuses and the resulting effects of employees incentives to an inspection game. *Journal of Business Economics*, *83*(7), 759–783.

Turner, L. E., & Weirich, T. R. (2006). A closer look at financial statement restatements. *The CPA Journal*, *76*(12), 12.

Verrecchia, R. E. (2001). Essays on disclosure. *Journal of accounting and economics*, *32*(1-3), 97–180.

Wans, N., He, L., & Sarath, B. (2014). Material weakness disclosures and restatement announcements. *2015 Canadian Academic Accounting Association (CAAA) Annual Conference. Available at SSRN: https://ssrn.com/abstract=2538934*.

Whitehouse, T. (2010a). Restatements, weaknesses drop again in 2009. *Compliance Week*.

Whitehouse, T. (2010b). SEC curious about drop in material weaknesses. *Compliance Week*, *7*(73), 62–63.

Wintoki, M. B., Linck, J. S., & Netter, J. M. (2012). Endogeneity and the dynamics of internal corporate governance. *Journal of Financial Economics*, *105*(3), 581–606.

Wolfe, C. J., Mauldin, E. G., & Diaz, M. C. (2009). Concede or deny: Do management persuasion tactics affect auditor evaluation of internal control deviations? *The Accounting Review*, *84*(6), 2013–2037.

Yermack, D. (1997). Good timing: CEO stock option awards and company news announcements. *The Journal of Finance*, *52*(2), 449–476.

Appendix A. Figures

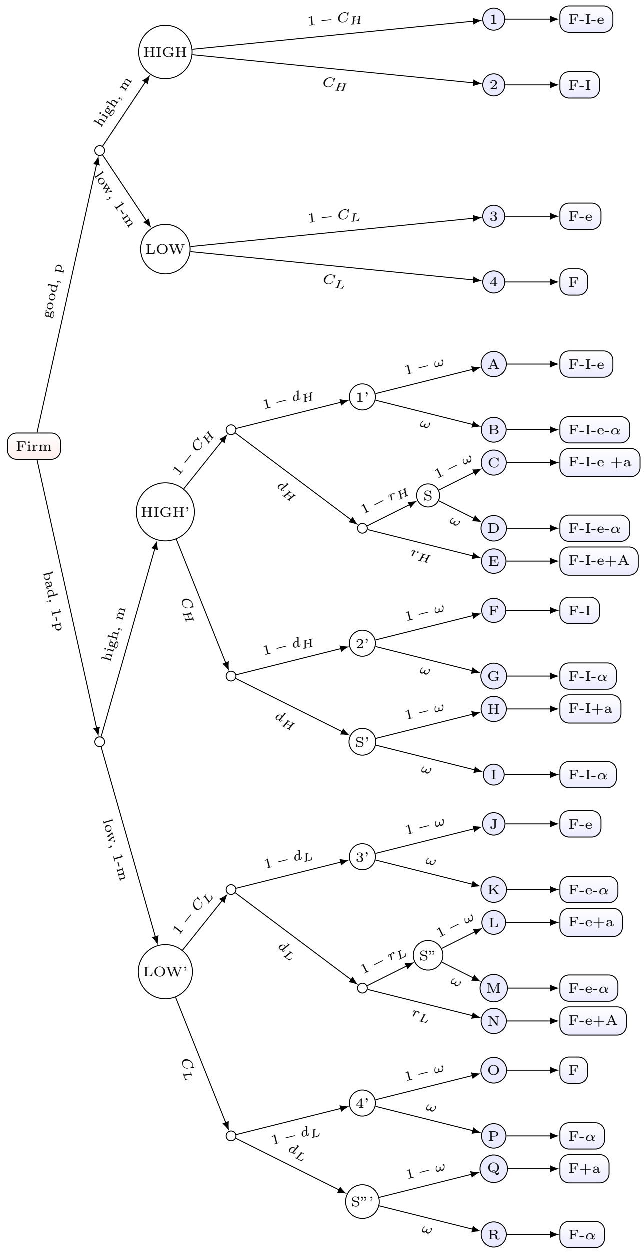


Figure A1. One Period Game Tree with Managerial Payoffs

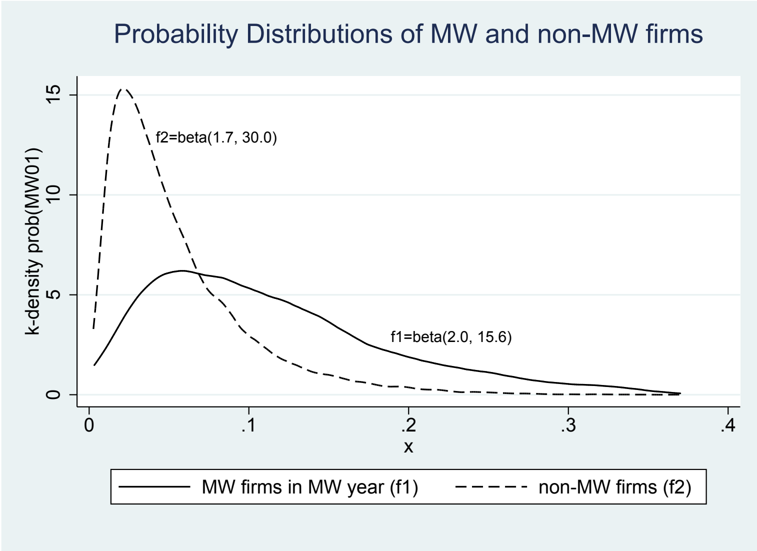


Figure A2. Probability of MW based on financial predictors: MW vs. non-MW firms

f_1 corresponds to a MW firm, f_2 corresponds to non-MW firm. Probability of having a MW is predicted using a panel probit regression with $marketcap$, $loss$, $big4$, $Audit\ Rotation$, $Audit\ Fees$ and $(Year-2004)$ as covariates. f_1 (f_2) is obtained by plotting kernel density of the actual data of MW firms in the year of MW (non-MW firms in all years). Both pdfs were also modeled as two-parameter beta distributions $B(\alpha, \beta)$ (estimates of α and β displayed) to a distribution of predicted probabilities on [0,1].

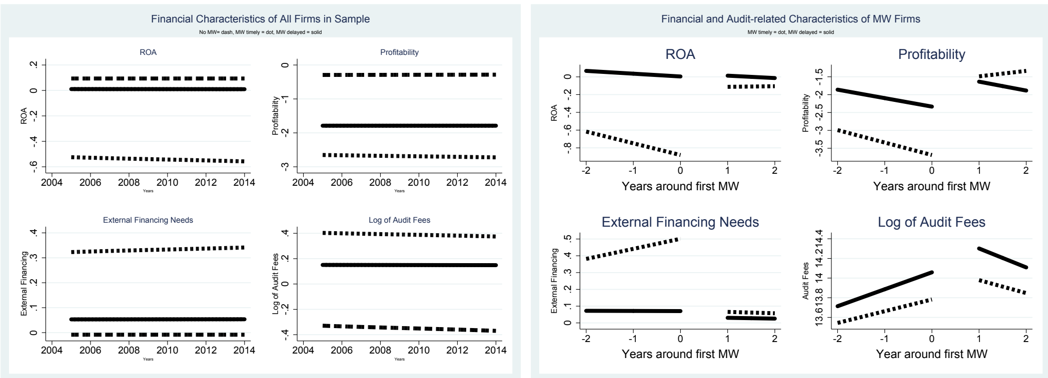


Figure A3. Variance in financial characteristics between control (non MW), Delaying and Timely firms

Illustration of various characteristics between non-MW, MW-delaying and timely disclosing firms. The left graph shows all firms. The right one shows delaying and timely firms around the year of first MW. Linear fits are calculated for pre-MW (Y-2, Y-1 and Y-0) and post-MW periods (Y1 and Y2). Y0 is the year of MW disclosure. Non-MW = dashed, Timely = dotted, Delayed = solid.

Appendix B. Definitions and Proofs

Definition B.1. A sequential equilibrium of a **one-period** disclosure game, in which some rational managers choose a disclosure strategy, consists of:

- $\sigma \in [g, b]$ indicates the state of the firm with respective prior probabilities of $p(g)$ and $1 - p(g)$
- the manager's message space is characterized by $\tilde{\sigma} \in [\tilde{g}, \tilde{b}]$
- $c_m = [c_H, c_L] \in [0, 1]$ is the strategy of the manager. The subscript denotes whether the manager decides to disclose the information in a timely manner (H) or delay the information (L); m refers to the manager's type
- $c_m \in argmaxE(\mathcal{Y}_M(m, c_m, \sigma))$, where $\mathcal{Y}_M(\cdot) = f(\mathcal{F}, I, e, A, a, \alpha)$ is the manager's payoff function at the beginning of the period, for each state, as defined in Equation 6
- c_m is sequentially rational in the sense of Kreps and Wilson (1982)
- the manager's and the owner's beliefs about the true state within the equilibrium are determined by the Bayes' rule

Proof. **Theorem 4.1** The manager provides decision-relevant information to the owner when the expected payoff with manager-provided information, that is:

$$p(\tilde{g})(\rho(g|\tilde{g})(V_g - F) + (1 - \rho(g|\tilde{g}))(V_b - F)) + (1 - p(\tilde{g}))(Z - F)$$

is strictly greater than the expected payoff without manager's information, that is:

$$p(g)V_g + (1 - p(g))V_b$$

□

Proof. **Theorem 4.2** The payoff of type H manager if he delays is:

$$A_{HD} = F - I + a(1 - p(g))(1 - \omega)d_H - \alpha(1 - p(g))\omega$$

The payoff of type H manager if he discloses timely is:

$$A_{HT} = F - I - e + a(1-p(g))(1-\omega)(d_H - k_H) + A(1-p(g))k_H - \alpha(1-p(g))\omega(1-k_H)$$

The respective payoffs of type L manager are:

$$A_{LD} = F + a(1-p(g))(1-\omega)d_L - \alpha(1-p(g))\omega$$

$$A_{LT} = F - e + a(1-p(g))(1-\omega)(d_L - k_L) + A(1-p(g))k_L - \alpha(1-p(g))\omega(1-k_L)$$

Conditions for PC equilibrium: Setting $A_{HT} > A_{HD}$ and $A_{LT} < A_{LD}$ yields Cond. 12. Since in this equilibrium $c_H = 1$ and $c_L = 0$, $\nu = 1 - mk_H < 1$, Cond. 10 holds. **Conditions for FC equilibrium**: Setting $\mathrm{A_{LT}} > \mathrm{A_{LD}}$ to induce Type L to disclose timely, and noting that. since $d_H > d_L$ and $r_H > r_L$, whenever Type L is induced to disclose timely, type H's incentive condition will also be satisfied, yields Cond. 11. Since within this equilibrium $c_H = c_L = 1$, $\nu = m(1-k_H) + (1-m)(1-k_L) = 1 - k_L - m(k_H - k_L) < 1$, Cond. 10 holds. To check the participation of both types, it must be true that whenever the compensation is set to cover operating costs of type H manager, type L must find it profitable to participate and not deviate. Therefore, it must be true that $A_{LT} > A_{HT}$, which, simplified, yields Cond. 13. **Conditions for NC equilibrium**: In NC equilibrium $c_H = c_L = 1$, so $\nu = 1$, $\rho(g|\tilde{g}) = p(g)$, thus, Cond. 10 does not hold, since the owner gains nothing by employing the manager. □

Appendix C. Tables of Results

Table C1. Sample Selection

	Attrition	Remaining Observations
Firm-years in Audit Analytics for years 2005-2014		118,363
Less duplicate cik-year observations	(42,841)	75,522
Less missing fiscal year	(180)	75,342
Less missing definitive ICFR opinion	(809)	74,533
Less missing SIC code	(232)	74,301
Less missing Compustat data for control variables	(20,565)	53,736
Less missing consecutive years	(22)	53,714
Total Sample		**53,714**

Graphical illustration of sample composition by industry using two-digit SIC code, year or size (using *marketcap*). All continuous variables are winsorized at the 0.01 level.

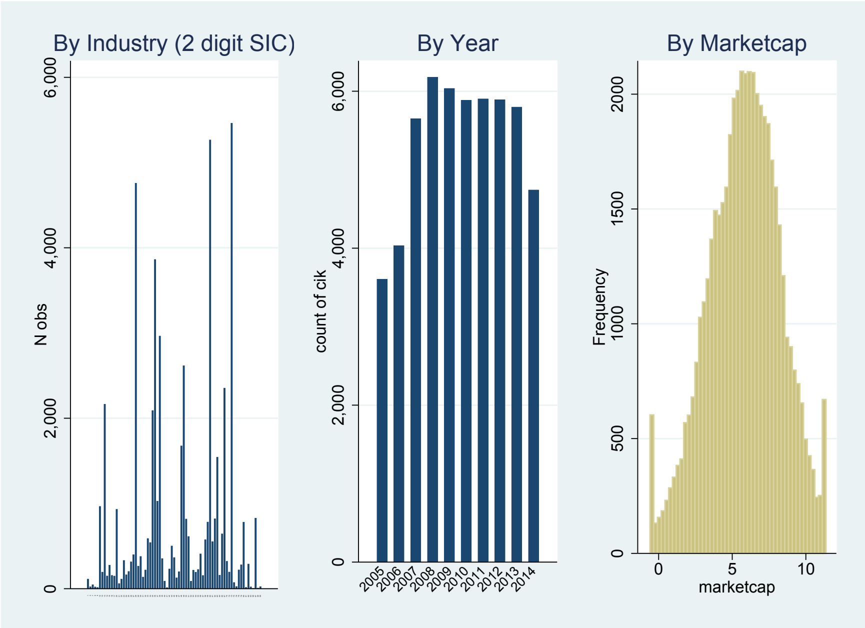


Table C2. Definitions of Variables

Variable Name	Description	Compustat Items
Financial Variables		
ROA	Earnings before interest, taxes and depreciation scaled by total assets	$ebitda/at$
$Profitability$	Net income (loss) scaled by Sales	$ni/sale$
TQ	Tobin's Q	$((csho * prcc\_f) + (dltt + dlc))/at$
$Aztile$	Decile of Altman (1968) Z-score	-
Audit-Related Variables		
$Audit\ Fees$	Logarithm of (audit fees scaled by total assets)	-
$Audit\ Rotation$	An indicator variable equal to one if the auditor changed in a prior year	-
$Core$	An indicator variable equal to 1 if MW or FSR disclosures are related to Revenue or Tax accounts	-
$Fraud$	An indicator variable equal to 1 if MW or FSR disclosures are related to Fraud	-
Dummy Variables		
D_i	An indicator variable equal to 1 is the firm is disclosing a MW in a "delayed" pattern	-
T_i	An indicator variable equal to 1 is the firm is disclosing a MW in a "timely" pattern	-
$Post_{it}$	An indicator variable equal to 1 is the firm-year observation corresponds to a "post-MW" period	-
Controls		
$marketcap$	A measure of firm's size, the logarithm of market capitalization	$log(prcc\_f * csho)$
$salegrowth$	A pressure associated with extreme sales growth, year over year change in Sales	$(sale_t - sale_{t-1})/sale_{t-1}$
$big4$	An indicator variable equal to 1 if a SOX 404 opinion supplied by Big4 auditors (DT, E&Y, KPMG, PWC)	au
$foreign$	An indicator variable equal to 1 is the firm pays foreign income taxes	$txfo > 0$
$tangibility$	Property, Plant and Equipment over total assets	$ppegt/at$
$leverage$	Total debt over total assets	dt/at
$bookmarket$	Total assets minus total liabilities scaled by market value	$(at - lt)/mkvalt$
$dividendpayer$	An indicator variable equal to one if a firm issues dividends	$dv > 0$
$finslack$	Cash over total assets	ch/at
age	Age of firm in years	using $ipodate$

Table C3. Summary Statistics

Covariate	N of obs	Mean	Std. Dev.	Min	Q1	Median	Q3	Max
Dependent Variables								
ROA	51,217	-.24	1.6	-14	.0053	.075	.14	.46
$Tobin's\ Q$	50,148	4.3	18	0.094	0.68	1.1	2	153
$Profitability$	50,405	-1.5	9.3	-80	-.053	.043	.11	.91
$Aztile$	36,816	5.5	2.9	1	3	5.5	8	10
$Audit\ Fees$	36.150	6.8	1.3	3.6	5.8	6.9	7.8	9.8
$Audit\ Rotation$	53,714	.32	.47	0	0	0	1	1
$Core$	53,714	.042	.2	0	0	0	0	1
$Fraud$	53,714	.0072	.085	0	0	0	0	1
Key Ind. Variables								
D	53,714	.018	.13	0	0	0	0	1
T	53,714	.53	.5	0	0	1	1	1
$Post$	53,714	.07	.26	0	0	0	0	1
$T \cdot Post$	53,714	.062	.24	0	0	0	0	1
Controls								
$marketcap$	50,537	5.8	2.5	-.59	4.2	5.9	7.5	11
$salegrowth$	45,475	1.2	5.1	-1	-.17	.1	.69	37
$big4$	53,714	.57	.5	0	0	1	1	1
$foreign$	53,714	.6	.49	0	0	1	1	1
$tangibility$	45,308	.49	.45	0	.13	.35	.78	2.1
$leverage$	53,714	.2	.31	0	0	.066	.3	2
$bookmarket$	53,714	.49	1.1	-5.2	.027	.39	.78	5.1
$dividendpayer$	53,714	.46	.5	0	0	0	1	1
$finslack$	52,107	.15	.19	0.00	.02	.07	.19	.95
age	26,815	12	6.6	0.74	6.6	12	16	29

Summary statistics are reported for all firms in the sample. All variables are constructed using data from Audit Analytics and Compustat. All continuous variables are winsorized at the 0.01 level. The number of observations varies due to data availability. Columns from left to right report the number of observations, mean, standard deviation, minimum value, first, second (median), and third quartile, and maximum value of the respective covariate.

Table C4. Pairwise Correlation Coefficients

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1.*ROA*	1.000																
2.*Tobin's Q*	-0.745*	1.000															
3.*Profitability*	0.432*	-0.311*	1.000														
4.*Aztile*	0.302*	-0.116*	0.165*	1.000													
5.*Audit Fees*	-0.198*	0.202*	-0.146*	0.066*	1.000												
6.*Audit Rotation*	-0.013*	0.005	-0.009*	-0.040*	-0.010*	1.000											
7.*core*	0.008	-0.015*	0.010*	-0.035*	0.119*	0.023*	1.000										
8.*fraud*	-0.007	0.003	-0.009*	-0.021*	0.038*	0.000	0.091*	1.000									
9.*marketcap*	0.313*	-0.181*	0.178*	0.371*	-0.524*	-0.127*	-0.033*	-0.014*	1.000								
10.*salegrowth*	-0.012*	0.012*	0.011*	-0.030*	0.035*	0.072*	0.008	0.016*	-0.010*	1.000							
11.*big4*	0.228*	-0.178*	0.142*	0.199*	-0.057*	-0.181*	-0.005	-0.009*	0.682*	-0.018*	1.000						
12.*foreign*	0.228*	-0.190*	0.181*	0.163*	-0.196*	0.010*	-0.002	-0.003	0.362*	-0.045*	0.225*	1.000					
13.*tangibility*	0.013*	-0.030*	0.046*	-0.188*	-0.216*	-0.012*	-0.021*	-0.012*	0.026*	-0.045*	0.021*	-0.007	1.000				
14.*leverage*	-0.161*	0.102*	-0.080*	-0.382*	-0.024*	-0.043*	0.010*	0.015*	-0.036*	0.034*	0.028*	-0.170*	0.167*	1.000			
15.*bookmarket*	0.235*	-0.191*	0.115*	0.108*	-0.084*	0.037*	-0.004	-0.002	-0.033*	-0.016*	-0.025*	0.104*	-0.025*	-0.277*	1.000		
16.*dividendpayer*	0.176*	-0.137*	0.147*	0.148*	-0.473*	-0.011*	-0.056*	-0.026*	0.403*	-0.081*	0.209*	0.236*	0.157*	-0.054*	0.053*	1.000	
17.*finslack*	-0.241*	0.233*	-0.218*	0.053*	0.479*	0.024*	0.001	-0.002	-0.208*	0.050*	-0.144*	-0.191*	-0.285*	-0.113*	-0.115*	-0.297*	1.000
18.*age*	-0.017*	0.023*	0.008	-0.003	-0.114*	-0.004	-0.020*	-0.004	0.046*	0.003	-0.017*	0.042*	0.106*	0.022*	-0.011	0.076*	-0.055*

The table reports the pairwise correlation coefficients of all variables used in the analysis. The correlation coefficients that are statistically significant at 5% are noted with a *. All continuous variables are winsorized at the 0.01 level.

Table C5. Univariate Tests

Panel A. Comparison of Control Sample (non-MW) and MW Sample

	Control	MW	Δ	p-value
ROA	0.09	-0.52	0.62	0.00
TQ	1.72	6.63	-4.92	0.00
Profitability	-0.29	-2.66	2.37	0.00
AZtil	6.20	4.92	1.27	0.00
Audit Fees	6.57	7.29	-0.72	0.00
Audit Tenure	0.25	0.37	-0.12	0.00
Core	0.01	0.07	-0.05	0.00
Fraud	0.00	0.01	-0.01	0.00
marketcap	7.33	4.48	2.84	0.00
salegrowth	1.00	1.30	-0.30	0.00
big4	0.87	0.32	0.55	0.00
foreign	0.70	0.51	0.18	0.00
tangibility	0.51	0.48	0.02	0.00
leverage	0.19	0.20	-0.01	0.00
bookmarket	0.48	0.51	-0.03	0.00
dividendpayer	0.61	0.34	0.27	0.00
finslack	0.11	0.18	-0.06	0.00

Panel B. Comparison of Timely and Delaying firms

	MW delayed	MW timely	Δ	p-value
ROA	0.01	-0.54	0.55	0.00
TQ	1.57	6.81	-5.24	0.00
Profitability	-1.79	-2.69	0.90	0.02
AZtile	5.58	4.90	0.68	0.00
Audit Fees	7.06	7.30	-0.24	0.00
Audit Rotation	0.29	0.37	-0.08	0.00
Core	0.12	0.06	0.05	0.00
Fraud	0.02	0.01	0.01	0.00
marketcap	6.02	4.43	1.60	0.00
salegrowth	0.89	1.31	-0.42	0.02
big4	0.69	0.31	0.38	0.00
foreign	0.63	0.51	0.12	0.00
tangibility	0.47	0.48	-0.01	0.59
leverage	0.22	0.20	0.02	0.11
bookmarket	0.71	0.50	0.21	0.00
dividendpayer	0.37	0.33	0.04	0.01
finslack	0.12	0.18	-0.06	0.00

The table reports the differences in means of financial characteristics between various sample groups using all observations. Two-group mean comparison tests examine whether significant differences exist between the means of the two groups. All continuous variables are winsorized at the 0.01 level.

Table C6. Regression Results - Static Panel Regression

	(1) ROA	(2) TQ	(3) Profitability	(4) aztile	(5) Audit Fees	(6) Audit Rotation	(7) Core	(8) Fraud
D	-0.00	0.69	-1.75*	0.26	0.08	0.00	0.17***	0.04***
	(-0.12)	(1.54)	(-1.67)	(1.24)	(1.10)	(0.13)	(7.11)	(3.37)
T	-0.40***	7.04***	-2.06***	0.60***	0.30***	0.04***	0.08***	0.01***
	(-11.42)	(14.64)	(-7.58)	(8.05)	(12.28)	(4.06)	(20.43)	(7.83)
$Post$	0.05**	-0.22	0.79	-0.14	0.20***	0.06**	-0.10***	-0.03*
	(2.16)	(-0.94)	(1.10)	(-0.78)	(4.22)	(2.24)	(-3.15)	(-1.75)
$T \cdot Post$	0.01	-0.30	-0.51	-0.08	-0.15***	-0.01	0.12***	0.03**
	(0.43)	(-1.10)	(-0.66)	(-0.40)	(-2.87)	(-0.44)	(3.84)	(1.99)
$marketcap$	0.07***	1.22***	0.18**	0.75***	-0.28***	-0.00	-0.00	0.00
	(6.13)	(8.23)	(2.46)	(46.05)	(-27.70)	(-0.94)	(-0.01)	(1.10)
$salegrowth$	0.01***	-0.05***	0.19***	-0.00	-0.00***	0.01***	0.00	0.00**
	(5.58)	(-3.59)	(10.76)	(-0.19)	(-4.54)	(12.95)	(0.67)	(1.98)
$big4$	0.05***	-1.45***	0.63***	-0.47***	0.15***	-0.11***	0.03***	-0.00
	(3.56)	(-10.24)	(2.69)	(-8.70)	(6.09)	(-12.12)	(5.45)	(-0.05)
$foreign$	0.09***	-0.92***	0.84***	-0.15***	-0.04***	0.01**	0.01***	0.00*
	(7.15)	(-9.21)	(5.53)	(-3.85)	(-3.18)	(2.52)	(2.86)	(1.71)
$tangibility$	-0.47***	4.93***	0.45	-0.48***	0.28***	0.03***	-0.01***	-0.00*
	(-6.09)	(5.66)	(1.15)	(-6.75)	(9.21)	(3.32)	(-2.77)	(-1.65)
$leverage$	-0.52***	3.37***	-1.92***	-2.11***	-0.35***	-0.02**	-0.01	0.00
	(-5.20)	(3.21)	(-3.14)	(-21.17)	(-9.19)	(-2.37)	(-1.16)	(1.26)
$bookmarket$	0.04***	0.03	0.08	-0.21***	-0.10***	0.01**	-0.00	0.00
	(2.94)	(0.31)	(1.08)	(-13.14)	(-8.45)	(2.53)	(-0.43)	(0.53)
$dividendpayer$	0.02	-0.49***	0.33***	0.07**	-0.07***	0.00	-0.00	-0.00
	(1.60)	(-4.84)	(2.72)	(2.05)	(-5.19)	(0.44)	(-0.95)	(-1.38)
$finslack$	-0.30***	5.69***	-3.64***	1.57***	0.54***	0.04**	-0.03***	-0.01**
	(-2.59)	(5.05)	(-4.56)	(11.57)	(14.05)	(2.43)	(-3.78)	(-2.19)
$Constant$	0.18	-9.30***	-5.40	2.08***	9.10***	0.27***	0.02	0.02
	(1.37)	(-3.12)	(-1.52)	(5.34)	(67.69)	(4.04)	(0.93)	(1.17)
N	36,265	36,269	35,906	30,873	26,973	36,425	36,425	36,425
R^2	0.12	0.03	0.10	0.33	0.66	0.05	0.04	0.01

t statistics in parentheses, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The table reports the results using static panel regression specified in Equation 17. All regressions performed using dummy variables for industry 2-digit SIC code and fiscal years. Results reported using cluster-robust standard errors at firm level. All continuous variables are winsorized at the 0.01 level. The number of observations in regressions varies due to data availability.

Table C7. Regression Results - Dynamic Panel Regression

	(1) ROA	(2) TQ	(3) Profitability	(4) Aztile	(5) Audit Fees	(6) Audit Rotation	(7) Core	(8) Fraud
$L.Y$	0.36***	0.37***	0.47***	0.46***	0.64***	0.09***	0.24***	0.02
	(7.11)	(5.79)	(11.63)	(35.44)	(49.75)	(14.80)	(16.97)	(0.81)
$L2.Y$	0.11**	0.14***	0.03	0.11***	0.15***	0.08***	0.04***	-0.03***
	(2.51)	(2.61)	(1.20)	(9.21)	(13.62)	(12.44)	(4.20)	(-3.45)
D	0.01	0.05	-1.11	-0.02	0.03	-0.03	0.14***	0.01
	(0.25)	(0.16)	(-1.61)	(-0.20)	(0.71)	(-0.87)	(5.83)	(1.31)
T	-0.08***	1.43***	-0.35**	0.08*	0.01	0.04***	0.04***	0.01***
	(-4.50)	(7.18)	(-2.46)	(1.88)	(1.52)	(3.99)	(11.51)	(5.97)
$Post$	0.04**	-0.20	0.66	0.07	0.01	0.06*	-0.11***	0.00
	(2.10)	(-1.13)	(1.50)	(0.62)	(0.17)	(1.70)	(-3.72)	(0.01)
$T \cdot Post$	0.03	-0.22	-0.54	-0.07	0.00	-0.01	0.12***	0.01
	(1.12)	(-1.12)	(-1.07)	(-0.58)	(0.02)	(-0.31)	(4.03)	(0.50)
$marketcap$	0.05***	0.39***	0.08	0.34***	-0.12***	0.00	-0.00*	0.00
	(5.81)	(4.62)	(1.56)	(28.62)	(-25.21)	(0.66)	(-1.87)	(0.16)
$salegrowth$	0.00***	-0.02	0.08***	0.00	-0.00	0.01***	0.00	0.00
	(3.17)	(-1.15)	(6.06)	(1.04)	(-0.48)	(11.59)	(0.53)	(0.78)
$big4$	0.01	-1.04***	0.36**	-0.45***	0.11***	-0.12***	0.02***	-0.00
	(0.60)	(-6.70)	(2.17)	(-10.38)	(9.30)	(-12.55)	(4.95)	(-0.83)
$foreign$	0.06***	-0.48***	0.57***	-0.10***	0.02***	0.02***	0.01**	0.00
	(5.98)	(-6.36)	(4.57)	(-2.87)	(2.88)	(3.10)	(2.33)	(1.34)
$tangibility$	-0.37***	3.77***	-0.01	-0.07	0.10***	0.02***	-0.01**	-0.00**
	(-5.57)	(5.07)	(-0.05)	(-1.47)	(8.45)	(2.66)	(-2.19)	(-2.46)
$leverage$	-0.33***	2.68***	-0.87*	-1.43***	-0.08***	-0.04***	-0.00	0.00
	(-3.12)	(2.59)	(-1.79)	(-19.38)	(-3.94)	(-3.24)	(-0.80)	(0.90)
$bookmarket$	0.03**	-0.19*	0.14*	-0.15***	-0.03***	0.00	0.00	0.00
	(2.57)	(-1.72)	(1.77)	(-11.29)	(-4.67)	(1.54)	(0.26)	(0.58)
$dividendpayer$	0.02	-0.24***	0.23**	-0.00	0.02***	0.00	0.00	-0.00
	(1.56)	(-2.58)	(2.19)	(-0.08)	(2.90)	(0.61)	(0.64)	(-0.78)
$finslack$	-0.11	2.01**	-2.59***	1.14***	0.15***	0.03	-0.02*	-0.01***
	(-1.01)	(1.97)	(-4.21)	(9.76)	(5.38)	(1.41)	(-1.94)	(-2.93)
$Constant$	0.01	-3.82***	-0.11	1.22***	2.10***	0.24***	-0.02	0.02
	(0.07)	(-3.43)	(-0.13)	(3.11)	(21.80)	(3.48)	(-1.01)	(1.05)
Ind./Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	24,999	24,826	24,694	20,545	18,508	25,314	25,314	25,314
R^2	0.57	0.56	0.52	0.73	0.94	0.12	0.10	0.01

t statistics in parentheses, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The table reports the results using dynamic panel regression specified in Equation 18. Two lags of dependent variables are included in all regressions. All regressions performed using dummy variables for industry 2-digit SIC code and fiscal years. Results reported using cluster-robust standard errors at firm level. All continuous variables are winsorized at the 0.01 level. The number of observations in regressions varies due to data availability.

Table C8. Robustness Tests - Panel Regressions with a Subsample

	(1) ROA	(2) TQ	(3) Profitability	(4) AZtile	(5) Audit Fees	(6) Audit Rotation	(7) Core	(8) Fraud
T	-0.35***	6.49***	-0.41	0.54***	0.15*	0.03	-0.04**	-0.01*
	(-5.90)	(8.69)	(-0.42)	(2.58)	(1.96)	(1.11)	(-2.31)	(-1.95)
$T \cdot Post$	0.06**	-0.69***	0.38	-0.25***	0.11***	0.05***	0.05***	0.01***
	(2.49)	(-3.75)	(1.24)	(-3.47)	(4.57)	(4.17)	(6.12)	(2.62)
$marketcap$	0.09***	1.42***	0.15	0.79***	-0.27***	0.00	-0.00	0.00
	(5.97)	(7.01)	(1.49)	(38.55)	(-17.20)	(0.59)	(-1.11)	(0.87)
$salegrowth$	0.01***	-0.08***	0.23***	0.00	-0.00	0.01***	0.00	0.00**
	(4.95)	(-3.33)	(10.35)	(0.56)	(-1.28)	(8.09)	(0.59)	(1.98)
$big4$	0.04**	-1.84***	0.68**	-0.56***	0.12***	-0.08***	0.03***	-0.00
	(2.04)	(-8.84)	(2.31)	(-8.51)	(3.59)	(-7.32)	(3.74)	(-0.89)
$foreign$	0.15***	-1.59***	1.11***	-0.11**	-0.07***	0.02**	0.01**	0.00
	(7.34)	(-9.08)	(4.69)	(-2.05)	(-3.45)	(2.21)	(2.43)	(1.64)
$tangibility$	-0.61***	6.14***	0.41	-0.56***	0.28***	0.03***	-0.02***	-0.00*
	(-6.01)	(5.34)	(0.76)	(-7.00)	(5.89)	(2.79)	(-2.73)	(-1.79)
$leverage$	-0.60***	3.96***	-2.27***	-1.51***	-0.29***	-0.02*	-0.01*	0.00
	(-4.94)	(3.12)	(-3.08)	(-16.55)	(-5.29)	(-1.81)	(-1.71)	(1.30)
$bookmarket$	0.05***	0.04	0.11	-0.15***	-0.09***	0.01**	-0.00	0.00
	(3.22)	(0.36)	(1.26)	(-8.77)	(-6.03)	(2.39)	(-0.77)	(0.51)
$dividendpayer$	0.04	-0.74***	0.53***	0.17***	-0.08***	0.01	-0.01	-0.00
	(1.36)	(-4.23)	(2.63)	(3.20)	(-3.90)	(0.68)	(-1.17)	(-1.40)
$finslack$	-0.45***	7.21***	-4.98***	1.36***	0.48***	0.05**	-0.04***	-0.01***
	(-2.92)	(4.67)	(-4.64)	(8.62)	(8.04)	(2.41)	(-3.84)	(-2.62)
$Constant$	0.12	-10.36***	-7.73*	1.73***	9.07***	0.25***	0.24***	0.06***
	(0.73)	(-2.75)	(-1.71)	(3.41)	(58.50)	(3.25)	(6.40)	(3.34)
Ind./Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	18,482	18,456	18,067	15,913	9,587	18,549	18,549	18,549
R^2	0.12	0.03	0.10	0.30	0.59	0.03	0.04	0.02

t statistics in parentheses, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The table reports the results using a subsample of observations, including "delayed" and "timely" firms only, excluding control (non-MW) firms. All regressions performed using dummy variables for industry 2-digit SIC code and fiscal years. Results reported using cluster-robust standard errors at firm level. Variables D, $Post$ and $D \cdot Post$ are omitted due to collinearity. All continuous variables are winsorized at the 0.01 level. The number of observations in regressions varies due to data availability.

Table C9. Robustness Tests - Propensity Score Matching

	(1) ROA	(2) TQ	(3) Profitability	(4) Aztile	(5) Audit Fees	(6) Audit Rotation	(7) Core	(8) Fraud
D	0.02	0.43*	-0.09	0.65**	-0.19**	-0.06	0.09**	0.01
	(0.69)	(1.76)	(-0.13)	(2.05)	(-1.98)	(-1.18)	(2.00)	(0.26)
$D \cdot Post$	-0.00	-0.23	1.29	-0.47	0.22**	-0.00	-0.12**	0.00
	(-0.03)	(-1.21)	(1.39)	(-1.59)	(2.25)	(-0.02)	(-2.34)	(0.07)
$marketcap$	0.05***	0.23***	0.26	0.66***	-0.44***	0.02	-0.01	-0.01
	(5.25)	(5.39)	(1.47)	(8.74)	(-13.12)	(1.56)	(-0.75)	(-1.45)
$salegrowth$	-0.00	0.01	0.16*	0.01	0.00	0.01***	-0.00	-0.00
	(-0.45)	(0.77)	(1.76)	(0.83)	(0.30)	(2.79)	(-0.73)	(-0.19)
$big4$	-0.06	-0.36*	-0.31	-0.62*	-0.02	-0.17***	0.00	0.01
	(-1.49)	(-1.91)	(-0.38)	(-1.88)	(-0.20)	(-3.56)	(0.11)	(0.31)
$foreign$	0.07***	-0.33**	1.33**	-0.05	-0.02	-0.09*	0.08**	0.04**
	(2.80)	(-2.35)	(2.04)	(-0.26)	(-0.28)	(-1.77)	(2.49)	(2.47)
$tangibility$	-0.02	0.33	-0.76	0.31	-0.01	0.07	-0.09*	0.00
	(-0.54)	(1.21)	(-1.02)	(1.14)	(-0.13)	(1.04)	(-1.95)	(0.12)
$leverage$	-0.22	0.81	0.39	-3.75***	-0.58***	-0.00	0.05	0.00
	(-1.58)	(1.60)	(0.36)	(-6.08)	(-3.29)	(-0.02)	(1.24)	(0.04)
$bookmarket$	-0.01	-0.16***	0.07	-0.23***	-0.11**	0.02*	0.01	0.00
	(-0.65)	(-3.67)	(0.56)	(-2.76)	(-2.59)	(1.67)	(0.56)	(0.24)
$dividendpayer$	0.05**	-0.37***	0.51	0.01	0.15*	0.02	-0.04	0.01
	(2.46)	(-3.23)	(1.34)	(0.03)	(1.76)	(0.53)	(-1.24)	(0.53)
$finslack$	-0.55***	2.77***	-10.05***	0.67	1.18***	0.07	-0.08	0.09
	(-3.58)	(3.15)	(-2.78)	(0.83)	(4.60)	(0.58)	(-0.84)	(1.17)
$Constant$	-0.08	0.54	-1.66	0.41	10.62***	0.99***	0.17	0.08
	(-0.96)	(0.92)	(-0.97)	(0.63)	(28.39)	(7.85)	(1.62)	(1.38)
Ind./Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	709	732	725	684	653	725	725	725
R^2	0.39	0.38	0.18	0.55	0.76	0.21	0.17	0.13

t statistics in parentheses, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The table reports the results using a matched sample of delayed and timely firms, matched on size ($logat$), age and one-digit industry code. OLS regression performed using a matched subsample using dummy variables for industry 2-digit SIC code and fiscal years. Variables T, $Post$ and $T \cdot Post$ are omitted due to collinearity. All continuous variables are winsorized at the 0.01 level. The number of observations in regressions varies due to data availability.

Is CEO Communication Fraud-Driven? Fraud Exposure as the Determinant of Language Choice in Codes of Conduct

Olga Bogachek[1†] Miles Gietzmann[1] Francesco Grossetti[1]

Abstract

When firms publish Codes of Conduct, they are either cognizant of the specific risks of fraud that employees may face or are addressing a mandatory disclosure requirement with minimal thought. If the former is true, for firms with subsidiaries in countries where business corruption is endemic, their Codes of Conduct should address the potential issues faced by employees operating in corrupt locations. This paper uses the Latent Dirichlet Allocation model on the Codes of Conduct of S&P 500 firms. We construct a quantitative index which measures the CEO's anti-fraud emphasis by automatically detecting and extracting fraud-related topics. We test and find a positive association between the CEO's anti-fraud emphasis and the risk (business corruption) exposure of the firm's foreign subsidiaries.

Keywords: fraud detection, code of conduct, LDA, business corruption

[†]: Corresponding author: `olga.bogachek@phd.unibocconi.it`
[1]: Bocconi University. Via Sarfatti 25, 20136 - Milan, Italy

1 Introduction

The adoption of the Sarbanes-Oxley Act of 2002 (SOX) mandated the firms to maintain and disclose their Codes of Ethics. Many firms adopted broader and more comprehensive documents, Codes of Conduct, which address a spectrum of topics, such as corruption, bribery, and overall unethical behavior of employees inside and outside of the firm. "*Don't be evil*", "*Integrity always in style*", "*No jerks*", are some of the mottos that firms use in their Codes.

When firms publish their Codes of Conduct, they are either cognizant of the specific risks of fraud that employees may face or are addressing this mandatory disclosure requirement with minimal thought.

If the former is true[1], they might be legitimately signaling a commitment to the anti-fraud policy when potential risks are heightened. We hypothesize that for firms with subsidiaries in countries where business corruption is endemic, their Codes of Conduct should address the potential issues faced by employees.

The purpose of this paper is to test the association between the CEO's anti-fraud language index (which we define as a measure of CEO's anti-fraud emphasis) and the firm's exposure to fraud, proxied by corruption indexes of subsidiaries' locations. We use linguistic characteristics of Codes of Conduct to estimate the prevalence of anti-fraud topics, which reflects the CEO's emphasis of the importance of anti-fraud measures within the firm. We find that CEO's anti-fraud language index is positively associated with the firm's fraud exposure, confirming our hypothesis that the firms take corruption risks seriously and reflect them in their communication with employees. At the same time, we find no association between audit fees and anti-fraud emphasis, suggesting that the language choice is not driven by the auditor.

We adopt an algorithmic approach by using *Latent Dirichlet Allocation* (LDA) model [Blei et al., 2003]. This allows us to estimate CEO's anti-fraud emphasis without directly observing it, avoiding field experiments, interviews with top level executives and generally costly and subjective surveys. We discuss the LDA approach in more detail in Section 5.

This paper is related to several streams of recent academic literature, and our contribution is multiple.

First, our paper is related to the existing literature on policy compliance and ethics. Organizations use various policies and communication from the top to address various heightened

[1]Recent examples demonstrate that the Codes of Conducts are taken seriously. Most recently, Intel's CEO resigned for violation of the non-paternalization clause of the firm's code, and Samsonite's CEO quit for falsely claiming to possess academic credentials on his resume.

risks. These policies form a key part of a firm's system of internal controls that protect firm's assets by prescribing appropriate behavior for employees [Crossler et al., 2016]. We show that the ethical tone from the top reflects the heightened business corruption risks, but not necessarily driven by the auditor procedures. Our paper may be of interest to policy setters in the area of ethical tone in internal controls and audit functions.

Second, our paper adds to the fraud-related research area, where a fundamental limitation is the partial observability (detection) of fraud. For example, the list of Accounting and Auditing Enforcement Releases (AAERs) from the Securities and Exchange Commission (SEC) website[2] only includes 11 civil lawsuits associated with the firms in our sample over the period of 2014-2017. Firms that engage in illegal business activities, but are not caught, remain unobserved, while empirical tests typically rely on samples of firms that are caught engaging in an activity [Dyck et al., 2013]. Our method provides a potential solution to this problem of "unobservability". While we cannot be sure that a firm is engaging in a fraudulent activity, our approach allows us to extract fraud-related topics to understand the amount of emphasis that a CEO puts on fraud. To obtain inherent fraud risk of the firm, we use the data on location of firm's subsidiaries. This allows us to estimate the aggregate level of unobserved fraud risk. We follow Garmaise and Liu [2005] and proxy the predisposition towards fraud with the World Bank's World Governance Institute Control of Corruption Index (WGI)[3] and Corruption Practices Index by Transparency International (CPI)[4]. We describe this in more detail in Section 5.

Third, we contribute to the literature on quantitative textual analysis in the context of accounting and economics, particularly, in the area of firm voluntary disclosures and relations between CEO's personal characteristics, beliefs and preferences and firm performance and policies. We employ a unique hand-collected data set of Codes of Conduct of S&P 500 companies, which allows us to obtain insights into documents which are typically targeted at an internal audience (i.e. employees). We assume that these documents are mainly focused on internally-important issues and exhibit a less legally scrutinized language. While there is substantial academic literature on trends in the textual characteristics of annual reports, particularly the Management's Discussion and Analysis (MD&A) section of the form 10-K, specific content and attributes of management anti-fraud communication such as a Code of Conduct have received less attention. To the best of our knowledge, this is the first paper that applies a topic modeling technique to the Codes of Conduct.

[2] `www.sec.gov/divisions/enforce/friactions.shtml`.
[3] World Bank's World Governance Index `info.worldbank.org/governance/wgi/`.
[4] Transparency International Corruption Practice Index `www.transparency.org`.

Fourth, we also contribute to vast economic literature on voluntary disclosures in the cheap talk setting [Crawford, 1998; Farrell and Rabin, 1996; Farrell, 1987; Crawford and Sobel, 1982]. We show that information content of a firm's anti-fraud communication may be used as a signaling device in the adverse selection game in the presence of heightened fraud risks.

The paper proceeds as follows. Section 2 provides a review of existing literature and background on the Codes of Conduct. Section 3 discusses research question and hypotheses. Section 4 describes the sample. Section 5 describes the research design and development of the measures used in our study. Section 6 reports the results, and Section 7 concludes.

2 Literature Review and Background

Extant policy compliance and ethics literature agrees that the ethical tone at the top of an organization is a key factor in establishing an effective internal control environment [Crossler et al., 2016; Pickerd et al., 2014]. Pickerd et al. [2014] define the tone from the top as "a culture of control consciousness, integrity, and ethical values from upper-level management and the board of directors". Cotton et al. [2016] emphasizes that the ethical culture fostered by the tone from the top will impact ethical decision making throughout the organization. Crossler et al. [2016] provides an extensive literature review on tone from the top and evidence that formal ethics policies influence employees' ethical decision making and even accounting decisions, such as financial reporting aggressiveness [Patelli and Pedrini, 2015]. Effectively, Crossler et al. [2016] offer a theoretical framework in which strong ethical emphasis of a policy results in higher policy compliance intentions. To support it, Balvers et al. [2016] find that the voluntary use by retail firms of customer satisfaction emphasis in their 10-Ks is a credible signal that is associated with higher subsequent customer satisfaction. Crossler et al. [2016] also show that a simple textual construct emphasizing ethical aspect of a policy effectively fosters policy compliance.

Many recent studies employed LDA topic modeling approach in the context of accounting and economics to examine the links between accounting measures, stock returns, firms performance and textual disclosures. Loughran and McDonald [2016] provide a detailed overview on general approaches in textual analysis and a review of the vast contemporary literature. In the area of fraud and compliance, Hoberg and Lewis [2017] use topic modeling and cosine similarity to show that fraudulent firms do not make qualitative disclosures that resemble their industry peers, but, instead, cluster with other fraudulent peers firms. Brown et al.

[2016] use LDA to find that the topics in financial statement disclosures are incrementally informative in predicting intentional misreporting. Dyer et al. [2017] find that the particular topics of fair value, internal controls and risk factor disclosures account for virtually all of the increase in 10-K disclosures from 1996 to 2013. Bandiera et al. [2017] employ LDA to the data on CEO daily activities during the work week to estimate a CEO behavioral index and to study its association with firm financial performance.

Another stream of literature explores the relations between CEO personal characteristics, beliefs and preferences derived from linguistic measures and firm performance and policies. Particularly, Larcker and Zakolyukina [2012] use quarterly conference calls narratives to develop linguistic classification models of deception. Gow et al. study the association between CEO "Big Five" personality traits with investment choice, financing choices and firm performance.

Fraud related and fraud predictive research is an important area of academic interest. We refer to Dechow et al. [2010] for a thorough review of existing literature on various measures as indications of "earnings quality", and to Hoberg and Lewis [2017] for a more recent bibliography on predictions of fraudulent behavior. Particularly, Hoberg and Lewis [2017] suggest that fraudulent firms tend to under-discuss factors that might explain potentially fraudulent accounting, but excessively discuss acquisitions, product lines and growth strategies. Hanley and Hoberg [2018] show that aggregate emerging risks may be obtained from the text of 10-K reports to predict heightened risk exposures. A fundamental limitation of fraud-related research is the limited or partial observability of fraud. Wang et al. [2010] use a bivariate probit method to deal with this limitation and find that fraud propensity increases with the level of investor beliefs about industry prospects but decreases when beliefs are extremely high.

A vast economic literature studies voluntary disclosures in the cheap talk setting as discussed in several papers [Crawford, 1998; Farrell and Rabin, 1996; Farrell, 1987; Crawford and Sobel, 1982; Grossman, 1981; Milgrom, 1981]. An extensive review can be found in Dranove and Jin [2010]. Economic agents may also attempt to overcome problems of hidden information by using commitment devices such as independent certification [Dranove and Jin, 2010; Viscusi, 1978]. In the area of discretionary risk reporting, Dobler [2008] finds that risk reporting depends on disclosure incentives.

The literature on business ethics and social responsibility examines the relationship between corporate Code of Ethics and corporate performance. Some studies have analyzed the contents of the codes of multinational firms. Carasco and Singh [2003] analyze the codes of 32 of the world's largest corporations and found that firms "were concerned with conduct both on behalf and against the firm, but concerns relating to the latter were more prominent". Helin and Sandström [2007] provide the most recent review of existing studies, mostly in the area of business ethics and social responsibility, and find that the evidence regarding the effectiveness of such codes is mixed. Earlier, Stevens [1994] concludes that corporate Codes of Ethics are "primarily "window-dressing" implemented to protect the firm from legal liability". Our results demonstrate that the focus areas of the management are not stale, but, instead, reflect the underlying business conditions. Canary and Jennings [2008] study the evolution of Codes of Ethics pre- and post- Sarbanes-Oxley (SOX) legislation in 2002 in the United States, and find that the Code structure has changed across time, with an increased emphasis on compliance in post-SOX codes.

2.1 Codes of Conduct

The Foreign Corrupt Practices Act of 1977[5] (FCPA) (15 U.S.C. §78dd-1, et seq.) prohibits U.S. companies from paying bribes to foreign government officials to obtain business or influence regulation. Corruption[6] reduces levels of investment and ultimately economic growth [Svensson, 2005]. At a firm level, corruptive and fraudulent behavior destroys value, especially in countries with weak corporate governance [Garmaise and Liu, 2005]. Fraud detection continues to be a high priority area for the SEC. Cases of serious misconduct receive media attention. For instance, in a recent example, Rolls-Royce paid about $170 million in U.S. penalties as part of an $800 million global resolution to anti-fraud investigations [Blanco, 2017].

In 1992 the Committee of Sponsoring Organizations of the Treadway Commission (COSO)[7] issued an Internal Control — Integrated Framework, which "quickly became the best-practice roadmap for designing, implementing and maintaining a system of internal control" [ACFE, 2016]. In 2013 COSO updated its guidance to include an explicit fraud-related principle: "the organization considers the potential for fraud in assessing risks to the achievement of objectives". According to Cotton et al. [2016], "the board of directors and top management and personnel at all levels of the organization — including every level of management, staff, and internal auditors — have responsibility for managing fraud risk."

[5] Karpoff et al. [2017] provides an extensive overview and historical background of the Act and its coverage in existing empirical literature.

[6] Svensson [2005] provides a definition of corruption and a extensive review of literature on its impact on growth.

[7] COSO is a joint initiative of five private-sector accounting and auditing associations organized in 1985, `www.coso.org`.

An important part of the anti-fraud and anti-corruption program is the Ethics and Compliance program, which, in particular, includes the firm's Code of Ethics and Conduct as well as business conduct training to employees, with main purpose being education and prevention of fraudulent and corruptive behavior [Ernst&Young, 2016].

The adoption of Sarbanes-Oxley Act of 2002, and in particular its Section 406[8], mandated the firms to maintain and disclose their Codes of Ethics. The SEC regulation[9] requires a company to disclose whether it has adopted "a Code of Ethics for its senior financial officers that applies to the company's principal financial officer and controller or principal accounting officer, or persons performing similar functions" [SEC, 2003]. If the company has not adopted a Code of Ethics, it must explain why not. While no standard template has been proposed by the SEC and the companies use their own judgment on designing and structuring their ethics communication, Item 406(b) of SEC [2002] defines "Code of Ethics" to mean written standards reasonably designed to deter wrongdoing and promote:

- honest and ethical conduct, including the ethical handling of actual or apparent conflicts of interest between personal and professional relationships;
- full, fair, accurate, timely, and understandable disclosure in reports, documents and in other public communications;
- compliance with applicable governmental laws, rules and regulations;
- the prompt internal reporting to an appropriate person or persons identified in the code of violations of the code;
- accountability for adherence to the code.

The SEC allows three different ways of satisfying the disclosure requirements of the Code of Ethics. First, it may be filed as an exhibit to the firm's annual report. Second, it may be made publicly available on the firm's website (provided that the firm indicates the appropriate website address in its annual report and its intention to post the Code of Ethics there). Third, public firms may undertake in their annual report to furnish a copy of the Code of Ethics to any person without charge upon request.

While the above guidance outlines the general set of requirements, the SEC strongly encouraged public companies to adopt codes that are broader and more comprehensive, and

[8] `www.sec.gov/about/laws/soa2002.pdf`
[9] `www.sec.gov/rules/final/33-8177.htm`

clarified that the Code of Ethics "may be a portion of a broader document that addresses additional topics or that applies to more persons" other than the officers required to be covered. While the vast majority of publicly-traded companies use the two terms interchangeably [Moberly, 2008], some companies bifurcate the Codes of Ethics and maintain separate[10] ones for different types of officers (e.g. Employee Code of Conduct vs. Principal Officer Code of Ethics). Our analysis focuses on the Code of Conduct as a broader type of communication in an attempt to capture more topics, higher textual variability and less legally scrutinized language, typically reserved for annual reports. If a firm distinguishes between the Code of Conduct for employees and the Code of Ethics for Senior Officers, the latter was left out of scope of our analysis.

While Code of Conduct is a popular title, firms use different names and refer to the document as the Code of Business Ethics, Code of Ethical Business Conduct and Standards of Ethics. Other creative but less popular variations include the Spirit Letter and Integrity Manual.

Regardless of the name, the Code of Conduct serves as a framework for ethical decision making within a firm. It is the key communication tool between management and employees that informs internal and external stakeholders about what is valued and accepted by a particular organization as ethical and integral. It is one of the elements of the anti-fraud system of internal controls along with other fraud detection mechanisms like reconciliations, access restrictions, segregation of duties etc. In many firms the Code of Conduct is a mandatory integration training, which all new employees must take at the beginning of their employment. Overall, a modern corporate Code of Ethics is a "statement setting down corporate principles, ethics, rules of conduct, codes of practice or company philosophy concerning responsibility to employees, shareholders, consumers, the environment, or any other aspects of society external to the company" [Langlois and Schlegelmilch, 1990].

> *"The Code of Conduct is the heart and soul of a company. Think of a Code of Conduct as an in-depth view of what an organization believes and how the employees of an organization see themselves and their relationship with each other and the rest of the world. The Code of Conduct paints a picture of how employees, customers, partners, and suppliers can expect to be treated as a result."* [Heathfield, 2018]

[10] A Code of Conduct can also be a document that details an organization's expectations and requirements of their vendors, suppliers, and partners, and lays the groundwork for the organization's relationship with its partners. This type of external communication, typically called the Supplier Code of Ethics, is, however, left out of scope of our research.

While Section 302 of SEC [2002] requires the Chief Executive Officer and the Chief Financial Officer to certify financial reports filed with the SEC, Section 406 does not require certification of corporate Codes of Ethics. Nevertheless, Orin [2008] finds that 31% of Codes of Ethics implemented by firms included certification signatures by upper management. Most recently, Davidson and Stevens [2012] study the effectiveness of a Code of Ethics in an experimental setting and find that it improves manager return behavior and investor confidence, but only when the code incorporates a public certification choice by the manager.

A subset of firms adds an acknowledgment form which employees need to sign after familiarizing themselves with the ethical guidance. One might expect that having an individual certify that he or she will adhere to the code will activate the social norm of promise-keeping. For example, in the audit context Carcello and Li [2013] find evidence that the engagement partner signature requirement leads to improved audit quality in the United Kingdom. We focus on the association between the content of the codes and the firm's fraud risk, and leave the discussion of effectiveness of the signature for future analysis.

3 Research Question and Hypotheses

Extant literature has shown that the textual content of the disclosure is informative and that a firm produces information about itself and the economy over time as discussed in Hoberg and Lewis [2017] and Hanley and Hoberg [2018]. The former suggest that fraudulent firms tend to under discuss factors that might explain potentially fraudulent accounting, but excessively discuss acquisitions, product lines and growth strategies. The latter show that aggregate emerging risks may be obtained from the text of 10-K reports to predict heightened risk exposures. In our paper, we investigate verbal clusters in CEO communication which might be consistent with specific concerns that the management has with regard to perceived or inherent business risks.

We conjecture that the verbal communication produced by the firms exposed to higher risks of fraud is different from the one produced by the firms with lower risks. To put this in context, we can think of a company which has operations in a highly corrupted country which, given the inherent nature of the business, might increase the chances of fraudulent behavior. Differences in the textual context of anti-fraud disclosures can be indicative of the signaling mechanism depending on whether a firm is a *peach* or a *lemon*[11] [Akerlof, 1970].

[11] In the jargon of Akerlof [1970], a *peach* is a well maintained car whereas a *lemon* is a car which is found to be defective after it has been purchased. In the context of our paper, a *lemon* is a firm less able to mitigate its fraud risks.

As investors are rational[12], it is reasonable to assume that without an additional signal on whether or not higher fraud risks can be mitigated, they[13] might downgrade their expectations of a given firm. The "goodness" of a firm can be defined as its ability to timely identify, evaluate, and mitigate fraud risks. This ability and commitment to foster anti-fraud compliance may be manifested in textual communication, where the firm acknowledges and addresses these higher fraud risks in its disclosures. For instance, a "good" firm cognizant of its heightened fraud exposure would put more emphasis on fraud risks management in its Code of Conduct in the hope that it will increase employee's compliance with the anti-fraud policies. If investors are not able to distinguish between peaches and lemons, then firms will be downgraded by the same amount thus leading to a pooling equilibrium in the spirit of Akerlof [1970]. We argue that additional fraud-emphasizing disclosure can be used as a device to sustain a separating equilibrium in which good firms are not downgraded [Crawford and Sobel, 1982].

We will provide detailed definitions of our key variables in Section 5, but in order to formulate the hypotheses, we offer the intuition here. To measure the degree of the CEO's anti-fraud emphasis, we construct the *CEO Anti-Fraud Language Index* ($\mathrm{FL}_{\mathrm{CEO}}$) inferred from the LDA topic modeling approach applied to the published Codes of Conduct. This allows us to build a picture of the topics that a firm concentrates on in its communication. Through an automated procedure, we pick up the topics related to fraud and measure their relative "weight" in the document. To measure a firm's fraud risk (business corruption) exposure, we look at the countries in which a firm has foreign subsidiaries. Each country is quantitatively characterized by publicly available corruption indexes, which allows us to construct a *Firm Fraud Risk Index* ($\mathrm{FR}_{\mathrm{Firm}}$).

In our first hypothesis, we expect a rational CEO to be cognizant of the specific risks of fraud that employees may face and, respectively, reflect them in the Code of Conduct. We expect that when the company has subsidiaries in countries with higher corruption (higher fraud risk), the CEO will put more anti-fraud emphasis in their communication with employees (higher CEO anti-fraud language index). Within this context, we can formulate our main hypothesis as follows:

Hypothesis 1: *CEO anti-fraud language index is positively associated with the firm's fraud risk level.*

[12] For a recent literature review on investor's rationality vs. irrationality debate see, for example, Mamun et al. [2015]

[13] We can also think of certification providers in the context of Viscusi [1978]; Dranove and Jin [2010].

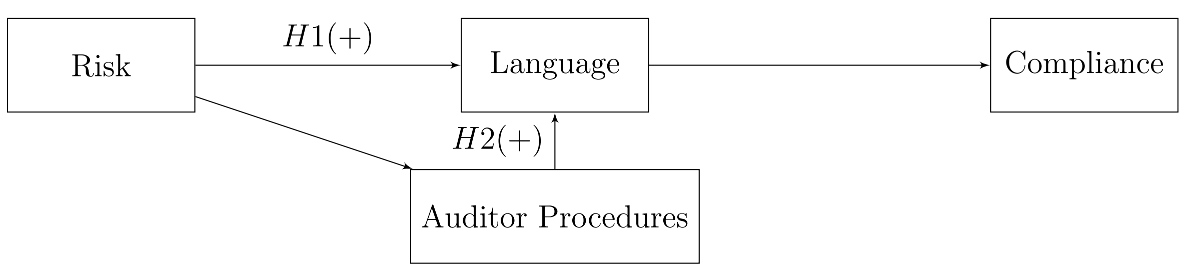
To assess this statement, we exploit fraud-related terminology in CEO communication. For instance, we expect a CEO to use specific lexicon in detailed paragraphs when discussing topics like working with foreign governments (such as "government", "violation", "gifts"), and/or specific terminology emphasizing the importance of integral anti-fraud behavior (such as "integrity", "safety", "values"). The linguistic content of 10-Ks can be more "legally scrutinized", or the CEO may be acting strategically by trying to conceal increased fraud risks. Therefore, we turn our attention to the communication targeted at firm's employees in the form of Codes of Conduct, where we assume that the management communicates their fraud-related concerns in a more linguistically liberated and less legally restrictive way. We expect that when the company has subsidiaries in countries with higher corruption (higher fraud tolerance), the CEO will put more emphasis on fraud-related topics (higher CEO anti-fraud language index).

While management rhetorics can provide credible signals about underlying firm attributes [Balvers et al., 2016], in our second hypothesis we also explore an alternative possible explanation. Increased business corruption risk may be assessed by the auditor in the form of increased audit procedures and, in turn, prompt the auditor to influence the firm's internal control procedures and tone from the top. Abnormally high audit fee levels, resulting from increased procedures, may also influence an auditor's judgment through economic bonding to the client[Blankley et al., 2012]. We address this potential causal effect of auditor's increased procedures in Hypothesis 2:

Hypothesis 2: *In presence of higher fraud risk level, CEO anti-fraud language index is positively associated with the firm's audit procedures.*

We adapt the concepts of the theoretical model of Crossler et al. [2016] to our setting and summarize our research model in Figure 1.

Figure 1: Theoretical Research Model



4 Sample Selection

4.1 Codes of Conduct Sample

In this study we focus on anti-fraud communication that is represented by the firm's Code of Conduct. We hand-collect Codes of Conduct for S&P 500 companies by downloading them from the companies' websites[14]. Out of 502 stocks listed on S&P 500 list in May 2017, one firm (Activision Blizard Inc) did not have a Code of Conduct posted on the website[15], and two documents (Lennar and Transdigm) were not machine-readable. Of 499 available documents 486 were in the PDF format and 13 in HTML format.

An initial visual examination of the documents enables us to determine a varying length between one and 120 pages, to detect both colored and monochromatic fonts with a variety of visual aids, flowcharts and diagrams. Some firms use case studies or scenarios to present the ethical dilemmas and better convey the recommended solutions to the audience. To summarize what we observe, we report descriptive statistics in Table 1. As these statistics were not obtained through an automated approach, we do not include them in our regressions, but instead publish them here to familiarize the reader with a typical content of a Code of Conduct. The mean length is 28 pages with a large usage of colored fonts (71%) and visual elements and diagrams to illustrate the message (66%). More than half of the companies (56%) include scenarios and questions and answers (Q&A) paragraphs with illustrative examples, though only a restricted number of them (14%) require a signature at the end of the document. We also include a representative example taken from Code of Conduct document for Intel Corporation (C)[16].

> [...] "*Intel strictly prohibits all forms of bribery. Intel's policy is to comply with anti-corruption laws and to accurately reflect all transactions in Intel's books and records. We must never offer or accept bribes or kickbacks and must not participate in or facilitate corrupt activity of any kind. Many countries' laws define facilitation payments made to government officials as bribes. We do not make facilitation payments on behalf of Intel to any government official. Intel's prohibition against offering, promising or paying bribes also applies to third parties who provide services or act on Intel's behalf, such as suppliers, agents, contrac-*

[14]Typically, these documents are publicly available and can be found on a firm's website in the section of "Investor Relations" or "Corporate Governance".

[15]We did not write to the Legal Department of the firm to request a printed version of the Code of Conduct.

[16]The document can be found at `www.intel.com/content/dam/www/public/us/en/documents/corporate-information/policy-code-conduct-corporate-information.pdf`.

tors, consultants and distributors. We must never engage a third party whom we believe may attempt to offer a bribe in connection with company business. Our anti-corruption expectations for third parties are set out in our Third Party Anti-Corruption Policy and Gifts, Meals, Entertainment and Travel ("GMET") Policy for Third Parties. When doing business with governments, consult with Legal to be certain you are aware of any special rules or laws that apply. Obtain the required approvals in our Worldwide Business Gifts, Meals, Entertainment, and Travel Policy ("GMET Policy") before providing anything of value to a Government Official." [...]

Table 1: Descriptive Statistics of Codes of Conduct

Descriptive Statistics of the visual characteristics of the Codes of Conduct are reported for 501 firms, which include our sample of 499 documents plus two non machine-readable documents. We summarize the results of the visual review of the documents. As these statistics were not obtained via an automated approach, we do not include them in our regressions.

Characteristic	**N. of Docs**	%
In color	357	71.3%
Contains visual aids and diagrams	332	66.3%
Contains scenarios and Q&A	280	55.9%
Contains acknowledgement form	69	13.8%
Total Documents (including non machine-readable)	501	100%

In order to correctly read and import the data, we develop several automated routines which take care of the preprocessing step in line with the standards of textual analysis [Loughran and McDonald, 2011; Muslu et al., 2014]. This consists of checking for the presence of inconsistent linguistic content such as symbols[17], non-ASCII characters, and URLs which do not provide any information on the context. The 499 documents have been then converted into a corpus in order to enable quantitative estimations. A corpus is a well structured and possibly very large object which is the base for quantitative textual mining [Loughran and McDonald, 2016; Henry and Leone, 2015].

4.2 Foreign Subsidiaries Sample

We collect firms foreign subsidiaries information from the most recent Exhibit 21 to the form 10-K filed prior to May 2017. We follow Dyreng and Lindsey [2009] and use a text search

[17] As given in the Unicode [S] class category.

program to scan Exhibit 21 to identify in which foreign countries subsidiaries material to firm's operations are located. The list of countries was obtained from the World Bank and Transparency International websites. The list consists of 175 countries for which both World Governance Index (WGI) and Corruption Perception Index (CPI) were calculated in 2016. We use WGI in our main specification and CPI as a robustness measure[18].

The disclosure rule (SEC rule 17 Item 601 of SEC Regulation S-K (§229.601)) requires that public firms list all significant subsidiaries in Exhibit 21, but clarifies that only significant subsidiaries must be disclosed, thus potentially omitting firm's factual presence in developing countries where they might not have a material operation [Gramlich and Whiteaker-Poe, 2013].

While the text search program does not perfectly identify all countries with material operations, it does enable us to study a much larger sample than would be possible if the data were read and coded by hand. We do not search for subsidiaries located in the USA. Instead, since all of our sample firms are U.S. public firms, we ensure that all of them have a subsidiary row corresponding to the country U.S.A. After excluding Georgia, Oman, Togo, Sudan, South Sudan, Cabo Verde and Republic of Congo and adding alternative country names such as Burma, England, Korea North and Korea South[19], our final list of countries consists of 171 country names, of which 167 are unique countries. Details on WGI and CPI corruption indexes are reported in Appendix A.

Importantly, the methodology presented in this paper allows us to identify the countries in which a subsidiary is located, but not the number of subsidiaries in a particular country. In our sample, 449 firms have disclosed foreign subsidiaries with an average of 13 countries per firm ranging from one to 89, 45 firms have disclosed only US-based subsidiaries and two firms have not disclosed any subsidiaries by not filing Exhibit 21[20].

[18]WGI index is available in two forms: estimate and rank. We used the rank version in our main regression. In unreported results we tested that our inferences remain unchanged when we use the estimate version. Both of the measures were normalized to a scale of [0, 1].

[19]There are some limitations to our approach. First, our text search program cannot distinguish between the country and the U.S. state of Georgia. We have excluded Georgia from the list of countries that we search for. Second, since the text search is employed on a text file with html tags, some country names are too simple and may be "false positively" found in the text: for example, Oman may be identified as a portion of the font tag "Times New Roman". We exclude the countries Oman and Togo from our search. Third, due to the lack of standard conventions, some countries allow for ambiguous spellings: for example, we search for both "United Kingdom" and "England", "North Korea" and "Korea North", "South Korea" and "Korea South" and "Burma" and "Myanmar". For the same reason we search for "Russia" instead of "Russian Federation" and "Czech" instead of "Czech Republic". Due to ambiguity between "Sudan" and "South Sudan", "Cape Verde" and "Cabo Verde" as well as "Republic of Congo" and "Democratic Republic of Congo", we exclude these countries from our list.

[20]We additionally verified the accuracy of the search by manually reviewing the firm's form 10-K for the

4.3 Financial and CEO Sample

Financial and CEO data to construct control variables is obtained from Compustat and Execucomp databases. Audit fees for 2017 (or the most recent year for which the data is available) were obtained from Audit Analytics.

Several studies of financial statement fraud find that firms that get involved in financial manipulations tend to be larger profitable firms [Dechow et al., 2011; Wang et al., 2010]. Certain industries also appear to have higher fraud concentration. Leverage has been used as a proxy for closeness to covenant restrictions and thus proneness to earnings management [Duke and Hunt III, 1990][21]. We proxy the auditor procedures by using the logarithm of audit fees. Following Blankley et al. [2012], we include an extensive set of financial covariates including current and other ratios. We also include industry controls since audit fees and anti-fraud emphasis may vary by industry. For our instrumental variable approach discussed later in Section 5.3, we include the data on the presence of mergers, calendar year end and segments. To control for internal control quality, we include variables defining the presence of a material weakness and auditor going concern opinion. All these variables are defined following Blankley et al. [2012] and are described in Table 4.

CEO personal characteristics including age, gender and total compensation could be affecting the style, personal preferences in communication and fraud-aversion. Cumming et al. [2015] show that having an optimal percentage (50%) of women on boards minimizes securities fraud. Ho et al. [2015] examine the relationship between CEO gender and accounting conservatism and find a positive association between the two. Borghans et al. [2009] study risk and ambiguity aversion and show that women are more risk averse than men. Dohmen et al. [2011] find that gender, age, height, and parental background have an economically significant impact on willingness to take risks. Jia et al. [2014] find positive association between CEO's facial masculinity and financial misreporting. Regulators and governance activists are pressuring firms to abolish CEO duality, that situation in which the CEO is also holding the position of the Chairman of the Board, although the academic evidence on relationship between CEO duality and firm performance is mixed [Yang and Zhao, 2014]. In our analyses we used CEO gender, age, total compensation and duality as control variables of CEO personal characteristics.

All the described datasets have been merged together in order to build a unique and

location of material subsidiaries. We performed this verification for the two firms which didn't file Exhibit 21, and for a small sample of firms with Exhibit 21.

[21] We acknowledge that leverage may be a relatively noisy proxy for closeness to covenants [Dichev and Skinner, 2002]

handy database on which we can run further analyses.

Most of the analyses discussed in this paper have been carried out through the `R` [R Core Team, 2017] open source programming language. In particular, data preprocessing has been carried out through the `data.table` package [Dowle and Srinivasan, 2017], textual data has been managed by `quanteda` [Benoit, 2018], LDA models have been computed with the `topicmodels` package [Hornik and Grün, 2011], and visualizations have been made possible by `ggplot2` [Wickham, 2009]. Instrumental variables regressions were performed by `ivreg` [Kleiber and Zeileis, 2008]. Panel regressions in subsection 6.3 were performed using StataMP (R) [StataCorp., 2017].

5 Research Design

5.1 Building the CEO Anti-Fraud Language Index

One of the key steps in this paper consists of defining a quantitative measure of CEO's anti-fraud emphasis, which we call CEO anti-fraud language index. In order to do that, we implement a Machine Learning (ML) algorithm called *Latent Dirichlet Allocation* (LDA) [Blei et al., 2003]. The LDA is an unsupervised approach which allows the identification of substructures called *topics* in a complex and possibly very large corpus of documents[22]. The LDA model assumes that the words contained in each document arise from a mixture of topics, each of which is a distribution over all the words detected in the corpus. The LDA is a generative probabilistic model in the sense that it uses the probability of different words to be contained in a set of documents at the same time. A researcher has to define a number of topics k a-priori, which is a typical step when applying unsupervised methods.

LDA allows us to estimate topic densities over a corpus of documents which are stored into a matrix called γ with dimensions of $N_d \times k$, where N_d is the number of documents in the corpus (e.g. 499 in our case). Word densities over a given topic are similarly stored into a matrix called β with dimensions of $k \times N_w$, where N_w is the total number of words detected in the corpus. In this paper we use the γ matrix as it quantifies the extent to which each of the topics is discussed in the individual Code of Conduct document. The results of the LDA

[22]For brevity purposes, we omit the technical details on the preprocessing routines which include the removal of stop words. In addition to the removal of standard English stop words, we included the lists of geographical locations, names, dates and numbers, currencies which we obtained from `https://www3.nd.edu/~mcdonald/Word_Lists.html`. Moreover, we also used the stop word list provided by Provalis Research in their text mining software WordStat(C) (`https://provalisresearch.com/products/content-analysis-software/`. Results are consistent using both sets of stop words.

allow the researcher to critically review and interpret the topics based on word's relevance. Each topic consists of all the words detected in the corpus with a specific order given by their relevance to the topic. Naturally, the outcome of the analysis changes with different k.

Choosing the appropriate number of topics is a challenging and sensitive step subjected to researcher's interpretation. There is a trade off between the goodness-of-fit of the model which improves with higher k and its interpretability [Chang et al., 2009]. In order to investigate the problem, we use an unbiased approach, typical of unsupervised methods, and run multiple LDA models for the following values of k:

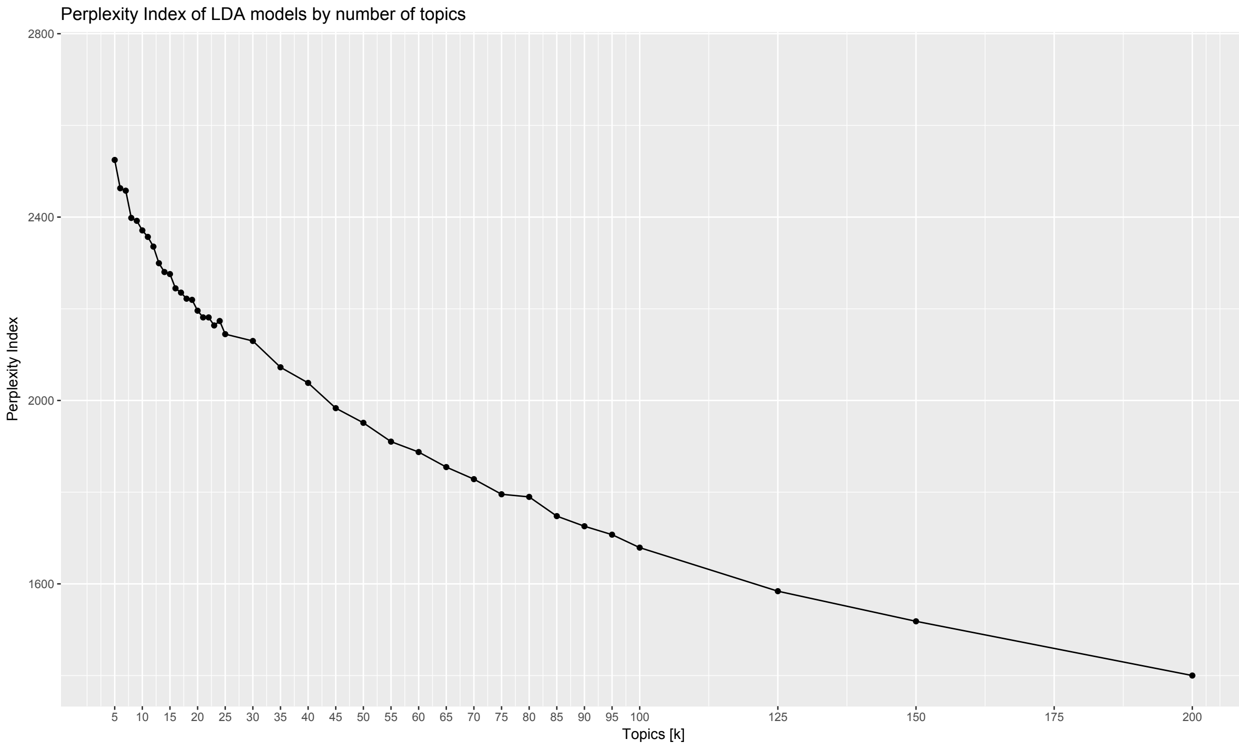
$$k \in \Big\{[2, \dots, 25], 30, 35, 40, 45, 50, 100, 125, 150, 200\Big\}, \tag{1}$$

where $[2, \dots, 25]$ represents a continuous sequence from 2 up to 25 with unit increments. In this regard, the described approach is intuitively akin to another very common unsupervised method for clustering called *k-means* [Hartigan and Wong, 1979; Forgy, 1965]. In order to find the best partition in the data, the researcher runs several models using a different number of partitions and then uses multiple measures to assess which is the best candidate.

One of the most common indexes to assess the appropriateness of the LDA is called *perplexity* introduced by Hornik and Grün [2011]; Newman et al. [2009]; Blei et al. [2003]. This measure is equivalent to the geometric mean per word likelihood and it takes into account how many times a word is detected in a given document. The final selection of the optimal k is usually done by visually inspecting the perplexity index versus the number of topics. In Figure 2, we show the perplexity plot. The rule of thumb for selecting the optimal k is to look for a sudden change in the slope of the function. We do not observe such a change, thus we limit the analysis to a lower range of k to be more confident in our interpretation of the topics. In Section 6, results are reported for $k = 3, 5, 10, 20$.

The next challenging step in our analysis is the lexical identification of fraud related topics. We naturally expect the communication on anti-fraud measures to contain directions about working with government officials, avoiding gifts and bribes to business contacts, to demonstrate integrity while conducting business and making decisions. To support our intuition, Canary and Jennings [2008] find that words "government" and "integrity" are among the top words related to ethical emphasis in Codes of Conduct. In an attempt to reconcile the intuition with the actual observations, for each k we extract the top 15 words in each topic to identify fraud related ones. As a result, we consistently confirm that the words "government" and "integrity" characterize the topics which are associated with fraud. In addition, we find the words "gift(s)" and "contact" are another pair of frequently used terms related to fraud.

Figure 2: Perplexity index versus the number of topics k for different LDA runs.

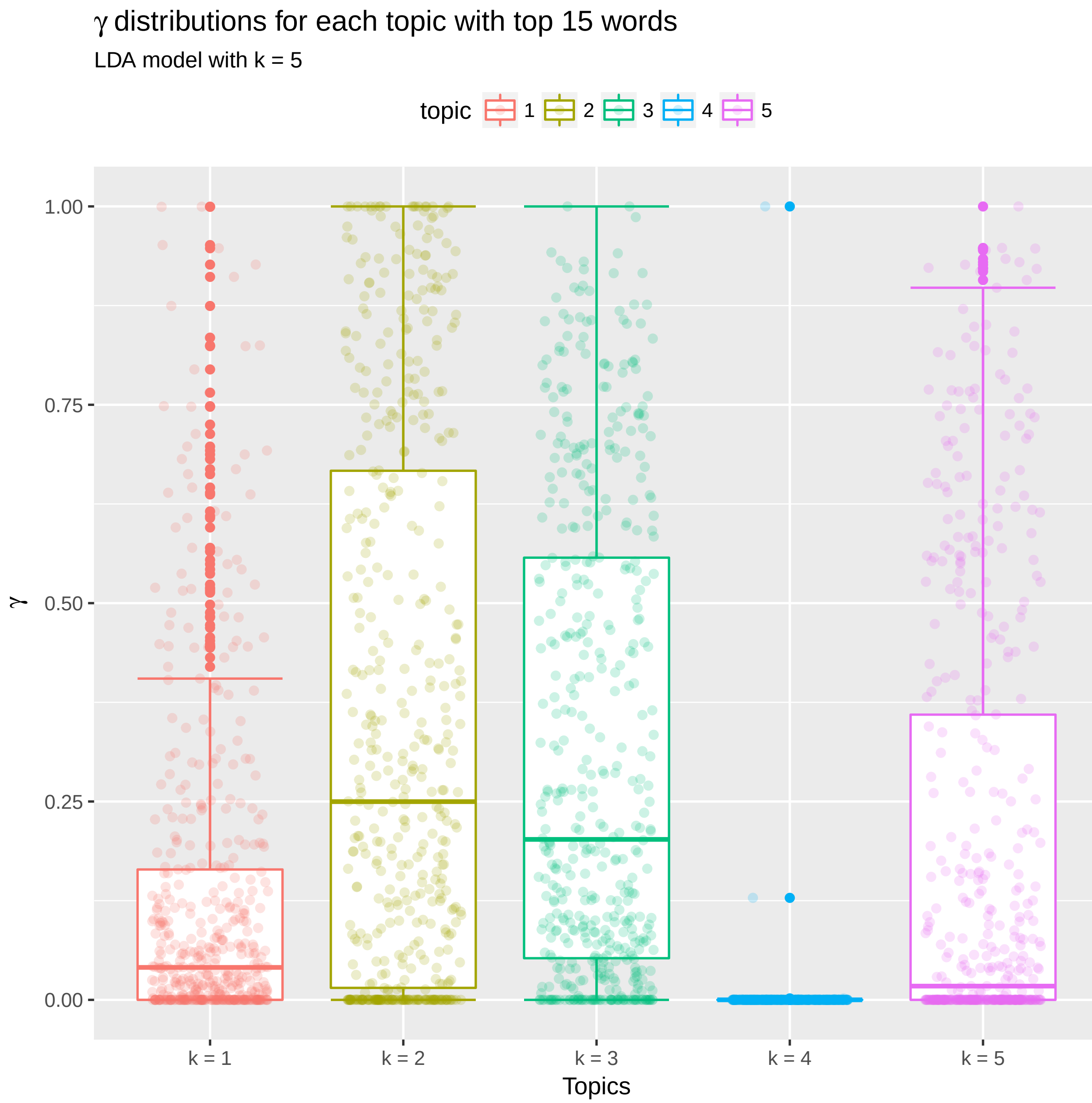


In Figure 3 and Table 2, we show the distributions of topic loadings γ given by the LDA model with $k = 5$, and the related top 15 words for each of the topics, respectively.

Table 2: Top 15 words for each topic given by the LDA with $k = 5$.

Topic 1	**Topic 2**	**Topic 3**	**Topic 4**	**Topic 5**
officer	integrity	employment	global	government
directors	government	public	pre	contact
financial	customers	officer	met	financial
officers	global	customers	pro	gifts
associates	contact	action	patients	customers
board	values	political	sur	office
director	products	government	fare	outside
committee	safety	gifts	ist	questions
associate	concerns	securities	au	property
violation	questions	property	ik	local
regulations	help	services	relative	services
executive	team	outside	nossas	integrity
accounting	local	regulations	pode	employment
securities	financial	time	delle	data
chief	public	entertainment	essere	public

Figure 3: Distributions of topic loadings γ given by the LDA model with $k = 5$.



We adopt two different criteria to automatically detect a fraud related topic. The algorithm uses different combinations of the four identified keywords as follows:

- *any:* choose a topic if any three of the four keywords appear in the top 15 words of the topic;
- *both:* choose a topic if both "government" and "integrity" appear in the top 5 words;

While we are not aware of prior literature using the exact same methodology, our approach resembles the one employed by Hanley and Hoberg [2018].

The CEO's Anti-Fraud Language Index for firm i is constructed as follows:

$$\mathrm{FL}_{\mathrm{CEO_i}} = \sum_{j \subseteq \tilde{k}} \gamma_{ij}, \tag{2}$$

where γ_{ij} is the $i-$th row, $j-$th column element of the γ matrix, and $\tilde{k}$ is the set of the detected fraud related topics. By construction, $\mathrm{FL}_{\mathrm{CEO_i}} \in [0,1]$ where low values of this score are associated with less fraud emphasis. In our paper, main results are reported in Table 6 and use the specification *any*. Robustness tests reported in Appendix B use the other specification.

5.2 Building the Firm Fraud Risk Index

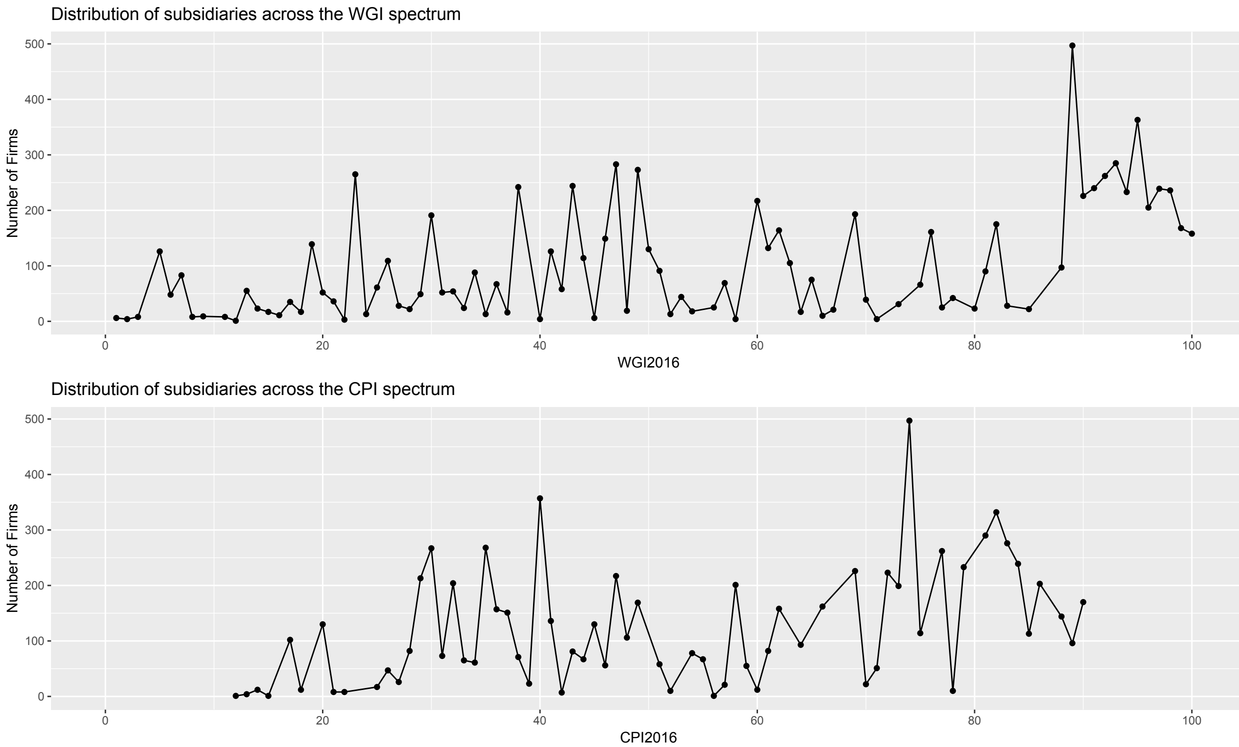
We hypothesize that firm's inherent predisposition towards fraud (e.g. governmental bribes and violations of the Foreign Corrupt Practices Act of 1977, are predetermined by the countries in which the firm is operating its business. We use country corruption index as a proxy of corruption practices in this country.

As described in Section 4, we collect information on all significant foreign subsidiaries of the firm using Exhibit 21 of the forms 10-K. Each country is characterized by the WGI and CPI. Both are composite indicators assessing a wide range of corruption data. According to Roca et al. [2010], CPI "has become the single most effective advocacy tool in the global fight against fraud, embezzlement and other abuses of public office for private gain". According to Rohwer [2009], WGI uses broader sources and a somewhat enhanced methodology. In this paper, we use WGI as our main specification and CPI as a robustness measure.

Figure 4 illustrates the distributions of firm subsidiaries across the corruption index spectra of WGI and CPI. Our methodology allows us to identify the countries in which a subsidiary is located, but not the number of subsidiaries in a particular country. For example, 357 firms in our sample have reported having a subsidiary in countries with CPI index of 40 (Brazil, India, China, Belarus). The graph also reflects that, by construction, all firms in our sample have a subsidiary in the US (with WGI of 90 and CPI of 74), as visually evidenced by the respective spikes in the graphs in Figure 4.
Since our sample includes large multinational companies operating in various countries, we aggregate this measure at a firm level in the following way.

Intuitively, the average firm fraud risk will then depends on the composition of its sub-

Figure 4: Distributions of subsidiaries across corruption index spectra.



sidiaries locations. We construct the *Firm Fraud Risk Index* ($\mathrm{FR}_{\mathrm{Firm}}$) the following way: we take the 100 minus the median of all corruption indexes for all countries in which the firm has subsidiaries, and normalize to $[0, 1]$ by dividing by 100. Low values are then associated with a lower fraud risk of a firm. For example, a value of one represents an extreme case of a firm with all of its subsidiaries located in the most corrupted countries. Conversely, a value of zero represents a firm with all of its subsidiaries located in the least corrupted countries.

We also construct an alternative aggregate measure of corruption risk. We assume that the riskiest country, from a corruption perspective, in which the firm agrees to operate, will characterize the firm's *tolerance* towards the amount of fraud risk that it is willing to take upon. We construct *Firm Fraud Tolerance Index* ($\mathrm{Tol}_{\mathrm{Firm}}$) by taking the minimum of all corruption indexes for all those countries in which the firm has subsidiaries. We also normalize this index to $[0, 1]$ by dividing by 100. Low values are then associated with riskier countries and vice versa. For example, a value of one represents a firm with subsidiaries located only in the least corrupted countries. Conversely, a value of zero, represents a firm with at least one subsidiary located in the most corrupted country.

Denoting the corruption index of country s by C_s, and the set of countries in which firm

i has subsidiaries by S_i, the fraud risk and fraud tolerance for firm i can be expressed as follows:

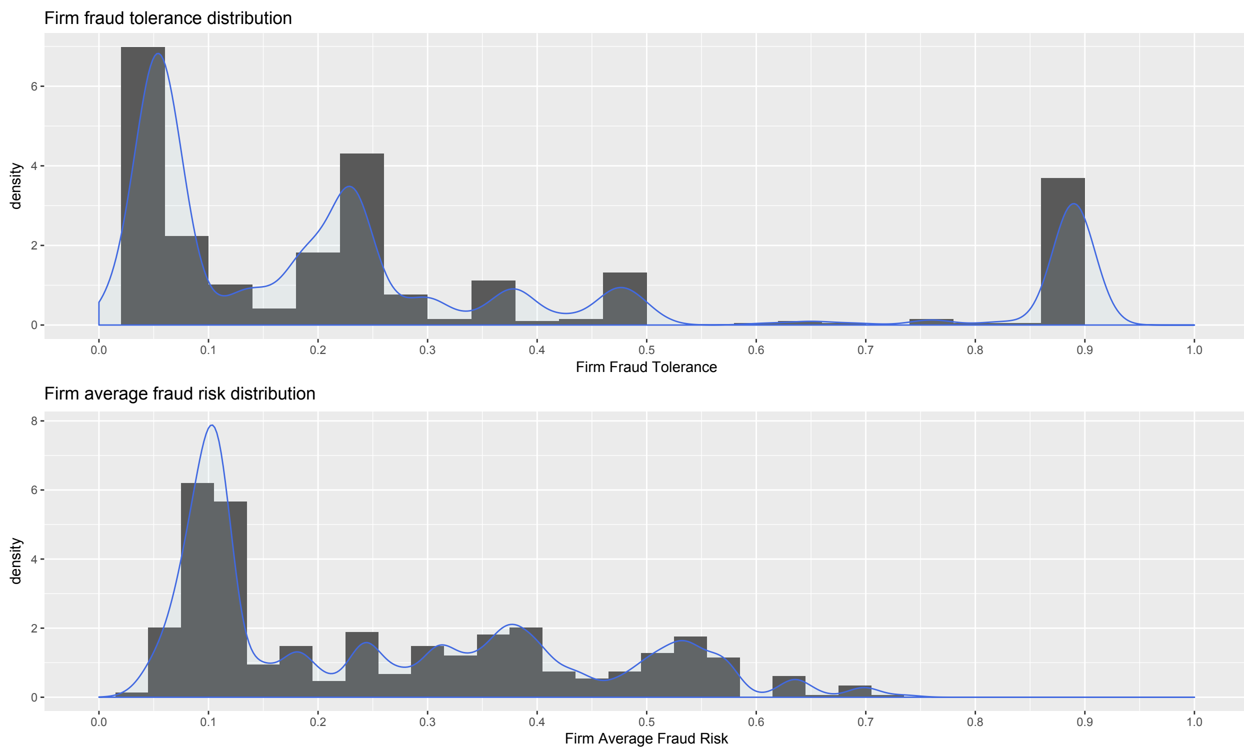
$$\mathrm{FR}_{\mathrm{Firm}_i} = 1 - \frac{\operatorname{median}_{s \in S_i} C_s}{100}, \tag{3}$$

$$\mathrm{Tol}_{\mathrm{Firm}_i} = \frac{\min_{s \in S_i} C_s}{100}. \tag{4}$$

Figure 5 illustrates the distributions of $\mathrm{Tol}_{\mathrm{Firm}}$ (top panel) and $\mathrm{FR}_{\mathrm{Firm}}$ (bottom panel).

In our regressions described in the subsequent sections, we focus on $\mathrm{FR}_{\mathrm{Firm}}$ as the main measure of the firm's corruption risk, and use $\mathrm{Tol}_{\mathrm{Firm}}$ in a robustness test.

Figure 5: Distributions of Firm Tolerance and Firm Fraud Risk Indexes. Solid line represents the smoothed density estimate.



5.3 Regression Model

In our model, we hypothesize that both anti-fraud emphasis and auditor procedures are simultaneously affected by the business corruption risk. We proxy the auditor procedures by using the logarithm of audit fees (LAF). To disentangle the relationship, we use an instrumental variable approach and estimate the below system of equations using two-stage least squares method. Equation 5 tests our Hypothesis 1 with the CEO's Anti-Fraud Language Index as the dependent variable. Our main independent variables are $\mathrm{FR}_{\mathrm{Firm}}$ and LAF, we also include several sets of covariates which we use as controls. Equation 6 instruments LAF using an audit fee model adapted from Blankley et al. [2012]. We put an emphasis on instrumenting LAF by fee determinants driven by the nature of the business, and not its financial results, which may have a direct effect on the choice of the language. In this stage, we regress LAF on variables controlling for merger and acquisition activity, business complexity and audit opinions. These determinants are associated with LAF [Blankley et al., 2012], but are unlikely to directly influence the anti-fraud emphasis of the CEO language, therefore, satisfying the exclusion restriction requirement. The predicted values of the LAF regression are then used in the second stage regression.

We can formalize our empirical model as follows:

$$\begin{cases} \mathrm{FL}_{\mathrm{CEO}} = \beta_0 + \beta_1 \mathrm{FR}_{\mathrm{Firm}} + \beta_2 \mathrm{LAF} + \beta_{3-10}\mathbf{C}_{\mathrm{FIN}} + \beta_{11-14}\mathbf{C}_{\mathrm{CEO}} + \beta_{15-27}\mathrm{IND} + \varepsilon & (5) \\ \mathrm{LAF} = \alpha_0 + \alpha_1 \mathrm{FR}_{\mathrm{Firm}} + \alpha_{2-6}\mathrm{C}_{\mathrm{Z}} + \alpha_{7-14}\mathbf{C}_{\mathrm{FIN}} + \alpha_{15-18}\mathbf{C}_{\mathrm{CEO}} + \alpha_{19-31}\ \mathrm{IND} + \varrho & (6) \end{cases}$$

where:

- $\mathrm{FL}_{\mathrm{CEO}}$ is the CEO's Anti-Fraud Language Index obtained through the LDA model;
- $\mathrm{FR}_{\mathrm{Firm}}$ is the Firm Fraud Risk Index defined in Equation 3[23];
- LAF is the logarithm of 2017[24] audit fees;
- $\mathbf{C}_{\mathrm{FIN}}$ includes firm size (logarithm of total assets, LTA), current ratio (CR), current assets scaled by total assets (CA_TA), the sum of accounts receivables and inventory scaled by total assets (ARINV), intangible assets scaled by total assets (INTANG), return on assets (ROA), leverage (LEV), and an indicator of Loss;
- $\mathbf{C}_{\mathrm{CEO}}$ includes CEO's age, gender, logarithm of total compensation, and a dummy marking the existence of CEO - Chairman duality;

[23]In robustness tests we use an alternative measure of $\mathrm{Tol}_{\mathrm{Firm}}$ discussed earlier.
[24]We used 2016 data if 2017 fees were not available.

- $\mathbf{C}_Z$ includes indicator variables of merger and acquisition activity (MERGER), December fiscal year end (BUSY), logarithm of number of business segments (SEG), disclosure of material weakness (MATWEAK) and going concern opinion (OPINION);

- IND includes industry fixed effects; industry membership follows Blankley et al. [2012] and is determined by SIC code as follows: agriculture (0100–0999), mining and construction (1000–1999, excluding 1300–1399), food (2000–2111), textiles and printing/publishing (2200– 2799), chemicals (2800–2824; 2840–2899), pharmaceuticals (2830–2836), extractive (1300–1399; 2900–2999), durable manufactures (3000–3999, excluding 3570–3579 and 3670– 3679), transportation (4000–4899), retail (5000–5999), services (7000–8999, excluding 7370–7379), computers (3570–3579; 3670–3679; 7370–7379), and utilities (4900–4999).

If our Hypothesis 1 is correct, we expect a significant positive coefficient β_1 on Firm Fraud Risk, indicating that higher fraud risk level is associated with higher anti-fraud emphasis. Analogously, a significant positive coefficient β_2 on LAF would indicate that in presence of increased fraud risk and after controlling for the business drivers of audit fee levels, audit fees are positively associated with anti-fraud emphasis of CEO's communication, supporting Hypothesis 2.

6 Results

In this section we present the main results of the paper including descriptive statistics, correlation matrix, model estimation, and robustness tests. All the tables can be found in Appendix B.

6.1 Descriptive Statistics

As mentioned in Section 4, out of 502 stocks listed on S&P 500 list in May 2017 one firm (Activision Blizard) did not have a Code of Conduct posted on the website, and two documents (Lennar and Transdigm) were not machine-readable. Of 499 available documents 486 were in the PDF format and 13 in HTML format. We were not able to produce LDA results for one document (Zyons Bankcorp), and four firms (Dupont, Arconic, Eaton and Mylan) were not found in Compustat. After the construction of the key variables of interest as described in

Section 5, we subsequently merge them with available financial and CEO information. Our final sample consists of 494 firms, for which CEO Anti-Fraud Language Index, Firm Fraud Risk and financial controls are calculated using available information, and 452 firms with CEO information in Execucomp available to construct our control variables.

Summary statistics for our sample is presented in Table 4. Our dependent variable $\mathrm{FL}_{\mathrm{CEO}}$ is reported for a different number of topics k (see subsection 5.1). An average CEO devotes from 30% ($k = 50$) up to 57% ($k = 10$) of the Code of Conduct to fraud related topics. The average Firm Tolerance Index is 0.28, and Fraud Risk index is 0.26. An average firm in our sample has a ratio of 9% in Return on Assets, 25% of intangibles-to-assets ratio, 28% in Leverage and 3.08 as current ratio; 10% of the firms reported a loss. An average CEO in our sample is 56 years old, 96% of them are males and 45% contemporaneously hold the position of the Chairman of the Board. Additionally, 5% of the firms in our sample reported M&A activity, 75% have a December fiscal year-end, and 2% disclosed a material weakness. Very few firms in our sample received a going-concern opinion.

Table 5 displays Pearson correlation coefficients between each covariate as well as the significance given as asymptotic p-values [Hollander et al., 1973]. FR_{Firm} and Tol_{Firm} are correlated with $\mathrm{FL}_{\mathrm{CEO}}$, foreshadowing our main results. By construction of the variables, FR_{Firm} is strongly negatively correlated with Tol_{Firm}, $ARINV$ is strongly positively correlated with CA_TA, and LTA is strongly positively correlated with LAF, which is a well-established relationship between audit fees and firm size. Interestingly, only CEO compensation is positively correlated with $\mathrm{FL}_{\mathrm{CEO}}$, but not CEO age or duality. Finally, LAF is strongly positively correlated with $\mathrm{FL}_{\mathrm{CEO}}$; this effect is mitigated by the use of instrumental variable approach and does not dominate our results.

6.2 Model Estimates

Table 6 reports the empirical results of model formalized in Equations (5) and (6) and provides consistent evidence to support Hypothesis 1. We find that the coefficient associated with firm fraud risk ($\mathrm{FR}_{\mathrm{Firm}}$) is positive and significant regardless of the LDA model considered up to $k = 20$. This is a robust result, which highlights a strong positive association between the CEO's Anti-Fraud Language index ($\mathrm{FL}_{\mathrm{CEO}}$) and the firm fraud risk ($\mathrm{FR}_{\mathrm{Firm}}$). This confirms that the CEO tends to emphasize fraud related topics more, when the firm operates in countries with higher corruption levels. Results also hold when controlling for financial and

CEO characteristics.

At the same time, we are not able to find convincing evidence to support Hypothesis 2. We find the coefficient β_2 of the audit fees (LAF) to be insignificant after instrumenting it with business complexity and risk drivers. We acknowledge that this insignificance may be attributed to the lack of auditor's skills or unclear guidance on verifying fraud-related language. For many companies, however, the Code of Conduct is a key entity-level control [25] and is, therefore, subject to annual audit procedures.

We also report results of the two standard specification tests in our empirical study. The weak instruments test for the null hypothesis that the instruments are weak is rejected, which allows us to conclude that the instruments are strong. The Sargan (1958) over-identifying restrictions test, tests whether instruments are valid with regard to an absence of correlation between instruments and model errors. Statistically insignificant Sargan statistics are consistent with the instruments being appropriate.

A note worth making concerns the choice of the number of topics. As discussed in Section 5.1, a researcher has to define a number of topics k a-priori. We believe that our results for $k = 5$ and 10 represent the "golden mean" of the analysis and enable us to balance the mathematical appropriateness of the estimation algorithm with an adequate level of human interpretation.

We acknowledge possible endogeneity issues. Codes of Conduct, being a key part of the system of internal controls, may affect the firm's corporate culture and induce more ethical behavior. A corporate culture change will, however, take time [Kotter, 2008]. To mitigate this concern, in Section 6.3 we also perform additional tests to understand the direction of causality, which indicate that the change in firm's risk tolerance levels drives changes in Code of Conducts. Overall, we strongly believe that the short-term nature of the firm's business affects the management communication style and not the other way around.

A reader may also argue that a concealing CEO will choose not to put emphasis on fraud risks, but instead avoid fraud-related terminology altogether. This might be true in the risk disclosure discussion to the investors, but should not affect CEO's communication to employees. We believe that the choice of Codes of Conduct (as opposed to 10-K reports) alleviates this concern. Alternatively, a CEO may choose to engage in a cheap talk and mention the fraud risks without underlying intentions to strengthen the firm's anti-fraud

[25] See, for example, `https://na.theiia.org/standards-guidance/Public%20Documents/Sarbanes-Oxley_Section_404_--_A_Guide_for_Management_2nd_edition_1_08.pdf`

controls. This is a valid consideration, as in our analysis we do not test the strength or effectiveness of anti-fraud controls (e.g. subsequent compliance with the anti-fraud policy), focusing on textual communication only. However, this does not change our inferences.

We also acknowledge that in our analysis we proxy fraud predisposition using firm' subsidiaries locations, without controlling for its monitoring policies, organizational structures or other internal controls considerations. We agree that additional information about firm's anti-fraud control structures would have allowed us to get better insights. We will consider using alternative fraud risk proxies for future research.

6.3 Robustness Tests

In this section, we conduct a series of sensitivity tests for our primary results.

First, In Table 7 we report results using an alternative measure of firm's fraud risk, Firm Fraud Tolerance Tol_{Firm}. We remind the reader that this variable is constructed in a way that lower values of this index are associated with riskier countries. Therefore, the negative and significant coefficient β_1 confirms our Hypothesis 1.

Our second test is performed using an alternative set of words to identify CEO's Anti-Fraud Language as described in subsection 5.1. Table 8 reports the results of the regression when the dependent variable $\mathrm{FL_{CEO}}$ is constructed using the *both* criterion.

Similar to the main results, our coefficient of interest on $\mathrm{FR_{Firm}}$ is positive with similar magnitude and strongly significant for k between 5 and 10. Our criterion now restricts the "fraud related" topic to contain both of the keywords. While lower numbers of k are unaffected, the results deteriorate for higher values of k. Naturally, the algorithm now selects very few topics as "fraud related", leading to an underestimation of the composite CEO's Anti-Fraud Language Index. Expectedly, the results for $k = 20$ lose their economic and statistical significance as the CEO's Anti-Fraud Language Index is virtually close to zero.

In addition to k, combination of keywords used to identify fraud related topics is also important. Including too many or too few terms leads to misspecification of the CEO's Anti-Fraud Language Index. The robustness of our results to different keyword combinations in the mid range of k gives us assurance that we selected the appropriate set of keywords to identify fraud topics. Automatic selection of the most representative set of keywords is not trivial and requires the inclusion of higher order effects to specifically capture the word structure when k is high.

Our third test uses an alternative firm fraud measure $\mathrm{FR}_{\mathrm{Firm}}$ based on the Transparency International Corruption Practice Index (i.e. CPI). These results are reported in Table 9, where our dependent variable $\mathrm{FL}_{\mathrm{CEO}}$ is estimated using the *any* criterion. Similarly to the main results of Table 6, the coefficient of $\mathrm{FR}_{\mathrm{Firm}}$ is also positive and strongly significant for k up to 20, and the coefficient of LAF is not statistically different from zero.

In our fourth test, we explicitly use a set of non-fraud-related keywords to investigate whether the relationship between $\mathrm{FL}_{\mathrm{CEO}}$ and $\mathrm{FR}_{\mathrm{Firm}}$ still holds. At the same time we do not consider topics which might be correlated with fraudulent behavior such as financial responsibility of officers/directors, but instead associated with customer orientation. The topic selection for the calculation of the CEO's Anti-Fraud Language Index is carried out through the *any* criterion requesting one of the following keywords in the top 5 words: "customer(s)", "help", "associate(s)". This approach leads to a $\mathrm{FL}_{\mathrm{CEO}}$ that no longer reflects the CEO anti-fraud emphasis. Hence we expect to see little to no association with $\mathrm{FR}_{\mathrm{Firm}}$ as k increases. The results are reported in Table 10 and as expected, the coefficient of $\mathrm{FR}_{\mathrm{Firm}}$ is no longer statistically or economically significant for k greater than 3.

To mitigate potential concerns that industry effects or peer firms behavior and preferences might influence the fraud-related content of CEO communication, we use matched peers methodology following the work of Hoberg and Lewis [2017][26]. This approach resembles their *Industry Similarity* measure which allows us to exclude fraud-related content that is similar between a given firm in our sample and its size-age-industry matched peers. We group all firms into clusters representing similarity based on firm's size (using LTA), firm's age (based on Compustat's variable $LINKDT$) and industry dummies IND. For each of the clusters we calculate the fraud content similarity as the mean of CEO's Anti-Fraud Language Index for all firms in a given bin. We de-mean the $\mathrm{FL}_{\mathrm{CEO}}$ by subtracting the cluster mean for each observation to obtain a measure of abnormal $\mathrm{FL}_{\mathrm{CEO}}$ in excess of the peer-driven one. We then use the abnormal $\mathrm{FL}_{\mathrm{CEO}}$ as the dependent variable in the regression model described in Section 5.3. To avoid collinearity, we omit LTA and IND from the set of covariates. CEO's Anti-Fraud Language Index is calculated according to our main criterion *any* using any three of the following four keywords: "government", "integrity", "gift(s)" and "contact" to identify a fraud-related topic. The results are reported in Table 11 and show that the coefficient of $\mathrm{FR}_{\mathrm{Firm}}$ remains consistently positive and statistically significant.

As an additional test, we also investigate whether potential hidden structures exist. To

[26]Please note that in our regressions we include controls for firm's size (LTA) and industry fixed-effects

verify this, we implement a *k-means* algorithm [Hartigan and Wong, 1979; Forgy, 1965] with 10 clusters. Results for FL_{CEO} computed through the LDA with $k = 5$ are reported in Figure 6. As we can observe, there are no identifiable clusters hence we have additional assurance to our previous robustness tests.

Finally, we thoroughly investigate the direction of causality. While we are not able to consistently collect historical Codes of Conduct prior to 2017[27], we proxy them in the following way. We collect Form 8-K Item 5.05[28] information from SEC Analytics to indicate if a firm made a disclosure of any substantive amendment to its Code of Ethics.[29] We construct the indicator variable $flag_{it}$, which takes value 1 if the firm i files Item 5.05 in year t, and zero otherwise.

We also collect Exhibit 21 for all years between 2005 and 2017 and extract subsidiaries information using the approach described in Section 5.2 to construct $\text{Tol}_{\text{Firm}_{it}}$ and $\text{FR}_{\text{Firm}_{it}}$. Finally, we create a variable $\Delta\text{Tol}_{\text{Firm}_{it}} = \text{Tol}_{\text{Firm}_{it}} - \text{Tol}_{\text{Firm}_{it-1}}$ indicating the net change in fraud tolerance level of firm i in year t as compared to prior year $t - 1$. This net change is indicative of a firm expanding its subsidiaries location profile to a more riskier country. After merging our original data set with financial information with $flag_{it}$ and historical $\text{Tol}_{\text{Firm}_{it}}$ and $\text{FR}_{\text{Firm}_{it}}$, we obtain a panel for 486 firms in our sample running for 13 years from 2005 until 2017, of a total of 5,776 observations, on which our further tests are performed.

Even though $flag_{it}$ is a binary variable and eventually lacks the power to characterize the document structures through topics as FL_{CEO}, we believe it provides a valid proxy to test. Below we report three models which aim at verifying the direction of causality. Equations (7) and (8) are panel regressions in which the outcome is modeled with a probit. Equation (9) is a panel Ordinary Least Square (OLS) model in which $flag_{it}$ is considered as a predictor while the net change in fraud tolerance is the outcome. Moreover, Equations (7) and (8) model the outcome in $t + 1$ and t, respectively. All the regressions were performed in Stata (R) with cluster-robust standard errors to avoid overstating t-statistics due to repeated observations of the same firm within the dataset [Petersen, 2009].

[27]Firms typically publish only the most current Code of Conduct on their websites.

[28]Item 5.05 - Amendments to the Registrant's Code of Ethics, or Waiver of a Provision of the Code of Ethics - requires disclosure of any substantive amendment to a company's Code of Ethics adopted by the SEC to satisfy Item 406 of Regulation S-K. [SEC, 2003]

[29]As discussed in Section 2.1, the companies may bifurcate their Codes of Ethics and Codes of Conduct, however, Item 5.05 is only filed to indicate amendments to the Code of Ethics.

$$\mathcal{P}(flag_{it+1}) = \alpha_0 + \alpha_1 \cdot \Delta\mathrm{Tol}_{\mathrm{Firm}_{it}} + \sum \boldsymbol{\alpha} \cdot \mathbf{Control}_{\mathrm{FIN}_{it}} + \mathrm{Tol}_{\mathrm{Firm}_{it}} + \mathrm{FR}_{\mathrm{Firm}_{it}}) + \rho_{it}, \tag{7}$$

$$\mathcal{P}(flag_{it}) = \beta_0 + \beta_1 \cdot \Delta\mathrm{Tol}_{\mathrm{Firm}_{it}} + \sum \boldsymbol{\beta} \cdot (\mathbf{Control}_{\mathrm{FIN}_{it}} + \mathrm{Tol}_{\mathrm{Firm}_{it}} + \mathrm{FR}_{\mathrm{Firm}_{it}}) + \epsilon_{it}, \tag{8}$$

$$\Delta\mathrm{Tol}_{\mathrm{Firm}_{it}} = \gamma_0 + \gamma_1 \cdot flag_{it-1} + \sum \boldsymbol{\gamma} \cdot \mathbf{Control}_{\mathrm{FIN}_{it}} + \mu_{it}, \tag{9}$$

where $\mathrm{Control}_{\mathrm{FIN}}$, $\mathrm{Tol}_{\mathrm{Firm}}$, $\mathrm{FR}_{\mathrm{Firm}}$ are described in Section 5.3.

If our hypothesis is correct and it is the nature of the business which is driving the ethics tone from the top, and not the other way around, we expect that expanding to a more corrupt country in year t will not have immediate effect on the anti-fraud communication channels. Conversely, it will take some time to adjust and be associated with a higher probability of changing the firm's Code of Conduct in year $t+1$, thus not the same year t. At the same time, we do not anticipate to see the tone drive any business changes and thus do not expect the coefficient γ_1 to be statistically different from zero.

The coefficients of interest are α_1, β_1 and γ_1. We find that α_1 is negative and weakly significant ($p < 0.1$), while β_1 and γ_1 are not statistically different from zero, indicating that when a firm's fraud tolerance change is negative (e.g. a firm expands to a more corrupt location), the probability of having a change in its next year Code of Ethics is higher. At the same time the Code of Conduct does not change immediately; the change in the Code doesn't appear to drive the change in fraud tolerance of the business which is consistent with our arguments above. Results of the models in Equations (7), (8), and (9) are reported in Table 12.

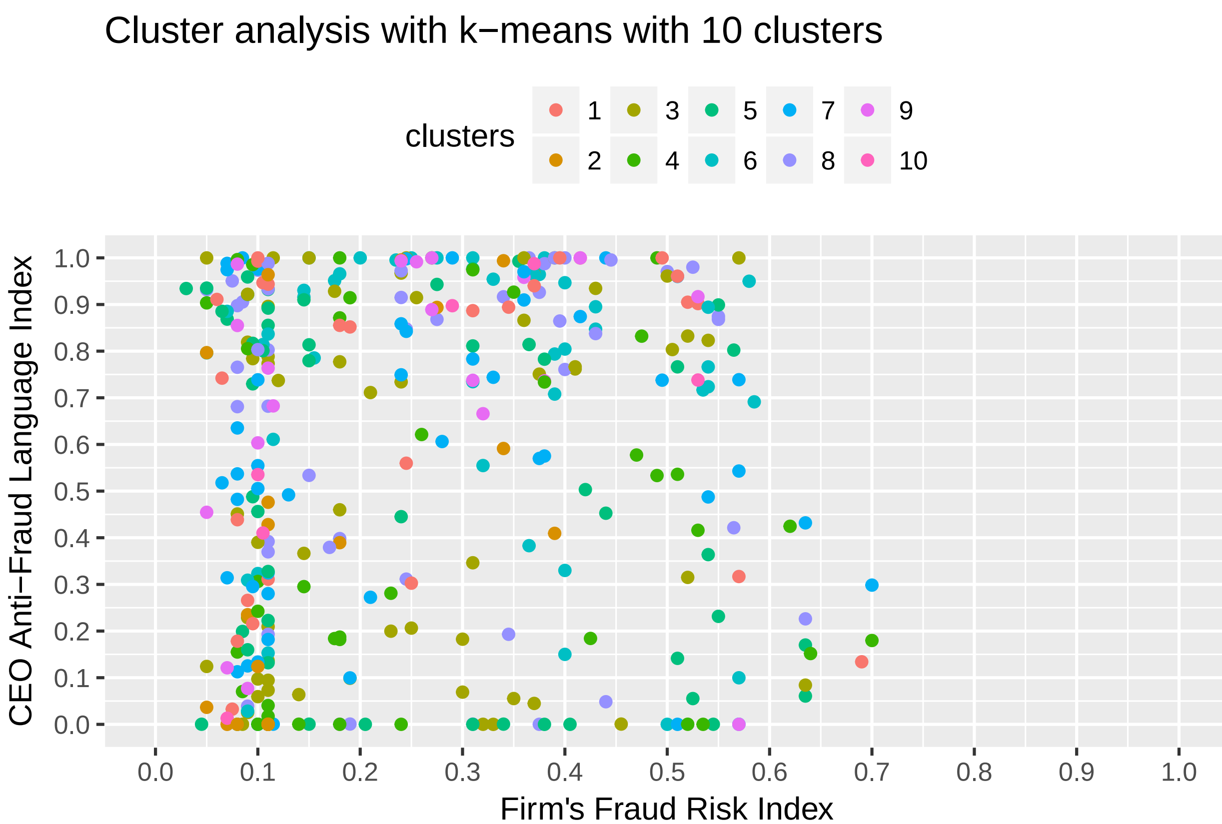
To conclude, these results reaffirm prior arguments on the relationship between firm's fraud tolerance and CEO's anti-fraud language index.

7 Conclusions

In this paper we investigate whether information content of a firm's anti-fraud communication reflects geographically-driven risk exposure that a firm faces.

We introduce a novel measure of CEO's anti-fraud language index which reflects the CEO's emphasis on the importance of anti-fraud communication within the firm by analyzing linguistic characteristics of Codes of Conduct of S&P 500 firms. This type of communication is targeted at firm employees, in which the management communicates their fraud-related

Figure 6: Visualization of clustering analysis using *k-means* algorithm with 10 clusters on top of the LDA model with $k = 5$.



concerns in a linguistically liberated and less legally restrictive way.

We use Latent Dirichlet Allocation topic modeling approach to develop a quantitative measure of CEO's Anti-Fraud Language Index. This approach allows us to infer CEO's personal preferences without directly observing them and thus avoid costly experiments, interviews or surveys.

An important step of our approach includes lexical identification of fraud-related topics using various combinations of fraud-related terms. This is done to balance the mathematical appropriateness of the estimation algorithm with an adequate level of human interpretation. We develop and adopt different criteria to automatically detect a fraud related topic. The robustness of our results gives us assurance that we select the appropriate set of keywords to identify fraud topics.

We estimate firm's fraud risk by using two different measures (FR_{Firm} and Tol_{Firm}) as well as two different corruption indexes of the countries in which the firm has subsidiaries. We perform robustness tests using non-fraud-related topics and peer matching methodology.

We find that CEO's anti-fraud language index is positively associated with the firm's fraud risk based on subsidiaries locations. This confirms our Hypothesis 1 that the CEO tends to emphasize fraud related topics more when the firm operates in countries with higher corruption. At the same time, we find no evidence in support of Hypothesis 2 that the fraud language is driven by the auditor procedures.

We acknowledge that this study is subject to several limitations. First, the dependent variable captures textual anti-fraud emphasis rather than actual anti-fraud policies. This methodology may be subject to a bias associated with the selection of fraud-related topics and the parameters of the algorithm[30]. However, utilizing an alternative set of words as well as a non-fraud vocabulary in our robustness tests allows us to partially mitigate this concern. Second, the instrumental variable approach to the auditor fees likely does not capture all of the relevant factors[31] that affect auditor procedures in practice. However, our inferences are formulated based on the coefficient of the fraud risk, and not the audit fees. The economic and statistical significance of our main coefficient remains robust to various additional specifications.

[30]We acknowledge that LDA topic modeling algorithm is subject to reviewer bias and construct validity considerations, and reliability and validity cannot be taken for granted. Maier et al. [2018] provide an excellent overview of the challenges presented by use of LDA in communication research which we fully concur with: appropriate pre-processing of the text collection, adequate selection of model parameters, including the number of topics to be generated, evaluation of the model's reliability and the process of validly interpreting the resulting topics. Our approaches are similar to those employed in related literature. Overall, as Maier et al. [2018] note, "a standard methodology for ensuring the reliability of the topics has yet to be developed in communication research".

[31]The full discussion of the economic determinants of audit fees is beyond the scope of this paper.

A List of Countries and 2016 Corruption Indexes

Table 3: List of countries with respective 2016 corruption indexes.

Country	ISO3	WGI	CPI	Country	ISO3	WGI	CPI
Afghanistan	AFG	3.37	15	China	CHN	49.04	40
Albania	ALB	41.35	39	Colombia	COL	44.23	37
Algeria	DZA	27.4	34	Comoros	COM	31.73	24
Angola	AGO	5.77	18	Costa Rica	CRI	75.48	58
Argentina	ARG	46.15	36	Côte d'Ivoire	CIV	33.65	34
Armenia	ARM	32.69	33	Croatia	HRV	62.5	49
Australia	AUS	93.27	79	Cuba	CUB	60.58	47
Austria	AUT	91.35	75	Cyprus	CYP	77.88	55
Azerbaijan	AZE	17.79	30	Czech Republic	CZE	67.79	55
Bahamas, The	BHS	82.69	66	Denmark	DNK	99.04	90
Bahrain	BHR	56.25	43	Djibouti	DJI	30.29	30
Bangladesh	BGD	21.15	26	Dominica	DMA	70.19	59
Barbados	BRB	87.98	61	Dominican Republic	DOM	22.6	31
Belarus	BLR	47.6	40	Ecuador	ECU	29.33	31
Belgium	BEL	92.31	77	Egypt, Arab Rep.	EGY	32.21	34
Benin	BEN	36.54	36	El Salvador	SLV	33.17	36
Bhutan	BTN	83.17	65	Eritrea	ERI	11.54	18
Bolivia	BOL	26.92	33	Estonia	EST	84.62	70
Bosnia and Herzegovina	BIH	37.02	39	Ethiopia	ETH	39.9	34
Botswana	BWA	80.29	60	Finland	FIN	99.52	89
Brazil	BRA	38.46	40	France	FRA	90.38	69
Brunei Darussalam	BRN	72.60	58	Gabon	GAB	24.52	35
Bulgaria	BGR	51.44	41	Gambia, The	GMB	22.12	26
Burkina Faso	BFA	53.37	42	Germany	DEU	93.75	81
Burundi	BDI	10.58	20	Ghana	GHA	50.96	43
Cambodia	KHM	8.17	21	Greece	GRC	56.73	44
Cameroon	CMR	11.06	26	Grenada	GRD	69.71	56
Canada	CAN	95.19	82	Guatemala	GTM	25.48	28
Central African Republic	CAF	9.13	20	Guinea	GIN	14.9	27
Chad	TCD	4.81	20	Guinea-Bissau	GNB	3.85	16
Chile	CHL	82.21	66	Guyana	GUY	45.19	34
Haiti	HTI	7.21	20	Malta	MLT	75.96	55
Honduras	HND	27.88	30	Mauritania	MRT	21.63	27
Hong Kong SAR, China	HKG	91.83	77	Mauritius	MUS	65.38	54

Hungary	HUN	61.06	48	Mexico	MEX	23.08	30
Iceland	ISL	95.67	78	Moldova	MDA	14.42	30
India	IND	47.12	40	Mongolia	MNG	35.58	38
Indonesia	IDN	42.79	37	Morocco	MAR	52.88	37
Iran, Islamic Rep.	IRN	25.96	29	Mozambique	MOZ	18.27	27
Iraq	IRQ	6.25	17	Montenegro	MON	54.33	45
Ireland	IRL	92.79	73	Myanmar	MMR	30.77	28
Israel	ISR	81.73	64	Namibia	NAM	65.87	52
Italy	ITA	59.62	47	Nepal	NPL	23.56	29
Jamaica	JAM	51.92	39	Netherlands	NLD	94.71	83
Japan	JPN	90.87	72	New Zealand	NZL	100	90
Jordan	JOR	64.42	48	Nicaragua	NIC	17.31	26
Kazakhstan	KAZ	20.67	29	Niger	NER	31.25	35
Kenya	KEN	16.83	26	Nigeria	NGA	13.46	28
Korea, Dem. Rep.	PRK	5.29	12	Norway	NOR	98.08	85
Korea, Rep.	KOR	66.83	53	Pakistan	PAK	19.23	32
Kuwait	KWT	50	41	Panama	PAN	36.06	38
Kyrgyz Republic	KGZ	12.02	28	Papua New Guinea	PNG	15.87	28
Lao PDR	LAO	15.38	30	Paraguay	PRY	25	30
Latvia	LVA	67.31	57	Peru	PER	43.27	35
Lebanon	LBN	13.94	28	Philippines	PHL	34.13	35
Lesotho	LSO	57.69	39	Poland	POL	76.44	62
Liberia	LBR	26.44	37	Portugal	PRT	80.77	62
Libya	LBY	2.88	14	Qatar	QAT	79.81	61
Lithuania	LTU	73.08	59	Romania	ROM	58.17	48
Luxembourg	LUX	97.6	81	Russian Federation	RUS	18.75	29
Macedonia, FYR	MKD	46.63	37	Rwanda	RWA	74.52	54
Madagascar	MDG	16.35	26	São Tomé and Principe	STP	55.29	46
Malawi	MWI	24.04	31	Saudi Arabia	SAU	62.98	46
Malaysia	MYS	61.54	49	Senegal	SEN	57.21	45
Kosovo	LWI	40.38	36	Sierra Leone	SLE	20.19	30
Maldives	MDV	28.85	36	Singapore	SGP	97.12	84
Mali	MLI	29.81	32	Slovak Republic	SVK	63.46	51
Slovenia	SVN	77.40	61	Trinidad and Tobago	TTO	48.56	35
Solomon Islands	SLB	43.75	42	Tunisia	TUN	53.85	41
Somalia	SOM	0.48	10	Turkey	TUR	50.48	41
South Africa	ZAF	60.1	45	Turkmenistan	TKM	4.33	22
Spain	ESP	68.75	58	Uganda	UGA	12.98	25
Sri Lanka	LKA	48.08	36	Ukraine	UKR	19.71	29
St. Lucia	LCA	70.67	60	United Arab Emirates	ARE	88.46	66

St. Vincent	VCT	74.04	60	United Kingdom	GBR	94.23	81
Suriname	SUR	44.71	45	United States of America	USA	89.90	74
Sweden	SWE	98.56	88	Uruguay	URY	89.42	71
Switzerland	CHE	96.15	86	Uzbekistan	UZB	10.1	21
Syrian Arab Republic	SYR	2.4	13	Venezuela, RB	VEN	6.73	17
Taiwan, China	TWN	78.85	61	Vietnam	VNM	41.83	33
Tajikistan	TJK	12.5	25	Serbia	SCG	45.67	42
Tanzania	TZA	35.1	32	Yemen, Rep.	YEM	0.96	14
Thailand	THA	40.87	35	Zambia	ZMB	42.31	38
Timor-Leste	TLS	34.62	35	Zimbabwe	ZWE	8.65	22

B Tables of Results

Table 4: Summary Statistics

Summary statistics are reported for our sample of 494 firm observations. Columns from left to right report the mean, the standard deviation, the minimum value, the first, second (median), and third quartile, and the maximum value of the respective covariate.

Covariate	**Mean**	**Std. Dev.**	**Min**	**Q1**	**Median**	**Q3**	**Max**
$\mathrm{FL_{CEO}}$ (k=3)	0.5	0.36	0	0.13	0.51	0.85	1
$\mathrm{FL_{CEO}}$ (k=3)	0.55	0.37	0	0.15	0.65	0.91	1
$\mathrm{FL_{CEO}}$ (k=10)	0.57	0.37	0	0.17	0.69	0.91	1
$\mathrm{FL_{CEO}}$ (k=20)	0.49	0.36	0	0.14	0.51	0.82	1
$\mathrm{FL_{CEO}}$ (k=50)	0.3	0.34	0	0	0.15	0.62	1
$\mathrm{Tol_{Firm}}$	0.28	0.29	0.01	0.05	0.19	0.38	0.89
$\mathrm{FR_{Firm}}$	0.26	0.18	0.03	0.11	0.21	0.39	0.74
log(Audit Fees) (LAF)	15.72	0.86	13.59	15.12	15.67	16.24	18.38
log(Total Assets) (LTA)	23.79	1.29	21.3	22.84	23.63	24.55	28.54
CR	3.08	18.14	0	0.97	1.42	2.08	371.57
CA_TA	0.34	0.22	0.01	0.16	0.32	0.49	0.99
ARINV	0.2	0.18	0	0.07	0.14	0.27	0.93
ROA	0.09	0.08	-0.27	0.04	0.08	0.13	0.46
LEV	0.28	0.18	0	0.15	0.26	0.39	1.65
INTANG	0.25	0.24	0	0.03	0.19	0.41	0.89
SEG	0.72	0.73	0	0	0.69	1.39	2.3
CEO Age	56.16	6.19	32	52.75	56	60	92
CEO Duality	0.45	0.5	0	0	0	1	1
log($\mathrm{CEO_{Comp}}$)	9.11	0.71	4.61	8.76	9.2	9.51	11.49
Rate of Occurrence							
LOSS	10%						
MERGER	5%						
BUSY	75%						
MATWEAK	2%						
OPINION	0%						
CEO GENDER	96%						

[32] *Variable Definitions:* LAF = logarithm of audit fees, LTA = logarithm of total assets. $\mathrm{FL_{CEO}}$ is the CEO's Anti-Fraud Language Index calculated for a different number of topics k. $\mathrm{Tol_{Firm}}$ and $\mathrm{FR_{Firm}}$ are firm's fraud characteristics based on the location of its subsidiaries. CR = current assets divided by current liabilities, CA_TA = current assets divided by total assets, ARINV = sum of accounts receivable and inventory divided by total assets, ROA = earnings before interest and taxes divided by total assets. LOSS = 1 if firm incurred a loss, 0 otherwise, LEV = long term debt divided by total assets, INTANG = intangible assets divided by total assets. CEO Duality = 1 if the CEO is the Chairman of the Board, 0 otherwise. MERGER = 1 if the firm reported an impact of a merger or acquisition on net income, 0 otherwise. BUSY = 1 if the firm's fiscal year is in December, 0 otherwise. SEG = logarithm of number of business segments. MATWEAK = 1 if the client receives a material weakness opinion, 0 otherwise. OPINION = 1 if the auditor issued a going concern opinion, 0 otherwise. CEO GENDER = 1 if CEO is male, 0 otherwise.

Table 5: Pearson Correlation Coefficients

Pearson Correlation Coefficients are reported for the main variables for our sample of 494 observations. FL_{CEO} is the CEO's Anti-Fraud Language Index calculated for a different number of topics k. Tol_{Firm} and FR_{Firm} are firm's fraud characteristics based on the location of its subsidiaries as described in Section 5. Financial and CEO covariates are obtained from Compustat, Execucomp and Audit Analytics. $*p<0.1$; $**p<0.05$; $***p<0.01$

.

	FL_{CEO} (k=3)	FL_{CEO} (k=5)	FL_{CEO} (k=10)	FL_{CEO} (k=20)	FL_{CEO} (k=50)	Tol_{Firm}	FR_{Firm}	log(Audit Fees)	log(Total Assets)	CR	CA_TA	ARINV	ROA	LEV	LOSS	INTANG	CEO Age	CEO Duality
FL_{CEO}	0.85***																	
FL_{CEO} k=10	0.90***	0.85***																
FL_{CEO} k=20	0.72***	0.68***	0.75***															
Tol_{Firm}	-0.21***	-0.25***	-0.21***	-0.19***	-0.09**													
FR_{Firm}	0.10**	0.09**	0.12***	0.14***	0.03	-0.56***												
log(Audit Fees)	0.38***	0.42***	0.40***	0.31***	0.16***	-0.21***	0.01											
log(Total Assets)	0.14***	0.12***	0.17***	0.12***	0.07	-0.01	0.11***	0.68***										
CR	-0.00	-0.01	-0.01	0.00	0.02	0.01	-0.00	-0.01	0.06									
CA_TA	-0.03	0.02	-0.02	-0.00	0.01	-0.03	-0.14***	0.08*	-0.03	-0.04								
ARINV	-0.06	-0.04	-0.05	-0.05	-0.01	-0.01	0.01	0.09**	0.08*	-0.04	0.74***							
ROA	0.02	0.05	0.01	0.04	0.00	-0.03	-0.12***	-0.16***	-0.40***	-0.03	0.29***	0.13***						
LEV	0.05	0.03	0.05	0.04	0.02	0.01	0.04	-0.08*	-0.17***	-0.11**	-0.33***	-0.24***	0.06					
LOSS	0.02	0.03	0.00	0.01	0.03	-0.02	0.11**	-0.02	0.00	-0.02	-0.18***	-0.18***	-0.45***	0.04				
INTANG	0.13***	0.19***	0.13***	0.07*	0.04	-0.13***	-0.15***	0.16***	-0.16***	-0.07*	-0.21***	-0.18***	0.13***	0.16***	-0.10**			
CEO Age	-0.07	-0.05	-0.04	-0.09*	-0.09**	0.01	0.04	0.03	0.08*	-0.01	-0.03	0.03	-0.02	0.00	-0.03	-0.02		
CEO Duality	0.09*	0.07	0.07	0.10**	0.03	-0.04	0.05	0.14***	0.15***	0.05	-0.01	0.03	-0.07	-0.06	-0.00	-0.07	0.34***	
log($CEO_{C}omp$)	0.14***	0.16***	0.17***	0.14***	0.00	-0.07	0.02	0.33***	0.19***	-0.02	0.10**	0.03	0.00	-0.00	0.05	0.10**	0.09**	0.26***

Table 6: CEO's Anti-Fraud Language Index Regressions

Results using main specification (criterion "any") for CEO's Anti-Fraud Language Index FL_{CEO} calculated for a different number of topics k. The dependent variable is constructed when the top 15 words contain any 3 of the following four words: "government", "integrity", "gift(s)", "contact". Key independent variable is firm's fraud risk FR_{Firm} based on the location of its subsidiaries as described in Section 5. Equations 5 and 6 are estimated using two-stage least squares. LAF is instrumented using variables MERGER, BUSY, SEG, MATWEAK and OPINION as described in Section 5.3. Results are reported for the second stage regression, different sets of financial, CEO and industry controls are included in the regressions where specified. $*p<0.1$; $**p<0.05$; $***p<0.01$.

	Dependent variable: CEO's Anti-Fraud Language Index FL_{CEO}											
	(k=3)			(k=5)			(k=10)			(k=20)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Fraud Risk	0.20^{**}	0.20^{**}	0.23^{**}	0.20^{**}	0.27^{***}	0.29^{***}	0.24^{***}	0.26^{***}	0.26^{***}	0.30^{***}	0.32^{***}	0.35^{***}
	(0.08)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.10)
LAF	0.13^{*}	0.03	0.07	0.09	−0.05	0.003	0.15^{*}	0.07	0.14	0.05	−0.02	−0.04
	(0.08)	(0.18)	(0.18)	(0.08)	(0.18)	(0.18)	(0.08)	(0.18)	(0.18)	(0.08)	(0.18)	(0.18)
LTA		0.06	0.03		0.12	0.08		0.06	0.004		0.09	0.09
		(0.10)	(0.10)		(0.11)	(0.10)		(0.10)	(0.10)		(0.10)	(0.10)
CR		0.001	0.001		0.001	0.001		0.0005	0.001		0.001	0.001
		(0.001)	(0.001)		(0.001)	(0.001)		(0.001)	(0.001)		(0.001)	(0.001)
CA_TA		0.14	0.06		0.34^{*}	0.26		0.21	0.13		0.28	0.30^{*}
		(0.19)	(0.17)		(0.19)	(0.17)		(0.19)	(0.17)		(0.19)	(0.18)
ARINV		−0.20	−0.13		-0.34^{**}	-0.29^{*}		-0.29^{*}	−0.23		-0.36^{**}	-0.39^{**}
		(0.16)	(0.16)		(0.17)	(0.16)		(0.16)	(0.16)		(0.16)	(0.16)
ROA		0.17	0.12		0.25	0.12		−0.04	−0.15		−0.01	−0.09
		(0.29)	(0.30)		(0.30)	(0.30)		(0.29)	(0.30)		(0.29)	(0.31)
LOSS		0.03	−0.01		0.11	0.08		0.03	−0.003		0.04	0.03
		(0.07)	(0.07)		(0.07)	(0.07)		(0.07)	(0.07)		(0.07)	(0.07)
LEV		0.10	0.06		0.12	0.07		0.08	0.03		0.11	0.09
		(0.10)	(0.10)		(0.11)	(0.10)		(0.10)	(0.10)		(0.10)	(0.10)
INTANG		0.08	0.05		0.22	0.16		0.06	0.01		0.07	0.10
		(0.18)	(0.16)		(0.18)	(0.17)		(0.18)	(0.16)		(0.18)	(0.17)
CEO Age			-0.005^{*}			−0.003			−0.003			-0.01^{**}
			(0.003)			(0.003)			(0.003)			(0.003)
CEO Duality			0.05			0.04			0.03			0.10^{***}
			(0.03)			(0.04)			(0.04)			(0.04)
CEO Gender			−0.01			0.01			0.0000			−0.03
			(0.08)			(0.08)			(0.08)			(0.08)
log(CEO Comp)			0.02			0.02			0.03			0.04
			(0.04)			(0.04)			(0.04)			(0.04)
IND	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Weak instruments	0	0.02	0.02	0	0.02	0.02	0	0.02	0.02	0	0.02	0.02
Sargan	0.21	0.65	0.66	0.13	0.92	0.94	0.26	0.79	0.87	0.09	0.77	0.8
Observations	494	494	452	494	494	452	494	494	452	494	494	452
Adjusted R^2	0.15	0.19	0.24	0.14	0.18	0.24	0.17	0.23	0.27	0.08	0.15	0.16

Table 7: Robustness Tests with Alternative Fraud Risk Measure - Fraud Tolerance

Results using main specification (criterion "any") for CEO's Anti-Fraud Language Index FL_{CEO} calculated for a different number of topics k. The dependent variable is constructed when the top 15 words contain any 3 of the following four words: "government", "integrity", "gift(s)", "contact". Key independent variable is firm's fraud tolerance Tol_{Firm} based on the location of its subsidiaries as described in Section 5. Equations 5 and 6 are estimated using two-stage least squares. LAF is instrumented using variables MERGER, BUSY, SEG, MATWEAK and OPINION as described in Section 5.3. Results are reported for the second stage regression, different sets of financial, CEO and industry controls are included in the regressions where specified. $*p<0.1$; $**p<0.05$; $***p<0.01$.

	Dependent variable: CEO's Anti-Fraud Language Index FL_{CEO}											
	(k=3)			(k=5)			(k=10)			(k=20)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Fraud Tolerance	-0.17^{**}	-0.15^{*}	-0.14^{*}	-0.26^{***}	-0.19^{**}	-0.17^{**}	-0.17^{**}	-0.14^{*}	-0.12	-0.20^{***}	-0.18^{**}	-0.19^{**}
	(0.07)	(0.08)	(0.08)	(0.07)	(0.08)	(0.08)	(0.07)	(0.08)	(0.08)	(0.07)	(0.08)	(0.08)
LAF	0.14^{*}	0.06	0.09	0.10	-0.02	0.02	0.16^{**}	0.10	0.16	0.06	0.001	-0.03
	(0.07)	(0.17)	(0.17)	(0.08)	(0.17)	(0.18)	(0.08)	(0.17)	(0.17)	(0.08)	(0.17)	(0.18)
LTA		0.05	0.03		0.10	0.07		0.04	-0.003		0.08	0.08
		(0.09)	(0.09)		(0.10)	(0.09)		(0.10)	(0.09)		(0.10)	(0.10)
CR		0.001	0.001		0.001	0.001		0.0005	0.001		0.001	0.001
		(0.001)	(0.001)		(0.001)	(0.001)		(0.001)	(0.001)		(0.001)	(0.001)
CA_TA		0.06	-0.01		0.24	0.18		0.12	0.05		0.19	0.20
		(0.18)	(0.16)		(0.18)	(0.16)		(0.18)	(0.16)		(0.18)	(0.17)
ARINV		-0.14	-0.07		-0.27^{*}	-0.21		-0.22	-0.16		-0.28^{*}	-0.29^{*}
		(0.15)	(0.15)		(0.16)	(0.16)		(0.16)	(0.16)		(0.16)	(0.16)
ROA		0.14	0.09		0.21	0.08		-0.07	-0.17		-0.04	-0.12
		(0.28)	(0.29)		(0.29)	(0.30)		(0.28)	(0.30)		(0.29)	(0.30)
LOSS		0.03	-0.01		0.11^{*}	0.08		0.03	-0.0003		0.04	0.03
		(0.07)	(0.07)		(0.07)	(0.07)		(0.07)	(0.07)		(0.07)	(0.07)
LEV		0.10	0.07		0.12	0.08		0.08	0.04		0.12	0.10
		(0.10)	(0.10)		(0.10)	(0.10)		(0.10)	(0.10)		(0.10)	(0.10)
INTANG		0.02	0.004		0.14	0.10		-0.01	-0.04		-0.001	0.04
		(0.16)	(0.15)		(0.17)	(0.16)		(0.17)	(0.16)		(0.17)	(0.16)
CEO Age			-0.004			-0.003			-0.002			-0.01^{**}
			(0.003)			(0.003)			(0.003)			(0.003)
CEO Duality			0.05			0.03			0.02			0.09^{**}
			(0.03)			(0.04)			(0.04)			(0.04)
CEO Gender			-0.01			0.01			-0.0002			-0.03
			(0.08)			(0.08)			(0.08)			(0.08)
log(CEO Comp)			0.02			0.02			0.03			0.04
			(0.04)			(0.04)			(0.04)			(0.04)
IND	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Weak instruments	0	0.01	0.01	0	0.01	0.01	0	0.01	0.01	0	0.01	0.01
Sargan	0.23	0.59	0.65	0.12	0.84	0.92	0.35	0.73	0.86	0.19	0.72	0.8
Observations	494	494	452	494	494	452	494	494	452	494	494	452
Adjusted R^2	0.16	0.21	0.25	0.18	0.22	0.26	0.17	0.24	0.27	0.09	0.16	0.17

Table 8: Robustness Tests with Alternative Fraud Identification Words

Results using an alternative specification for CEO's Anti-Fraud Language Index reporting the results of the regression when the dependent variable FL_{CEO} is constructed when the top 5 words contain *both* of the words "government" and "integrity" calculated for a different number of topics k. Key independent variable is firm's fraud risk FR_{Firm} based on the location of its subsidiaries as described in Section 5. Equations 5 and 6 are estimated using two-stage least squares. LAF is instrumented using variables MERGER, BUSY, SEG, MATWEAK and OPINION as described in Section 5.3. Results are reported for the second stage regression, different sets of financial, CEO and industry controls are included in the regressions where specified. $*p<0.1$; $**p<0.05$; $***p<0.01$.

	Dependent variable: CEO's Anti-Fraud Language Index FL_{CEO}											
	(k=3)			(k=5)			(k=10)			(k=20)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Fraud Risk	0.20**	0.20**	0.23**	0.19**	0.25***	0.27***	0.16**	0.18**	0.16**	−0.02	0.03	0.02
	(0.08)	(0.09)	(0.09)	(0.09)	(0.09)	(0.10)	(0.07)	(0.08)	(0.08)	(0.05)	(0.05)	(0.05)
LAF	0.13*	0.03	0.07	−0.001	−0.12	−0.07	0.06	0.10	0.15	0.10**	0.15	0.11
	(0.08)	(0.18)	(0.18)	(0.08)	(0.19)	(0.18)	(0.07)	(0.15)	(0.16)	(0.04)	(0.10)	(0.09)
LTA		0.06	0.03		0.13	0.09		−0.02	−0.05		−0.08	−0.05
		(0.10)	(0.10)		(0.11)	(0.10)		(0.09)	(0.09)		(0.06)	(0.05)
CR		0.001	0.001		0.001	0.001		−0.0003	−0.0003		−0.0001	−0.0001
		(0.001)	(0.001)		(0.001)	(0.001)		(0.001)	(0.001)		(0.0005)	(0.0004)
CA_TA		0.14	0.06		0.31	0.19		0.17	0.14		−0.12	−0.04
		(0.19)	(0.17)		(0.20)	(0.18)		(0.16)	(0.15)		(0.11)	(0.09)
ARINV		−0.20	−0.13		−0.27	−0.19		−0.28**	−0.27*		−0.002	−0.05
		(0.16)	(0.16)		(0.17)	(0.16)		(0.14)	(0.14)		(0.09)	(0.08)
ROA		0.17	0.12		0.12	0.10		−0.16	−0.30		−0.06	−0.06
		(0.29)	(0.30)		(0.30)	(0.31)		(0.25)	(0.27)		(0.17)	(0.16)
LOSS		0.03	−0.01		0.08	0.04		0.04	−0.01		−0.09**	−0.07*
		(0.07)	(0.07)		(0.07)	(0.07)		(0.06)	(0.06)		(0.04)	(0.04)
LEV		0.10	0.06		0.09	0.005		0.14	0.09		−0.02	0.03
		(0.10)	(0.10)		(0.11)	(0.10)		(0.09)	(0.09)		(0.06)	(0.05)
INTANG		0.08	0.05		0.25	0.19		0.003	−0.01		−0.15	−0.09
		(0.18)	(0.16)		(0.19)	(0.17)		(0.15)	(0.15)		(0.10)	(0.09)
CEO Age			−0.005*			−0.001			0.001			0.001
			(0.003)			(0.003)			(0.003)			(0.001)
CEO duality			0.05			0.002			0.01			0.02
			(0.03)			(0.04)			(0.03)			(0.02)
CEO gender			−0.01			0.05			0.10			−0.12***
			(0.08)			(0.08)			(0.07)			(0.04)
log(CEO Comp)			0.02			0.06*			0.03			−0.02
			(0.04)			(0.04)			(0.03)			(0.02)
IND	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Observations	494	494	452	494	494	452	494	494	452	494	494	452
Adjusted R^2	0.15	0.19	0.24	0.003	0.01	0.10	0.05	0.12	0.12	-0.11	-0.10	0.0002

Table 9: Robustness Tests with an Alternative Country Corruption Measure

Results using main specification (criterion "any") for CEO's Anti-Fraud Language Index FL_{CEO} calculated for a different number of topics k with an alternative country corruption measure by Transparency International Corruption Perception Index. The dependent variable is constructed when the top 15 words contain any 3 of the following four words: "government", "integrity", "gift(s)", "contact". Key independent variable is firm's fraud risk FR_{Firm} based on the location of its subsidiaries as described in Section 5. Equations 5 and 6 are estimated using two-stage least squares. *LAF* is instrumented using variables MERGER, BUSY, SEG, MATWEAK and OPINION as described in Section 5.3. Results are reported for the second stage regression, different sets of financial, CEO and industry controls are included in the regressions where specified. $*p<0.1$; $**p<0.05$; $***p<0.01$.

	Dependent variable: CEO's Anti-Fraud Language Index FL_{CEO}											
	(k=3)			(k=5)			(k=10)			(k=20)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Fraud Risk	0.30***	0.29**	0.30**	0.33***	0.39***	0.39***	0.34***	0.35***	0.32***	0.41***	0.43***	0.47***
	(0.11)	(0.11)	(0.12)	(0.11)	(0.12)	(0.12)	(0.11)	(0.12)	(0.12)	(0.11)	(0.12)	(0.12)
LAF	0.13*	0.03	0.07	0.09	−0.05	0.001	0.15*	0.07	0.14	0.05	−0.02	−0.05
	(0.08)	(0.17)	(0.18)	(0.08)	(0.18)	(0.18)	(0.08)	(0.18)	(0.18)	(0.08)	(0.18)	(0.18)
LTA		0.06	0.03		0.12	0.08		0.06	0.004		0.09	0.09
		(0.10)	(0.10)		(0.10)	(0.10)		(0.10)	(0.10)		(0.10)	(0.10)
CR		0.001	0.001		0.001	0.001		0.0005	0.0005		0.001	0.001
		(0.001)	(0.001)		(0.001)	(0.001)		(0.001)	(0.001)		(0.001)	(0.001)
CA_TA		0.14	0.06		0.34*	0.26		0.20	0.12		0.28	0.29*
		(0.19)	(0.17)		(0.20)	(0.17)		(0.19)	(0.17)		(0.19)	(0.18)
ARINV		−0.19	−0.13		−0.34**	−0.29*		−0.28*	−0.22		−0.36**	−0.38**
		(0.16)	(0.16)		(0.17)	(0.16)		(0.16)	(0.16)		(0.16)	(0.16)
ROA		0.16	0.11		0.24	0.11		−0.05	−0.16		−0.02	−0.10
		(0.29)	(0.30)		(0.30)	(0.30)		(0.29)	(0.30)		(0.29)	(0.31)
LOSS		0.03	−0.01		0.11	0.08		0.03	−0.002		0.04	0.03
		(0.07)	(0.07)		(0.07)	(0.07)		(0.07)	(0.07)		(0.07)	(0.07)
LEV		0.10	0.06		0.12	0.07		0.08	0.03		0.11	0.08
		(0.10)	(0.10)		(0.11)	(0.10)		(0.10)	(0.10)		(0.10)	(0.10)
INTANG		0.08	0.05		0.22	0.17		0.06	0.01		0.07	0.11
		(0.18)	(0.16)		(0.19)	(0.17)		(0.18)	(0.17)		(0.18)	(0.17)
CEO age			−0.005*			−0.003			−0.003			−0.01**
			(0.003)			(0.003)			(0.003)			(0.003)
CEO duality			0.05			0.04			0.03			0.10***
			(0.03)			(0.04)			(0.04)			(0.04)
CEO gender			−0.01			0.01			0.0001			−0.03
			(0.08)			(0.08)			(0.08)			(0.08)
log(CEO Comp)			0.02			0.02			0.03			0.04
			(0.04)			(0.04)			(0.04)			(0.04)
IND	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Observations	494	494	452	494	494	452	494	494	452	494	494	452
Adjusted R^2	0.15	0.19	0.24	0.15	0.19	0.24	0.17	0.24	0.27	0.09	0.15	0.16

Table 10: Robustness Tests with Non-Fraud Related Topics

Results using main specification (criterion "any") for CEO's Anti-Fraud Language Index FL_{CEO} with an alternative set of non fraud related keywords calculated for a different number of topics k. The dependent variable is constructed when the top 5 words contain any of the following three words: "customer(s)", "help", "associate(s)". Key independent variable is firm's fraud risk FR_{Firm} based on the location of its subsidiaries as described in Section 5. Equations 5 and 6 are estimated using two-stage least squares. LAF is instrumented using variables MERGER, BUSY, SEG, MATWEAK and OPINION as described in Section 5.3. Results are reported for the second stage regression, different sets of financial, CEO and industry controls are included in the regressions where specified. *p<0.1; **p<0.05; ***p<0.01.

	Dependent variable: CEO's Anti-Fraud Language Index FL_{CEO}											
	(k=3)			(k=5)			(k=10)			(k=20)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Fraud Risk	0.20**	0.20**	0.23**	0.01	0.01	0.0004	0.15*	0.13	0.11	0.03	0.06	0.11
	(0.08)	(0.09)	(0.09)	(0.01)	(0.01)	(0.002)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)
LAF	0.13*	0.03	0.07	0.01	0.01	0.0000	0.10	0.23	0.21	0.01	0.04	0.14
	(0.08)	(0.18)	(0.18)	(0.01)	(0.02)	(0.003)	(0.07)	(0.16)	(0.16)	(0.07)	(0.17)	(0.17)
LTA		0.06	0.03		−0.01	0.0001		−0.12	−0.12		−0.02	−0.06
		(0.10)	(0.10)		(0.01)	(0.002)		(0.09)	(0.09)		(0.10)	(0.09)
CR		0.001	0.001		0.0000	0.0000		0.001	0.001		0.0005	0.0004
		(0.001)	(0.001)		(0.0001)	(0.0000)		(0.001)	(0.001)		(0.001)	(0.001)
CA_TA		0.14	0.06		0.002	−0.001		−0.19	−0.15		0.03	0.04
		(0.19)	(0.17)		(0.03)	(0.003)		(0.17)	(0.16)		(0.18)	(0.16)
ARINV		−0.20	−0.13		0.0004	0.001		−0.09	−0.09		0.13	0.07
		(0.16)	(0.16)		(0.02)	(0.003)		(0.14)	(0.14)		(0.15)	(0.15)
ROA		0.17	0.12		−0.08**	0.001		−0.05	−0.10		0.16	0.09
		(0.29)	(0.30)		(0.04)	(0.01)		(0.26)	(0.27)		(0.27)	(0.29)
LOSS		0.03	−0.01		−0.003	0.001		−0.07	−0.08		0.04	0.05
		(0.07)	(0.07)		(0.01)	(0.001)		(0.06)	(0.06)		(0.06)	(0.07)
LEV		0.10	0.06		0.001	0.001		−0.09	−0.10		−0.13	−0.13
		(0.10)	(0.10)		(0.01)	(0.002)		(0.09)	(0.09)		(0.10)	(0.10)
INTANG		0.08	0.05		−0.005	−0.001		−0.18	−0.14		−0.03	−0.09
		(0.18)	(0.16)		(0.02)	(0.003)		(0.16)	(0.15)		(0.17)	(0.16)
CEO age			−0.005*			0.0000			−0.01**			−0.002
			(0.003)			(0.0001)			(0.003)			(0.003)
CEO duality			0.05			0.0005			0.03			0.06*
			(0.03)			(0.001)			(0.03)			(0.03)
CEO gender			−0.01			−0.001			0.002			0.01
			(0.08)			(0.001)			(0.07)			(0.07)
log(CEO Comp)			0.02			−0.001			0.02			−0.06*
			(0.04)			(0.001)			(0.03)			(0.03)
IND	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Observations	494	494	452	494	494	452	494	494	452	494	494	452
Adjusted R^2	0.15	0.19	0.24	-0.02	0.01	-0.04	-0.03	0.07	0.10	-0.003	0.04	0.02

Table 11: Robustness Tests with Size-Age-Industry Clusters

using "abnormal" CEO's Anti-Fraud Language Index (in excess of peer-driven) using matched peer methodology with clusters based on firm's size, age and industry sector. Details of the approach are described in Section 6.3. Key independent variable is firm's fraud risk FR_{Firm} based on the location of its subsidiaries as described in Section 5. Equations 5 and 6 are estimated using two-stage least squares. LAF is instrumented using variables MERGER, BUSY, SEG, MATWEAK and OPINION as described in Section 5.3. Results are reported for the second stage regression, financial and CEO controls are included in the regressions where specified. LTA, and Industry Sector dummies are omitted from the set of covariates to avoid collinearity; $**p<0.05$; $***p<0.01$.

	Dependent variable: CEO's Anti-Fraud Language Index FL_{CEO}											
	(k=3)			(k=5)			(k=10)			(k=20)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Fraud Risk	0.19** (0.08)	0.22** (0.09)	0.23*** (0.09)	0.25*** (0.08)	0.31*** (0.09)	0.32*** (0.09)	0.25*** (0.08)	0.30*** (0.09)	0.29*** (0.09)	0.26*** (0.08)	0.31*** (0.08)	0.34*** (0.09)
LAF	0.01 (0.07)	−0.0005 (0.07)	0.03 (0.08)	−0.02 (0.07)	−0.02 (0.07)	0.01 (0.08)	0.03 (0.07)	0.01 (0.07)	0.06 (0.08)	−0.02 (0.07)	−0.02 (0.07)	−0.02 (0.08)
CR		0.001 (0.001)	0.001 (0.001)		0.001 (0.001)	0.001 (0.001)		0.0005 (0.001)	0.0005 (0.001)		0.001 (0.001)	0.001 (0.001)
CA_TA		0.15 (0.12)	0.12 (0.12)		0.26** (0.12)	0.23** (0.12)		0.22* (0.12)	0.17 (0.12)		0.30** (0.12)	0.33*** (0.12)
ARINV		−0.17 (0.13)	−0.18 (0.13)		−0.24* (0.13)	−0.24* (0.13)		−0.25* (0.13)	−0.26* (0.13)		−0.29** (0.12)	−0.37*** (0.13)
ROA		−0.17 (0.31)	−0.07 (0.30)		−0.23 (0.31)	−0.15 (0.30)		−0.26 (0.31)	−0.11 (0.31)		−0.26 (0.31)	−0.25 (0.30)
LOSS		−0.01 (0.06)	−0.02 (0.06)		0.04 (0.06)	0.03 (0.06)		−0.01 (0.06)	−0.02 (0.06)		−0.01 (0.06)	−0.03 (0.06)
LEV		0.07 (0.09)	0.07 (0.09)		0.08 (0.09)	0.08 (0.09)		0.08 (0.09)	0.07 (0.09)		0.09 (0.09)	0.08 (0.09)
INTANG		0.08 (0.09)	0.08 (0.09)		0.14 (0.10)	0.12 (0.09)		0.09 (0.10)	0.07 (0.09)		0.08 (0.09)	0.10 (0.09)
CEO age			−0.003 (0.003)			−0.001 (0.003)			−0.001 (0.003)			−0.004 (0.003)
CEO duality			0.03 (0.03)			0.02 (0.03)			0.01 (0.03)			0.07** (0.03)
CEO gender			−0.01 (0.07)			−0.004 (0.07)			−0.01 (0.07)			−0.02 (0.07)
log(CEO Comp)			0.01 (0.03)			0.005 (0.03)			0.02 (0.03)			0.02 (0.03)
Observations	489	489	448	489	489	448	489	489	448	489	489	448
Adjusted R^2	0.02	0.002	0.03	-0.02	-0.01	0.03	0.04	0.03	0.06	-0.003	0.001	0.02

Table 12: Robustness Tests Using Changes to Code (Item 5.05) Panel Data

Results using panel regressions with Item 5.05 information described in Section 6.3. $Flag_{it}$ is an indicator variable equal to one if the firm i files Form 8-K Item 5.05 in year t. L and F are lag and lead operators respectively. ΔTol_{Firm} is the net change in Tol_{Firm} of firm i in year t as compared to prior year. Details of the approach are described in Section 6.3. Coefficients of control covariates are supressed for brevity purposes. Eq.7 (1) and Eq.8 (2) are modeled using panel probit. Eq.9 (3) uses panel OLS. These regressions were performed in Stata (R) with cluster-robust standard errors.

Dependent Variable	(1) $p(F.Flag)$	(2) $p(Flag)$	(3) ΔTol_{Firm}
ΔTol_{Firm}	-0.509*	-0.197	-
	(-1.80)	(-0.64)	-
$L.Flag$	-	-	0.001
	-	-	(0.07)
Tol_{Firm}	0.155	0.327	-
	(0.49)	(1.24)	-
FR_{Firm}	0.171	0.287	-
	(0.41)	(0.78)	-
$Control_{FIN}$	Yes	Yes	Yes
Ind. Sector	Yes	Yes	Yes
Observations	3,931	4,283	4,283

Note: t-statistics in parentheses
$*$p<0.1; $**$p<0.05; $***$p<0.01

References

ACFE (2016). Joining forces to manage fraud risk: The Association of Certified Fraud Examiners partners with COSO. `www.acfe.com/uploadedFiles/ACFE_Website/Content/documents/2016-Q4-IRG.pdf`.

Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, pages 488–500.

Balvers, R. J., Gaski, J. F., and McDonald, B. (2016). Financial disclosure and customer satisfaction: Do companies talking the talk actually walk the walk? *Journal of Business Ethics*, 139(1):29–45.

Bandiera, O., Hansen, S., Prat, A., and Sadun, R. (2017). CEO behavior and firm performance. *NBER Working Paper No. 23248*.

Benoit, K. (2018). *Quanteda: Quantitative Analysis of Textual Data.* 10.5281/zenodo.1004683, `http://quanteda.io`, `R` package version 0.99.22 edition.

Blanco, K. (2017). Acting Assistant Attorney General Kenneth Blanco keynote address on FCPA and Anti-Bribery Convention. `wp.nyu.edu/compliance_enforcement/2017/11/10/`.

Blankley, A. I., Hurtt, D. N., and MacGregor, J. E. (2012). Abnormal audit fees and restatements. *Auditing: A Journal of Practice & Theory*, 31(1):79–96.

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3:993–1022.

Borghans, L., Heckman, J. J., Golsteyn, B. H., and Meijers, H. (2009). Gender differences in risk aversion and ambiguity aversion. *Journal of the European Economic Association*, 7(2-3):649–658.

Brown, N. C., Crowley, R. M., and Elliott, W. B. (2016). What are you saying? Using topic to detect financial misreporting. *Available at SSRN: https://ssrn.com/abstract=2803733.*

Canary, H. E. and Jennings, M. M. (2008). Principles and influence in codes of ethics: A centering resonance analysis comparing pre-and post-Sarbanes-Oxley Codes of Ethics. *Journal of Business Ethics*, 80(2):263–278.

Carasco, E. F. and Singh, J. B. (2003). The content and focus of the codes of ethics of the world's largest transnational corporations. *Business and Society Review*, 108(1):71–94.

Carcello, J. V. and Li, C. (2013). Costs and benefits of requiring an engagement partner signature: Recent experience in the United Kingdom. *The Accounting Review*, 88(5):1511–1546.

Chang, J., Gerrish, S., Wang, C., Boyd-Graber, J. L., and Blei, D. M. (2009). Reading tea leaves: How humans interpret topic models. *Advances in Neural Information Processing Systems*, pages 288–296.

Cotton, D. L., Johnigan, S., and Givarz, L. (2016). Committee of Sponsoring Organizations of the Treadway Commission (COSO) Fraud Risk Management Guide. `http://www.coso.org/documents/COSO-Fraud-Risk-Management-Guide-Executive-Summary.pdf`.

Crawford, V. (1998). A survey of experiments on communication via cheap talk. *Journal of Economic theory*, 78(2):286–298.

Crawford, V. P. and Sobel, J. (1982). Strategic information transmission. *Econometrica: Journal of the Econometric Society*, pages 1431–1451.

Crossler, R. E., Long, J. H., Loraas, T. M., and Trinkle, B. S. (2016). The impact of moral intensity and ethical tone consistency on policy compliance. *Journal of Information Systems*, 31(2):49–64.

Cumming, D., Leung, T. Y., and Rui, O. (2015). Gender diversity and securities fraud. *Academy of Management Journal*, 58(5):1572–1593.

Davidson, B. I. and Stevens, D. E. (2012). Can a Code of Ethics improve manager behavior and investor confidence? An experimental study. *The Accounting Review*, 88(1):51–74.

Dechow, P., Ge, W., and Schrand, C. (2010). Understanding earnings quality: A review of the proxies, their determinants and their consequences. *Journal of Accounting and Economics*, 50(2):344–401.

Dechow, P. M., Ge, W., Larson, C. R., and Sloan, R. G. (2011). Predicting material accounting misstatements. *Contemporary Accounting Research*, 28(1):17–82.

Dichev, I. D. and Skinner, D. J. (2002). Large–sample evidence on the debt covenant hypothesis. *Journal of Accounting Research*, 40(4):1091–1123.

Dobler, M. (2008). Incentives for risk reporting—a discretionary disclosure and cheap talk approach. *The International Journal of Accounting*, 43(2):184–206.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3):522–550.

Dowle, M. and Srinivasan, A. (2017). *`data.table`: Extension of `data.frame`. R package version 1.10.4*. `http://CRAN.R-project.org/package=data.table`.

Dranove, D. and Jin, G. Z. (2010). Quality disclosure and certification: Theory and practice. *Journal of Economic Literature*, 48(4):935–963.

Duke, J. C. and Hunt III, H. G. (1990). An empirical examination of debt covenant restrictions and accounting-related debt proxies. *Journal of Accounting and Economics*, 12(1-3):45–63.

Dyck, I., Morse, A., and Zingales, L. (2013). How pervasive is corporate fraud? *Rotman School of Management Working Paper No. 2222608*.

Dyer, T., Lang, M., and Stice-Lawrence, L. (2017). The evolution of 10-k textual disclosure: Evidence from latent dirichlet allocation. *Journal of Accounting and Economics*, 64(2-3):221–245.

Dyreng, S. D. and Lindsey, B. P. (2009). Using financial accounting data to examine the effect of foreign operations located in tax havens and other countries on US multinational firms' tax rates. *Journal of Accounting Research*, 47(5):1283–1316.

Ernst&Young (2016). Effective implementation of COSO's new anti-fraud guidance. `https://www.ey.com/Publication/vwLUAssets/ey-effective-implementation-of-coso-thought-leadership/$FILE/ey-effective-implementation-of-coso-thought-leadership.pdf`.

Farrell, J. (1987). Cheap talk, coordination, and entry. *The RAND Journal of Economics*, pages 34–39.

Farrell, J. and Rabin, M. (1996). Cheap talk. *Journal of Economic Perspectives*, 10(3):103–118.

Forgy, E. W. (1965). Cluster analysis of multivariate data: Efficiency vs. interpretability of classifications. *Biometrics*, 21:768–769.

Garmaise, M. J. and Liu, J. (2005). Corruption, firm governance, and the cost of capital. *Available at SSRN: https://ssrn.com/abstract=644017*.

Gow, I., Kaplan, S., Larcker, D., and Zakolyukina, A. CEO personality and firm policies. *NBER Working Paper No. 22435*.

Gramlich, J. and Whiteaker-Poe, J. (2013). Disappearing subsidiaries: the cases of Google and Oracle. *Available at SSRN: https://ssrn.com/abstract=2229576*.

Grossman, S. J. (1981). The informational role of warranties and private disclosure about product quality. *The Journal of Law and Economics*, 24(3):461–483.

Hanley, K. W. and Hoberg, G. (2018). Dynamic interpretation of emerging risks in the financial sector. *Available at SSRN: https://ssrn.com/abstract=2792943*.

Hartigan, J. A. and Wong, M. A. (1979). Algorithm as 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1):100–108.

Heathfield, S. M. (2018). Code of Conduct: Develop, integrate and disseminate a code of conduct in your organization. `www.thebalance.com/code-of-conduct-1918088`.

Helin, S. and Sandström, J. (2007). An inquiry into the study of corporate Codes of Ethics. *Journal of Business Ethics*, 75(3):253–271.

Henry, E. and Leone, A. J. (2015). Measuring qualitative information in capital markets research: Comparison of alternative methodologies to measure disclosure tone. *The Accounting Review*, 91(1):153–178.

Ho, S. S., Li, A. Y., Tam, K., and Zhang, F. (2015). CEO gender, ethical leadership, and accounting conservatism. *Journal of Business Ethics*, 127(2):351–370.

Hoberg, G. and Lewis, C. (2017). Do fraudulent firms produce abnormal disclosure? *Journal of Corporate Finance*, 43:58–85.

Hollander, M., A Wolfe, D., and Chicken, E. (1973). The one-way layout. *Nonparametric Statistical Methods, Third Edition*, Wiley Online Library:202–288.

Hornik, K. and Grün, B. (2011). topicmodels: An R package for fitting topic models. *Journal of Statistical Software*, 40(13):1–30.

Jia, Y., Lent, L. V., and Zeng, Y. (2014). Masculinity, testosterone, and financial misreporting. *Journal of Accounting Research*, 52(5):1195–1246.

Karpoff, J. M., Lee, D. S., and Martin, G. S. (2017). Foreign bribery: Incentives and enforcement. *Available at SSRN: https://ssrn.com/abstract=1573222*.

Kleiber, C. and Zeileis, A. (2008). *Applied econometrics with R*. Springer Science & Business Media.

Kotter, J. P. (2008). *Corporate culture and performance*. Simon and Schuster.

Langlois, C. C. and Schlegelmilch, B. B. (1990). Do corporate Codes of Ethics reflect national character? Evidence from Europe and the United States. *Journal of International Business Studies*, 21(4):519–539.

Larcker, D. F. and Zakolyukina, A. A. (2012). Detecting deceptive discussions in conference calls. *Journal of Accounting Research*, 50(2):495–540.

Loughran, T. and McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, 66(1):35–65.

Loughran, T. and McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, 54(4):1187–1230.

Maier, D., Waldherr, A., Miltner, P., Wiedemann, G., Niekler, A., Keinert, A., Pfetsch, B., Heyer, G., Reber, U., Häussler, T., et al. (2018). Applying LDA topic modeling in communication research: Toward a valid and reliable methodology. *Communication Methods and Measures*, 12(2-3):93–118.

Mamun, M. A., Syeed, M., and Yasmeen, F. (2015). Are investors rational, irrational or normal? *Journal of Economic & Financial Studies*, 3(04):01–15.

Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics*, pages 380–391.

Moberly, R. (2008). Protecting whistleblowers by contract. *University of Colorado Law Review*, 79:975.

Muslu, V., Radhakrishnan, S., Subramanyam, K., and Lim, D. (2014). Forward-looking MD&A disclosures and the information environment. *Management Science*, 61(5):931–948.

Newman, D., Asuncion, A., Smyth, P., and Welling, M. (2009). Distributed algorithms for topic models. *Journal of Machine Learning Research*, 10(Aug):1801–1828.

Orin, R. M. (2008). Ethical guidance and constraint under the Sarbanes-Oxley Act of 2002. *Journal of Accounting, Auditing & Finance*, 23(1):141–171.

Patelli, L. and Pedrini, M. (2015). Is tone at the top associated with financial reporting aggressiveness? *Journal of Business Ethics*, 126(1):3–19.

Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *The Review of Financial Studies*, 22(1):435–480.

Pickerd, J. S., Summers, S. L., and Wood, D. A. (2014). An examination of how entry-level staff auditors respond to tone at the top vis-à-vis tone at the bottom. *Behavioral Research in Accounting*, 27(1):79–98.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria - `http://www.R-project.org/`.

Roca, T., Orme, W., and Brown, J. (2010). Fear and loathing of the corruption perception index: Does transparency international penalize press freedom? *Available at SSRN: https://ssrn.com/abstract=1694211*.

Rohwer, A. (2009). Measuring corruption: A comparison between the transparency international's corruption perceptions index and the World Bank's worldwide governance indicators. *CESifo DICE Report*, 7(3):42–52.

SEC (2002). Sarbanes-Oxley Act of 2002.

SEC (2003). Final rule: Disclosure required by sections 406 and 407 of the Sarbanes-Oxley Act of 2002.

StataCorp. (2017). Stata Statistical Software: Release 15.

Stevens, B. (1994). An analysis of corporate ethical code studies: “Where do we go from here?”. *Journal of Business Ethics*, 13(1):63–69.

Svensson, J. (2005). Eight questions about corruption. *The Journal of Economic Perspectives*, 19(3):19–42.

Viscusi, W. K. (1978). A note on "lemons" markets with quality certification. *The Bell Journal of Economics*, pages 277–279.

Wang, T. Y., Winton, A., and Yu, X. (2010). Corporate fraud and business conditions: Evidence from IPOs. *The Journal of Finance*, 65(6):2255–2292.

Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis*. Springer Science & Business Media.

Yang, T. and Zhao, S. (2014). CEO duality and firm performance: Evidence from an exogenous shock to the competitive environment. *Journal of Banking & Finance*, 49:534–552.

The Value Relevance of Tone: Evidence from 2006-2015 10-Ks.

Olga Bogachek[1]† Francesco Grossetti[1]

Abstract

We study whether tone measures of MD&A sections of firms' 10-K reports exhibit value relevance over time in addition to the twelve accounting amounts and ten industry indicators studied in the literature. We find that absolute tone and change in tone are positively correlated with the equity price, while abnormal tone is largely discounted by investors. We also find that the equity price is negatively associated with the average number of words per sentence of the MD&A section of 10-K, suggesting that investors appreciate readable and succinct information.

Keywords: value relevance, textual tone, Ohlson model, generalized additive model.

†: Corresponding author: `olga.bogachek@phd.unibocconi.it`
[1]: Bocconi University. Via Sarfatti 25, 20136 - Milan, Italy

1 Introduction

In this study we examine the association between tone characteristics of firm disclosures and the stock price of the company. Prior research has shown that it's not just the numbers, but the associated narratives which are value relevant to the investors, and qualitative disclosures help market participants interpret accounting information [Barth et al., 2017; Huang et al., 2013]. One possible presentation of 10-K may produce significantly more timely relevant information than another 10-K despite similar accounting amounts being discussed, supporting *incremental information* argument of disclosure tone [Merkl-Davies and Brennan, 2007].

We study whether tone and average words per sentence measures of the management discussion and analysis (MD&A) section of annual 10-K reports have exhibited value relevance over time, in addition to standard accounting amounts and industry sectors. Our research covers the period from 2006 to 2015, allowing for a span of 10 years and a large universe of 10-K reports. During this time, value relevant information availability has increased, enriching the information environment for the investors. The economy has also shifted towards more knowledge-based assets and operations, which are more difficult to quantify [Barth et al., 2017].

Our results show that tone and year over year change in tone are positively associated with the equity price, providing empirical evidence that the tone of financial disclosures is value-relevant. We also find that the average words per sentence of the MD&A section of 10-K is negatively associated with equity price. These results suggest that while the investors appreciate the increased information content (tangible or intangible) provided to them in the MD&A section, they want this information to be readable and succinct.

We contribute to the growing literature on qualitative disclosure landscape and value relevance. Our value relevance model was developed closely following the work of Barth et al. [2017]. We follow the definition of value relevance adopted in extant literature and, in particular, follow [Barth et al., 2001] We expand this definition as we add tone characteristics. Therefore, the information we consider comprises twelve accounting amounts, ten industry indicators, three different tone measures, as well average words per sentence of the MD&A section of 10-K .

In our paper, we view equity price as jointly determined by economic fundamentals and the disclosure tone. The tone, in turn, is both economically-driven and discretionary-driven. Accordingly, we decompose tone into a non-discretionary component driven by firm financial fundamentals, and a discretionary component (*abntone*), which could reflect other factors such as managerial incentives, impression management and noise [Huang et al., 2013].

We also contribute to the stream of literature on tone measures and the role of tone in managerial disclosures. We use a positive and negative word classification proposed by Loughran and McDonald [2011]. We review contemporary literature on textual disclosures in Section 2.

We address several important limitations of existing literature. First, prior research has shown that word classification in business communication is dictionary-dependent, and alternative dictionary lists may produce various classification results. We alleviate this concern by adding robustness tests using another widely-used business dictionaries such as Henry [2008]. Second, most value relevance researchers base their inferences on linear regression results, and do not account for non-linearities in the relation between equity price and accounting and textual measures. We address this concern by using generalized additive model approach and show how accounting for non-linearities allows to obtain a better goodness-of-fit even in a linear model.

The remainder of the paper proceeds as follows. Section 2 explains how our study relates to prior research. Section 3 develops the hypotheses. Section 4 describes our research model, the data and sample construction process. Section 5 presents the findings. Section 6 provides a summary and concluding remarks. All tables are presented in the Appendix.

2 Literature Review

There is vast literature on value relevance. We refer to [Barth et al., 2001; Holthausen and Watts, 2001; Francis and Schipper, 1999; Beisland, 2009; Lo and Lys, 2000a; Liu and Ohlson, 2000] as recent extensive reviews on the usefulness of financial information and the association between financial measures and share prices, as well as an overview and criticism of Ohlson and Feltham-Ohlson model. Ota [2003] provides a comprehensive review of both price and return models commonly used in contemporary literature, and explains the pitfalls and differences in obtained results.[1]. [Choi et al., 2006] expanded the baseline model by using a conservatism-correction term based on the properties of past realizations of residual income. Gietzmann and Ostaszewski [2003] provide a theoretical justification for why the functional relationship between earnings and value will be non linear and develop a real options based residual income theory based on Tobin's Q theory.

[1]The problem related with the price model is often referred to as "scale effects" and those with the return model are termed "accounting recognition lag" and "transitory earnings". [Ota, 2003]

Many of the studies commonly find that there has been a decrease in the value relevance of earnings measures, which may have implications for market efficiency and resource allocations [Balachandran and Mohanram, 2011; Collins et al., 1997; Lev and Zarowin, 1999]. Francis and Schipper [1999] found different evidence, but confirm that that after controlling for the scale effects, there are indications of decreasing value relevance. Lev [2018] attributes the deterioration of the usefulness of financial information to the change in valuation model and failure to adjust asset recognition rules to the fundamental shift in value-creating resources to intangible assets. However, Barth et al. [2017] found that the value relevance of the total accounting information does not decrease. [Givoly et al., 2017] find that the information content of accounting numbers to debt holders has increased over time. [Barth et al., 2016] also study the value relevance role of accruals as an additional financial measure and find that depending on their types, accruals may contain value-relevant information about future cash flows.

Decreasing value relevance of accounting measures has led many researchers to explore the value relevance of non-financial information, generally concluding that accounting information alone has only a limited ability to explain a firm's market value and its variations. Matolcsy and Wyatt [2008] provide evidence that the interactions of current earnings with technological innovation are associated with market value. Rajgopal et al. [2003] find that network advantages are important intangible assets that are valued by the market, in spite of not being recognized in financial statements. Lourenço et al. [2014] provide an overview of literature on value relevance of corporate sustainability and environmental performance, and show that the reputation for sustainability leadership has an effect on the market valuation of the firm. Cahan et al. [2016] observe a positive relation between unexpected corporate social responsibility (CSR) disclosure and firm value, and Kaspereit and Lopatta [2016] reveal a positive effect of sustainability performance on firm value of European firms.

Fast-growing and technology-based industries demonstrate value relevance of non-traditional measures. Amir and Lev [1996]; Riley Jr et al. [2003]; Hughes [2000] find that certain non-financial performance indicators are more informative to investors than traditional accounting measures in wireless communication, airline and utility firms. Wang et al. [2016] shows that corporate reputation adds to market values. Aaker and Jacobson [2001] find that changes in brand attitude are associated contemporaneously with stock return and lead accounting financial performance in high-tech markets. Caserio and Napoli [2017] find that Codes of Ethics, acting as a "summarization" of ethical variables, are also value-relevant.

The investors acknowledge the issue of disconnected pieces of financial and non-financial information, and look for ways to improve the quality of available information. For example, the framework of integrated reporting was introduced to promote a more cohesive and efficient approach to corporate reporting by connecting financial and sustainability information. Baboukardos and Rimmel [2016] find strong evidence of a sharp increase of the earnings' valuation coefficient after the mandatory adoption of an integrated reporting approach on the Johannesburg Stock Exchange.

An important non-financial measure that possesses a great amount of information content and provides an important narrative to interpret accounting numbers is the text. The text of 10-K reports, earnings releases, conference calls and other disclosures can be used as an obfuscation[2] mechanism to induce desired reaction from the market. Loughran and McDonald [2016] review recent literature on textual disclosures landscape, increasing use of new measures for narrative disclosures [Lang and Stice-Lawrence, 2015], and discuss how well they work in empirical studies. Dyer et al. [2017] document that over the period between 1996 and 2013, 10-K disclosures increased in length [Loughran and McDonald, 2014], boilerplate language [Lang and Stice-Lawrence, 2015] and stickiness [Brown and Tucker, 2011] and decreased in readability [Miller, 2010], specificity [Hope et al., 2016] and the relative prevalence of informative numbers in the text ("hard" information) [Blankespoor, 2012]. Ball et al. studied the topics prevailing in MD&A. Dyer et al. [2017] finds that that the topics of fair value, internal controls, and risk factor disclosures account for virtually all of the increase in length of the 10-Ks. Hoberg and Lewis [2017] find that fraudulent firms' 10-K disclosures are different from counterfactuals.

Recent academic research demonstrates that the information content of earnings press releases has increased in recent decade [Collins et al., 2009; Landsman and Maydew, 2002; Lo and Lys, 2000b; Francis et al., 2002a,b]. Davis et al. [2006] find that managers use optimistic/pessimistic language choice in earnings releases to provide credible information about expected future firm performance to the market, and the market responds to managers' language usage. Readability of disclosure information is associated with a more rapid response to the information content of 10-K filings and less readable 10-K reports are associated with higher stock price crash risk [Kim et al., 2018]. Frankel et al. [2017] find using a combination of machine-learning and traditional text-based measures that analysts revise their forecasts based on relevant information in prior quarter conference calls. Capital markets reaction is more pronounced to more specific disclosures [Hope et al., 2016]. Lee et al. [2014] show that

[2]For instance, 10-Ks of firm's with lower earnings are longer and harder to read, and in general bad news disclosures are less readable than good news [Li, 2008; Dyer et al., 2017; Asay et al., 2018]), however, complex language can simply reflect the provision of complex information such as technical disclosures, and in fact the information component of the disclosure may reduce the information asymmetry [Bushee et al., 2018].

using text from 8-K disclosures boosts prediction accuracy of stock prices over 10% (relative) over a strong baseline that incorporates many financially-rooted features, and the impact is most important in the short term. Kogan et al. [2009] uses support vector regression to predict stock volatility based on 10-K texts. Engelberg [2008] shows that linguistic information has greater long-term predictability for asset prices than quantitative information.

The recent literature on information content of disclosure tone is also evolving. Bakarich et al. [2019] show that firms' 10-K disclosures become less complex, less ambiguous, and more optimistic as they progress from the introduction to the maturity stage, at which clarity and readability peak. Durnev and Mangen [2011] show that the tone of restatement text carries new information to the investors. Files et al. [2018] provide evidence that firms alter the tone of financial disclosures following negative events such as fraudulent activities. In the area of abnormal tone, Huang et al. [2013] provide an extensive literature review. They also construct a measure of tone management from residuals of a tone model that controls for firm quantitative fundamentals and find that abnormal tone predicts negative future earnings and cash flows, and has a positive stock return effect at the earnings announcement. Loughran and McDonald [2011] linked the tone of the annual report with 10-K filing returns, trading volume, and post-filing stock market volatility while Feldman et al. [2010] related changes in tone in the MD&A section to contemporaneous returns around SEC filing date. At the same time Loughran and McDonald [2017]; Drake et al. [2015] show that 10-K reports on SEC EDGAR are accessed very infrequently, especially those of small firms.

An important inspiration to our analysis is the study by Feldman et al. [2010], who show that the tone change of the MD&A section has incremental information content beyond accruals and earnings surprises. They show that tone is more informative when the information environment surrounding the firm (size and analyst following) is weaker. Their study also emphasizes that the incremental contribution of using qualitative information is not large. We contribute to their analysis.

3 Hypothesis Development

There are a number of reasons underlying the belief that the tone of corporate disclosures can be a tool for a manager either to improve understanding or to obfuscate firm's fundamentals.

Signaling theory can be traced back to the works of Akerlof et al. [1970]. When applied to situations of information asymmetry between a company's insiders and outsiders, the basic premise of signaling theory is that the managers of a high-quality firm want to signal the firm's value to its stakeholders [Magness, 2009] and may use corporate narratives to truthfully convey additional value relevant information [Healy and Palepu, 2001]. One way to do so it to use disclosure rhetoric to facilitate the interpretation of quantitative information. A firm's qualitative characteristics of its earnings press releases, 10-K reports and conference calls may signal value-relevant information to investors. This is known as *incremental information* argument and assumes that managers have no economic incentive to engage in opportunistic reporting [Merkl-Davies and Brennan, 2007].

In alternative view, opportunistic managers might obfuscate failures and emphasize successes to enhance their reputation and compensation, avoiding negative consequences of poor performance. The main argument of *impression management* is that managers use the discretion in corporate communication to manipulate public impression of the company, rather than conveying truthful information. Arena et al. [2015] provide an extensive overview of the two competing views and underlying theories.

Managers can adopt different discretionary strategies in their corporate communication: thematic content, visual and structural presentation of the disclosures, as well as the use of language and verbal tone [Arena et al., 2015].

Disclosure tone, or the use of optimistic vs. pessimistic language, is an important element of communication and affects how markets process information, perceive it and understand Morris et al. [2005]. However, lack of regulation and difficulties in measuring tone allows managers to use it flexibly, resulting in mixed results on the information content of disclosure tone.

According to incremental information theory, the tone of the disclosures may help resolve investor uncertainty and influence the share price response [Huang et al., 2013]. Obviously, the optimism or pessimism of the tone may be affected by various economic drivers. To alleviate this concern, we use a benchmark tone that controls for firm's fundamentals; our abnormal tone measure will capture the effects orthogonal to the underlying quantitative, industry or seasonal components.

Length and readability are found to play an important role in information dissemination. Li [2008] find that the annual reports of firms with lower earnings are more complex (have a higher Fog index and are longer). More complex annual reports are found to significantly reduce small investor trading behavior [Miller, 2010]. Lawrence [2013] find that investors hold shares of companies with better written 10-Ks, but also reported that financially literate individuals (e.g., accountants) are found to be unaffected by the 10-K's readability or document length. Loughran and McDonald [2017] find that since retail investors do not frequently access companies' filings, they are unlikely being affected by the readability of the annual report in the days around the filing date. Lehavy et al. [2011]; Loughran and McDonald [2014] find that as the readability of the annual report decreases, analysts have more difficulty incorporating value relevant information into their earnings forecasts.

Overall, clear communication has been identified as one of the most important intangible resources that provide a firm sustainable competitive advantage. Companies with good communication reputation will see the cost of contracts with governments, suppliers, community representatives and other stakeholders reduced and other benefits. Firms gain clarity by adding value-relevant information in the text of disclosures, adding to the length of MD&A section. At the same time, they do so in a concise and succinct way, avoiding wordy and unnecessary rhetoric, which may, in turn, obfuscate the information.

Based on the above discussion, we expect that equity price will be higher for firms with higher disclosure tone (as measured by absolute tone and year over year change in tone), thus, supporting *incremental information* argument of the disclosure tone. At the same time, we expect the abnormal component of the tone (the one not associated with the underlying financials) to be largely ignored or discounted by the investors. We also expect the equity price to be negatively associated with less readable disclosures. The underlying argument is that the tone and readability of disclosures has information content for investors, and they consider it when valuing securities. Our main hypotheses are formulated as follows:

Hypothesis 1. *Tone and year over year change in tone of MD&A section of 10-K reports are value relevant and are positively associated with equity price. Abnormal tone (tone not associated with the underlying financial information) is not value-relevant for the investors.*

Hypothesis 2. *Average number of words per sentence in MD&A section of 10-K reports is value relevant and negatively associated with equity price.*

4 Research Design

4.1 Abnormal Tone Model

Previous literature measures qualitative textual characteristics of financial reports using various software packages and algorithms, such as Diction (c) or Latent Dirichlet Allocation [Blei et al., 2003]. Loughran and McDonald [2011] argue that word classification for business communication should differ from traditional dictionaries, and compile an alternative words list for describing positive and negative tone in financial communication. We use their word list to classify the frequency of optimistic versus pessimistic words appearing in the MD&A section of annual 10-K report, and construct a variable *tone* as the frequency difference between the positive and negative words scaled by total words in an MD&A report. This is a typical approach widely used in prior literature [Huang et al., 2013].

To determine the influence of abnormal tone on the value relevance, we use the following model to estimate the tone orthogonal to the financial and accounting characteristics[3]. In this stage, we regress the *tone* on equity price P as well as twelve accounting amounts described below. All measures are described in Section 4.2.

We include twelve accounting amounts that prior research [Barth et al., 2017] identifies as value relevant. They include earnings NI, equity book value BVE, operating cash flow CF, cash holdings $CASH$, dividends to common shareholders DIV, research and development expense RD, recognized intangible assets $INTAN$ (including capitalized software), advertising expense ADV, special items SPI, other comprehensive income OCI, revenue REV and total assets ASS. All accounting measures are deflated by shares outstanding. FF is a set of indicator variables for the ten Fama French industry groups. All variables are described in Section 7.1.

We also include ten Fama-French industry sectors FF and perform cross sectional regressions by year. The abnormal tone model we use for this stage is:

$$tone_j = \alpha_0 + \delta_1 P_j + \alpha_1 NI_j + \alpha_2 BVE_j + \alpha_3 CF_j + \alpha_4 CASH_j + \alpha_5 DIV_j + \alpha_6 RD_j + \alpha_7 INTAN_j + \alpha_8 ADV_j + \alpha_9 SPI_j + \alpha_{10} OCI_j + \alpha_{11} REV_j + \alpha_{12} ASS_j + \sum \alpha_{13} \cdot FF_j + \epsilon_j \tag{1}$$

The selected determinants allow to control for information about the firm fundamentals to obtain abnormal tone. Table 3 reports the estimation results of Regression 1 in cross sectional regressions by year. Table 4 reports the estimation results of Regression 1 as a panel. Consistent with Huang et al. [2013], we find that *tone* is positively associated with profitability and equity price. Abnormal tone, *abntone*, is the residual of Regression 1, and will be later used as an independent variable in our second stage analysis in Section 4.2. By construction, *abntone* is designed to be unrelated to key firm fundamentals and industry sectors.

[3] Our model is adapted from recent prior studies (e.g. Huang et al. [2013])

4.2 Model

In the second stage of our analysis, we expand the Ohlson [Ohlson, 1995] and Feltham-Ohlson [Feltham and Ohlson, 1995] price models by including accounting amounts and tone measures.

We include three different disclosure tone measures ($tone\_measure$) in our regressions.

1. variable *tone* is constructed as the frequency difference between the positive and negative words scaled by total words in an MD&A report following Loughran and McDonald [2011] and using their dictionary; $tone = (\#$ positive words $+ \#$ negative words$)$ $/$ total non-numerical words;

2. $delta\_tone$ is a year-over-year change in $tone$ for each firm; and is defined as $delta\_tone_{jt} = tone_{jt} - tone_{jt-1}$;

3. abnormal tone *abntone*, is the residual of Regression 1. We take the residuals from the abnormal tone model discussed in previous section and include them as an independent variable in Equation 2.

We also include the average number of words per sentence *read* as a readability measure which incorporates disclosure length and wordiness. Loughran and McDonald [2014] show that widely used readability measures such as Fog index include a proportion of complex (three syllables and more) words, which may be a noisy measure in financial communication. In unreported results we also included the logarithm of the number of sentences as an alternative measure of readability, and it yielded numerically identical results. All continuous financial variables were winsorized at 0.01 level. All variables are described in details in Appendix 7.1.

Our main value relevance model can be summarized as follows:

$$\begin{aligned} P_{jt} = \beta_0 + \beta_1 NI_{jt} + \beta_2 BVE_{jt} + \beta_3 CF_{jt} + \beta_4 CASH_{jt} + \beta_5 DIV_{jt} + \beta_6 RD_{jt} + \\ \beta_7 INTAN_{jt} + \beta_8 ADV_{jt} + \beta_9 SPI_{jt} + \beta_{10} OCI_{jt} + \beta_{11} REV_{jt} + \beta_{12} ASS_{jt} + \\ \delta_1 tone\_measure_{jt} + \delta_2 read_{jt} + \sum \beta_{13} \cdot FF_j + \sum \beta_{14} \cdot YY_t + \epsilon_{jt} \end{aligned} \tag{2}$$

Finally, we test whether the *tone* exhibits value relevance in cross-sectional analysis by year, and run the following regressions for all years in our sample (2006 through 2015).

$$\begin{aligned} P_j = \beta_0 + \beta_1 NI_j + \beta_2 BVE_j + \beta_3 CF_j + \beta_4 CASH_j + \beta_5 DIV_j + \beta_6 RD_j + \\ \beta_7 INTAN_j + \beta_8 ADV_j + \beta_9 SPI_j + \beta_{10} OCI_j + \beta_{11} REV_j + \beta_{12} ASS_j + \\ \delta_1 tone_j + \delta_2 read_j + \sum \beta_{13} \cdot FF_j + \epsilon_j \end{aligned} \tag{3}$$

4.3 Data and Sample Construction

We obtain the text of MD&A sections of annual 10-K reports from SEC EDGAR database and calculate the number of positive and negative words in each document using the methodology described in Loughran and McDonald [2011]. This allows us to obtain the tone measures for each document. Our MD&A samples consists of 61,884 firm-year observations. We remove duplicate entries (multiple 10-Ks in a year), incomplete years (e.g. 2005 and 2016) and merge them with firm historical financial data from Compustat. In order to obtain the match, we do the following. First, for each CIK we obtain the firm's fiscal year-end month. We add three months to obtain the deadline, by which the 10-K should have been submitted to SEC (we used the maximum deadline for non-accelerated filers defined in https://www.sec.gov/fast-answers/answers-form10khtm.html as 90 days). If the *filing date* on SEC report was filed within the required deadline, we mark this 10-K as relating to prior fiscal year. If the *filing date* exceeds the deadline, we mark the form relating to current fiscal year. We then merge MD&A and Compustat data using *cik* and year.

Following [Barth et al., 2017], we perform the following adjustments. We eliminate firms with missing data on cik, $fyear$, ib, ceq, $prcc\_f$, $csho$, at, lag_at, $revt$ and sic. We replace missing values with zeros for re, lag_re, act, che, lct, dlc, txp, dvc, xrd, $intan$, xad and spi. We winsorize all continuous variables at 0.01 level by year, and impute missing for cf as $(ib-(act-lct))/csho$. Equity price is obtained from CRSP three months after fiscal year end ($prcc\_q$). This results in 53,915 firm-year observations. We exclude financial firms. We omit missing values from our regressions and limit the years to 2006-2015. The intersection of our data constraints results in a sample 22,316 firm-year observations used in panel regressions. Definitions for all of our variables are included in the Appendix 7.1. Section 5.1 provides descriptive sample statistics.

5 Results

5.1 Descriptive Statistics

Table 1 provides descriptive sample statistics. The average firm has a normalized (e.g. deflated by shares outstanding) net earnings of 0.77 and book value of 9.87, which are substantively comparable with the ones reported by Barth et al. [2017]. The average equity price in our sample is 24.83, which is slightly higher than 18.14 reported Barth et al. [2017]. The average *tone* is -0.405 meaning that the disclosure tone in MD&A sections is generally relatively pessimistic, consistent with Loughran and McDonald [2011]. *delta_tone* is -0.006, and *abntone* is 0.0001. Table 2 displays Pearson correlation coefficients between each covariate as well as the significance given as asymptotic p-values.

Figure 1 reports group means and confidence intervals for the three tone measures by years (*fiscal_year*, left) and Fama French industry groups (*FFIND*, right). We find that the average annual tone exhibits more variability than the other tone measures. For example, the absolute *tone* was relatively high in 2006-2007 prior to financial crisis, dropped during the crisis years and is slowly gaining back its positivity, remaining generally pessimistic. The *abntone* is also showing a similar trend with less variability. When grouped by Fama French industry sectors, telephone and TV firms (group 6) exhibit lowest average *tone*, while health care (group 8) have the highest average tone among all industries. The effect is much less pronounced for *delta_tone* or *abntone*.

5.2 Model Estimates

The first stage in our multivariate analysis involves generating the residuals representing *abntone* from the abnormal tone model described above. Table 3 reports the estimation results of Equation 1 as cross sectional regressions by years, and Table 4 reports the estimation results using a panel regression. In the panel regressions, consistent with Petersen [2009], we perform a clustered, robust regression since we have multiple observations of individual firms within the data. The model goodness of fit (R^2 of 2%) is low and consistent with Huang et al. [2013]. Similar to existing literature, we find that *tone* is positively associated with profitability *ni* and equity price P.

We next evaluate the evidence concerning our Hypothesis 1 and Hypothesis 2.

Table 5 reports the empirical results of model formalized in Equation 2. The coefficients of *tone* and *delta_tone* are positive and statistically significant, supporting our Hypothesis 1 and providing empirical evidence that tone is incrementally value-relevant. The coefficient of *abntone* is not statistically diffrent from zero, indicating that the abnormal component of the tone (that is not associated with the underlying financial information) is largely discounted by the investors. We also find that our readability measure of the MD&A section of 10-K (*read*) is negatively associated with equity price. These results suggest that while the investors appreciate the increased amount of information provided to them in the MD&A section, they want this information to be readable and succinct, consistent with diminishing information value of unreadable disclosures. This supports our Hypothesis 2. For comparison purposes in Column (1) of Table 5 we report the results of the value relevance regression with accounting amounts only. The coefficient of earnings (NI) is consistently at 2.47, and book value of equity (BVE) is at 0.52.

Regressions in Table 5 performed in R using package *plm* with cluster-robust standard errors to avoid overstating t-statistics due to repeated observations of the same firm within the dataset [Petersen, 2009].

Table 6 reports the results of annual cross-sectional linear regressions formalized in Equation 3. The coefficient of tone measure *tone* is economically and statistically significant, and is positively associated with the equity price across all years in our sample. Overall, these results provide additional evidence to support our prior inferences.

We also analyze the resulting R^2 as a goodness of fit of the model described in Equation 3. We conclude that adding tone measures improves the R^2 by less than 1%, suggesting that while the tone of the MD&A section is value relevant, its incremental information value to the investors is minimal compared to the quantitative information.

Finally, Table 7 reports the empirical results of model in Equation 2 using an alternative dictionary [Henry, 2008]. The coefficients of *tone* and *delta_tone* are positive and statistically significant, confirming our prior findings in Table 5 and supporting our Hypothesis 1. Our average words per sentence measure of the MD&A section of 10-K (*read*) is negatively associated with equity price, supporting Hypothesis 2.

5.3 Generalized Additive Models

When non-linearities are present, traditional linear models often fail. We attempt to address the issue of non-linearity in our research design by using *generalized additive models* (GAM), automatic flexible statistical methods that may be used to identify and characterize nonlinear regression effects [Friedman et al., 2001]. In the regression setting, a generalized additive model has the form

$$E(Y|X_1, X_2, ..., X_p) = \alpha + f_1(X_1) + f_2(X_2) + ... + f_p(X_p) \tag{4}$$

where X_1;X_2;... ;X_p represent predictors and Y is the outcome; the f_j's are unspecified smooth (nonparametric) functions. Friedman et al. [2001] provide an overview of this class of models and algorithms used for fitting them. We implement GAM in R using the 'gam' package [Fasiolo et al., 2018; Wood et al., 2016].

To demonstrate how accounting for non-linearities may improve the fit of the model, we estimate the following regression using only the last year of our sample, 2015. Our benchmark is the R^2 in the last column of Table 6, 64%. We fit the following gam model:

$$p = s(ni) + s(bve) + s(cf) + s(cash) + s(div) + s(rd) + s(intan) + s(adv) + s(spi) + oci + s(rev) + s(at) + s(tone) + s(read) \tag{5}$$

where $s(x)$ indicates that the fitting is done by estimating smooth splines with various degrees of freedom. Table 8 reports the approximate significance of smooth terms. The covariate rd is confirmed to be linear ($edf = 1$) while *tone* is best fit with a cubic polynomial ($edf = 3.43$). Key accounting fundamentals, *ni* and *bve*, appear to be best fit using a spline of order eight. The resulting adjusted R^2 of the GAM model is 71% and is significantly improved over the benchmark case in Table 6 (improvement of 7%). Figure 2 visually demonstrates the shape of splines and the resulting degrees of freedom. For each of the independent variables included in the model a respective plot shows a relationship between the variable itself (x-axes) and the response value (y-axes) based on smoothing splines with degrees of freedom specified on the y-axes. Most of the covariates (except for *oci* and already mentioned rd) need to enter the model with degrees of freedom higher than one up to eight. Particularly, negative *ni* appears to have very low impact on p, but its association dramatically increases when *ni* is positive. Low values of *bve* exhibit a similar association. On the contrary, *spi* in its mid-range values has no impact on the equity price, but its extreme positive and extreme negative tail values are negatively associated with p. Extreme negative values of *tone* appear to have no association with p. Increasing *tone* is positively associated with equity price up to a certain point, after which the market appears to penalize overly positive tonality of the disclosure, disproving *impression management* theory and providing evidence in support of *incremental information* contained in financial disclosures.

6 Conclusions

MD&A sections of 10-K reports are an important source of accounting information. They provide not only quantitative information, but also contain qualitative text to help investors interpret the former. Prior literature has shown that the tone in qualitative disclosures is an important mechanism to influence investor's assessments of the value of the firm. In this study we analyze whether the tone measure of the MD&A sections of annual 10-K reports is associated with the equity price of the firm. Particularly, we research whether the tone and average words per sentence of the disclosure exhibit value relevance, in additional to other accounting and amounts shown to be value-relevant over time.

We find that both *tone* and *delta_tone* are positively associated with the equity price, providing empirical evidence that the tone of financial disclosures is value-relevant, thus, supporting *incremental information* argument of the disclosure tone. The coefficient of *abntone* is not significant, indicating that the abnormal component of the tone (that is not associated with the underlying financial statements) is largely igrnored or discounted by the investors. We also find that the equity price is negatively associated with our readability measure, consistent with diminishing information value of unreadable disclosures. These results suggest that while the investors appreciate the increased amount of information provided to them in the MD&A section, they want this information to be readable and succinct. Additionally, we address the concern of non-linearities in value relevance research by using generalized additive model approach and showing how accounting for non-linearities allows to obtain a better fit of the model.

7 Appendix

7.1 Variables Description

VARIABLE	COMPUSTAT ITEMS
$TONE$	(# positive words + # negative words) / total non-numerical words
$DELTA\_TONE$	$delta\_tone_{jt} = tone_{jt} - tone_{jt-1}$
$READ$	number of words divided by the number of sentences
P	equity price three months after fiscal year end (prcc_q)
NI	$ib/csho$
BVE	$ceq/csho$
CF	$(act - che - lct - dlc - txp)/csho$ [4]
$CASH$	$che/csho$
DIV	$dvc/csho$
RD	$xrd/csho$
$INTAN$	$intan/csho$
ADV	$xad/csho$
SPI	$spi/csho$
OCI	$(re - lag.re + dvc - ib)/csho$
REV	$revt/csho$
ASS	$at/csho$
$LOSS$	an indicator variable if $ni < 0$

FF includes **Fama-French industry** fixed effects and is determined by SIC code as follows: Consumer Non-Durables - Food, Tobacco, Textiles, Apparel, Leather, Toys (0100-0999, 2000-2399, 2700-2749, 2770-2799, 3100-3199, 3940-3989);Consumer Durables -Cars, TV's, Furniture, Household Appliances (2500-2519, 2590-2599, 3630-3659, 3710-3711, 3714-3714, 3716-3716, 3750-3751, 3792-3792, 3900-3939, 3990-3999); Manufacturing -Machinery, Trucks, Planes, Chemicals, Off Furn, Paper, Com Printing (2520-2589, 2600-2699, 2750-2769, 2800-2829, 2840-2899, 3000-3099, 3200-3569, 3580-3621, 3623-3629, 3700-3709, 3712-3713, 3715-3715, 3717-3749, 3752-3791, 3793-3799, 3860-3899); Energy -Oil, Gas, and Coal Extraction and Products (1200-1399, 2900-2999); High-Tech Business Equipment - Computers, Software, and Electronic Equipment (3570-3579, 3622-3622, 3660-3692, 3694-3699, 3810-3839, 7370-7372 , 7373-7373, 7374-7374, 7375-7375, 7376-7376, 7377-7377, 7378-7378, 7379-7379, 7391-7391,8730-8734); Telecom -Telephone and Television Transmission (4800-4899); Shops - Wholesale, Retail, and Some Services (Laundries, Repair Shops) (5000-5999, 7200-7299, 7600-7699); Health - Healthcare, Medical Equipment, and Drugs (2830-2839, 3693-3693, 3840-3859, 8000-8099); Utilities (4900-4949);Other - Mines, Constr, BldMt, Trans, Hotels, Bus Serv, Entertainment, Finance.

7.2 Tables of Results

Table 1: Descriptive Statistics

Descriptive statistics are reported for our complete sample of 22,316 observations. Columns from left to right report the mean, the standard deviation, the minimum value, the first, the third quartile, and the maximum value of the respective covariate. Textual characteristics are reported for the main dictionary by Loughran and McDonald [2011]. All financial variables are winsorized at 0.01 level.

Statistic	N	Mean	St. Dev.	Min	Q1	Q3	Max
tone	22,316	−0.405	0.214	−1.000	−0.542	−0.296	1.000
delta_tone	22,316	−0.006	0.139	−1.601	−0.060	0.054	1.750
abntone	22,316	0.0001	0.201	−0.686	−0.123	0.100	1.455
read	22,316	26.386	5.843	3.150	24.187	28.306	302.000
ni	22,316	0.769	2.305	−14.288	−0.251	1.729	10.506
bve	22,316	9.871	10.369	−9.831	2.421	14.154	63.698
cf	22,316	0.200	5.225	−37.943	−0.836	1.832	17.820
cash	22,316	2.849	3.703	0.000	0.500	3.697	25.579
div	22,316	0.299	0.634	0	0	0.3	5
rd	22,316	0.417	0.730	0.000	0.000	0.563	4.999
intan	22,316	4.915	8.679	0.000	0.062	5.694	65.216
adv	22,316	0.190	0.521	0	0	0.1	4
spi	22,316	−0.292	0.969	−13	−0.2	0	2
oci	22,316	−0.197	1.199	−9.639	−0.154	0.021	5.128
rev	22,316	24.269	33.463	0.000	4.195	30.420	236.241
at	22,316	24.716	28.145	0.004	5.482	33.807	199.142
P	22,316	24.830	27.946	0.240	5.128	34.720	193.694

Figure 1: Tone Measures means and confidence intervals by groups.

Group means and confidence intervals for the three tone measures are reported by years (left) and Fama French industry groups ($FFIND$, right).

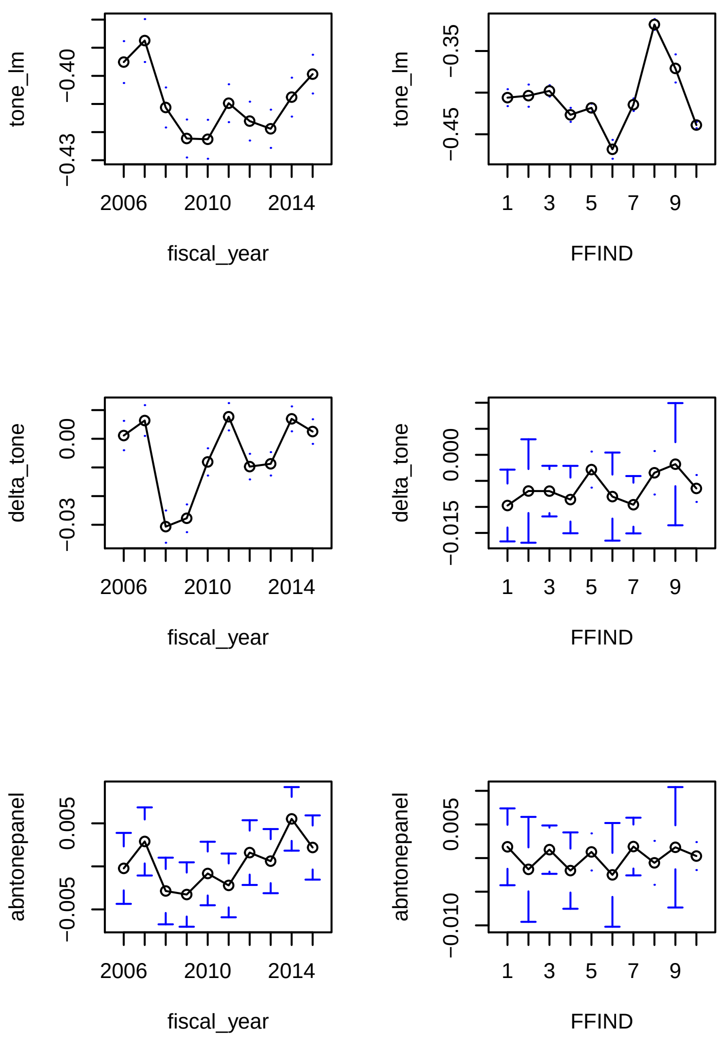


Figure 2: Results of Generalized Additive Models

Results of applying generalized additive modeling (GAM) as described in Equation 5. For each of the independent variables included in the model a respective plot shows a relationship between the variable itself (x-axes) and the response value (y-axes) based on smoothing splines with degrees of freedom specified on the y-axes.

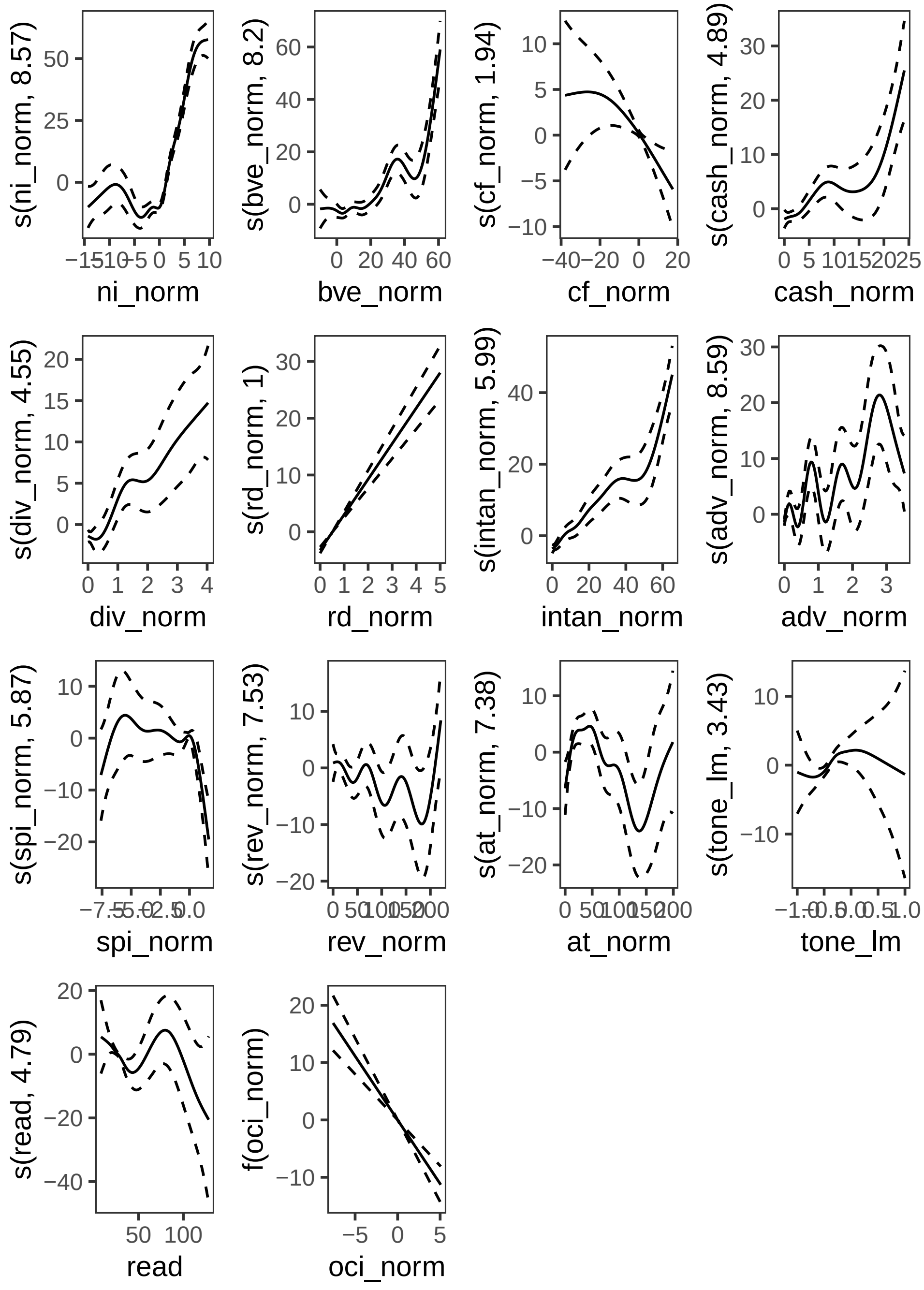


Table 2: Pearson Correlation Matrix

Pearson correlation coefficients are reported for the main variables for our sample. All variables are described in Appendix. All financial variables are winsorized at 0.01 level. Significance at p<0.05 level is denoted by *.

	tone	delta_tone	abntone	read	ni	bve	cf	cash	div	rd	intan	adv	spi	oci	rev	at
delta_tone	0.28*															
abntone	-0.01	0.00														
read	-0.02*	-0.02*	0.00													
ni	0.09*	0.06*	0.00	-0.07*												
bve	-0.03*	-0.02*	0.01	-0.03*	0.52*											
cf	0.02*	0.00	0.00	-0.02*	0.04*	0.08*										
cash	-0.01	0.01	0.01	0.01	0.31*	0.47*	-0.20*									
div	0.02*	0.00	-0.01	-0.04*	0.36*	0.35*	-0.07*	0.15*								
rd	0.04*	0.01	0.01	0.02*	0.07*	0.11*	0.03*	0.37*	0.01							
intan	-0.08*	-0.01*	0.00	-0.02*	0.28*	0.49*	-0.02*	0.18*	0.20*	0.10*						
adv	0.00	-0.02*	0.00	-0.02*	0.13*	0.11*	-0.08*	0.13*	0.09*	-0.01	0.16*					
spi	0.13*	0.10*	-0.01	0.00	0.40*	-0.09*	0.01	-0.07*	-0.05*	-0.06*	-0.20*	-0.07*				
oci	0.01	0.02*	0.00	0.02*	-0.15*	-0.08*	0.01*	-0.11*	-0.04*	-0.10*	-0.11*	-0.10*	0.11*			
rev	-0.05*	0.00	0.00	-0.02*	0.40*	0.56*	0.02*	0.35*	0.28*	0.00	0.39*	0.22*	-0.16*	-0.16*		
at	-0.07*	-0.02*	0.00	-0.01*	0.42*	0.77*	-0.19*	0.47*	0.39*	0.06*	0.60*	0.17*	-0.20*	-0.15*	0.71*	
p	0.05*	0.02*	0.00	-0.04*	0.62*	0.65*	-0.04*	0.49*	0.42*	0.27*	0.44*	0.21*	-0.02*	-0.18*	0.46*	0.60*

Table 3: Regression Results for Abnormal Tone Model - Cross Section

Cross section regression results by year to determine abnormal tone reported for Equation 1. Abnormal tone, *abntone*, is the residual of Regression 1, and is later used as an independent variable in our second stage analysis in Section 4.2. All financial variables are winsorized at 0.01 level. *p<0.1; **p<0.05; ***p<0.01

	Dependent variable: TONE									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
P	0.001***	0.002***	0.002***	0.001***	0.001***	0.001***	0.001**	0.001***	0.001**	0.0004**
	(0.0004)	(0.0004)	(0.0005)	(0.0004)	(0.0003)	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0002)
ni	0.003	0.003	0.01**	0.01***	0.01	0.004	0.01	0.004	0.01*	0.003
	(0.01)	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.002)
bve	0.001	−0.001	−0.002**	−0.001	−0.001	−0.0004	−0.0002	−0.001	−0.0002	−0.0002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
cf	0.001	0.003**	0.0002	0.001	0.0004	0.0001	−0.001	−0.0002	−0.001	0.0000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
cash	−0.001	−0.0004	−0.002	−0.002	0.0003	−0.002	0.0005	−0.002	0.0005	0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
div	0.03**	0.02*	0.03***	0.02**	0.02*	−0.004	−0.004	−0.003	−0.0004	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
rd	0.01	0.01	0.01	0.01	−0.01	−0.01	−0.003	−0.002	0.0004	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
intan	−0.002	−0.002**	−0.001	−0.002**	−0.001	−0.001*	−0.001*	−0.001**	−0.002***	−0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
adv	0.03***	0.02**	0.002	0.01	0.003	0.01	0.01	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
spi	0.04***	0.04***	0.01*	0.02***	0.04***	0.03***	0.02***	0.03***	0.02***	0.02***
	(0.01)	(0.01)	(0.004)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.005)
oci	0.01	−0.01**	0.002	0.003	0.001	−0.01*	0.005	0.004	0.002	0.002
	(0.01)	(0.005)	(0.003)	(0.01)	(0.01)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)
rev	−0.0004	−0.0001	−0.0001	−0.0001	−0.0002	−0.0001	−0.0003	−0.0001	0.0000	0.0002
	(0.0003)	(0.0003)	(0.0002)	(0.0003)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
at	−0.001	−0.0002	−0.0003	−0.001	−0.0005	−0.0003	−0.0002	−0.0001	−0.0004	−0.001*
	(0.001)	(0.0005)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0003)	(0.0003)
FF	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,173	2,184	2,193	2,381	2,324	2,323	2,123	2,124	2,206	2,285
Adjusted R^2	0.06	0.07	0.08	0.07	0.06	0.07	0.05	0.06	0.07	0.10

Table 4: Regression Results for Abnormal Tone Model - Panel Regression

Panel regression results by year to determine abnormal tone. All financial variables are winsorized at 0.01 level. *p<0.1; **p<0.05; ***p<0.01

	Dependent variable:
P	0.0004***
	(0.0001)
ni	0.01***
	(0.001)
bve	0.0000
	(0.0003)
cf	−0.001
	(0.0004)
cash	−0.001**
	(0.001)
div	−0.01**
	(0.003)
rd	0.01
	(0.004)
intan	−0.001**
	(0.0003)
adv	0.001
	(0.01)
spi	0.01***
	(0.001)
oci	0.001
	(0.001)
rev	0.0001
	(0.0001)
at	−0.0001
	(0.0002)
FF	Yes
YY	Yes
Observations	22,316
R^2	0.02

Table 5: Regression Results for Value Relevance of Tone - Panel Regression

Main regression results to estimate Equation 2. Regressions performed in R using package plm with cluster-robust standard errors to avoid overstating t-statistics due to repeated observations of the same firm within the dataset [Petersen, 2009]. Standard errors reported in parentheses. *p<0.1; **p<0.05; ***p<0.01

	Dependent variable: P			
	(1)	(2)	(3)	(4)
tone		2.49^{***}		
		(0.66)		
delta_tone			1.49^{**}	
			(0.59)	
abntone				0.19
				(0.42)
read		-0.04^{*}	-0.05^{**}	-0.05^{**}
		(0.02)	(0.02)	(0.02)
ni	2.47^{***}	2.44^{***}	2.46^{***}	2.47^{***}
	(0.07)	(0.07)	(0.07)	(0.07)
bve	0.52^{***}	0.52^{***}	0.53^{***}	0.52^{***}
	(0.03)	(0.03)	(0.03)	(0.03)
cf	0.04	0.04	0.04	0.04
	(0.03)	(0.03)	(0.03)	(0.03)
cash	1.08^{***}	1.08^{***}	1.08^{***}	1.08^{***}
	(0.05)	(0.05)	(0.05)	(0.05)
div	3.71^{***}	3.71^{***}	3.70^{***}	3.70^{***}
	(0.25)	(0.25)	(0.25)	(0.25)
rd	5.78^{***}	5.76^{***}	5.77^{***}	5.77^{***}
	(0.34)	(0.34)	(0.34)	(0.34)
intan	0.27^{***}	0.28^{***}	0.27^{***}	0.27^{***}
	(0.03)	(0.03)	(0.03)	(0.03)
adv	5.16^{***}	5.16^{***}	5.16^{***}	5.17^{***}
	(0.51)	(0.51)	(0.51)	(0.51)
spi	-1.62^{***}	-1.63^{***}	-1.64^{***}	-1.62^{***}
	(0.12)	(0.12)	(0.12)	(0.12)
oci	−0.12	−0.12	−0.12	−0.12
	(0.08)	(0.08)	(0.08)	(0.08)
rev	0.08^{***}	0.08^{***}	0.08^{***}	0.08^{***}
	(0.01)	(0.01)	(0.01)	(0.01)
at	0.07^{***}	0.07^{***}	0.07^{***}	0.07^{***}
	(0.02)	(0.02)	(0.02)	(0.02)
Observations	22,203	22,203	22,203	22,203
R^2	0.33	0.34	0.33	0.33
Adjusted R^2	0.18	0.18	0.18	0.18

Table 6: Regression Results for Value Relevance of Tone - Cross Section

Cross sectional regression results to estimate Equation 3. Standard errors reported in parentheses. *p<0.1; **p<0.05; ***p<0.01

	Dependent variable: P																			
	2006		2007		2008		2009		2010		2011		2012		2013		2014		2015	
tone		3.84***		5.88***		4.31***		2.78***		3.75***		5.48***		3.67**		6.70***		5.69**		4.12*
		(1.10)		(1.19)		(0.95)		(1.06)		(1.38)		(1.57)		(1.80)		(2.14)		(2.22)		(2.27)
read		−0.18***		−0.02		−0.02		−0.03		−0.02		−0.02		0.05		0.15*		0.14		−0.18***
		(0.06)		(0.06)		(0.04)		(0.04)		(0.04)		(0.04)		(0.06)		(0.08)		(0.09)		(0.06)
ni	6.38***	6.32***	5.76***	5.67***	2.30***	2.24***	3.87***	3.82***	5.16***	5.12***	5.84***	5.78***	5.61***	5.59***	6.91***	6.90***	6.53***	6.49***	2.98***	2.95***
	(0.23)	(0.23)	(0.21)	(0.21)	(0.12)	(0.12)	(0.17)	(0.17)	(0.23)	(0.23)	(0.24)	(0.24)	(0.26)	(0.27)	(0.30)	(0.31)	(0.29)	(0.30)	(0.22)	(0.22)
bve	0.40***	0.39***	0.64***	0.64***	0.54***	0.54***	0.73***	0.73***	0.70***	0.70***	0.63***	0.63***	0.66***	0.66***	0.42***	0.42***	0.48***	0.48***	0.57***	0.56***
	(0.06)	(0.06)	(0.06)	(0.06)	(0.04)	(0.04)	(0.04)	(0.04)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
cf	0.09	0.08	−0.04	−0.05	−0.20***	−0.20***	−0.06	−0.06	−0.12	−0.12	−0.24***	−0.24***	−0.16*	−0.16*	0.08	0.08	−0.42***	−0.41***	−0.23***	−0.23***
	(0.07)	(0.07)	(0.07)	(0.07)	(0.05)	(0.05)	(0.06)	(0.06)	(0.08)	(0.08)	(0.07)	(0.07)	(0.08)	(0.08)	(0.09)	(0.09)	(0.09)	(0.09)	(0.08)	(0.08)
cash	1.16***	1.16***	0.51***	0.50***	0.89***	0.89***	0.81***	0.81***	0.54***	0.54***	0.50***	0.51***	0.54***	0.54***	1.02***	1.03***	1.15***	1.14***	1.29***	1.29***
	(0.10)	(0.10)	(0.11)	(0.11)	(0.09)	(0.09)	(0.09)	(0.09)	(0.11)	(0.11)	(0.12)	(0.12)	(0.13)	(0.13)	(0.14)	(0.14)	(0.15)	(0.15)	(0.15)	(0.15)
div	4.55***	4.44***	3.97***	3.82***	3.71***	3.55***	6.02***	5.93***	6.20***	6.11***	5.75***	5.74***	3.40***	3.41***	4.55***	4.50***	4.59***	4.63***	6.90***	6.78***
	(0.62)	(0.62)	(0.54)	(0.54)	(0.43)	(0.43)	(0.53)	(0.53)	(0.59)	(0.59)	(0.66)	(0.65)	(0.49)	(0.49)	(0.75)	(0.75)	(0.75)	(0.75)	(0.69)	(0.69)
rd	4.77***	4.70***	4.66***	4.56***	2.89***	2.83***	5.35***	5.33***	5.21***	5.22***	4.62***	4.62***	6.64***	6.63***	7.54***	7.50***	9.41***	9.36***	6.97***	6.92***
	(0.50)	(0.50)	(0.50)	(0.50)	(0.37)	(0.37)	(0.43)	(0.43)	(0.55)	(0.55)	(0.52)	(0.52)	(0.56)	(0.56)	(0.65)	(0.65)	(0.67)	(0.66)	(0.63)	(0.63)
intan	0.32***	0.33***	0.04	0.05	0.24***	0.24***	0.12***	0.13***	0.16***	0.16***	0.19***	0.20***	0.21***	0.22***	0.33***	0.35***	0.44***	0.45***	0.57***	0.58***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
adv	1.39**	1.27**	0.29	0.17	2.53***	2.50***	3.05***	3.02***	3.34***	3.31***	3.83***	3.78***	4.50***	4.45***	4.67***	4.54***	5.69***	5.64***	4.80***	4.76***
	(0.58)	(0.58)	(0.54)	(0.54)	(0.46)	(0.45)	(0.58)	(0.58)	(0.69)	(0.68)	(0.61)	(0.61)	(0.74)	(0.74)	(0.84)	(0.84)	(0.85)	(0.85)	(0.80)	(0.80)
spi	−5.00***	−5.09***	−4.94***	−5.11***	−1.13***	−1.15***	−2.84***	−2.87***	−4.14***	−4.28***	−4.16***	−4.31***	−3.06***	−3.17***	−4.37***	−4.64***	−5.44***	−5.59***	−0.67	−0.71
	(0.58)	(0.58)	(0.50)	(0.50)	(0.16)	(0.16)	(0.32)	(0.32)	(0.69)	(0.69)	(0.55)	(0.55)	(0.58)	(0.58)	(0.86)	(0.86)	(0.73)	(0.73)	(0.50)	(0.50)
oci	−0.51*	−0.51*	−0.59**	−0.53**	−0.04	−0.05	−0.95***	−0.95***	−1.52***	−1.51***	−0.95***	−0.90***	−0.65*	−0.66*	−0.99***	−1.01***	−0.62*	−0.65*	−3.19***	−3.17***
	(0.28)	(0.28)	(0.25)	(0.25)	(0.15)	(0.15)	(0.27)	(0.27)	(0.37)	(0.37)	(0.32)	(0.31)	(0.36)	(0.36)	(0.37)	(0.37)	(0.35)	(0.35)	(0.35)	(0.35)
rev	−0.002	−0.0002	−0.06***	−0.06***	0.03***	0.03***	−0.01	−0.01	−0.04**	−0.04**	−0.04**	−0.03**	−0.05***	−0.05***	−0.03*	−0.03*	0.03	0.02	0.02	0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
at	0.002	0.004	0.06**	0.06**	−0.08***	−0.07***	0.01	0.01	0.02	0.02	−0.01	−0.01	0.09***	0.09***	0.09***	0.09***	0.03	0.03	0.06*	0.06**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,173	2,173	2,184	2,184	2,193	2,193	2,381	2,381	2,324	2,324	2,323	2,323	2,123	2,123	2,124	2,124	2,206	2,206	2,285	2,285
Adjusted R^2	0.71	0.71	0.66	0.66	0.59	0.59	0.70	0.70	0.68	0.68	0.67	0.67	0.68	0.68	0.69	0.69	0.69	0.69	0.64	0.64

Table 7: Robustness Tests - Alternative Dictionary

Robustness regression results to estimate Equation 2 using an alternative dictionary. Regressions performed in R using package plm with cluster-robust standard errors to avoid overstating t-statistics due to repeated observations of the same firm within the dataset [Petersen, 2009]. Standard errors reported in parentheses. *p<0.1; **p<0.05; ***p<0.01

	Dependent variable: P			
	(1)	(2)	(3)	(4)
tone		7.64***		
		(1.42)		
delta_tone			4.88***	
			(1.29)	
abntone				1.31
				(0.86)
read		−0.06**	−0.05**	−0.05**
		(0.02)	(0.02)	(0.02)
ni	2.47***	2.45***	2.46***	2.47***
	(0.07)	(0.07)	(0.07)	(0.07)
bve	0.52***	0.52***	0.53***	0.52***
	(0.03)	(0.03)	(0.03)	(0.03)
cf	0.04	0.04	0.04	0.04
	(0.03)	(0.03)	(0.03)	(0.03)
cash	1.08***	1.09***	1.08***	1.08***
	(0.05)	(0.05)	(0.05)	(0.05)
div	3.71***	3.70***	3.70***	3.69***
	(0.25)	(0.25)	(0.25)	(0.25)
rd	5.78***	5.79***	5.80***	5.78***
	(0.34)	(0.34)	(0.34)	(0.34)
intan	0.27***	0.27***	0.27***	0.27***
	(0.03)	(0.03)	(0.03)	(0.03)
adv	5.16***	5.12***	5.19***	5.17***
	(0.51)	(0.51)	(0.51)	(0.51)
spi	−1.62***	−1.63***	−1.63***	−1.62***
	(0.12)	(0.12)	(0.12)	(0.12)
oci	−0.12	−0.12	−0.12	−0.12
	(0.08)	(0.08)	(0.08)	(0.08)
rev	0.08***	0.08***	0.08***	0.08***
	(0.01)	(0.01)	(0.01)	(0.01)
at	0.07***	0.07***	0.07***	0.07***
	(0.02)	(0.02)	(0.02)	(0.02)
Observations	22,203	22,203	22,203	22,203
R^2	0.33	0.34	0.34	0.33
Adjusted R^2	0.18	0.18	0.18	0.18

Table 8: Results of Fitting a Generalized Additive Model

Regression results using generalized additive model (GAM) for Year 2015 using Equation 5. Results below represent approximate significance of smooth terms. Number of observations is 2,285. Adjusted R^2 is 0.711.

	edf	Ref.df	F	p-value
s(ni)	8.57	8.94	60.56	0.000
s(bve)	8.19	8.81	14.77	0.000
s(cf)	1.93	2.47	4.43	0.006
s(cash)	4.89	5.94	6.83	0.000
s(div)	4.55	5.52	7.60	0.000
s(rd)	1.00	1.00	133.55	0.000
s(intan)	5.99	7.11	22.30	0.000
s(adv)	8.58	8.94	5.89	0.000
s(spi)	5.87	6.97	4.67	0.000
s(rev)	7.53	8.45	2.20	0.024
s(at)	7.38	8.36	4.18	0.000
s(tone)	3.43	4.37	2.24	0.060
s(read)	4.79	5.90	2.91	0.008

References

Aaker, D. A. and Jacobson, R. (2001). The value relevance of brand attitude in high-technology markets. *Journal of Marketing Research*, 38(4):485–493.

Akerlof, G. et al. (1970). The market for lemons. *Quarterly Journal of Economics*, 84(3):488–500.

Amir, E. and Lev, B. (1996). Value-relevance of nonfinancial information: The wireless communications industry. *Journal of Accounting and Economics*, 22(1-3):3–30.

Arena, C., Bozzolan, S., and Michelon, G. (2015). Environmental reporting: Transparency to stakeholders or stakeholder manipulation? An analysis of disclosure tone and the role of the board of directors. *Corporate Social Responsibility and Environmental Management*, 22(6):346–361.

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

Baboukardos, D. and Rimmel, G. (2016). Value relevance of accounting information under an integrated reporting approach: A research note. *Journal of Accounting and Public Policy*, 35(4):437–452.

Bakarich, K. M., Hossain, M., and Weintrop, J. (2019). Different time, different tone: Company life cycle. *Journal of Contemporary Accounting & Economics*, 15(1):69–86.

Balachandran, S. and Mohanram, P. (2011). Is the decline in the value relevance of accounting driven by increased conservatism? *Review of Accounting Studies*, 16(2):272–301.

Ball, C., Hoberg, G., and Maksimovic, V. Disclosure, business change and earnings quality. *Available at SSRN: https://ssrn.com/abstract=2260371*.

Barth, M. E., Beaver, W. H., and Landsman, W. R. (2001). The relevance of the value relevance literature for financial accounting standard setting: another view. *Journal of Accounting and Economics*, 31(1-3):77–104.

Barth, M. E., Clinch, G., and Israeli, D. (2016). What do accruals tell us about future cash flows? *Review of Accounting Studies*, 21(3):768–807.

Barth, M. E., Li, K., and McClure, C. G. (2017). Evolution in value relevance of accounting information. *Stanford University Graduate School of Business Research Paper No. 17-24. Available at SSRN: https://ssrn.com/abstract=2933197*.

Beisland, L. A. (2009). A review of the value relevance literature. *The Open Business Journal*, 2(1):7–27.

Blankespoor, E. A. (2012). *The impact of investor information processing costs on firm disclosure choice: Evidence from the XBRL mandate.* PhD thesis, University of Michigan.

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3:993–1022.

Brown, S. V. and Tucker, J. W. (2011). Large-sample evidence on firms' year-over-year MD&A modifications. *Journal of Accounting Research*, 49(2):309–346.

Bushee, B. J., Gow, I. D., and Taylor, D. J. (2018). Linguistic complexity in firm disclosures: Obfuscation or information? *Journal of Accounting Research*, 56(1):85–121.

Cahan, S. F., De Villiers, C., Jeter, D. C., Naiker, V., and Van Staden, C. J. (2016). Are CSR disclosures value relevant? Cross-country evidence. *European Accounting Review*, 25(3):579–611.

Caserio, C. and Napoli, F. (2017). Value relevance and Codes of Ethics: An empirical analysis of Italian listed companies. *International Journal of Business Governance and Ethics*, 12(1):1–20.

Choi, Y.-S., O'Hanlon, J. F., and Pope, P. F. (2006). Conservative accounting and linear information valuation models. *Contemporary Accounting Research*, 23(1):73–101.

Collins, D. W., Li, O. Z., and Xie, H. (2009). What drives the increased informativeness of earnings announcements over time? *Review of Accounting Studies*, 14(1):1–30.

Collins, D. W., Maydew, E. L., and Weiss, I. S. (1997). Changes in the value-relevance of earnings and book values over the past forty years. *Journal of Accounting and Economics*, 24(1):39–67.

Davis, A. K., Piger, J. M., Sedor, L. M., et al. (2006). Beyond the numbers: An analysis of optimistic and pessimistic language in earnings press releases. Technical report, Federal Reserve Bank of St. Louis.

Drake, M. S., Roulstone, D. T., and Thornock, J. R. (2015). The determinants and consequences of information acquisition via edgar. *Contemporary Accounting Research*, 32(3):1128–1161.

Durnev, A. and Mangen, C. (2011). The real effects of disclosure tone: Evidence from restatements. *Available at SSRN: https://ssrn.com/abstract=1650003*.

Dyer, T., Lang, M., and Stice-Lawrence, L. (2017). The evolution of 10-K textual disclosure: Evidence from Latent Dirichlet Allocation. *Journal of Accounting and Economics*, 64(2-3):221–245.

Engelberg, J. (2008). Costly information processing: Evidence from earnings announcements. *AFA 2009 San Francisco Meetings Paper*.

Fasiolo, M., Nedellec, R., Goude, Y., and Wood, S. N. (2018). Scalable visualisation methods for modern Generalized Additive Models. *arXiv preprint arXiv:1809.10632*.

Feldman, R., Govindaraj, S., Livnat, J., and Segal, B. (2010). Management's tone change, post earnings announcement drift and accruals. *Review of Accounting Studies*, 15(4):915–953.

Feltham, G. A. and Ohlson, J. A. (1995). Valuation and clean surplus accounting for operating and financial activities. *Contemporary Accounting Research*, 11(2):689–731.

Files, R., Holcomb, A., Martin, G. S., and Mason, P. (2018). Damage Control: Changes in Disclosure Tone after Financial Misconduct. *Available at SSRN: https://ssrn.com/abstract=3209501*.

Francis, J. and Schipper, K. (1999). Have financial statements lost their relevance? *Journal of Accounting Research*, 37(2):319–352.

Francis, J., Schipper, K., and Vincent, L. (2002a). Earnings announcements and competing information. *Journal of Accounting and Economics*, 33(3):313–342.

Francis, J., Schipper, K., and Vincent, L. (2002b). Expanded disclosures and the increased usefulness of earnings announcements. *The Accounting Review*, 77(3):515–546.

Frankel, R. M., Jennings, J. N., and Lee, J. A. (2017). Using Natural Language Processing to Assess Text Usefulness to Readers: The Case of Conference Calls and Earnings Prediction. *Available at SSRN: https://ssrn.com/abstract=3095754*.

Friedman, J., Hastie, T., and Tibshirani, R. (2001). *The elements of statistical learning*, volume 1. Springer series in statistics, New York.

Gietzmann, M. and Ostaszewski, A. (2003). An alternative to the Feltham-Ohlson valuation framework: using q-theoretic income to predict firm value. *Cass Business School Research Paper*.

Givoly, D., Hayn, C., and Katz, S. (2017). The changing relevance of accounting information to debt holders over time. *Review of Accounting Studies*, 22(1):64–108.

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(1-3):405–440.

Henry, E. (2008). Are investors influenced by how earnings press releases are written? *The Journal of Business Communication (1973)*, 45(4):363–407.

Hoberg, G. and Lewis, C. (2017). Do fraudulent firms produce abnormal disclosure? *Journal of Corporate Finance*, 43:58–85.

Holthausen, R. W. and Watts, R. L. (2001). The relevance of the value-relevance literature for financial accounting standard setting. *Journal of Accounting and Economics*, 31(1-3):3–75.

Hope, O.-K., Hu, D., and Lu, H. (2016). The benefits of specific risk-factor disclosures. *Review of Accounting Studies*, 21(4):1005–1045.

Huang, X., Teoh, S. H., and Zhang, Y. (2013). Tone management. *The Accounting Review*, 89(3):1083–1113.

Hughes, K. (2000). The value relevance of nonfinancial measures of air pollution in the electric utility industry. *The Accounting Review*, 75(2):209–228.

Kaspereit, T. and Lopatta, K. (2016). The value relevance of SAM's corporate sustainability ranking and GRI sustainability reporting in the European stock markets. *Business Ethics: A European Review*, 25(1):1–24.

Kim, C. F., Wang, K., and Zhang, L. (2018). Readability of 10-k reports and stock price crash risk. *Contemporary Accounting Research*.

Kogan, S., Levin, D., Routledge, B. R., Sagi, J. S., and Smith, N. A. (2009). Predicting risk from financial reports with regression. In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 272–280. Association for Computational Linguistics.

Landsman, W. R. and Maydew, E. L. (2002). Has the information content of quarterly earnings announcements declined in the past three decades? *Journal of Accounting Research*, 40(3):797–808.

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

Lawrence, A. (2013). Individual investors and financial disclosure. *Journal of Accounting and Economics*, 56(1):130–147.

Lee, H., Surdeanu, M., MacCartney, B., and Jurafsky, D. (2014). On the importance of text analysis for stock price prediction. In *LREC*, pages 1170–1175.

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

Lev, B. (2018). The deteriorating usefulness of financial report information and how to reverse it. *Accounting and Business Research*, 48(5):465–493.

Lev, B. and Zarowin, P. (1999). The boundaries of financial reporting and how to extend them. *Journal of Accounting research*, 37(2):353–385.

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics*, 45(2-3):221–247.

Liu, J. and Ohlson, J. A. (2000). The feltham-ohlson (1995) model: empirical implications. *Journal of Accounting, Auditing & Finance*, 15(3):321–331.

Lo, K. and Lys, T. (2000a). The ohlson model: contribution to valuation theory, limitations, and empirical applications. *Journal of Accounting, Auditing & Finance*, 15(3):337–367.

Lo, K. and Lys, T. Z. (2000b). Bridging the gap between value relevance and information content. *Sauder School of Business Working Paper*.

Loughran, T. and McDonald, B. (2011). When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of Finance*, 66(1):35–65.

Loughran, T. and McDonald, B. (2014). Measuring readability in financial disclosures. *The Journal of Finance*, 69(4):1643–1671.

Loughran, T. and McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, 54(4):1187–1230.

Loughran, T. and McDonald, B. (2017). The use of EDGAR filings by investors. *Journal of Behavioral Finance*, 18(2):231–248.

Lourenço, I. C., Callen, J. L., Branco, M. C., and Curto, J. D. (2014). The value relevance of reputation for sustainability leadership. *Journal of Business Ethics*, 119(1):17–28.

Magness, V. (2009). Environmental disclosure in the mining industry: A signaling paradox? In *Sustainability, Environmental Performance and Disclosures*, pages 55–81. Emerald Group Publishing Limited.

Matolcsy, Z. P. and Wyatt, A. (2008). The association between technological conditions and the market value of equity. *The Accounting Review*, 83(2):479–518.

Merkl-Davies, D. M. and Brennan, N. M. (2007). Discretionary disclosure strategies in corporate narratives: incremental information or impression management? *Journal of Accounting Literature*, 27:116–196.

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

Morris, M. W., Sheldon, O. J., Ames, D. R., and Young, M. J. (2005). Metaphor in stock market commentary: Consequences and preconditions of agentic descriptions of price trends. *Unpublished manuscript, Columbia University*.

Ohlson, J. A. (1995). Earnings, book values, and dividends in equity valuation. *Contemporary Accounting Research*, 11(2):661–687.

Ota, K. (2003). The impact of price and return models on value relevance studies: A review of theory and evidence. *Accounting Research Journal*, 16(1):6–20.

Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *The Review of Financial Studies*, 22(1):435–480.

Rajgopal, S., Venkatachalam, M., and Kotha, S. (2003). The value relevance of network advantages: The case of e–commerce firms. *Journal of Accounting Research*, 41(1):135–162.

Riley Jr, R. A., Pearson, T. A., and Trompeter, G. (2003). The value relevance of non-financial performance variables and accounting information: The case of the airline industry. *Journal of Accounting and Public Policy*, 22(3):231–254.

Wang, D. H.-M., Yu, T. H.-K., and Chiang, C.-H. (2016). Exploring the value relevance of corporate reputation: A fuzzy-set qualitative comparative analysis. *Journal of Business Research*, 69(4):1329–1332.

Wood, S. N., Pya, N., and Säfken, B. (2016). Smoothing parameter and model selection for general smooth models. *Journal of the American Statistical Association*, 111(516):1548–1563.