APPLYING TEXTUAL ANALYSIS TO AUDITING

by

YUE LIU

A Dissertation submitted to the

Graduate School-Newark

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Management

written under the direction of

Dr. Miklos A. Vasarhelyi and Dr. Kevin Moffitt

and approved by

_____

_____

_____

_____

Newark, New Jersey

May, 2019

ABSTRACT OF THE DISSERTATION

Applying Textual Analysis to Auditing

By YUE LIU

Dissertation Directors:

Dr. Miklos A. Vasarhelyi and Dr. Kevin Moffitt

Textual analysis is the process of extracting useful information from unstructured text data. In the domain of accounting, auditing, and finance, textual analysis is still an emerging area (Loughran and McDonald 2016). Motivated by the potential benefits of applying textual analysis to auditing, this dissertation consists of three essays on using textual analysis to improve the understanding of annual report review process, audit fee decision, and internal control risks.

The first essay uses text mining to uncover the intensity of SEC comment letters and its association with the probability of restatement of 10-K filings. Specifically, it utilizes the Loughran and McDonald strong and weak modal word lists (Loughran and McDonald 2011) and measure the intensity of initial SEC comment letters based on the use of strong/weak modal language. The paper finds a positive association between the intensity of comment letter and the probability of restatement of the reviewed 10-K filing.

The second essay examines whether the qualitative disclosures of earnings press releases provides additional information about audit risks that relates to audit fee decision. We find that a more abnormally negative tone of earnings in press releases is associated with higher audit fees, showing that the abnormal tone of press releases can be a signal of the client's business risk. We also find the abnormal tone as a proxy of

opportunistic disclosure behaviors when the abnormal tone is extremely positive. In addition, the association between an abnormally negative tone and audit fees is moderated by the credibility of the disclosure and by auditor's experience with the client.

The third essay investigates whether the risk factor disclosures on internal control over financial reporting (ICFR-related risk factors) complement the mandatory SOX 404 disclosures of material weaknesses in internal control. We find that the ICFR-related risk factor disclosure incrementally predicts future related adverse consequences. The results also suggest that firms with ICFR-related risk factors are likely to have higher audit fees. In addition, the contents of ICFR-related risk factors can help financial statement users assess the severity of internal control issues. These findings suggest that ICFR-related risk factors reflect potential internal control deficiencies.

Acknowledgments

Firstly, I would like to express my deepest appreciation to my dissertation chairs, Dr. Miklos A. Vasarhelyi and Dr. Kevin Moffitt, for their valuable guidance, strong encouragement, and continuous support during my Ph.D. studies. Without their help, this dissertation would not have been possible. I would also like to express my sincere gratitude to my dissertation committee members: to Dr. Alexander Kogan for his insightful advice and continuous support, which help me become a better researcher, and to Dr. Ingrid Fisher, for her encouraging comments and positive feedback that motivated me in fine tuning my research. I also own thanks to Dr. Chanta Thomas and Dr. Soo Hyun Cho for their valuable suggestions on my research. In addition, I appreciate the support and encouragement from Dr. Dan Palmon, who has helped me improve my teaching skills.

Secondly, I would like to express my deep gratitude to the China Scholarship Council and my undergraduate school, Southwestern University of Finance and Economics, for supporting me both financially and spiritually during my PhD studies.

I am indebted to my dear parents, Zuchi Liu and Xiaomin Zeng, my dear grandmother, Guangxiu Wu, and my boyfriend, Shuheng (Sean) Sun. Their unconditional love, whole-hearted support, and continuous faith in me encouraged me to achieve my goals and fulfill my dreams.

Finally, I want to give my deepest thanks to my dearest friends, especially to Kexing Ding, Xuan Peng, Tiffany Chiu, Yunsen Wang, Qiao Li, Feiqi Huang, Ting Sun, Zhaokai Yan, Lu Zhang, Kyunghee Yoon, and all my friends in the Rutgers AIS group, for their warm friendship and academic support.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Textual analysis is the process of extracting useful information from unstructured text data. A wide range of texts can be available, such as annual reports, quarterly reports, earnings press releases, newsletters, SEC comment letters, purchase reviews, contracts, emails, messages, and so on. The texts could contain valuable information for decision making of their users, and textual analysis provides a way to identify the valuable information and transfer the unstructured texts into data sets that can be directly used for further statistical analysis. Textual analysis is still an emerging area in the domain of accounting, auditing, and finance (Loughran and McDonald 2016). Studies have been examining the contents of qualitative disclosures and their association with firm evaluation, market reaction, earnings management, as well as audit processes (e.g., Li 2008; Campbell et al. 2014; Moon and Swanquist 2017). In the auditing domain, a lot are left to be explored about how textual analysis can be applied to uncover specific patterns that help audit assessment and decisions and understand the value of qualitative information in auditing. Therefore, this dissertation contains three essays on the applications of textual analysis to the auditing field. Specifically, it examines the qualitative disclosures of SEC comment letters, earnings press releases, and risk factor disclosures in 10-K filings and their association with 10-K restatement, audit fees, and internal control deficiencies. Chapter one provides an introduction of the background, motivation, and research topics of the dissertation thesis. Chapter two, three, and four are the three essays on three applications of textual analysis to auditing, respectively. The fifth chapter provides a conclusion of the findings and implications.

The history of extracting patterns from texts is long-standing, but in the accounting and auditing domain, textual analysis is still an emerging area (Loughran and McDonald 2016). With the rapidly increasing computing power and the online availability of texts from a wide range of resources including Securities and Exchange Commission (SEC) filings, news articles, earnings press releases, texts from social media and so on that can be used for accounting and auditing related decision makings, there is an enormous potential for applying textual analysis to the field of accounting and auditing. Textual analysis can be used to determine the sentiment of texts, identify topics of texts, measure similarity between documents and classify documents into various categories, and examine readability of the information, etc. With its broad capability, it is interesting to see how textual analysis can be applied to various texts in the accounting and auditing domain to parse particular patterns that are valuable to the information users. Motivated by this, the dissertation incorporates three essays that apply textual analysis to the content of SEC comment letters, earnings press releases, and risk factor disclosures, respectively, and examine how the extracted information can be used to help understand audit-related issues and improve the understanding of audit risks.

Specifically, SEC comment letters are the correspondence from SEC staff to SEC filers about the review of disclosure filings. The SEC periodically reviews the public companies' filings to evaluate the adequacy of disclosures and the filings' compliance with Generally Accepted Accounting Principles (GAAP). For filings that are perceived to be deficient, SEC staff will issue comment letters to the filers (Cassell, Dreher, and Myers 2013). In 2005, the correspondence of comment letters between SEC and filers became publicly available on the EDGAR database for filings made after August 1, 2004.

Two types of letters are released: 1) the “upload” type, which refers to SEC staff’s comments on the reviewed filings, and 2) the “corresp” type, which is the filer’s responses to the comments. The SEC staff makes suggestions for companies to remedy perceived disclosure deficiencies in their filings, but comment letters do not result in definitive consequences. Also, they are likely to vary in severity (Gietzmann, Marra, and Pettinicchio 2016). Using initial comment letters for 10-K filings in year 2004-2015, the first essay develops a measure of the intensity of SEC comment letters based on the use of modal words in the comment letters, and examines whether the intensity of SEC comment letters can be an indicator of the probability of restatement of the reviewed 10-K filings.

Audit fees are determined by two main factors. The first factor is the costs that occur to auditors if they fail to identify material misstatements, and the second factor is the audit effort associated with the audit engagement (Stice 1991). Prior studies have investigated various proxies for possible determinants of audit fees, such as auditor size (Palmrose 1986), client size and auditor expertise (Carcello et al. 2002; Carson and Fargher 2007), etc. Furthermore, some studies have recently examined financial reporting and voluntary disclosures as a proxy of litigation risks and the auditor’s effort (Krishnan, Pevzner, and Sengupta 2012; Yang Yu, Liu, and Wu, 2018). However, scholars have rarely investigated whether and how the management qualitative voluntary disclosures can be associated with audit fee decisions. Qualitative management disclosures have the potential to provide additional information to quantitative disclosures on audit risks and audit efforts which relates to audit fees and are worth exploring. Therefore, the second essay investigates whether the abnormal tone of management voluntary earnings press

releases can be used as an indicator for client business risks as well as litigation risks to estimate audit fee decisions.

Public firms are required by the SEC to discuss risk factors that are likely to affect the firm's performance and securities in a separate section called "Item 1A" in 10-K. The SEC tries to restrict risk factor disclosures by making firms to "avoid generic risk factor disclosure that could apply to any company", but it's hard in practice to prove whether firms are following suggestions. The internal control reporting has long been stressed by the regulators. Sarbanes-Oxley Act Section 404 (SOX 404) and Auditing Standard 2201 (PCAOB 2017) make requirement that management and auditors must disclose material weaknesses in internal controls over financial reporting (ICFR). In addition to this mandatory assessment on ICFR, firms can also disclose ICFR-related risk factors in Item 1A. However, compared to the strong requirement on material weakness disclosures, the disclosure of ICFR-related risk factors is relatively voluntary in nature. Significant deficiencies that do not achieve material level are typically not disclosed in ICFR reports, and therefore the disclosure of ICFR-related risk factors can be a complementary source for stakeholders to learn more about the internal control status of a firm. Since the ICFR-related risk factor disclosure is relatively voluntary, it is interesting to study how reliable the disclosure is, and whether it complements the SOX 404 material weakness disclosure by indicating higher internal control risks for firms that disclose ICFR-related risk factors. Hence, the third essay is aimed to examine whether the disclosure of ICFR-related risk factor is informative of potential internal control risk and whether it can be used to predict future material weaknesses and restatement of the 10-K filing.

Chapter 2: Text Mining to Uncover the Intensity of SEC Comment Letters and Its Association with the Probability of 10-K Restatement

2.1. Introduction

Securities and Exchange Commission (SEC) comment letters are the correspondence sent from the SEC staff to filers regarding the filers' public information disclosure. The SEC periodically reviews the filings of public companies to evaluate the adequacy of disclosures and the filings' compliance with Generally Accepted Accounting Principles (GAAP), and issues comment letters to companies whose filings are perceived to be deficient (Cassell et al. 2013). In 2005, the SEC began releasing these comment letters to the public on the EDGAR database for filings made after August 1, 2004.

Two types of letters are released: the "upload" type contains SEC staff comments on the filing, and the "corresp" type is the filer's responses to the comments. The SEC staff makes suggestions for companies to remedy the perceived disclosure deficiencies in their filings (Bozanic, Dietrich, and Johnson 2017). For example, the companies may be requested to revise the disclosure or to provide additional information that is deemed necessary by the SEC staff. The comment letter usually contains multiple comment topics.

Companies may react differently in responding to the SEC comments: they may make the suggested changes, attempt to avoid making substantial changes, make a confidential treatment request, or negotiate with the SEC staff to reduce the number of required changes (Bozanic et al. 2017). Also, there can be several rounds of comments and responses between the SEC staff and the filer before the staff issues a final letter

confirming that the review is complete. Although the comment letters do not represent the official opinions of the SEC (SEC, 2011b), they have significant impact on the firm's disclosure (Bozanic et al. 2017).

Comment letters differ substantively from regulatory investigatory forms and are characterized as inquisitorial rather than definitive, meaning that the letters raise questions about the disclosure without accusing the registrant of wrongdoing or suggesting that legal action is imminent. They also vary in severity (Gietzmann et al. 2016). The severity of SEC comment letters is a reflection of the perceived deficiencies in the filings, and it may provide information about disclosure quality and cost of comment remediation. Analysis of the severity of comment letters is important because some companies expect to receive a comment letter at least once every three years, based on factors such as firm size or complexity of operations, and they evaluate the success of the review process based on the severity of the comment letter (Cassell et al. 2013). Prior studies have used the number of comment topics (Cassell et al. 2013) and conversation time (the number of days between the initial letter and the "no further comment" letter) (Chen, Johnston, and Ramnath 2010; Gietzmann and Pettinicchio 2014; Gietzmann et al. 2016) as proxies for the severity of comment letters, and have found significant associations between severity and restatement, some firm/auditor/governance characteristics (Cassell et al. 2013), improvement in disclosure quality (Chen et al. 2010), and Chief Financial Officer (CFO) turnover (Gietzmann et al. 2016). Different from these severity measures, the intensity of comment letters captures the degree of strong and weak modality by looking at the content of comment letters directly. Intensity of comment letters reflects what the modal words try to capture in comment letters and

therefore more directly reflects the SEC reviewers' perceptions of deficiencies in filings compared to the indirect measures of severity used in past research.

This study uses text mining to examine the intensity of SEC comment letters. The measure of intensity is based on Loughran and McDonald (2011) "modal strong" (e.g. always, must, unequivocal) and "modal weak" (e.g. almost, perhaps, might) word lists. If the comment letter contains more model strong words, the overall intensity of the letter will be higher. Empirical analysis is conducted on initial comment letters related to 10-K filings. Results show that there is a positive association between the strong-word ratio of the comment letter and the probability of restatement of the reviewed 10-K filings, indicating that the SEC staff uses strong modality to express stronger concerns in the reviewed filings. Further, a process of word list modification is conducted to better measure the intensity of comment letters. By reading and scoring randomly selected sentences for the use of each word in the Loughran and McDonald (2011) word lists, we find that some words do not function as modal words. These are deleted to create modified modal word lists. Using these lists, the association between strong word ratio and restatement of reviewed 10-K filings is still positive and significant, and the results becomes slightly, albeit not significantly, stronger.

This chapter contributes to the literature by introducing a direct measure of the intensity of SEC comment letters based on the strong/weak modality of the letter. The severity measures used by prior studies (Chen et al. 2010; Cassell et al. 2013; Gietzmann et al. 2016), including the conversation time and the number of comment topics, are indirect measures which don't look at the content of comment letters. Further, this chapter tried to modify the Loughran and McDonald word lists to identify modal words that are

more relevant to intensity of comment letters. The modification process shows that some words in the word lists are not used as modal words, and therefore, modification is necessary. The remaining part of this chapter is organized as follows: Section 2.2 provides background of this study and reviews the prior literature, Section 2.3 introduces the intensity measure and research method, Section 2.4 presents the empirical results, and Section 2.5 illustrates the modification of the word lists and discusses the results using the modified word lists. Section 2.6 concludes and provides potential related future research directions.

2.2. Background and Literature

2.2.1 SEC Comment Letter

According to the Sarbanes-Oxley Act (SOX) of 2002 Section 408, the SEC Division of Corporation Finance shall review registrants' filings at least once every three years to evaluate the filings' compliance with SEC and GAAP disclosure requirements and to ensure the quality of information presented to stakeholders. As a result of the review, the SEC staff will issue a comment letter to companies whose filings are regarded as deficient and could be improved in some way. The company receiving the comment letter is required to respond within 10 business days. Upon receiving the response, the SEC staff will review the response letter and determine whether the response is satisfactory. If not found satisfactory, additional comment letters will be issued to the filer until the SEC is satisfied with the response and issues a "no further comment" letter.

For 10-K filings, the review is conducted by a review team that consists of any combination of two accountants and two attorneys based on the review history and

availability of resources. The review team collaborates on making the comments (Bozanic et al. 2017).

The SEC comment letter does not represent the official opinion of the SEC (SEC 2011b) and is inquisitorial in nature. However, failure to comply can result in serious consequences, such as a definitive restatement of the reviewed filing (Cassell et al. 2013). In addition, comment letter remediation is a costly process in terms of time and resources. The more severe the comment letter, the higher the cost of remediation can be.

2.2.2 Related Research in the Literature

Prior research on SEC comment letters has focused on the determinants and consequences of receiving comment letters for IPO filings, 8-K filings, 10-K filings, etc. Chen et al. (2010) investigate the content and determinants of SEC comment letters related to 10-Ks and 10-Qs and the impact of letter resolution. They use content analysis to confirm that comments focused on disclosure and identify five characteristics of firms that are significantly related to the probability of receiving a comment letter: previous year restatement, share of industry revenue, Price to Earnings (P/E) disparity, firm age, and cash flow volatility. They also find that comment letters could improve disclosure quality as reflected by a decrease in abnormal market reactions. Further, higher severity of a comment letter (measured by the duration of letter period) could improve disclosure quality more.

Boone, Linthicum, and Poe (2013) study the association between the characteristics of accounting standards (i.e., rules and accounting estimates) and the likelihood that a standard will be an issue of concern during the SEC reviewing process.

The study finds that rule-based characteristics in GAAP increase the probability of raising an SEC comment.

Cassell et al. (2013) conduct a more comprehensive study on SEC comment letters. The study examines firm, auditor and governance characteristics that are associated with the receipt of comment letters, and which affect the cost of remediation. The measures for the extent of comments (the number of topics), types of comments, and cost of remediation in terms of resolution time and number of rounds between the firm and SEC staff all provide some information about the severity of a comment letter, and factors such as restatement, company size, auditor type, Chief Executive Officer (CEO) chair duality, etc. are significantly related to these proxies and letter severity.

Gietzmann and Pettinicchio (2014) find that audit fees increase after the client receives a comment letter, suggesting that auditors reassess the reputation and litigation risk of the client based on the receipt of SEC comment letters. Gietzmann et al. (2016) investigate the impact of comment letter review on CFO turnover in a dynamic model and introduce comment letter severity measured by conversation time as a moderator. They find an increase in CFO turnover once a firm received a highly intense comment letter.

All prior studies that investigate the severity of comment letters choose proxies such as the number of topics or the conversation time, but none of them uses textual analysis to measure intensity. Yet the content of comment letters can provide some information about the intensity of a comment letter. Specifically, the modal nature of some comment letters could be stronger than others, and thus may indicate higher intensity.

To identify the words that may suggest higher intensity, Loughran and McDonald (2011) modal word lists are used. The lists contain words that are perceived to be "modal strong" in the financial environment, including: always, best, clearly, definitely, definitively, highest, lowest, must, never, strongly, unambiguously, uncompromising, undisputed, undoubtedly, unequivocal, unequivocally, unparalleled, unsurpassed and will. On the other hand, words that indicate a weaker modality may also be used in the same comment letter and could have an impact on the overall modal tone of the comment letter. "modal weak" words on the list include the following: almost, apparently, appeared, appearing, appears, conceivable, could, depend, depended, depending, depends, may, maybe, might, nearly, occasionally, perhaps, possible, possibly, seldom, seldomly, sometimes, somewhat, suggest, suggests, uncertain and uncertainly.

2.3. Measure Development and Research Method

2.3.1 Intensity Measure

Cassell et al. (2013) argue that a large proportion of all comment letters are related to 10-K filings, and these letters are important to investors making investment decisions because 10-K filings contain highly significant financial information. Therefore, this study focuses on comment letters related to 10-K filings. Also, for simplicity, only initial letters sent by the SEC are used for empirical analysis.

Based on the "modal strong" list, the strong-word ratio in each initial letter is calculated as the number of strong words divided by the total number of words in the text (although some comment letters deal with more than one filing and may include comments on filings other than 10-K, the strong-word ratios on the overall letter level are

still used due to inability to identify exact comments on 10-K filings). Therefore, the intensity of a comment letter is measured by its strong-word ratio, and the higher the ratio, the higher the intensity. Similarly, the weak-word ratio is also calculated as the number of weak words divided by the total number of words in the clean text[1] in the comment letters.

2.3.2 Research Model and Sample Selection

Prior research suggests that the severity of a comment letter can be positively related to a restatement of the reviewed filing. To be more specific, the probability that the comment letter would result in a restatement is positively associated with the severity of the original comment letter, as measured by the number of topics included. Therefore, it is reasonable to expect a positive association between the strong-word ratio of the comment letter and restatement of the reviewed 10-K filing. Since both strong and weak words can be used in the same comment letter, the weak-word ratio is also included as a control variable. In addition, several other control variables used in prior research are included in the analysis: internal control weakness, company size, company age, profitability, bankruptcy measure, sales growth, merger and acquisition indicator, restructuring indicator, litigation indicator (J. Francis, Philbrick, and Schipper 1994), auditor type, and auditor resign/dismiss indicator. Year effect and industry effect are also controlled.

To test the association between the level of intensity of comment letter and probability of restatement, the following model (1) is estimated using logistic regression:

[1] The clean text is obtained by removing the header which includes the company information and the footer of comment letters.

$$\begin{aligned} \text{restatement} = {} & \beta_0 + \beta_1 \text{strongratio} + \beta_2 \text{weakratio} + \beta_3 \text{icw} + \beta_4 \text{sz} + \beta_5 \text{companyage} + \\ & \beta_6 \text{loss} + \beta_7 \text{bankruptcyrank} + \beta_8 \text{growth} + \beta_9 \text{m\&a} + \beta_{10} \text{restructuring} + \beta_{11} \text{litigation} \\ & + \beta_{12} \text{big4} + \beta_{13} \text{second-tier} + \beta_{14} \text{resign} + \beta_{15} \text{dismiss} + \text{yeardummy} + \text{industrydummy} \end{aligned} \quad (1)^2$$

Data used in the empirical analysis are obtained from SEEKINF[3] website, Compustat, and Audit Analytics. Variable definitions are as follows:

restatement = An indicator variable equal to 1 if the company filed a 10-K restatement in year t, and equal to 0 otherwise

strongratio = Comment letter's strong-word ratio, calculated by the count of strong words used in the comment letter divided by the total number of words in the clean text of the comment letter.

weakratio = Comment letter's weak-word ratio, calculated by the count of weak words used in the comment letter divided by the total number of words in the clean text of the comment letter.

icw = An indicator equal to 1 if the internal control audit opinion or the management certification as reported in Audit Analytics is qualified for a material weakness in year t, and equal to 0 otherwise.

sz = A proxy for company size, as measured by the natural log of market capitalization, which is calculated by the product of CSHO and PRCC_F in Compustat.

companyage = The total number of years (through year t) for which total assets are reported.

---

[2] We have also examined the prediction capability of strongratio and weakratio by using different classification models (logistic regression, SVM, decision tree) to predict the restatement through 10-fold cross validation. We find that for logistic regression and decision tree, prediction with strongratio and weakratio has a little higher AUC than prediction without the two variables (i.e. 0.578 vs 0.575, and 0.505 vs. 0.499, respectively), but the difference is not significant.

[3] https://www.seekinf.com/

loss = An indicator equal to 1 if the earnings before extraordinary items (IB) is negative in year t, and equal to 0 otherwise.

bankruptcyrank = The decile rank of the company's Altman's Z-score[4]. Companies in the decile ranked from 10 to 1, with those having the poorest financial health assigned a value of 10 and companies with the highest financial health assigned a value of 1. growth = The percentage change in annual revenue from year t-1 to year t.

m&a = An indicator equal to 1 for non-zero acquisitions or mergers as reported on a pre-tax basis (AQP) in year t, and equal to 0 otherwise.

restructuring = An indicator equal to 1 for non-zero restructuring costs as reported on a pre-tax basis (RC) in year t, and equal to 0 otherwise.

litigation = An indicator equal to 1 if the company is in a highly litigious industry (SIC 2833-2836, 3570-3577, 3600-3674, 5200-5961, or 7370-7374) and equal to 0 otherwise.

big4 = An indicator equal to 1 if the auditor is a Big 4 audit firm, and equal to 0 otherwise.

second-tier = An indicator equal to 1 if the auditor is a second-tier audit firm (i.e., BDO Seidman, Crowe Horwath, Grant Thornton, or McGladrey & Pullen), and equal to 0 otherwise.

resign = An indicator equal to 1 if the auditor resigned in year t and equal to 0 otherwise.

[4] Altman's Z-score is calculated by 1.2*(working capital/total assets)+1.4*(retained earnings/total assets)+3.3*(earnings before interest and taxes/total assets)+0.6*(market value of equity/book value of total liabilities)+ 0.99*(sales/total assets) (Altman 1968).

dismiss = An indicator equal to 1 if the auditor was dismissed in year t and equal to 0 otherwise.

year dummy = Indicator variables for each fiscal-year represented in the sample.

industry dummy = Indicator variables for each industry represented by the first two digits of the SIC code.

The following sample selection steps are taken:

SEC comment letters for all companies from 2004-2015 from the SEEKINF website	84,928
Less: letters not about 10-K filings or not containing necessary information	(32,464)
Less: not initial letters	(34,705)
Sample letter: clean initial letters related to 10-K filings	17,759
Less: duplicates[5]	(60)
Less: observations not able to match Compustat and Audit Analytics or with missing value for any variable in the model	(7,567)
Final sample:	10,132

In the merged dataset of Compustat and Audit Analytics, there are 5,196 observations for restatements in total, and in the final sample, there are 1,095 observations for restatements. Therefore, 4,101 observations are restatements without a

[5] For some companies in some years, there are 10-Ks for different fiscal year ends in the same year (different month/day but same year). All these observations are dropped.

comment letter. The proportion of restatements with/without a comment letter is 0.211/0.789.

2.4 Results

Summary statistics for the sample are presented in Table 1. Mean value (0.003) of strongratio shows that strong words are not frequently used in the comment letters. This is reasonable because the SEC uses a template for comment letters and the language used is mostly formal. However, compared with the mean value, the standard deviation (0.002) of strongratio shows that there is variation among comment letters in the use of strong words. The situation is similar for the weak-word ratio.

Table 2 presents a correlation matrix of the variables, which shows a positive correlation between restatement and strongratio and a negative correlation between restatement and weakratio. Table 3 presents the regression result for model (1). The result shows that the probability of restatement of the 10-K filings is positively associated with the strong-word ratio of the comment letters related to the 10-K filings (coefficient=48.394, t-value=2.21, significant at 5%). That is, 10-K filings that receive comment letters of higher intensity are more likely to have restatement. It suggests that the strong or weak modal nature of SEC comment letters may reflect perceived deficiencies in the reviewed filings, and therefore are related to the probability of restatement of the reviewed filing.

Table 1. Summary statistics for the sample (Chapter2)

Variable	Obs	Mean	Std	Median	Min	Max
restatement	10132	0.108	0.310	0	0	1
strongratio	10132	0.003	0.002	0.003	0	0.013
weakratio	10132	0.009	0.003	0.008	0	0.024
icw	10132	0.095	0.294	0	0	1
sz	10132	6.549	2.289	6.702	-4.096	13.348
companyage	10132	21.735	16.212	17	1	64
loss	10132	0.304	0.460	0	0	1
bankruptcyrank	10132	4.486	2.721	4	1	10
growth	10132	0.425	8.134	0.067	-9.286	438.000
m&a	10132	0.158	0.365	0	0	1
restructuring	10132	0.308	0.462	0	0	1
litigation	10132	0.300	0.458	0	0	1
big4	10132	0.726	0.446	1	0	1
secondtier	10132	0.069	0.254	0	0	1
resign	10132	0.014	0.119	0	0	1
dismiss	10132	0.050	0.218	0	0	1

Table 2. Correlation matrix of the variables (Chapter 2)

Panel A: Pearson Correlation Matrix – Part1								
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
(1) restatement	1							
(2) strongratio	0.02	1						
(3) weakratio	-0.01	0.11	1					
(4) icw	0.04	-0.03	-0.02	1				
(5) sz	-0.01	-0.02	0.09	-0.27	1			
(6) companyage	-0.01	0.01	0.04	-0.12	0.37	1		
(7) loss	0.01	0.02	-0.08	0.17	-0.43	-0.2	1	
(8) bankruptcyrank	0	0	-0.03	0.15	-0.25	0.03	0.38	1
(9) growth	0	-0.02	-0.02	0.04	-0.03	-0.03	0.02	-0.01
(10) m&a	-0.02	0.01	0.01	-0.05	0.17	0.04	-0.06	0.01
(11) restructuring	-0.01	0.01	0.03	-0.04	0.2	0.19	0.04	0.16
(12) litigation	-0.01	0	-0.03	0.01	-0.02	-0.15	0.1	0.05
(13) big4	0.02	-0.01	0.07	-0.22	0.62	0.22	-0.24	-0.09
(14) secondtier	0.01	0.02	-0.02	0.02	-0.09	-0.05	0.03	-0.07
(15) resign	0.01	-0.02	-0.02	0.1	-0.11	-0.04	0.05	0.02
(16) dismiss	0.02	-0.01	-0.02	0.12	-0.12	-0.06	0.06	0.04
Panel B: Pearson Correlation Matrix – Part2								
	(9)	**(10)**	**(11)**	**(12)**	**(13)**	**(14)**	**(15)**	**(16)**
(9) growth	1							
(10) m&a	-0.01	1						
(11) restructuring	-0.02	0.14	1					
(12) litigation	0.01	0.04	0.07	1				
(13) big4	-0.04	0.1	0.23	0.01	1			
(14) secondtier	0.02	-0.04	-0.05	0.02	-0.44	1		
(15) resign	0	-0.03	-0.05	0.02	-0.16	0.07	1	
(16) dismiss	0.02	-0.02	-0.04	-0.02	-0.15	0.06	-0.03	1

Table 3. Result for regression (1) (Chapter 2)

	coefficient	t-value	p-value
strongratio	48.394**	2.21	0.03
weakratio	-20.146*	-1.69	0.09
icw	0.298**	2.78	0.01
sz	-0.034	-1.47	0.14
companyage	0.001	0.41	0.68
loss	0.115	1.36	0.17
bankruptcyrank	-0.034**	-2.15	0.03
growth	-0.001	-0.32	0.75
m&a	0.083	0.83	0.41
restructuring	0.011	0.14	0.89
litigation	-0.056	-0.42	0.68
big4	0.248**	2.13	0.03
secondtier	0.211	1.42	0.16
resign	0.094	0.36	0.72
dismiss	0.211	1.49	0.14
constant	-1.955*	-1.85	0.07
year dummy		controlled	
industry dummy		controlled	
n		10075	
Pseudo R2		0.0315	

*, ** represent significance at the 0.1 and 0.05 level. Estimated with logistic regression.

2.5. Word List Modification and Discussion

The Loughran and McDonald (2011) “modal strong” and “modal weak” word lists were originally established for textual analysis in financial applications, especially financial disclosures. However, comment letters may be different from other financial filings and may have unique wording features. Therefore, the word lists may not be perfectly suitable to capture the intensity of comment letters. To investigate this problem, a word list modification process is conducted to evaluate whether the words in the word

lists could suggest higher or lower intensity as expected, and to eliminate words that do not.

2.5.1 Modification Methodology

First, for each word in the word lists, 50 sample sentences that have used the word are randomly selected from all the comment letters in our sample and are extracted into an Excel file. Second, the sample sentences for each word are read and scored to evaluate whether the word is relevant to the intensity of comment letters, meaning that it is used to indicate strong/weak modality as intended to. If the word in the sample sentence is irrelevant to intensity, the sentence is scored as 0, and if the word is relevant, the sentence is scored as 1. For this step, two coders work independently on scoring, and differences in scoring are solved by discussions between the coders. Table 4 presents some sample sentences and their scores. From the scored sample sentences, we notice that some words are never used, including: uncompromising, undisputed, undoubtedly, seldom, seldomly, definitely, conceivable and uncertainly. Some words are sometimes relevant and sometimes irrelevant, such as: could, apparently, etc. Some words are mostly irrelevant, including: occasionally, unsurpassed, unparalleled, lowest, highest, depends, depending, depend, almost and uncertain. We eliminate the words that are never used and the words for which all the sample sentences are scored as 0 from the original word lists to create the modified word lists. By this method the strong word ratio and weak word ratio will change because while the denominator stays the same the numerator will change since the words to be counted as strong/weak words will be different. This will also change the regression result. Since it is a strict criterion to eliminate words, 50 sample sentences are

large enough to make the decision. Using these modified lists, we re-run the regression for model (1), and the results are shown in Table 5.

Table 4. Sample sentences and scores (Chapter 2)

Sample Sentences	Score
The legality opinion opines upon shares to be issued by the company in a **best** efforts offering.	0
You state that **occasionally** as an agent you procure material and equipment on behalf of your clients for which you do not record revenues and costs.	0
There is no authority on point governing the validity of this allocation, and it is possible that the IRS **could** successfully challenge it.	0
In addition, please explain why these estimates were not **apparently** corrected in your restatements or whether you underestimated accruals for remediation efforts undertaken in 2004.	1
If you choose not to include these payments, a footnote to the table should **clearly** identify the excluded items and provide any additional information that is material to an understanding of your cash requirements.	1
We believe your disclosures **could** be improved by using tables to present dollar and percentage changes in amounts, rather than including such information in narrative form.	1

2.5.2 Discussion

Results in Table 5 show that by using the modified word lists, the t-value for strongratio improves slightly, suggesting that the effect is stronger. This improvement is consistent with the argument that the Loughran and McDonald (2011) lists are not perfect for identifying strong and weak words in comment letters, and that proper modifications are needed. However, Chi-square tests show that the changes in the coefficients for strongratio (p-value=0.622) and weakratio (p-value=0.842) are not significant.

In addition, in order to find out whether some strong words are more important in comment letters, the usage rate of each strong word between the restatement group and non-restatement group is calculated as the ratio of each word used in restatement observations divided by the ratio of the word used in non-restatement observations. The larger the usage rate, the more important the word in the sense that it's more likely to result in restatement if the word is used. Table 6 presents the usage rates for each word in the modified strong word list. The words "never", "unambiguously", and "strongly" may be more important according to the results. In addition, to check the co-occurrence of words, a term-document matrix is created (count of each word for each document) and correlation matrix is calculated based on this matrix. Table 7 presents the correlation matrix. The words are not strongly correlated, and the words with the highest correlation are "unequivocal" and "unequivocally", for which the correlation coefficient is 0.5.

Table 5. Regression result using modified word lists (Chapter 2)

	coefficient	t-value	p-value
strongratio	48.951**	2.24	0.03
weakratio	-19.938*	-1.67	0.10
icw	0.298**	2.78	0.01
sz	-0.034	-1.47	0.14
companyage	0.001	0.41	0.68
loss	0.115	1.36	0.17
bankruptcyrank	-0.034**	-2.15	0.03
growth	-0.001	-0.32	0.75
m&a	0.083	0.82	0.41
restructuring	0.011	0.14	0.89
litigation	-0.056	-0.42	0.67
big4	0.249**	2.14	0.03
secondtier	0.211	1.42	0.16
resign	0.093	0.35	0.72
dismiss	0.211	1.49	0.14
constant	-1.960*	-1.85	0.06
year dummy		controlled	
industry dummy		controlled	
n		10075	
Pseudo R2		0.0315	

*, ** represent significance at the 0.1 and 0.05 level. Estimated with logistic regression.

Table 6. usage rates of each word between restatement group and non-restatement group (Chapter 2)

	always	best	clearly	definitively	must	never
usage rates betweem two groups	0.990	0.974	1.057	0.635	1.046	2.153
	strongly	unambiguously	unequivocal	unequivocally	will	
usage rates betweem two groups	1.651	2.063	0.000	0.000	1.001	

Table 7. Correlation matrix of the strong words (Chapter 2)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1) always	1										

(2) best	0.05	1									
(3) clearly	0.03	0.03	1								
(4) definitively	0.00	0.00	0.03	1							
(5) must	0.04	0.03	0.13	0.03	1						
(6) never	0.03	-0.01	0.03	0.00	0.03	1					
(7) strongly	0.00	0.05	0.03	0.00	0.07	0.00	1				
(8) unambiguously	0.00	0.00	-0.01	0.00	-0.01	0.00	0.00	1			
(9) unequivocal	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	1		
(10) unequivocally	0.00	0.02	0.02	0.00	0.06	0.08	0.00	0.00	0.50	1	
(11) will	0.07	0.11	0.22	0.03	0.29	0.02	0.06	0.02	0.07	0.06	1

2.6. Conclusion and Future Research

Research on the comment letter should be of interest to stakeholders, such as investors who use it to measure the firm's compliance and reporting quality, or managers who are impacted by the costs associated with the SEC review process (Cassell et al. 2013). This study develops an intensity measure for SEC comment letters based on the modal (strong or weak) nature of words used in the comment letters. Specifically, Loughran and McDonald (2011) "modal strong" and weak word lists are used to identify words in the comment letters and the strong-word ratio is calculated to measure the level of intensity of comment letters. In contrast to severity measures in prior literature, the intensity measure is based on the content of comment letters directly. Empirical analysis on a sample of initial comment letters related to 10-K filings shows that the strong-word ratio in the comment letter is positively associated with the probability of restatement of the reviewed 10-K filings. The result indicates that the SEC staff may use modal strong words to express higher concerns on some reviewed filings.

In addition, we argue that the Loughran and McDonald (2011) word lists may not be perfectly suitable for identifying strong/weak words in comment letters because

comment letters are different from financial filing documents. Therefore, we evaluate each word's relevance to intensity of comment letters and modified the word lists accordingly. Using the modified word lists, we again find a positive and significant association between strong word ratio and restatement of reviewed 10-K filings with slightly stronger results.

This study has several limitations. First, some comment letters review several different filings simultaneously, but in this study, the strong-word ratio is calculated on the overall letter, rather than on specific 10-K related comments, due to our inability to identify comments specific to 10-K filings. However, even the overall intensity of a comment letter is found to be positively related to the probability of restatement of 10-K filings, leaving open the possibility that the effect could be stronger at the 10-K related comment level. Second, even though we have modified the word lists, we only eliminate irrelevant words and it is possible that some words not in the word lists may suggest higher or lower intensity. Future research could identify these words to further improve the modal word lists for this context.

Chapter 3: Audit Fees and Voluntary Disclosure Tone

3.1. Introduction

Audit fees are mainly determined by the costs that auditors might incur if they fail to identify material misstatements and by the audit effort associated with the engagement (Stice 1991). Based on those linkages, previous studies have suggested various audit risk proxies such as indicators related to audit complexity, inherent risk, client business risk, and corporate governance (Hay, Knechel, and Wong 2006). Nevertheless, those proxies are frequently developed based on numerical information.

This study examines the association between the abnormally positive and negative tone of voluntary qualitative disclosures, which cannot be explained by concurrent, relevant financial quantitative indicators, and audit fee decisions. We also examine if the potential audit fee premium or discount for abnormal tone is conditioned on management's prior disclosure behaviors. Prior studies find that the tone of earnings press releases can be used to measure management opportunistic behaviors (e.g., Huang, Teoh, and Zhang 2014) as well as signaling regarding a firm's future performance (e.g., Davis, Piger, and Sedor 2012), both of which are associated with audit risk. However, scholars have rarely investigated whether and how the tone of management qualitative voluntary disclosures is associated with audit fee decisions. Our study fills this void by using enhanced versions of the audit fee model that incorporates both quantified qualitative management disclosures and financial variables (Hay et al. 2006).

Firms issuing overly positive earnings press releases are likely to pose greater inherent risks. Auditing standards state that overly optimistic press releases can be red flags for misstatement since management might face strong pressure to meet targets set in those press releases (AICPA 2002). Prior studies indicate that deceptive qualitative disclosures can indicate financial reporting risk; management that discloses less credible qualitative disclosures might tend to disclose misstated financial statements (Humpherys, Moffitt, Burns, Burgoon, and Felix 2011; Patelli and Pedrini 2015). In addition, the tone of earnings press releases can be used to measure the level of management strategic behavior since management is more likely to strategically utilize vague qualitative disclosures than specific quantitative disclosures (Huang et al. 2014). Thus, we predict that abnormally positive earnings press releases are associated with higher audit risk, thereby increasing audit fees.

On the other hand, the tone of earnings press releases might provide credible signals to evaluate audit risk especially related to client business risk. Client business risk arises when the client's economic condition will weaken in the short and/or long term, and prior studies often suggest financial indicators such as proxies for client business risks such as ROA (Colbert, Luehlfing, and Alderman 1996). Prior studies present earnings press release language as an incremental informative signal to understand and predict a firm's performance (Henry 2008; Kothari, Li, and Short 2009; Demers and Vega 2011). However, the auditing standard states that the value of information used to evaluate audit risk is largely associated with the reliability of the information, and internally generated information is generally considered less reliable than externally

generated information (AICPA 2006). Based on the findings of prior studies, we expect that qualitative earnings press releases can be used as a proxy for client business risks if managers offer credible information. Studies show that negative news disclosure tends to be more credible than positive news disclosure (Williams 1996; Hutton, Miller, and Skinner 2003) since management has strong incentives to disclose positive news (McNichols 1989). Accordingly, we expect that abnormally negative tone of earnings press releases can be used as a proxy of client business risks.

To investigate whether abnormally negative or positive press releases are associated with audit fees, we measure the abnormal tone of management earnings press releases that cannot otherwise be explained by relevant financial indicators (hereafter, "abnormal tone"[1]). To evaluate abnormal tone, we use a similar method to the model proposed by Huang et al. (2014)[2] and measure abnormal tone using the residuals from a cross-sectional regression of the tone of management disclosures along a series of tone determinants introduced by Li (2010). Since abnormally positive and negative tones provide different insights regarding audit risk, and since we assume that firms disclosing overly positive tone and negative tone have different risk profiles, we subset cases disclosing abnormally positive tone and abnormally negative tone and examine the

[1] Huang et al. (2014) refer abnormal tone to abnormal "positive" tone. It is important to note that abnormal positive tone is able to be negative. To avoid a possible confusion, we refer abnormal positive tone, which is the residual of the tone model proposed by Huang et al. (2014) as "abnormal tone" in this paper. In addition, we consider positive and negative abnormal tones separately, depending on whether they are larger or smaller than 0.

[2] While Huang et al. (2014) use the annual earnings press releases to calculate tone, this paper uses all earnings press releases over the year to calculate the average tone for the fiscal year because annual press releases may not contain all the information in the quarterly press releases.

associations between positive (negative) abnormal tone and audit fees. Furthermore, if management has previously issued abnormally positive earnings forecasts, auditors tend to think that current earnings forecasts are less credible (Feng and Li 2014). Hence, we also examine whether the association between abnormal tone and audit fees varies depending on the credibility of management disclosures.

The results are partially consistent with the arguments developed above. Using firms that issued earnings press releases via 8-K filings from 2006 to 2015, we could not find a statistically meaningful association between the abnormally positive tone of press releases and audit fees except where it is extremely positive. On the other hand, we find a negative association between the abnormally negative tone of press releases and audit fees. These findings suggest that an abnormally negative tone in earnings press releases can proxy for client business risks. However, except in extreme cases, we could not find any obvious evidence to support the usefulness of an abnormally positive tone as an indicator of opportunistic managerial disclosure behaviors for understanding audit fee decisions. These inconsistent results may be influenced by the tone of press releases in general. Since managers may have greater opportunity to influence the language utilized in press releases as compared to other filings, the tone of press releases tends to be more positively biased (Davis and Tama-Sweet 2012). Therefore, when abnormal tone is slightly positive, it might not deliver enough insights to measure managerial opportunistic disclosure behavior. These findings can imply that abnormal tone can only proxy for audit risk in certain cases.

The findings in this paper relate to the current line of inquiry concerning the implications of qualitative information on a firm's fundamentals. Prior studies have highlighted the importance of qualitative disclosures which might offer meaningful insights that quantitative information might not provide. Despite an extensive literature on the usefulness of qualitative information in understanding a firm's fundamentals, little is known about its function as a legitimate risk indicator for external auditors. In clarifying this association, we emphasize the importance of understanding management qualitative disclosures.

In addition, this study may add to the literature on auditing by highlighting the tone of press releases as an important proxy to measure audit risk. Academics have gauged how auditors have estimated audit risk by using publicly available information, even though auditors clearly evaluate a wider scope of information, including private information, to assess risks. To overcome this intrinsic limitation, a variety of determinants have been introduced. As progressively advanced research methods have been employed, various novel ways to measure audit risk have been utilized, such as using CEO narcissism as a proxy of inherent and control risks (Judd, Olsen, and Stekelberg 2017) and a measure based on the risk factor section in 10-K filings (Yang et al. 2018). This study adds to this literature by measuring the tone of management voluntary disclosure, which is the outcome of management's complex strategic formula, as a distinct insight to measure audit risk.

The remainder of this chapter is divided into six sections: Section 3.2 summarizes the literature related to the topics presented in this paper and develops the hypotheses,

Section 3.3 describes the research design and sample selection, Section 3.4 explains results, Section 3.5 contains additional analyses, Section 3.6 presents robustness tests, and Section 3.7 provides conclusions and limitations.

3.2. Related Literature and Hypothesis Development

3.2.1. Abnormal Tone of Earnings Press Releases and Audit Fees

Earnings press releases are generally recognized as an important outlet for managers to communicate firm performance to stakeholders, and the information covered by earnings press releases has expanded over time (Davis et al. 2012). Earnings press releases, containing both qualitative and quantitative information, are voluntary disclosures, and there is little regulation over the disclosures. Because of these attributes, prior studies show that examining the contents of earnings press releases offers a chance to measure signals managers send to their stakeholders regarding their expectation about future performance and suggest the incremental values of earnings press releases in the financial market (Bowen, Davis, and Matsumoto 2005; Levi 2007; Riley 2011; Tsileponis, Stathopoulos, and Walker 2016).

Audit fees are the outcomes of an auditor's economic decision-making process regarding risk assessments of their clients. Audit fees are comprised of two main factors: effort and expected loss (Simunic 1980). These two factors are negatively correlated, so expected losses tend to decrease as audit production or effort increases. Certain aspects, such as inherent risk, can also be linked to audit fees because these factors require auditors to make greater effort and use more resources in performing audit procedures

(Seetharaman, Gul, and Lynn 2002). Auditing standards indicate that auditors should evaluate indicators that clients may have strong incentives to commit financial statement fraud, such as high client business risks and heavy pressures to meet expectations or requirements (AU 316.85 A2) (AICPA 2002). We believe that the tone of earnings press releases can be a reasonable proxy of audit risk, which will lead to higher audit fees. This study does not argue that publicly issued earnings press releases are used by auditors when they formulate audit fee decisions, since they are obviously able to access private information from managers. On the other hand, as Feng and Li (2014) note, we also insist that if managers publicly issue earnings press releases containing earnings forecasts, auditors are likely to collect consistent private information from managers with the contents of earnings press releases. If there is inconsistency between the two, the credibility of any private information offered by managers might be reduced. In this case, auditors are likely to consider a client as high risk. Consequently, managers tend to offer consistent private information to auditors. We therefore posit that the contents of earnings press releases can be a proxy for information provided by managers to auditors.

Firms issuing abnormally positive earnings press releases are likely to pose greater inherent risks due to highlighted financial reporting risks. When management has strong pressure to meet expectations (potentially indicated by the expectations management used in overly optimistic press releases), it has an incentive to commit fraud resulting in misstatements (AICPA 2002). In addition, prior studies find the association between deceptive disclosure language and financial reporting risks. Management disclosing deceptive qualitative information might tend to disclose deceptive quantitative

information, as well (Humpherys et al. 2011; Patelli and Pedrini 2015). Moreover, the tone of qualitative earnings press releases might offer a unique setting to gauge audit risk. Since it is easier to make vague qualitative disclosure as opposed to quantitative disclosure, management frequently uses strategically crafted qualitative disclosures. Since qualitative voluntary disclosures are often thought to be less material in causing shareholder litigations (Palmiter 2017) management may concludes that it has more freedom in formulating qualitative disclosures than quantitative disclosures. Huang et al. (2014) show a positive association between abnormal tone of earnings press releases, which cannot be explained by concurrent, relevant financial quantitative indicators, and the possibility of future earnings restatement. Thus, we predict that abnormally positive earnings press releases are associated with high audit risk, leading to the following hypothesis:

H1: Audit fees are positively associated with the abnormally positive tone of earnings press releases.

Auditing standards indicate that auditors should evaluate factors associated with client business risk to evaluate financial reporting risks (AICPA 2002). Financial statements are more susceptible to material misstatement when client business risk is high. In such instances, a client does not have sufficient resources to provide reliable reporting and might have a motivation to manipulate financial reports to hide poor performance (Stanley 2011). Prior studies often suggest numerical measurements, such as ROA, as proxies of client business risk (Hay et al. 2006).

Qualitative earnings press release data may provide incremental insight into client business risks that quantitative measurements cannot deliver. Numerical indicators measuring a firm's business risks are highly correlated with each other (Li 2006), but management qualitative disclosure often provides risk information not captured by numerical information. Prior studies examine whether qualitative management disclosures provide incremental value to understand a firm's fundamentals (Henry 2008; Kothari et al. 2009; Demers and Vega 2011). For instance, Davis et al. (2012) find a significant positive (negative) association between the level of optimistic (pessimistic) tone in earnings press releases and future ROA.

On the other hand, auditing standards state that the value of information in assessing audit risk is largely associated with the information's reliability, which is affected by its source and nature (AICPA 2006). In general, internally generated information is less reliable than externally generated information, making management earnings forecasts inappropriate as general audit evidence, and requiring auditors to evaluate the credibility of prospective financial information (PCAOB 2001). Krishnan et al. (2012) report that auditors perceive earnings forecasts, often included as the quantitative element of earnings press releases, as a risk incentive that might actually be related to fraud risks since management frequently releases forecasts for strategic purposes. Based on prior studies associating qualitative earnings press releases and a firm's fundamentals, we expect that qualitative earnings press releases can proxy for client business risks if managers offer credible information. Studies show that negative news disclosures tend to be more credible than positive news disclosures (Williams 1996;

Hutton et al. 2003) since management has strong incentives to disclose positive news (McNichols 1989). Accordingly, we argue that abnormally negative tone of earnings press releases can be used as a proxy of client business risks and predict that overly pessimistic tone of earnings press releases is associated with high audit risk. This leads to the following hypothesis:

H2: Audit fees are negatively associated with the abnormally negative tone of earnings press releases.

As previously discussed, the credibility of earnings press releases determines their value as a proxy for client business risk. The auditing standards state that a comparison between prospective financial information in prior periods and corresponding actual results can be used for credibility assessment (AICPA 2002). If auditors perceive a large discrepancy between the contents of earnings press releases previously issued and corresponding actual results, then they are likely to find internally generated information less credible (Feng and Li 2014). In addition, Francis and Krishnan (1999) suggest that auditors tend to react conservatively to unverified audit evidence. These findings imply that even if managers issue truthful earnings press releases, auditors might suspect their credibility when they perceive previously issued press releases to be overly optimistic. This leads to a third hypothesis:

H3: The association between the abnormal tone of earnings press releases and audit fees is weaker if managers previously issue abnormally positive earnings press releases.

3.3. Research Design

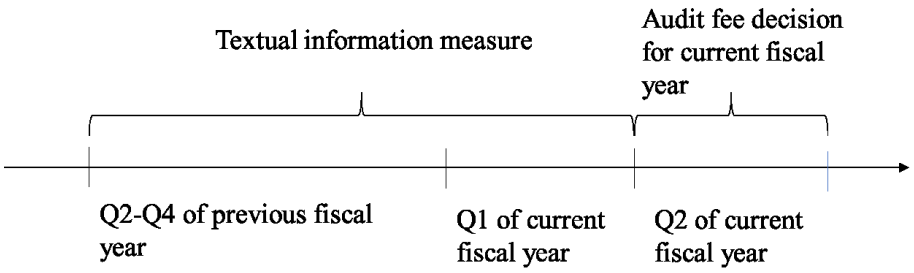
3.3.1. Sample Selection

We obtain quarterly earnings press releases from 2006 to 2015 via the SEEKINF database (www.seekinf.com)[3] to create the sample for this study. The firms' financial information is gathered from Compustat, the stock return data can be found in CRSP, and the information on analysts' earnings forecast is obtained from the I/B/E/S database. Quarterly textual data is first matched with the CRSP/Compustat (Fundamentals Quarterly) and then merged with the database by Central Index Key (CIK) and announcement dates. The tone of quarterly earnings press releases is calculated using a dictionary-based method that will be discussed in the following section, and the average tone[4] for each fiscal year is calculated and used as the overall annual tone of earnings press releases for that fiscal year. Figure 1 presents the timeline of textual information. All continuous variables, other than returns and variables that take logarithms, are winsorized at the 1% and 99% level.

Figure 1. Timeline of textual information

[3] The SEEKINF database provides links for downloading all press releases, and we limit the sample to earnings press releases by requiring that the press release is filed on the earnings announcement date and the press release contains the key word "earning(s)".

[4] We calculate average tone of management press releases during the period from the second quarter of previous fiscal year and the first quarter of current fiscal year and match it to current fiscal year because audit fee decisions are typically made in the second quarter of the current fiscal year (Hackenbrack, Jenkins, and Pevzner 2014).



Panel A of Table 1 shows the sample selection process. In total, there are 17,753 observations in the tone regression, and 14,598 observations in the audit fee regression, for the fiscal years 2006 to 2015. Panel B of Table 1 presents the sample composition by industry[5], Panel C presents the sample composition by fiscal year, and Panel D shows the distribution of earnings press releases during the sample period. Firms issue between 3 and 4 earnings press releases each year on average.

Table 8. Sample information (Chapter 3)

Panel A: Sample Selection

	Firm-Year Observations
Compustat & CRSP Merged database in fiscal year 2006-2015, excluding observations with blank cik and missing fiscal year	56,127
Less: merge data with RET & STDRET	(388)
Less: merge segment data	(0)
Less: merge data to calculate age	(1)
Less: merge with press data	(24,253)
Less: merge with IBES	(10,837)
Less: missing variables to estimate abnormal tone	(2,895)
Total number of observations for tone regression	17,753
Less: missing variables for audit fee model	(3,156)
Final sample for examining the association between abnormal tone and audit fees	14,598

[5] Here, the Fama-French 12 industry classification is used: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/det_12_ind_port.html

Panel B: Sample Composition by Fama and French 12 industry classification

Industry	Number of observations	Percentage
Customer non-durables	810	5.55
Consumer durables	361	2.47
Manufacturing	1,546	10.59
Oil, Gas, and Coal Extraction and Products	707	4.84
Chemicals and Allied Products	438	3.00
Business Equipment	2,477	16.97
Telephone and Television Transmission	410	2.81
Utilities	604	4.14
Wholesale, Retail, and Some Services	1,646	11.28
Healthcare, Medical Equipment, and Drugs	1,194	8.18
Finance	2,752	18.85
Other	1,653	11.32
Total	14,598	100

Panel C: Sample Composition by Fiscal Year

Fiscal Year	Number of observations	Percentage
2006	1,431	9.80
2007	1,483	10.16
2008	1,466	10.04
2009	1,486	10.18
2010	1,509	10.34
2011	1,466	10.04
2012	1,475	10.10
2013	1,419	9.72
2014	1,435	9.83
2015	1,428	9.78
Total	14,598	100

Panel D: Distribution of Press Releases Issued by Fiscal Year

Fiscal Year	Mean number of press releases	standard deviation	Total number of press releases
2006	3.19	0.98	8,607
2007	3.22	1.00	9,040
2008	3.33	0.95	9,198
2009	3.39	0.93	9,192
2010	3.44	0.89	9,538
2011	3.47	0.87	9,726
2012	3.49	0.86	9,780
2013	3.51	0.85	10,132
2014	3.51	0.85	10,711
2015	3.49	0.85	10,821
Total	3.41	0.91	96,745

3.3.2. Proxy for Textual Information

Both statistical and rule-based (dictionary-based) methods have been used in previous studies to measure the tone of textual information (Li 2006; Loughran and McDonald 2011). While there are some benefits to statistical approaches, there is a multi-faceted rationale for favoring a rule-based approach. First, a dictionary-based method requires the investigator to follow explicit rules in order to generate the desired results, and it is therefore easier to understand and follow compared to such statistical methods as the Naïve Bayes Classifier which is discussed by Li (2010). Second, supervised statistical methods require a large sample of labeled observations. This requires considerable manual work to train the classifier, and accuracy cannot always be assured. Third, dictionaries that have been applied in other studies (Bonini, Pavesi, and Scotti 2011; Rogers, Van Buskirk, and Zechman 2011) are readily available. In particular, Loughran and McDonald (2011) have developed a series of word lists that can be applied to textual analysis in financial areas. Their lists offer six categories (positive, negative, uncertainty, litigious, strong modal, and weak modal words), and were last updated in March 2015[6].

Based on Loughran and McDonald's (2011) positive and negative word lists[7], the tone of an earnings press release (TONE) is measured as the difference between the number of positive words and negative words divided by the total number of words (excluding stop words[8]) in the release. In the same way that Loughran and McDonald

[6] See https://www3.nd.edu/~mcdonald/Word_Lists.html.

[7] We use the latest version of Loughran and McDonald sentiment word lists from their website.

[8] Stop words are words that are commonly used in texts such as pronouns, prepositions, and conjunctions etc. Stop words are identified using the Natural Language Tool Kit (NLTK) package in Python.

(2011) deal with negation, we count a positive word within three words after negation words (i.e., no, not, none, neither, never, nobody) as a negative word.

As Huang et al. (2014) indicate, an abnormal tone of an earnings press release (ABTONE) is measured as the residual from cross-sectional regression of TONE on a series of tone determinants introduced by Li (2010). These determinants capture information related to the fundamentals, performance, operational risks, and complexity of the firm. Industry[9] and year fixed effects are also included in the regression. Specifically, the model is presented below, and variable definitions are provided in Appendix B.

$$TONE_t = \alpha + \beta_1 EARN_t + \beta_2 RET_t + \beta_3 SIZE_t + \beta_4 BTM_t + \beta_5 STDRET_t + \beta_6 STDEARN_t + \beta_7 AGE_t + \beta_8 BUSSEG_t + \beta_9 GEOSEG_t + \beta_{10} LOSS_t + \beta_{11} \Delta EARN_t + \beta_{12} AFE_t + \beta_{13} AF_t + IND + YEAR + \varepsilon \quad (2)$$

Lobo and Zhao (2013) suggest that audit decisions are affected by the pre-audit probability of misstatements. Therefore, pre-audit misstatement risk should be controlled when analyzing the association between abnormal tone and audit fee. As with Lobo and Zhao (2013), pre-audit misstatement risk is measured by PSCORE, which is calculated as the predicted probability from the misstatement detection model (Dechow, Ge, Larson, and Sloan 2011) following the logistic specification shown below. All variables are defined in Appendix B.

$$RESTATEMENT_t = \alpha + \beta_1 TOTALACCURAL_t + \beta_2 \Delta REC_t + \beta_3 \Delta INV_t + \beta_4 SOFTASSET_t + \beta_5 \Delta CSALE_t + \beta_6 \Delta ROA_t + \beta_7 ISSUANCE_t + \beta_8 \Delta EMP_t + \beta_9 LEASE_t + \beta_{10} ABRET_t + \beta_{11} LAGABRET_t + \varepsilon \quad (3)$$

[9] Industry fixed effect is specified using the two-digit SIC code.

In order to analyze how audit fee decisions are associated with abnormal tone in earnings press releases, the following audit fee model is applied. Quantitative factors, as suggested by recent studies (Francis, Reichelt, and Wang 2005; Krishnan, Sami, and Zhang 2005; Ghosh and Pawlewicz 2009; Choi, Kim, Kim, and Zang 2010; Stanley 2011), are controlled. Definitions of variables in the model are provided in Appendix B.

$$\begin{aligned} LNFEE_t = {} & \alpha + \beta_1 ABTONE_{t-1} + \beta_2 PSCORE_t + \beta_3 ROAEARNING_t + \beta_4 SIZE_t + \beta_5 INVREC_t + \beta_6 BUSSEG_t \\ & + \beta_7 FOREIGN_t + \beta_8 MERGE_t + \beta_9 SPECIAL_t + \beta_{10} LEVERAGE_t + \beta_{11} CURRENTRATIO_t + \beta_{12} LOSS_t \\ & + \beta_{13} BTM_t + \beta_{14} GROWTH_t + \beta_{15} BIG4_t + \beta_{16} RESIGNATION_t + \beta_{17} DISMISSAL_t + \beta_{18} GC_t + \beta_{19} IW_t \\ & + \beta_{20} TENURE_t + \beta_{21} INDEXPERT_t + IND + YEAR + \varepsilon \end{aligned} \quad (4)$$

[10, 11]

In order to test H1 and H2, the sample is divided into two subsamples to test the associations between audit fees and the abnormally positive and negative earnings press release tones. If H1 is true, the association between the abnormally positive tone and audit fees should be positive. If H2 is true, the association between the abnormally negative tone and audit fees should be negative.

H3 examines whether the association between abnormal tone and audit fees is moderated by the credibility of earnings press releases. As Feng and Li (2014) do, we consider that auditors regard earnings press releases as being more credible when the previous year's earnings press releases lack a highly abnormal tone, and are less credible when the previous year's earnings press releases contain a highly abnormal tone. To test

[10] We exclude the predicted tone from Model (1) because it is not of interest to us, and because it is highly correlated with control variables (e.g., *SIZE*, *BUSSEG*, etc.), it is not necessary in Model (3). In addition, since predicted tone and abnormal tone provide different dimensions of management qualitative information, the results are generally consistent even if we add predicted tone to audit fee model.

[11] We use t-1 to annotate the timing of tone in order to show that the tone is measured right before audit fee decision. Some people may argue that audit fee decision is made even before the current fiscal year. Therefore, as a robustness test, we also use previous year abnormal tone as the explanatory variable in the audit fee regressions, and the main results are consistent with our results.

H3, two new variables are generated based on ABTONE: ABTONE_G, and ABTONE_B. Specifically, if the prior year's abnormal tone is among the top 25% of the entire earning press releases, we collect as the full sample, we consider it as less credible. Therefore, ABTONE_B equals ABTONE if ABTONE > = (P75) ABTONE (i.e., 0.0041), and 0 otherwise. Similarly, ABTONE_G equals ABTONE if ABTONE < (P75) ABTONE (i.e., 0.0041), and 0 otherwise. ABTONE_G and ABTONE_B capture abnormal tone in credible/non-credible situations separately, so using both in regression (4) allows us to evaluate how auditors deal with more credible and less credible information. If H3 is true, then only the coefficient of ABTONE_B should be positive and significant, or the coefficient of ABTONE_B should be higher than the coefficient of ABTONE_G for the subsample with abnormally positive tone (indicating higher audit fee premium when the firm continues to act opportunistic). On the other hand, only the coefficient of ABTONE_G should be negative and significant for the subsample with abnormally negative tone (indicating audit fee discount for firms with less pessimistic earnings press releases only when the press releases are trustful).

Stanley (2011) uses a change specification of the audit fee model to control for unobserved client factors that affect audit fees. Using the same method, this paper tests whether the results are robust for the change specification. Specifically, in the change specification, all variables are the first-order difference in the level variable between successive years, and prior year mispricing is also controlled by adding the prior period's unexpected audit fee in the regression. Unexpected audit fee is the residual from Regression (4).

3.4. Results

Table 9 indicates the estimated results for Regression (2). The results are generally similar to Huang et al. (2014) in that tone is found to be more positive for smaller, growing, more profitable firms, and for firms with fewer business segments. Unlike Huang et al. (2014), but consistent with Li (2010), we find that tone is more positive for firms with lower stock return volatility and lower earnings volatility.

Table 9. Estimating expected tone (Chapter 3) [12]

		TONE		Huang et al. (2014)	
Variables	+/-	Coefficient	t-value	Coefficient	t-value
EARN	+	0.0137***	11.63	0.0011**	2.47
RET	+	-0.0001	-0.58	0.0000	0.01
SIZE	-	-0.0003***	-3.93	-0.0002***	-3.34
BTM	-	-0.0016***	-6.64	-0.0013***	-4.52
STDRET	+	-0.0081***	-3.30	0.0690***	7.58
STDEARN	-	-0.0057**	-2.70	0.0000	-0.05
AGE	-	-0.0000	-1.45	-0.0003	-1.63
BUSSEG	-	-0.0002	-0.74	-0.0006***	-4.44
GEOSEG	+	-0.0003	-1.20	0.0002	0.79
LOSS	-	-0.0038***	-14.66	-0.0013***	-4.48
ΔEARN	-	-1.0338***	-6.85	-0.0012	-1.19
AFE	+	-0.0014***	-3.84	0.0008***	3.10
AF	-	0.0057	1.42	-0.0001	-0.30
Observations		17,753		14,475	
Adjusted R-squared		0.2486		0.0441	
Fixed effects		Industry & Year		Industry & Year	
Cluster		Firm & Year		Firm & Year	
Years covered		2006-2015		1997-2007	

*, **, ***denote the significance levels of 0.1, 0.05, and 0.01 (two-tailed), respectively, using t statistics adjusted for firm and year clustering. This table reports

[12] Although some of our coefficients have different sign than Huang et al. (2014), they are in the same direction as Li (2010).

the results of regression (1) explaining tone of earnings press release. Constant is absorbed by the fixed effects. The sample is composed of 17,753 firm-year observations covering fiscal year 2006-2015. Variable definitions are provided in Appendix B.

Panel A of Table 10 presents descriptive statistics for variables in the main regression. On average, audit fees for firms in the sample are about $3,428,963. By construction, the mean value of ABTONE is almost 0 (mean = 0.0003). However, the standard deviation is relatively large compared to the mean (std = 0.007), showing considerable variation, similar to Huang et al. (2014). Panel B of Table 10 presents the univariate comparison between subsamples with positive ABTONE and negative ABTONE observations. Compared to firms with a negative ABTONE, firms with a positive ABTONE have lower audit fees and have lower probability of misstatement, lower leverage, and higher sales growth. They are more likely to disclose special items, less likely to be audited by Big 4 firms, and less likely to receive going-concern opinions. The result indicates that the two subsamples have different risk profiles. Thus, it makes sense to analyze them separately.

Table 10. Descriptive statistics (Chapter 3)

Panel A: Descriptive Statistics for Dependent and Independent Variables

Variables	Mean	Median	Std.	P25	P75
AUDIT_FEES	3,428,963	1,603,000	6,484,108	837,000	3,568,800
LNFEE	14.381	14.287	1.078	13.638	15.088
TONE	-0.005	-0.004	0.008	-0.009	0.001
ABTONE	< 0.001	< 0.001	0.007	-0.004	0.004
PSCORE	0.036	0.036	0.007	0.031	0.040
ROAEARNING	0.076	0.074	0.098	0.026	0.123

SIZE	7.761	7.665	1.762	6.536	8.876
INVREC	0.278	0.227	0.212	0.109	0.390
BUSSEG	1.191	1.000	0.466	0.693	1.609
FOREIGN	0.750	1.000	0.433	1.000	1.000
MERGE	0.269	< 0.001	0.444	< 0.001	1.000
SPECIAL	0.150	< 0.001	0.357	< 0.001	< 0.001
LEVERAGE	0.578	0.576	0.245	0.399	0.758
CURRENTRATIO	2.144	1.615	1.668	1.000	2.595
LOSS	0.187	< 0.001	0.390	< 0.001	< 0.001
BTM	0.597	0.490	0.497	0.286	0.774
GROWTH	0.084	0.059	0.232	-0.024	0.156
BIG4	0.869	1.000	0.338	1.000	1.000
RESIGNATION	0.004	< 0.001	0.060	< 0.001	< 0.001
DISMISSAL	0.029	< 0.001	0.167	< 0.001	< 0.001
GC	0.004	< 0.001	0.062	< 0.001	< 0.001
IW	0.035	< 0.001	0.183	< 0.001	< 0.001
TENURE	2.004	2.079	0.614	1.792	2.398
INDEXPERT	0.337	< 0.001	0.473	< 0.001	1.000

Panel B: Univariate Comparison between Positive ABTONE and Negative ABTONE

Variables	Positive ABTONE Sample Mean	Negative ABTONE Sample Mean	Difference
LNFEE	14.353	14.412	-0.059***
PSCORE	0.035	0.036	-0.001***
ROAEARNING	0.077	0.076	< 0.001
SIZE	7.776	7.744	0.033
INVREC	0.277	0.280	-0.003
BUSSEG	1.189	1.192	-0.003
FOREIGN	0.745	0.756	-0.010*
MERGE	0.270	0.269	0.001
SPECIAL	0.157	0.143	0.015***
LEVERAGE	0.567	0.588	-0.021***
CURRENTRATIO	2.159	2.128	0.032
LOSS	0.186	0.187	< 0.001

BTM	0.598	0.597	< 0.001
GROWTH	0.102	0.065	0.037***
BIG4	0.863	0.874	-0.011**
RESIGNATION	0.004	0.003	<0.001
DISMISSAL	0.028	0.029	-0.001
GC	0.003	0.005	-0.003***
IW	0.034	0.036	-0.002
TENURE	2.005	2.002	0.003
INDEXPERT	0.336	0.338	-0.002

*, **, ***denote the significance levels of 0.1, 0.05, and 0.01 (two-tailed). This table presents differences in means between positive ABTONE and negative ABTONE observations on the variables employed in the audit fee regression. All variables are defined in Appendix B.

Table 11 presents the correlation matrix for variables in audit fee model. The correlation between ABTONE and LNFEE is negative and significant at the 0.01 level (r = -0.03, p-value = 0.00). However, it should be noted that abnormal positive and negative tones may have opposite associations with audit fees. Therefore, the negative association does not indicate that there is a negative association between over-optimistic or over-pessimistic disclosures and audit fees.

Table 11. Pearson correlation (Chapter 3)

Panel A: Pearson Correlation Matrix – Part1								
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
(1) LNFEE	1							
(2) ABTONE	**-0.03**	1						
(3) PSCORE	**0.23**	-0.05	1					
(4) ROAEARNING	**0.11**	**0.02**	**0.03**	1				
(5) SIZE	**0.74**	0.01	0.02	**0.07**	1			
(6) INVREC	**-0.19**	0	**0.04**	**-0.05**	0.01	1		
(7) BUSSEG	**0.34**	0.01	**0.11**	**0.06**	**0.25**	**-0.06**	1	
(8) FOREIGN	**0.27**	-0.01	**0.04**	**0.04**	**0.24**	**0.15**	**0.11**	1
(9) MERGE	**0.13**	0.01	**0.06**	**-0.03**	**0.09**	**-0.04**	**0.06**	**0.1**
(10) SPECIAL	**0.01**	0.02	**-0.02**	**-0.02**	**0.04**	0.01	**0.03**	0
(11) LEVERAGE	**0.17**	**-0.07**	**-0.05**	**-0.18**	**0.43**	**0.25**	**0.07**	**0.07**
(12) CURRENTRATIO	**-0.21**	**0.02**	-0.01	**0.06**	**-0.4**	**-0.1**	**-0.13**	-0.01
(13) LOSS	**-0.08**	**-0.02**	**0.06**	**-0.53**	**-0.2**	**-0.06**	**-0.07**	**-0.09**
(14) BTM	**-0.08**	0	**0.03**	**-0.33**	**0.1**	**0.23**	**0.03**	0
(15) GROWTH	**-0.08**	**0.1**	**-0.41**	**0.13**	**-0.09**	**-0.04**	**-0.05**	**-0.05**
(16) BIG4	**0.44**	**-0.02**	**0.11**	**0.11**	**0.27**	**-0.24**	**0.11**	**0.04**
(17) RESIGNATION	**-0.03**	0	-0.01	**-0.01**	**-0.02**	0.01	-0.01	0
(18) DISMISSAL	**-0.08**	0	**-0.03**	**-0.04**	**-0.06**	**0.04**	-0.01	-0.01
(19) GC	-0.01	**-0.02**	0	**-0.14**	**-0.04**	**-0.02**	-0.01	**-0.02**
(20) IW	-0.01	-0.01	**0.02**	**-0.09**	**-0.09**	0.01	**-0.02**	-0.01
(21) TENURE	**0.29**	0	**0.08**	**0.08**	**0.26**	-0.1	**0.08**	**0.07**
(22) INDEXPERT	**0.25**	0.01	**0.06**	**0.08**	**0.18**	-0.08	**0.1**	**0.03**

Panel B: Pearson Correlation Matrix – Part2								
	(9)	**(10)**	**(11)**	**(12)**	**(13)**	**(14)**	**(15)**	**(16)**
(9) MERGE	1							
(10) SPECIAL	-0.01	1						
(11) LEVERAGE	0	**0.02**	1					
(12) CURRENTRATIO	-0.01	-0.01	-0.61	1				
(13) LOSS	-0.01	**-0.08**	0.05	**0.03**	1			
(14) BTM	**-0.02**	0.01	0.05	**-0.09**	**0.25**	1		
(15) GROWTH	**0.11**	-0.01	-0.11	**0.08**	**-0.13**	**-0.16**	1	
(16) BIG4	0.01	0.01	0.01	**-0.07**	**-0.06**	**-0.14**	**-0.05**	1
(17) RESIGNATION	-0.01	-0.01	0	0	**0.02**	0.01	0	**-0.06**

(18) DISMISSAL	0.01	0.01	0.01	0.01	**0.04**	**0.02**	**0.03**	**-0.1**
(19) GC	**-0.02**	-0.01	0.06	**-0.02**	**0.12**	**0.04**	0	0
(20) IW	0	0	-0.02	**0.03**	**0.11**	**0.03**	0.01	**-0.03**
(21) TENURE	**0.08**	0	0.04	**-0.06**	**-0.1**	**-0.09**	**-0.1**	**0.4**
(22) INDEXPERT	-0.01	0	0.02	**-0.05**	**-0.05**	**-0.04**	**-0.03**	**0.28**

Panel C: Pearson Correlation Matrix – Part3

	(17)	**(18)**	**(19)**	**(20)**	**(21)**	**(22)**
(17) RESIGNATION	1					
(18) DISMISSAL	-0.01	1				
(19) GC	0	**0.03**	1			
(20) IW	**0.06**	**0.04**	**0.07**	1		
(21) TENURE	**-0.17**	**-0.44**	**-0.02**	**-0.07**	1	
(22) INDEXPERT	-0.01	**-0.02**	0	-0.01	0.14	1

This table presents the Pearson correlation matrix among the variables. Numbers in bold denote the significance levels of 0.05 (two-tailed). All variables are defined in Appendix B.

Table 12 presents the estimation results of Regression (4). Panel A shows the results for testing H1 and H2. In Panel A, the result of the audit fee regression in terms of ABTONE in Model (1) uses the full sample, Model (2) uses the subsample of observations with abnormally positive tone, and Model (3) uses the subsample of observations with abnormally negative tone. Overall, the result for Model (1) in Panel A suggests that abnormal tone is generally negatively related to audit fees. However, taking a closer look at the results for the subsamples, only an abnormally negative tone is associated with audit fees, but an abnormally positive tone is not related with audit fees. As shown in Model (2) and Model (3) of Panel A, the coefficient of ABTONE when abnormal tone is positive (coef = -0.91, p-value = 0.69), not significant, and of a much smaller magnitude than the coefficient of ABTONE when abnormal tone is negative (coef = -5.70, p-value = 0.00). These results support H2 and not H1, indicating that

abnormally negative earnings press releases rather than abnormally positive earnings press releases can signal higher client business risks that relate to higher audit fees.

Panel B of Table 12 shows the results for testing H3. In Panel B, Model (1) presents the results for the subsample with abnormally positive tone, and Model (2) presents the results for the subsample with abnormally negative tone. The results show that only ABTONE_G in the abnormally negative tone subsample has a significant and negative coefficient (coef. = 5.55, p-value = 0.00). This result supports H3, indicating that only an abnormally negative tone with high perceived credibility can serve as a good proxy for the client business risk. To be more specific, firms with highly abnormally positive tone in the previous year are less likely to have lower audit fees in the current year even if the current disclosure tone is less pessimistic.

Table 12. Regression results (Chapter 3)

Panel A: Association between Abnormal Tone and Audit Fees

	(1)		(2)		(3)	
VARIABLES	coef.	tstat	coef	tstat	coef	tstat
ABTONE	-4.8071***	-4.38	-0.9143	-0.40	-5.6991***	-3.20
PSCORE	7.7798***	4.62	9.4812***	3.96	5.7565***	3.08
ROAEARNING	-0.2918***	-2.89	-0.2808**	-2.12	-0.2871**	-2.32
SIZE	0.4986***	55.76	0.4988***	44.46	0.4968***	46.17
INVREC	0.2128**	2.22	0.1359	1.14	0.2602**	2.16
BUSSEG	0.1439***	6.38	0.1529***	5.28	0.1337***	5.08
FOREIGN	0.2138***	9.20	0.2070***	7.06	0.2263***	8.02
MERGE	0.0413***	2.93	0.0367**	2.03	0.0452**	2.52
SPECIAL	-0.0128	-1.08	-0.0002	-0.01	-0.0302*	-1.84
LEVERAGE	-0.0420	-0.78	-0.1143	-1.56	0.0377	0.61
CURRENTRATIO	-0.0182***	-2.61	-0.0192**	-2.25	-0.0166*	-1.88

LOSS	0.1806***	9.87	0.1923***	7.63	0.1615***	7.19
BTM	-0.0944***	-4.85	-0.1178***	-4.68	-0.0650***	-2.69
GROWTH	0.0274	0.90	0.0878**	2.02	-0.0455	-1.21
BIG4	0.3245***	11.62	0.3569***	10.48	0.2959***	7.79
RESIGNATION	-0.0314	-0.38	-0.0529	-0.50	0.0146	0.11
DISMISSAL	-0.1221***	-4.45	-0.1471***	-3.70	-0.0912**	-2.46
GC	0.1955**	2.53	0.3022***	2.65	0.0914	0.96
IW	0.2472***	9.43	0.2338***	6.21	0.2708***	7.61
TENURE	-0.0182	-1.25	-0.0434**	-2.42	0.0058	0.32
INDEXPERT	0.0550***	2.99	0.0512**	2.24	0.0623***	2.81
Observations	14,598		7,595		7,003	
Adjusted R-squared	0.8124		0.8153		0.8142	
Fixed effects	Firm & Year		Firm & Year		Firm & Year	
Cluster	Firm		Firm		Firm	

*, **, ***denote the significance levels of 0.1, 0.05, and 0.01 (two-tailed), respectively using t statistics adjusted for firm clustering (Petersen 2009). This table reports the results of regression (4) explaining audit fees (LNFEE). Constant is absorbed by the fixed effects. The sample is composed of firm-year observations covering fiscal years 2006-2015. The estimation of Model (1) is based on the audit fee model with ABTONE, the estimation of Model (2) is based on a subsample of observations with abnormally positive tone, and the estimation of Model (3) is based on a subsample of observations with abnormally negative tone. Variable definitions are provided in Appendix B.

Panel B: Association between Abnormal Tone and Audit Fees Moderated by Perceived Credibility

	(1)		(2)	
VARIABLES	coef	tstat	coef	tstat
ABTONE_G	-0.0507	-0.02	-5.5523***	-3.08
ABTONE_B	-2.3456	-0.78	-11.5511	-0.75
PSCORE	9.5080***	3.97	5.7702***	3.08
ROAEARNING	-0.2774**	-2.09	-0.2870**	-2.32
SIZE	0.4992***	44.43	0.4968***	46.14
INVREC	0.1369	1.15	0.2601**	2.16
BUSSEG	0.1531***	5.29	0.1338***	5.08
FOREIGN	0.2072***	7.06	0.2264***	8.02
MERGE	0.0366**	2.02	0.0452**	2.52
SPECIAL	-0.0007	-0.04	-0.0301*	-1.84
LEVERAGE	-0.1153	-1.58	0.0380	0.62
CURRENTRATIO	-0.0192**	-2.25	-0.0166*	-1.88
LOSS	0.1919***	7.62	0.1615***	7.19

BTM	-0.1186***	-4.72	-0.0650***	-2.69
GROWTH	0.0872**	2.01	-0.0458	-1.22
BIG4	0.3570***	10.48	0.2958***	7.79
RESIGNATION	-0.0537	-0.51	0.0144	0.11
DISMISSAL	-0.1469***	-3.70	-0.0910**	-2.45
GC	0.3023***	2.64	0.0911	0.96
IW	0.2338***	6.20	0.2704***	7.60
TENURE	-0.0430**	-2.40	0.0059	0.32
INDEXPERT	0.0515**	2.25	0.0622***	2.81
Observations	7,595		7,003	
Adjusted R-squared	0.8153		0.8141	
Fixed effects	Industry & Year		Industry & Year	
Cluster	Firm		Firm	

*, **, ***denote the significance levels of 0.1, 0.05, and 0.01 (two-tailed), respectively using t statistics adjusted for firm clustering (Petersen 2009). This table reports the results of Regression (4) explaining audit fees (LNFEE). Constant is absorbed by the fixed effects. The sample is composed of firm-year observations covering fiscal years 2006-2015. The estimation of Model (1) is for the audit fee model with ABTONE_G and ABTONE_B, based on a subsample of observations with abnormally positive tone, and the estimation of Model (2) is based on a subsample of observations with abnormally negative tone. Variable definitions are provided in Appendix B.

Table 13 presents the estimation results of the change specification in the audit fee model for the subsample of observations with an abnormally negative tone. Models (1) and (2) are the results of audit fee regression in terms of ΔABTONE. The results suggest that changes in an abnormal tone are negatively related to changes in audit fees, regardless of whether the abnormal tone is positive or negative. This indicates that audit fees are more likely to relate to changes in tone, rather than the direction of the change.

Table 13. Association between changes in abnormal tone and changes in audit fees (Chapter 3)

Dependent Variable	Δ LNFEE

Sample	ABTONE > 0		ABTONE < 0	
	(1)		(2)	
Variable	coef.	t-stat.	coef.	t-stat.
ΔABTONE	-1.5496***	-2.84	-1.3057**	-2.40
ΔPSCORE	1.8217***	3.17	0.6517	1.19
ΔROAEARNING	-0.2719***	-4.26	-0.2260***	-3.62
ΔSIZE	0.3406***	18.02	0.3094***	16.21
ΔINVREC	0.0964	1.25	0.1022	1.29
ΔBUSSEG	0.0298***	3.00	0.0178*	1.75
ΔFOREIGN	0.0257***	4.79	0.0328***	5.29
ΔMERGE	0.0033	0.75	-0.0024	-0.46
ΔSPECIAL	0.0914**	2.23	0.1970***	4.45
ΔLEVERAGE	-0.0108***	-2.87	-0.0069	-1.59
ΔCURRENTRATIO	0.0549***	6.45	0.0325***	4.43
ΔLOSS	0.0067	0.76	0.0016	0.19
ΔBTM	0.0521***	3.21	0.0184	1.07
ΔGROWTH	0.1728***	3.77	0.1925***	4.29
ΔBIG4	-0.0388	-0.86	0.0586	0.86
ΔRESIGNATION	-0.0666***	-3.87	-0.0448***	-2.85
ΔDISMISSAL	-0.0324	-0.70	0.0225	0.52
ΔGC	0.0372**	2.15	0.0435***	2.66
ΔIW	0.0418***	3.26	0.0591***	4.98
ΔTENURE	0.0181**	2.53	0.0187**	2.52
ΔINDEXPERT	-0.0795***	-14.62	-0.1080***	-16.55
$UNEXPECTEDFEE_{t-1}$	0.0073	0.49	0.0137	0.74
Observations	5,152		4,824	
Adjusted R-squared	0.2671		0.2720	
Fixed effects	Industry & Year		Industry & Year	
Cluster	Firm		Firm	

*, **, ***denote the significance levels of 0.1, 0.05, and 0.01 (two-tailed), respectively using t statistics adjusted for firm clustering (Petersen 2009). This table reports the change specification results for the audit fee model. Constant is absorbed by the fixed effects. The sample is composed of firm-year observations covering fiscal years 2006-2015. The estimation of Model (1) is based on the subsample with abnormally positive tone, and the estimation of Model (2) is based on the subsample with abnormally negative tone.

3.5. Additional Analyses

Two additional analyses are performed. The first tests whether audit fees are associated with an abnormally positive tone in extreme cases. Since earnings press releases are positively biased, abnormal positive tone might not offer an insight to gauge managers' strategic disclosure behavior. Accordingly, we test whether such an association is consistent even if managers issue extremely over-optimistic earnings press releases. By generating an indicator of EXPOSABTONE, which captures the extremely abnormally positive tone, we test whether audit fees are positively associated with EXPOSABTONE. Specifically, EXPOSABTONE is a dummy variable that equals 1 if ABTONE > (P95) ABTONE (i.e., 0.0105). The result is presented in TABLE 7. As expected, the findings show that an abnormally positive tone can be an indicator of management opportunistic behavior and thus financial reporting risk in the extreme case.

Table 14. Association between extremely abnormally positive tone and audit fees (Chapter 3)

	(1)	
VARIABLES	coef	tstat
EXPOSABTONE	0.0454**	2.24
PSCORE	7.7140***	4.55
ROAEARNING	-0.2453**	-2.48
SIZE	0.4953***	55.96
INVREC	0.1896**	1.97
BUSSEG	0.1421***	6.35
FOREIGN	0.2121***	9.37
MERGE	0.0452***	3.24
SPECIAL	-0.009	-0.76
LEVERAGE	-0.039	-0.72
CURRENTRATIO	-0.0171**	-2.5

LOSS	0.1802***	9.84
BTM	-0.0942***	-4.9
GROWTH	0.0003	0.01
BIG4	0.3402***	12.3
RESIGNATION	-0.0191	-0.23
DISMISSAL	-0.1331***	-4.9
GC	0.2261***	2.97
IW	0.2476***	9.69
TENURE	-0.0236*	-1.67
INDEXPERT	0.0568***	3.12
Observations	14,598	
Adjusted R-squared	0.8089	
Fixed effects	Industry & Year	
Cluster	Firm	

*, **, ***denote the significance levels of 0.1, 0.05, and 0.01 (two-tailed), respectively using t statistics adjusted for firm clustering (Petersen 2009). Constant is absorbed by the fixed effects. The sample is composed of firm-year observations covering fiscal years 2006-2015. The estimation of Model (1) is based on the audit fee model with EXPOSABTONE. Variable definitions are provided in Appendix B.

In the second analysis, we test whether audit experience with the client moderates the association between abnormal tone and audit fees. Specifically, two variables are used to measure audit experience with the client: 1) INITIAL, which measures whether it is the initial year of audit for the auditor; and 2) TENURE, which measures how long the auditor has audited the client. Adding the interaction of the audit experience and ABTONE in regression (3) allows us to examine whether audit experience affected the association between abnormal tone and audit fees. TABLE 8 shows the results for the regression with interactions for the subsample of abnormally negative tone observations. The coefficient of the interaction between ABTONE and INITIAL is negative and

significant at the 0.1 level (coef = -11.15, p-value = 0.082). The coefficient of the interaction between ABTONE and TENURE is positive and significant at the 0.05 level (coef = 4.95, p-value = 0.046). Both results show that the negative association between an abnormally negative tone and audit fees is less pronounced for auditors with more experience with the clients, indicating that the usefulness of publicly available earnings press releases as a proxy for audit risk is affected by previous auditor-client experience.

Table 15. Association between abnormal tone and audit fees moderated by audit experience (Chapter 3)

	(1)		(2)	
VARIABLES	coef	tstat	coef	tstat
ABTONE	-5.2789***	-2.92	-15.6264***	-3.11
INITIAL	-0.1559***	-2.82		
ABTONE*INITIAL	-11.1456*	-1.74		
TENURE	0.0061	0.33	0.0325	1.50
ABTONE*TENURE			4.9448**	2.00
PSCORE	5.7265***	3.06	5.6961***	3.05
ROAEARNING	-0.2872**	-2.32	-0.2868**	-2.32
SIZE	0.4967***	46.21	0.4963***	46.22
INVREC	0.2597**	2.16	0.2608**	2.17
BUSSEG	0.1339***	5.09	0.1341***	5.10
FOREIGN	0.2270***	8.05	0.2272***	8.08
MERGE	0.0455**	2.54	0.0452**	2.52
SPECIAL	-0.0303*	-1.85	-0.0306*	-1.87
LEVERAGE	0.0373	0.61	0.0374	0.61
CURRENTRATIO	-0.0166*	-1.88	-0.0169*	-1.91
LOSS	0.1610***	7.17	0.1611***	7.19
BTM	-0.0647***	-2.67	-0.0639***	-2.64
GROWTH	-0.0466	-1.24	-0.0461	-1.23
BIG4	0.2958***	7.78	0.2968***	7.81
RESIGNATION	0.1005	0.74	0.0043	0.03
DISMISSAL			-0.0948**	-2.55
GC	0.0922	0.98	0.0886	0.94
IW	0.2705***	7.60	0.2700***	7.59

INDEXPERT	0.0623***	2.82	0.0626***	2.83
Observations	7,003		7,003	
Adjusted R-squared	0.8143		0.8143	
Fixed effects	Industry & Year		Industry & Year	
Cluster	Firm		Firm	

*, **, ***denote the significance levels of 0.1, 0.05, and 0.01 (two-tailed), respectively using t statistics adjusted for firm clustering (Petersen 2009). This table presents the estimation results of the moderation effects of audit experience on the association between abnormally negative tone and audit fees. Constant is absorbed by the fixed effects. The sample is composed of firm-year observations covering fiscal years 2006-2015 with abnormally negative tone. The estimation of Model (1) is based on the audit fee model with ABTONE and ABTONE*INITIAL, and the estimation of Model (2) is based on the audit fee model with ABTONE and ABTONE*TENURE. Variable definitions are provided in Appendix B.

We also test whether audit firm size, auditors' expertise, and pre-audit risk modify the association between abnormal tone and audit fees, but we find no statistically significant results. In un-tabulated results, the coefficient of the interaction between BIG4 and ABTONE is 1.209 and not significant (p-value = 0.66), and the coefficient of the interaction between INDEXPERT and ABTONE is 1.639 and not significant (p-value = 0.42). For pre-audit misstatement risk, we generate a categorical variable (PREAUDIT) indicating low, moderate, or high pre-audit misstatement risk based on 5% percentile and 95% percentile of PSCORE. The coefficient of the interaction between PREAUDIT and ABTONE is 0.583 and not significant (p-value = 0.81).

3.6. Robustness Tests

H3 suggests that the value of information provided by the abnormal tone of the earnings press releases should depend on the credibility of the press releases. Our chosen way to measure the credibility of earnings press releases is to determine whether previous

issued earnings press releases have a highly abnormally positive tone. As a robustness test, we use another way to measure credibility: whether the firm has an annual report restatement in the previous year. The intuition is that firms with restatements are less likely to be trusted. Specifically, we develop ABTONE_GR and ABTONE_BR to represent accurate inaccurate management disclosures based on previous year restatement of the annual report. ABTONE_GR equals ABTONE if $RESTATEMENT_{t-1}$=0, and 0 otherwise. Similarly, ABTONE_BR equals ABTONE if $RESTATEMENT_{t-1}$=1, and 0 otherwise. The results of running the fee regression separately on subsamples with abnormally positive and negative tones are presented in Table 16, which shows that for the subsample with abnormally positive tone, when the firm had annual report restatement the previous year, audit fees are higher when the firm continues to have an abnormally positive tone in the current year, although the result is only marginally significant. On the other hand, for the subsample with abnormally negative tone, when the firm did not have restatement the previous year, audit fees are lower when the abnormal tone is less negative. The results indicate that an abnormally positive tone can signal higher financial reporting risk when there is already some evidence of the firm's untrustful disclosing behavior, and the value of an abnormally negative tone as a proxy of client business risk is stronger when the firm is believed to disclose truthfully.

Table 16. Association between Abnormal Tone and Audit Fees Based on Previous Restatement (Chapter 3)

	(1)		(2)	
Variable	coef.	t-stat.	coef.	t-stat.
ABTONE_GR	-1.2380	-0.54	-5.8199***	-3.21

ABTONE_BR	26.1797*	1.73	-11.7253	-0.84
PSCORE	9.4531***	3.96	5.7588***	3.08
ROAEARNING	-0.2808**	-2.12	-0.2868**	-2.32
SIZE	0.4988***	44.50	0.4967***	46.15
INVREC	0.1369	1.16	0.2607**	2.16
BUSSEG	0.1519***	5.24	0.1337***	5.08
FOREIGN	0.2063***	7.04	0.2263***	8.03
MERGE	0.0359**	1.99	0.0452**	2.52
SPECIAL	0.0001	0.01	-0.0303*	-1.85
LEVERAGE	-0.1170	-1.60	0.0377	0.61
CURRENTRATIO	-0.0191**	-2.25	-0.0166*	-1.88
LOSS	0.1920***	7.61	0.1614***	7.19
BTM	-0.1193***	-4.73	-0.0650***	-2.68
GROWTH	0.0868**	2.00	-0.0453	-1.21
BIG4	0.3548***	10.44	0.2958***	7.79
RESIGNATION	-0.0566	-0.53	0.0147	0.11
DISMISSAL	-0.1456***	-3.66	-0.0906**	-2.44
GC	0.3010***	2.60	0.0912	0.96
IW	0.2309***	6.12	0.2709***	7.61
TENURE	-0.0434**	-2.42	0.0060	0.32
INDEXPERT	0.0511**	2.24	0.0623***	2.82
Observations	7,594		7,003	
Adjusted R-squared	0.8155		0.8142	
Fixed effects	Industry & Year		Industry & Year	
Cluster	Firm		Firm	

*, **, ***denote the significance levels of 0.1, 0.05, and 0.01 (two-tailed), respectively using t statistics adjusted for firm clustering (Petersen 2009). This table presents the estimation results of the audit fee regression using previous restatement to measure the credibility of current earnings press releases. Constant is absorbed by the fixed effects. The sample is composed of firm-year observations covering fiscal years 2006-2015. The dependent variable is LNFEE, and variables of interest are ABTONE_GR and ABTONE_BR, which are composed based on whether the firm has annual report restatement in the previous year. The estimation of Model (1) is based on the subsample of earnings press releases with abnormally positive tone, and the estimation of Model (2) is based on the subsample of earnings press releases with abnormally negative tone.

As another robustness test, we add error in management forecasts as an additional control variable in regression (3). Krishnan et al. (2012) find that audit fees are positively related to management forecast error. To test whether the effect of an abnormal tone in earnings press releases is dominated by management forecast error, we re-run regression (3) with management forecast error as a control variable. The results, presented in Table 17, are consistent with the results of the regression without management forecast error, suggesting that an abnormal tone can be a distinct indicator to measure audit risk. Specifically, abnormally negative earning press releases signal higher audit risk.

Table 17. Association between abnormal tone and audit fees controlling for management forecast error (Chapter 3)

	(1)		(2)		(3)	
VARIABLES	coef	tstat	coef	tstat	coef	tstat
ABTONE	-5.1755***	-4.61	-0.5122	-0.21	-5.3490***	-2.80
MFERROR	0.0007***	6.53	0.0008***	6.04	0.0005***	4.13
PSCORE	7.8615***	4.56	9.4119***	4.00	5.7639***	2.89
ROAEARNING	-0.3368***	-3.05	-0.3151**	-2.12	-0.3806***	-2.91
SIZE	0.4772***	49.05	0.4708***	38.38	0.4850***	41.37
INVREC	0.3558***	3.56	0.3000**	2.46	0.3535***	2.70
BUSSEG	0.1463***	6.51	0.1562***	5.55	0.1340***	5.07
FOREIGN	0.2449***	10.30	0.2461***	8.23	0.2480***	8.35
MERGE	0.0385***	2.60	0.0486***	2.65	0.0221	1.13
SPECIAL	-0.0151	-1.16	-0.0004	-0.02	-0.0403**	-2.16
LEVERAGE	-0.0092	-0.17	-0.0652	-0.92	0.0668	1.08
CURRENTRATIO	-0.0163**	-2.37	-0.0167**	-2.05	-0.0159*	-1.81
LOSS	0.1436***	7.42	0.1427***	5.31	0.1352***	5.69
BTM	-0.1139***	-5.72	-0.1363***	-5.21	-0.0826***	-3.31
GROWTH	0.0172	0.53	0.0776*	1.74	-0.0489	-1.18
BIG4	0.2489***	7.07	0.3118***	6.98	0.1868***	4.06
RESIGNATION	0.0139	0.15	-0.0316	-0.28	0.0944	0.60

DISMISSAL	-0.1202***	-3.64	-0.1315***	-2.69	-0.1070**	-2.44
GC	0.1019	1.06	0.2145	1.49	0.0425	0.37
IW	0.2302***	7.83	0.2303***	5.56	0.2483***	6.17
TENURE	-0.0090	-0.56	-0.0339*	-1.68	0.0174	0.87
INDEXPERT	0.0544***	3.06	0.0439*	1.95	0.0669***	3.10
Observations	11,496		5,968		5,528	
Adjusted R-squared	0.7969		0.7997		0.7994	
Fixed effects	Industry & Year		Industry & Year		Industry & Year	
Cluster	Firm		Firm		Firm	

*, **, ***denote the significance levels of 0.1, 0.05, and 0.01 (two-tailed), respectively using t statistics adjusted for firm clustering (Petersen 2009). This table presents the estimation results of the audit fee regression with management forecast error as an additional control variable. Constant is absorbed by the fixed effects. The sample is composed of firm-year observations covering fiscal years 2006-2015. The estimation of Model (1) is based on the audit fee model with the full sample, the estimation of Model (2) is based on the audit fee model with the subsample of observations with abnormally positive tone, and the estimation of Model (3) is based on the audit fee model with the subsample of observations with abnormally negative tone. Variable definitions are provided in Appendix B.

Thus, our results show that the tone of qualitative earnings press releases can serve as a proxy for audit risk. Specifically, a more abnormally negative tone of earnings press releases signals higher audit risk and is positively related to audit fees. On the other hand, for earnings press releases with an abnormally positive tone, only the presence of an extremely abnormally positive tone can signal higher audit risk and relates to higher audit fees. In addition, the association between audit fees and the abnormal tone of earnings press releases is moderated by auditors' experience with their clients. We also show that qualitative disclosures provide a distinct measure for audit risk in addition to quantitative management forecasts.

3.7. Conclusions

This chapter investigates whether the tone of qualitative earnings press releases provides incremental value in understanding of the process of setting audit fees. We find that the tone of a press release that cannot be measured by relevant financial indicators (i.e., abnormal tone) is associated with audit fees only if the abnormal tone is abnormally negative. On the other hand, we could not find a meaningful association between the abnormally positive tone and audit fees on average. These results suggest that in general abnormal positive tone might not be a signal of management strategic behavior, but an abnormally negative tone can indicate client business risk. However, when the abnormal tone is extremely positive, the association between audit fees and abnormal tone becomes clearer. Moreover, we find evidence that the association between audit fees and abnormal tone is different depending on the credibility of previous earnings press releases.

This result might be consistent with the findings of prior studies and the statement of the accounting standard which highlight a tendency toward positively biased management press releases. Therefore, the tone of earnings press releases provides incremental value in understanding audit fee decisions when managers issue abnormally pessimistic earnings press release or extremely optimistic press releases.

Chapter 4: Does the Internal Control Risk Factor Disclosure Complement SOX 404 Disclosure of Material Weaknesses in Internal Control?

4.1. Instruction

Understanding internal control of a firm is essential to stakeholders because the effective internal control is the foundation for more complete and reliable financial statements (AICPA 2014). Ashbaugh-Skaife, Collins, and Lafond (2009) argue that internal control issues can increase information risk and even influence the capital market. Prior studies have documented that material internal control weakness is associated with lower accruals quality, less efficient investment, and higher audit fees (Hoitash, Hoitash, and Bedard 2008; Ashbaugh-Skaife et al. 2009; Cheng, Dhaliwal, and Zhang 2013). Section 404 of Sarbanes–Oxley Act (SOX 404) and Auditing Standard 2201 (Public Company Accounting Oversight Board (PCAOB) 2017) require that management and auditors disclose material weaknesses in internal controls over financial reporting (ICFR) (Raghunandan and Rama 2006). In addition to this mandatory assessment on ICFR, firms can also disclose ICFR-related risk factors in section Item 1A – Risk Factors of annual reports. Only a small portion of firms disclose material weaknesses. Many firms with significant internal control deficiencies not rising to the level of materiality may discuss these issues in the risk factor section. Unlike material weakness disclosures, there are no specific requirements on the disclosure of ICFR-related risk factors in Item 1-A (Coleman and Plante 2017).

Although the Securities and Exchange Commission (SEC) highlights that firms should "avoid generic risk factor disclosure that could apply to any company" (SEC 2010), it is practically difficult to evaluate whether firms are including trivial issues in the ICFR-related risk factor disclosure. This is because managers are not required to quantify the likelihood of the risks (Campbell et al. 2014) and there is no ex-post settling up of the disclosure (Schrand and Elliott 1998). Thus, ICFR-related risk factor disclosure is relatively more flexible compared to the disclosure of material weaknesses. Firms have several possible options to disclose ICFR-related risk factor. First, firms may want to reduce the litigation risk by including all possible risks even if some risks are too trivial and not firm-specific (Campbell et al. 2014), which makes the disclosure boilerplate and lacks informativeness. Second, firms with significant deficiencies in internal controls that do not reach the level of materiality may choose to provide such information in the risk factor section. This disclosure can provide additional insight for the statement users in addition to material internal control weakness disclosure (e.g., Kravet and Muslu 2013; Campbell et al. 2014; Bao and Datta 2014). Third, firms may also conceal immaterial internal control issues to avoid potential negative consequences of disclosing ICFR-related risk factors. The majority of academic research on ICFR reporting has focused on the material weakness disclosure under SOX 404; limited attention has been paid to the role of Item 1-A risk factor disclosure related to ICFR. (Coleman and Plante 2017) claim that many firms are using the risk factor section to enhance their ICFR reporting and that there might be additional information about the firm's internal control in risk factor disclosures. Notably, the article points out that firms typically do not disclose significant deficiencies that do not rise to a material level in SOX 404 reports but are likely to

include them as ICFR-related risk factors. As a result, the risk factor section may be a potential source to provide insight into the firm's internal control over financial reporting.

This study attempts to investigate whether firms meaningfully disclose ICFR-related risk factors to provide information about deficiencies in internal controls and whether the disclosure of ICFR-related risk factors can complement SOX 404 disclosure of material weaknesses in internal control to predict related adverse outcomes in the future. Focusing on the risk factor section in public firms' annual reports from 2005 to 2016, the study finds that the ICFR-related risk factor disclosures are positively associated with restatements of the current period's annual report and are predictive of future periods' adverse outcomes including restatement and material internal control weaknesses. Also, the inclusion of ICFR-related risk factors is positively related to audit fees when the firm does not disclose material weaknesses in internal control. This study further examines the contents of ICFR-related risk factors by investigating four characteristics including length, specificity, ranking, and level of boilerplate and finds that the risk factors that are longer, more specific, and positioned toward the beginning of Item 1A indicate more severe internal control risk.

The contribution of this study is three-fold. First, it provides evidence that the disclosure of ICFR-related risk factors is overall informative and provides incremental information about the firm's potential internal control deficiencies. Under the requirement of SOX 404, firms must disclose material internal control weaknesses. In addition to that, the risk factor section may be another source for investors to learn about the firm's internal control status, especially when the internal control deficiencies do not rise to the materiality level. However, the risk factor section is hard to assess due to its

high flexibility. Therefore, it is essential to examine the effectiveness and informativeness of ICFR-related risk factor disclosures. Second, the paper contributes to the auditing literature by showing that ICFR-related risk factor disclosure contains useful information about internal control issues and can predict restatement of the corresponding annual report as well as future restatements and material weaknesses in internal control. This may be useful for the regulators and auditors in decision-making. Third, the paper contributes to the literature that studies the informativeness of risk factor disclosure (e.g., Abraham and Shrives 2014). It provides supporting evidence that risk factor disclosure reveals firms' risk types and risk level and can serve as indicators of future adverse consequences.

The remaining part of this chapter is organized as follows. Section 4.2 presents the review of related literature and hypotheses development. Research design and empirical results are provided in Section 4.3 and Section 4.4, respectively. Section 4.5 discusses additional analysis and Section 4.6 concludes.

4.2. Related Literature and Hypothesis Development

The Securities and Exchange Commission (SEC) has long emphasized the importance of firms providing informative disclosure to investor instead of merely conforming to regulations. Substantial efforts have been made to alleviate the consequences of information asymmetry between firms and investors, including many regulations and scrutiny (Bamber and McMeeking 2016; Elshandidy et al. 2018). After a series of accounting scandals in the early 21st century, there is growing attention in firms' internal control and the disclosure of the internal control deficiencies. Since 2002, effective under Sarbanes-Oxley Act Section 302 (SOX 302), management is required to

evaluate the effectiveness of internal control and indicate material weakness in their SEC filings (Rice and Weber 2012). Later in 2004, Sarbanes-Oxley Act Section 404 further requires external auditors to attest to the accuracy of the management assessment of internal control (Rice and Weber 2012). Under SOX 404, any material weakness in internal control must be disclosed in the firm's annual report, usually in section Item 9A, while internal control issues that do not reach material level are not required to be disclosed under SOX 404 (Coleman and Plante 2017). In 2005, SEC requires public firms to discuss crucial factors that may affect the firm and its security in a separate section called "Item 1A - risk factors" (SEC 2005). With this requirement, firms may discuss immaterial internal control issues in Item 1A. To improve the relevance and informativeness of risk factor disclosure, Regulation S-K 503 (2018) emphasizes that firms should "not present risks that could apply to any issuer".

However, it is hard in practice to evaluate whether firms disclose significant factors, or they simply include symbolic, generic factors that are overly generalized. To reduce the potential litigation risk of omitting important risk factors in their disclosure, firms may merely disclose all possible risks without assessing the likelihood of the risk affecting the firm (Campbell et al. 2014). Since managers are not required to quantify the probability of the risks (Campbell et al. 2014), the disclosure of ICFR-related risk factors is relatively more flexible and less definitive compared to SOX 404 disclosure of material weaknesses. Despite the growing interest in firms' overall risk disclosure especially via SOX 404 material weaknesses disclosure, academic understanding of firms' risk factor disclosure is still quite limited. This study intends to fill the research gap by focusing on

the role of ICFR-related risk factors and how they reflect firms financial reporting deficiencies.

The paper first examines the association between ICFR-related risk factor disclosure and firms' current- and future-period financial restatements. Firms with internal control deficiencies are more likely to contain deficiencies in financial reports and experience subsequent restatements of the annual report (e.g., Li and Wang 2006; Hoitash et al. 2008; Nagy 2010; Rice and Weber 2012). Hence, the first hypothesis is stated as follows:

H1: Disclosure of ICFR-related risk factors is positively related to the restatement of the current year's annual report.

Risk factor disclosures are forward-looking. Specific types of qualitative risk disclosures are found to be informative of future outcomes (Gaulin 2017; Hail, Muhn, and Oesch 2019; Beatty, Cheng, and Zhang 2018; Campbell et al. 2018). For example, Campbell et al. (2018) document that the disclosure of tax risk factors can predict firms' future cash flows. Gaulin (2017) shows a positive association between disclosure of energy risks and future energy prices; the disclosure of strategic alliance risks is also positively related to goodwill impairments. Similarly, this study argues that if ICFR-related risk factor disclosure truly reflects potential internal control deficiencies, it should be predictive of future restatements of annual reports and material weaknesses in internal controls. The argument leads to the following hypotheses:

H2a: Disclosure of ICFR-related risk factors is positively related to future restatements of 10-Ks.

H2b: Disclosure of ICFR-related risk factors is positively related to future material weaknesses in internal control.

Auditors have the responsibility to collect evidence from all possible sources to evaluate the firm's effectiveness of internal control over financial reporting (Auditing Standard 2201) and therefore should understand the firm's internal control status comprehensively. Even though a firm may not have material weaknesses, it is still possible that it has immaterial internal control deficiencies. This potential risk should be captured and considered by auditors during the audit planning process. Audit risk model explains audit risk as the product of three components: inherent risk, control risk, and detection risk. Based on this model, if the internal control risk is high, auditors need to reduce the detection risk by increasing audit effort to maintain audit risk at a tolerable level. If the internal control risk increases, auditors need to make more audit effort to reduce the detection risk and keep audit risk at an acceptable level. Consistent with the prediction, Hogan and Wilkins (2008) find that audit effort, proxied by audit fees, is significantly higher for firms with internal control deficiencies. Therefore, if the disclosure of ICFR-related risk factors effectively reflects internal control weakness over financial reporting, it should be associated with higher audit fees. The third hypothesis is then presented as the following:

H3: Disclosure of ICFR-related risk factors is positively related to the current year's audit fees.

4.3. Research Method

We extract Item 1A sections of annual reports from the year 2005 to 2016 and split them into individual risk factors using the method described in Campbell et al.

(2014) and Gaulin (2017). ICFR-related risk factors are identified by searching the heading of each risk factors with the following keywords and phrases[1]: "internal control(s)", "material weakness(es)", "restatement(s)", "reporting obligation(s)", and "reporting deficiency(ies)". The data set is then merged with Compustat and Audit Analytics data sets to form the final sample in the analysis. Table 18 presents the sample selection process. The final sample includes 30,429 firm-year observations, among which 1615 observations disclose material weaknesses and 28,726 observations do not disclose material weaknesses.

Table 18. Sample selection (Chapter 4)

	Firm-Year Observations
Compustat & CRSP Merged database in fiscal year 2005-2016, excluding observations with blank cik and missing fiscal year	66787
Less: merge with risk factor data	(0)
Less: merge with data from Audit Analytics	(30814)
Less: missing variables for restatement model	(5544)
Final sample for restatement model	30429
Final model for audit fee model	30341
Final sample for material weaknesses model	28844

In order to test the first hypothesis, we create three variables to identify different situations: 1) ONLYICRF, which is a dummy variable equal to 1 if the firm only discloses ICFR-related risk factors, 2) ONLYMW, which is a dummy variable equal to 1 if the firm only discloses material weaknesses in internal control, and 3) ICFR_MW, which is a dummy variable equal to 1 if the firm discloses both ICFR-related risk factors

[1] There is no prior study providing the keyword list for identifying internal control risks. Doyle, Ge, and McVay (2007) use the keyword "material weakness" to identify material weakness disclosures, but internal control deficiencies can have more keywords. Therefore, we manually read 100 randomly extracted risk factor disclosures with the keyword "internal control" or "material weakness(es)" to extend the list of keywords and phrases.

and material weaknesses. This setting also avoids the potential multicollinearity issue caused by the possible high correlation between disclosure of ICFR-related risk factors and materiality disclosure. The following regression model based on Dechow et al. (2011) is used to test H1:

$$RESTATEMENT_t = \alpha + \beta_1 ONLYICRF_t + \beta_2 ONLYMW_t + \beta_3 ICRF\_MW_t + \beta_4 TOTALACCURAL_t + \beta_5 \Delta REC_t + \beta_6 \Delta INV_t + \beta_7 SOFTASSET_t + \beta_8 \Delta CSALE_t + \beta_9 \Delta ROA_t + \beta_{10} ISSUANCE_t + \beta_{11} \Delta EMP_t + \beta_{12} LEASE_t + \beta_{13} ABRET_t + \beta_{14} LAGABRET_t + IND + YEAR + \varepsilon \quad (5)$$

The one-period lagged values of $ONLYICFR_t$, $ONLYMW_t$, and $ICFR\_MW_t$ are used in the material misstatement prediction model to test H2a, as shown in regression (6). Regression (7) is based on Judd et al. (2017) to examine the association between inclusion of ICFR-related risk factors and future likelihood of the firm having material weaknesses (H2b).

$$RESTATEMENT_t = \alpha + \beta_1 ONLYICRF_{t-1} + \beta_2 ONLYMW_{t-1} + \beta_3 ICRF\_MW_{t-1} + \beta_4 TOTALACCURAL_t + \beta_5 \Delta REC_t + \beta_6 \Delta INV_t + \beta_7 SOFTASSET_t + \beta_8 \Delta CSALE_t + \beta_9 \Delta ROA_t + \beta_{10} ISSUANCE_t + \beta_{11} \Delta EMP_t + \beta_{12} LEASE_t + \beta_{13} ABRET_t + \beta_{14} LAGABRET_t + IND + YEAR + \varepsilon \quad (6)$$

$$MW_t = \alpha + \beta_1 ONLYICRF_{t-1} + \beta_2 ONLYMW_{t-1} + \beta_3 ICRF\_MW_{t-1} + \beta_4 LNASSET_t + \beta_5 AUDITLAG_t + \beta_6 LOSS_t + \beta_7 ZSCORE_t + \beta_8 MERGER_t + \beta_9 GROWTH_t + \beta_{10} MTB_t + \beta_{11} RESTRUCTURE + + \beta_{12} AGE_t + \beta_{13} OVERCONFIDENT_t + IND + YEAR + \varepsilon \quad (7)$$

H3 is testing whether the disclosure of ICFR-related risk factors is reflective of potential internal control deficiencies that could increase control risk and thus lead to higher audit fees. It predicts a positive contemporaneous association between audit fee and the inclusion of ICFR-related risk factors. The following audit fee model based on Hay, Knechel, and Wong (2006) is used in this analysis:

$$LNFEE_t = \alpha + \beta_1 ONLYICRF_t + \beta_2 ONLYMW_t + \beta_3 ICRF\_MW_t + \beta_4 PSCORE_t + \beta_5 ROAEARNING_t + \beta_6 LNASSET_t + \beta_7 INVREC_t + \beta_8 BUSSEG_t + \beta_9 FOREIGN_t + \beta_{10} MERGE_t + \beta_{11} SPECIAL_t + \beta_{12} LEVERAGE_t + \beta_{13} CURRENTRATIO_t + \beta_{14} LOSS_t + \beta_{15} BTM_t + \beta_{16} GROWTH_t + \beta_{17} BIG4_t + \beta_{18} RESIGNATION_t + \beta_{19} DISMISSAL_t + \beta_{20} GC_t + \beta_{21} TENURE_t + \beta_{22} INDEXPERT_t + IND + YEAR + \varepsilon \quad (8)$$

Change specification as suggested by Stanley (2011) is also used to test the robustness of the association between audit fees and ICFR-related risk factor disclosures. In the change specification, dependent and independent variables are the first-order difference from Model (8). We include unexpected audit fees as the residual from Model (8) to control for mispricing of the previous year.

4.4. Results and Discussion

Panel A of Table 19 shows the summary statistics of the test sample. The mean value of ONLYICRF is 0.18, indicating that 18% of the observations in the sample disclose ICFR-related risk factor but not material weaknesses. In comparison, only 2% of the observations disclose only material weaknesses and 3% of the observations disclose material weaknesses and ICFR-related risk factor at the same time. The relationship between disclosure of ICFR-related risk factors and material weaknesses is also presented in Panel B of Table 19. Only a small portion of firms disclose material weaknesses (1631 out of 30429 cases) while considerably more firms (6676 cases) disclose ICFR-related risk factors, and in many cases (5626 out of 6676 cases), firms only disclose ICFR-related risk factors.

Table 19. Descriptive statistics (Chapter 4)

Panel A: summary statistics for variables in the main analysis

Variables	Mean	Median	Std	P1	P25	P75	P99
Variables in restatement regression							
ONLYICRF	0.18	0.00	0.39	0.00	0.00	0.00	1.00
ONLYMW	0.02	0.00	0.14	0.00	0.00	0.00	1.00

ICRF_MW	0.03	0.00	0.18	0.00	0.00	0.00	1.00
RESTATEMENT	0.10	0.00	0.30	0.00	0.00	0.00	1.00
TOTALACCURAL	0.02	0.01	0.14	-0.47	-0.03	0.06	0.47
ΔREC	0.00	0.00	0.04	-0.13	-0.01	0.01	0.12
ΔINV	0.00	0.00	0.02	-0.09	0.00	0.00	0.09
SOFTASSET	0.59	0.62	0.27	0.04	0.39	0.82	0.98
ΔCSALE	0.12	0.06	0.39	-0.64	-0.03	0.18	2.33
ΔROA	0.00	0.00	0.14	-0.50	-0.03	0.02	0.61
ISSUANCE	0.94	1.00	0.23	0.00	1.00	1.00	1.00
ΔEMP	-0.04	-0.02	0.27	-1.20	-0.11	0.06	0.72
LEASE	0.80	1.00	0.40	0.00	1.00	1.00	1.00
ABRET	0.02	-0.03	0.44	-0.83	-0.24	0.20	1.83
LAGABRET	0.04	-0.02	0.45	-0.78	-0.22	0.20	1.97
Variables in material weaknesses regression							
MW	0.05	0.00	0.22	0.00	0.00	0.00	1.00
LNASSET	7.02	6.97	1.87	2.97	5.73	8.21	11.80
AUDITLAG	100.75	96.00	33.80	57.00	83.00	113.00	309.00
LOSS	0.27	0.00	0.44	0.00	0.00	1.00	1.00
AZSCORE	1.17	2.00	0.95	0.00	0.00	2.00	2.00
MERGE	0.23	0.00	0.42	0.00	0.00	0.00	1.00
GROWTH	0.12	0.06	0.39	-0.64	-0.03	0.18	2.36
MKBK	3.02	1.90	60.43	-12.72	1.18	3.33	28.37
RESTRUCTURE	0.00	0.00	0.01	-0.04	0.00	0.00	0.00
AGE	19.66	15.00	17.16	1.00	8.00	26.00	84.00
OVERCONFIDENT	0.25	0.00	0.43	0.00	0.00	1.00	1.00
Variables in audit fee regression							
LNFEE	14.01	13.95	1.12	11.51	13.27	14.69	16.98
PSCORE	0.10	0.10	0.01	0.07	0.09	0.11	0.12
ROAEARNING	0.04	0.06	0.17	-0.82	0.02	0.11	0.37
LNASSET	7.02	6.97	1.87	2.97	5.73	8.21	11.80
INVREC	0.28	0.22	0.22	0.00	0.10	0.40	0.85
BUSSEG	1.14	1.00	0.45	0.69	0.69	1.39	2.20
FOREIGN	0.68	1.00	0.47	0.00	0.00	1.00	1.00
MERGE	0.23	0.00	0.42	0.00	0.00	0.00	1.00
SPECIAL	0.15	0.00	0.36	0.00	0.00	0.00	1.00
LEVERAGE	0.57	0.56	0.27	0.07	0.36	0.77	1.37
CURRENTRATIO	2.43	1.70	2.28	0.43	1.00	2.83	13.31
LOSS	0.27	0.00	0.44	0.00	0.00	1.00	1.00
BTM	0.60	0.49	0.54	-0.61	0.27	0.78	3.21
GROWTH	0.12	0.06	0.39	-0.64	-0.03	0.18	2.36
BIG4	0.78	1.00	0.41	0.00	1.00	1.00	1.00
RESIGNATION	0.01	0.00	0.09	0.00	0.00	0.00	0.00

DISMISSAL	0.04	0.00	0.19	0.00	0.00	0.00	1.00
GC	0.02	0.00	0.12	0.00	0.00	0.00	1.00
TENURE	1.84	1.95	0.70	0.00	1.39	2.40	2.83
INDEXPERT	0.25	0.00	0.43	0.00	0.00	1.00	1.00

Panel B: Relationship between disclosure of ICFR-related risk factors and material weaknesses[2]

		MW		
		0	1	Total
IC_RF	0	23172	581	23753
	1	5626	1050	6676
	Total	28798	1631	30429

Notes: This table presents the descriptive statistics for the sample in our analysis. The sample contains 30,341 firm-year observations for fiscal year 2005-2016. Panel A shows the summary statistics (including mean, median, standard deviation, 1 percentile, 25 percentile, 75 percentile, and 99 percentiles) of variables used in the main regression analysis. Variable definitions are provided in Appendix C. Panel B shows the 2 by 2 panel of the occurrence of material weaknesses disclosure and ICFR-related risk factor disclosure.

Figure 2 compares the distribution of firms that disclose ICFR-related risk factors and firms that disclose material weaknesses over time. The percentage of firms disclosing ICFR-related risk factors is much higher than the percentage of firms disclosing material weaknesses. Besides, even though the percentage of firms disclosing material weaknesses stays stable, the percentage of firms disclosing ICFR-related risk factors is increasing over time, especially after 2009.

Figure 2. Disclosure of ICFR-related risk factors and disclosure of material weaknesses

[2] The correlation between IC_RISK and MW in the sample is 0.2305, and significant at 0.01 level.

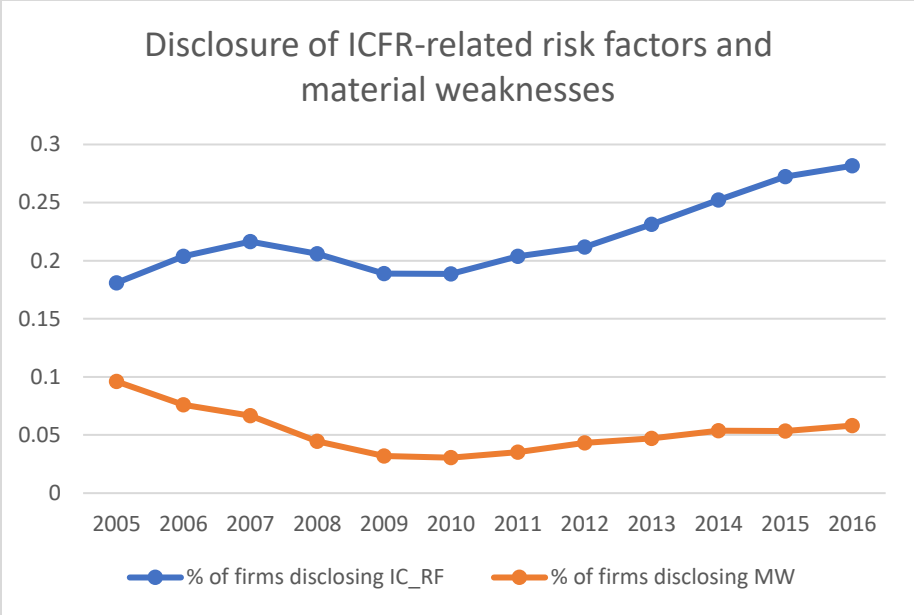


Table 20 presents the result for H1. Consistent with the predictions in H1, the coefficients of all three indicators (ONLYICRFt, ONLYMWt, ICRF_MWt) are positive and significant. Particularly, the coefficient of ONLYICRFt is much smaller than the coefficient of ONLYMWt. The difference is 2.02 and significant at 0.01 level (chi-square=310.21, p-value<0.01). This is consistent with the intuition that the risk factor disclosure indicates less serious issues (material) than material weaknesses. Interestingly, the coefficient of ICFR_MWt is smaller than the coefficient of ONLYMWt. The difference is 0.89 and significant at 0.01 level (chi-square=52.82, p-value<0.01), suggesting that firms only disclose material weaknesses are more likely to restate the annual report than firms that disclose internal control issues in both sections.

Table 20. ICFR-related risk factor disclosure in predicting restatement of the corresponding annual report (Chapter 4)

	RESTATEMENT	
VARIABLES	coef	tstat
$ONLYICRF_t$	0.1675**	2.10

$ONLYMW_t$	2.1900***	22.69
$ICRF\_MW_t$	1.2956***	14.65
TOTALACCURAL	0.2061	1.31
ΔREC	-1.0572*	-1.79
ΔINV	-0.4382	-0.52
SOFTASSET	0.5287***	3.31
ΔCSALE	0.0424	0.73
ΔROA	-0.1284	-0.97
ISSUANCE	0.2049*	1.77
ΔEMP	-0.0265	-0.31
LEASE	-0.0714	-0.61
ABRET	-0.0422	-0.90
LAGABRET	-0.0740	-1.64
Observations	30429	
Fixed effects	Industry & Year	
Cluster	Firm	
Pseudo R-squared	0.0647	

Notes: This table presents the result of regression to test H1. The dependent variable RESTATEMENT is a dummy variable that equals 1 if the firm restates the 10-K filing for the current fiscal year, and 0 otherwise. The variables of interest are $ONLYICRF_t$, $ONLYMW_t$, and $ICRF\_MW_t$, representing the case of firm only disclosing ICFR-related risk factors, firms only disclosing material weaknesses, and firms disclosing both ICFR-related risk factors and material weaknesses for the current fiscal year, respectively. Definitions for the control variables are provided in Appendix C. Logistic regression is used for estimation. Industry and year fixed effects are controlled. Standard errors are clustered by firm. t-statistics are in parentheses and ***, **, * indicate two-tailed significance at 1%, 5%, and 10% level respectively.

Results for H2a and H2b are presented in Table 21. Coefficients for all three indicators ($ONLYICRF_{t-1}$, $ONLYMW_{t-1}$, $ICRF\_MW_{t-1}$) are positive and significant in both restatement prediction and material weaknesses prediction, suggesting that each case is informative of the likelihood of future restatement of 10-Ks and material weaknesses in internal control. Specifically, for subsequent period's restatement prediction, the coefficient of $ONLYICRF_{t-1}$ is significantly smaller than the coefficient of $ONLYMW_{t-1}$ (difference=0.539, chi-square= 11.23, p-value<0.00); and the coefficient of $ONLYMW_{t-1}$

is larger than the coefficient of $ICRF\_MW_{t-1}$, but the difference is insignificant. For future material weaknesses prediction, the coefficient of $ONLYICRF_{t-1}$ is also smaller compared to the coefficients of $ONLYMW_{t-1}$ and $ICRF\_MW_{t-1}$, which is consistent with the idea that risk factor disclosures are not as definitive and severe as material weaknesses disclosures.

Table 21. ICFR-related risk factor disclosure in predicting related future outcomes (Chapter 4)

	(1)		(2)	
	RESTATEMENT		MW	
VARIABLES	coef	tstat	coef	tstat
$ONLYICRF_{t-1}$	0.1454*	1.71	0.3267***	3.84
$ONLYMW_{t-1}$	0.6840***	4.69	2.1003***	17.63
$ICRF\_MW_{t-1}$	0.6369***	5.58	2.4411***	26.19
TOTALACCUR AL	0.1360	0.77		
ΔREC	-1.2092*	-1.78		
ΔINV	-0.6105	-0.64		
SOFTASSET	0.6089***	3.41		
ΔCSALE	0.0422	0.63		
ΔROA	-0.2593*	-1.78		
ISSUANCE	0.1564	1.29		
ΔEMP	-0.0958	-0.98		
LEASE	-0.0313	-0.23		
ABRET	-0.1217**	-2.32		
LAGABRET	-0.0803	-1.60		
LNASSET			-0.0840***	-3.63
AUDITLAG			0.0086***	14.22
LOSS			0.5859***	7.56
AZSCORE			0.1055***	2.62
MERGE			0.1553*	1.93
GROWTH			0.0347	0.47
MKBK			0.0001	0.41
RESTRUCTURE			-3.7259	-0.97
AGE			-0.0045*	-1.78
OVERCONFIDENT			0.1391**	2.07
Observations	26,004		28,844	
Pseudo R-squared	0.0371		0.159	
Fixed effects	Industry & Year		Industry & Year	

Cluster	Firm	Firm
Notes: This table presents the results of regressions to test H2a and H2b. The dependent variable in regression (1) is RESTATEMENT, which is a dummy variable that equals 1 if the firm restates the 10-K filing for the current fiscal year, and 0 otherwise. The dependent variable in regression (2) is MW, which is a dummy variable that equals 1 if the firm discloses material weaknesses in internal control for the current fiscal year, and 0 otherwise. The variables of interest are $ONLYICRF_{t-1}$, $ONLYMW_{t-1}$, and $ICRF\_MW_{t-1}$, representing the case of firm only disclosing ICFR-related risk factors, firms only disclosing material weaknesses, and firms disclosing both ICFR-related risk factors and material weaknesses for the previous year, respectively. Definitions for the control variables are provided in Appendix C. Logistic regressions are used for estimation. Industry and year fixed effects are controlled. Standard errors are clustered by firm. t-statistics are in parentheses and ***, **, * indicate two-tailed significance at 1%, 5%, and 10% level respectively.		

Table 22 presents the result for H3. The coefficient of $ONLYICRF_t$ is 0.0852 and significant at 0.01 level (t-value=5.33 and p-value<0.01), the coefficient of ONLYMWt is 0.1665 and significant at 0.01 level (t-value=6.29 and p-value<0.01), and the coefficient of ICRF_MWt is 0.4338 and significant at 0.01 level (t-value=17.76 and p-value<0.01). Compared to firms that disclose neither material weakness nor ICFR-related risk factors, audit fees are about 8.52% higher for firms that only discloses ICFR-related risk factors, 16.65% higher for firms only discloses material weaknesses in internal control, and approximately 43.38% higher for firms that discloses both ICFR-related risk factors and material weaknesses. The result is consistent with the prediction in H3 that disclosure of ICFR-related risk factors indicates significant internal control risk over financial reporting, which may lead to greater audit efforts and thus is associated with higher audit fees. Table 23 shows the result of the change model. The result suggests that the argument is robust to the change specification, showing that the inclusion of ICFR-related risk factors is positively related to audit fee increases.

Table 22. Association between disclosure of ICFR-related risk factors and audit fees (Chapter 4)

	LNFEE	
VARIABLES	coef	tstat
$ONLYICRF_t$	0.0852***	5.33
$ONLYMW_t$	0.1665***	6.29
$ICRF\_MW_t$	0.4338***	17.76
PSCORE	4.1731***	5.90
ROAEARNING	-0.2667***	-5.88
LNASSET	0.4757***	70.23
INVREC	0.1625**	2.42
BUSSEG	0.1198***	6.74
FOREIGN	0.2171***	12.93
MERGE	0.0498***	4.53
SPECIAL	0.0082	0.83
LEVERAGE	-0.0167	-0.45
CURRENTRATIO	-0.0183***	-5.19
LOSS	0.1585***	12.32
BTM	-0.0857***	-6.55
GROWTH	-0.0524***	-5.72
BIG4	0.3884***	20.21
RESIGNATION	-0.0923**	-2.31
DISMISSAL	-0.1138***	-6.41
GC	0.1870***	5.70
TENURE	-0.0315***	-3.14
INDEXPERT	0.0842***	5.26
Observations	30,341	
Adjusted R-squared	0.8145	
Fixed effects	Industry & Year	
Cluster	Firm	

Notes: This table presents the result of regression to test H3. The dependent variable is LNFEE, which is the natural logarithm of audit fee for the current fiscal year. The variables of interest are $ONLYICRF_t$, $ONLYMW_t$, and $ICRF\_MW_t$, representing the case of firm only disclosing ICFR-related risk factors, firms only disclosing material weaknesses, and firms disclosing both ICFR-related risk factors and material weaknesses for the current fiscal year, respectively. Definitions for the control variables are provided in Appendix C. Ordinary Least Square (OLS) regression is used for estimation. Industry and year fixed effects are controlled. Standard errors are clustered by firm. t-statistics are in parentheses and ***, **, * indicate two-tailed significance at 1%, 5%, and 10% level respectively.

Table 23: Association between change of disclosure of ICFR-related risk factors and change of audit fees (Chapter 4)

	ΔLNFEE	
VARIABLES	coef	tstat
ΔONLYICRF	0.0542***	7.30
ΔONLYMW	0.0174*	1.88
ΔICRF_MW	0.1211***	11.15
ΔPSCORE	2.4353***	12.65
ΔROAEARNING	-0.1415***	-6.71
ΔLNASSET	0.2886***	33.51
ΔINVREC	0.1417***	3.95
ΔBUSSEG	0.0279***	3.15
ΔFOREIGN	0.0267***	5.39
ΔMERGE	0.0294***	9.50
ΔSPECIAL	-0.0037	-1.38
ΔLEVERAGE	0.0991***	5.26
ΔCURRENTRATIO	-0.0049***	-3.30
ΔLOSS	0.0385***	10.22
ΔBTM	-0.0084**	-2.17
ΔGROWTH	0.0063	1.45
ΔBIG4	0.2380***	14.77
ΔRESIGNATION	-0.0338*	-1.72
ΔDISMISSAL	-0.0699***	-9.21
ΔGC	0.0534***	3.59
ΔTENURE	0.0401***	7.66
ΔINDEXPERT	0.0324***	6.50
$UNEXPECTEDFEE_{t-1}$	-0.1017***	-34.78
Observations	24,878	
Adjusted R-squared	0.2700	
Fixed effects	Industry & Year	
Cluster	Firm	

Notes: This table presents the result of changed setting for audit fee model to test H3. The dependent variable is ΔLNFEE, which is the change of LNFEE from the previous fiscal year to the current fiscal year. The variables of interest are ΔONLYICRF, ΔONLYMW, and ΔICRF_MW, the change of ONLYICRF, ONLYMW, and ICRF_MW from the previous fiscal year to the current fiscal year, respectively. Control variables are the first order difference of the control variables in the level model of audit fee model. Previous year unexpected audit fee $UNEXPECTEDFEE_{t-1}$ is also controlled. Ordinary Least Square (OLS) regression is used for estimation. Industry and year fixed effects are controlled. Standard errors are clustered by firm. t-statistics are in parentheses and ***, **, * indicate two-tailed significance at 1%, 5%, and 10% level respectively.

4.5. Content Analyses

This section focuses on the contents of ICFR-related risk factor disclosure. The evidence so far suggests that the inclusion of ICFR-related risk factors in 10-K generally reflects internal control risk that potentially relates to a high probability of restatement, high chance of having material weaknesses in the future, and high audit fees. However, the ICFR-related risk factor disclosures may vary in informativeness and importance. In this study, four characteristics are chosen to examine the contents of ICFR-related risk factors.

The first measure takes into account the relative position of ICFR-related risk factor disclosure in Item 1A. The SEC describes on its official website that firms "generally list the risk factors in order of their importance (SEC 2011a)". (Chin, Liu, and Moffitt 2018) find that the relative position or rank of credit risk in risk factor disclosure is informative of the credit risk level. This study extends their argument and conjecture that the ICFR-related risk factor is more significant if it is positioned towards the beginning of Item 1A. Similar to Chin et al. (2018), we measure the rank of ICFR-related risk factor (IC_RANK) as (1-original rank of ICFR-related risk factor divided by the total number of risks in Item 1A)[3]. We next analyze the details of risk factor disclosures. Some firms state that they cannot assure that there will not be internal control issues due to the inherent limitations of internal control in general. These firms may simply mention ICFR-related risk factors in their reports to reduce potential litigation risk. Some firms, on the other hand, explain why they face internal control risk. Possible reasons may

[3] When firms disclose more than one ICFR-related risk factors, only the risk factor with the highest rank (toward the top of the disclosure) is used to calculate the four measures because it's likely to be the most significant one.

include recent restatement or discovery of material weaknesses. These disclosures are more meaningful and informative of the internal control deficiencies. Examples of more and less informative ICFR-related risk factor disclosures are provided in Appendix D. Building on this view, we develop the second characteristic measure as the length (LENGTH) of the ICFR-related risk factor disclosure, calculated as the number of words excluding stop words[4] (Gaulin 2017). It aims to capture the verbosity of the disclosure. The third measure captures the level of disclosure detail. Hope, Hu, and Lu (2016) develop a specificity measure and document that more specific risk factor disclosures are more informative about the underlying risks. Following their method, the specificity (SPECIFICITY) of ICFR-related risk factors is defined as the number of specific words[5] divided by the length of the risk factor. Lastly, we measure the level of boilerplate (BOILERPERCENT) as the percentage of words from sentences that contain at least one of the 4-word-phrases that appear in most of the disclosures in a given disclosure (Lang and Stice-Lawrence 2015; Dyer, Lang, and Stice-Lawrence 2017).

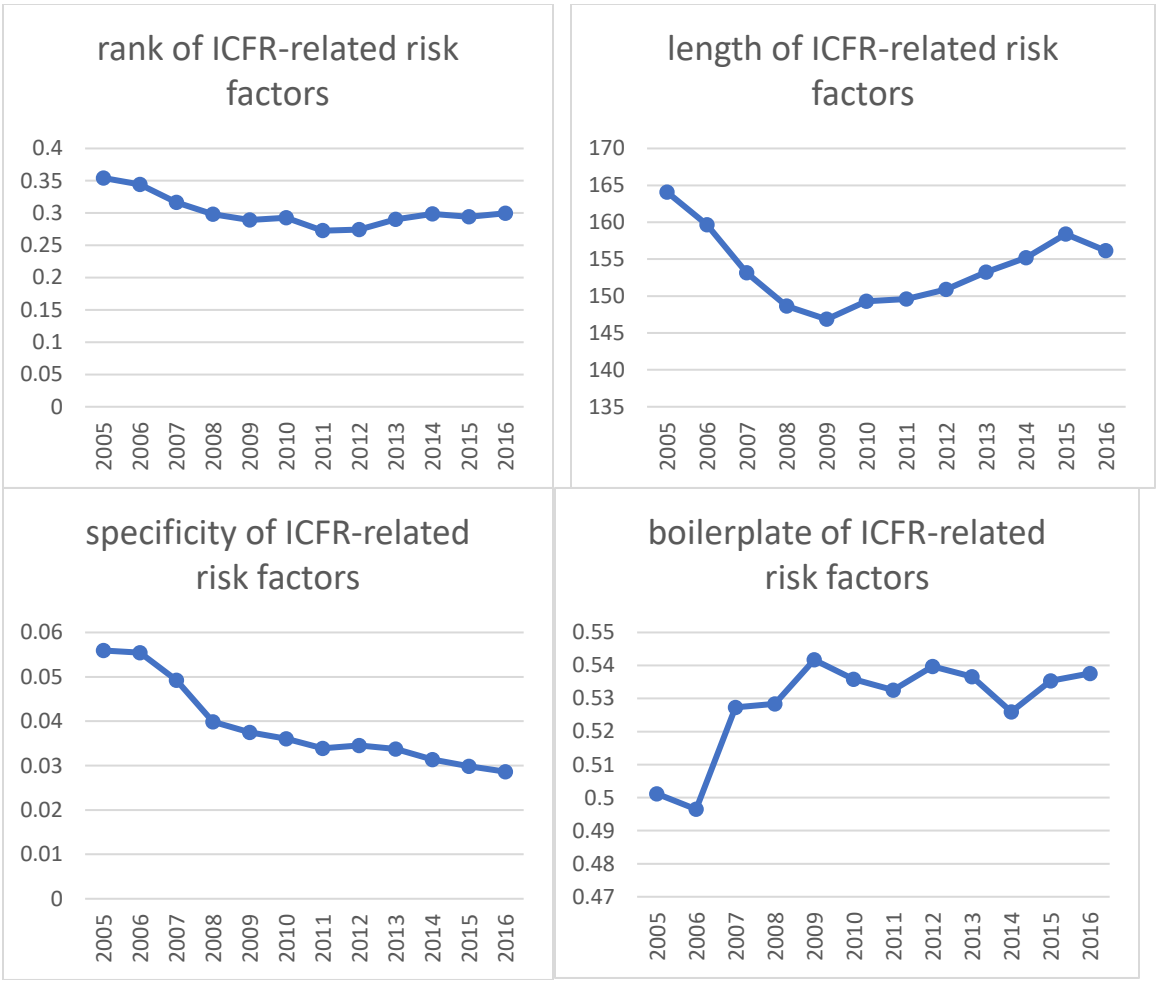
Figure 3 presents the changes in average values of the four measures over time. There is a slight decrease in the average of IC_RANK and a noticeable decrease in the mean value of SPECIFICITY, while the average of BOILERPERCENT seems to increase significantly from 2005 to 2007 and then fluctuate a little after 2007. These

[4] Stop words are the most commonly used words in English. The "stopwords" package provided by the Natural Language Toolkit (NLTK) in Python program is used to get the list of stop words. The list contains 153 words, including pronouns such as "me", "you", prepositions such as "in", "at", conjunctions such as "and", "but", and many other commonly used words.

[5] Specific words are identified the Stanford NER (Named Entity Recognizer) program and include seven categories 1) a person's name, 2) name of a location, 3) name of an organization, 4) percentages, 5) money values, 6) date, and 7) time.

trends may suggest that although more firms are disclosing ICFR-related risk factors over the past years, the contents of the disclosures are becoming less meaningful.

Figure 3. Characteristics of ICFR-related risk factors over time



To statistically test the effects of ICFR-related risk factor disclosure contents, we re-run the regressions in the main tests with the four content measures. Table 24 reports the regression results. LENGTH is positively and significantly associated with concurrent as well as future restatements of 10-K, meaning that firms that provide longer ICFR-related risk factor disclosures are likely to restate current and subsequent annual reports. Firms that place ICFR-related risk factors towards the beginning of Item 1A and provide more detailed discussions on ICFR-related risk factors are more likely to have material weaknesses in the future. Additionally, IC_RANK, SPECIFICITY, and LENGTH are positively associated with audit fees. Overall, the results support the view that the position and details of ICFR-related risk factors are indicative of the severity of internal control risks over financial reporting.

Table 24. Additional analysis on the contents of ICFR-related risk factors (Chapter 4)

	(1)		(2)		(3)		(4)	
	RESTATEMENT		RESTATEMENT		MW		LNFEE	
VARIABLES	coef	tstat	coef	tstat	coef	tstat	coef	tstat
IC_RANK_t	0.0354	0.18					0.1895***	4.64
$SPECIFICITY_t$	1.595	1.38					1.1169***	4.35
$LENGTH_t$	0.0018***	2.78					0.0006***	3.78
$BOILERPERCENT_t$	-0.1403	-0.73					0.037	0.96
IC_RANK_{t-1}			0.0061	0.03	0.9970***	5.11		
$SPECIFICITY_{t-1}$			-0.5266	-0.4	2.6483**	2.13		
$LENGTH_{t-1}$			0.0019***	2.72	0.0019***	2.87		
$BOILERPERCENT_{t-1}$			-0.0685	-0.32	-0.0692	-0.36		
Control variables	Controlled		Controlled		Controlled		Controlled	
Observations	6,542		5,534		5,914		6,615	
Fixed effects	Industry & Year		Industry & Year		Industry & Year		Industry & Year	
Cluster	Firm		Firm		Firm		Firm	
Adjuested R-squared (Pseudo R-squared)	0.0672		0.0789		0.122		0.751	

Notes: This table presents the result for the additional analysis on the contents of ICFR-related risk factors. The dependent variable in regression (1) and (2) are both RESTATEMENT, a dummy variable

that equals 1 if the firm restates the 10-K filing for the current fiscal year, and 0 otherwise. The dependent variable in regression (3) is MW, which is a dummy variable that equals 1 if the firm discloses material weaknesses in internal control for the current fiscal year, and 0 otherwise. The dependent variable in regression (4) is LNFEE, which is the natural logarithm of audit fee for the current fiscal year. The variables of interest in regression (1) and (4) are IC_RANK_t, $SPECIFICITY_t$, $LENGTH_t$, and $BOILERPERCENT_t$. And the variables of interest in regression (2) and (3) are IC_RANK_{t-1}, $SPECIFICITY_{t-1}$, $LENGTH_{t-1}$, and $BOILERPERCENT_{t-1}$. Variable definitions are given in Appendix C. Logistic regressions are used for estimation of regression (1), (2), and (3). Ordinary Least Square (OLS) regression is used for estimation of regression (4). Industry and year fixed effects are controlled. Standard errors are clustered by firm. t-statistics are in parentheses and ***, **, * indicate two-tailed significance at 1%, 5%, and 10% level respectively.

4.6. Conclusion

This study investigates the disclosure of ICFR-related risk factors and provides evidence that the risk factor disclosure contains incremental information about the potential internal control deficiencies of the firm. Specifically, ICFR-related risk factor disclosure is predictive of the restatement of current annual reports and future adverse consequences including restatement and material weaknesses in internal control. It is also positively related to audit fees, which capture audit efforts. Collectively, these findings suggest that ICFR-related risk factors are indicative of potential flaws in internal control over financial reporting. Furthermore, the contents of ICFR-related risk factor disclosure may be used to infer the severity of firms' internal control risk. Specifically, higher ranked, more specific and more extended ICFR-related risk factor disclosures indicate higher internal control risk over financial reporting.

To summarize, this study shows that ICFR-related risk factors do complement the SOX 404 material weaknesses disclosure by providing additional information about the firm's potential internal control issues. This information is valuable because most of the time firms with internal control deficiencies do not have to disclose as long as the issues are not material. In that case, the study argues that stakeholders can learn about potential

internal control problems in the risk factor section. The result suggests that the risk factor disclosure generally provides useful information on the internal control issues.

There are several opportunities for future research. This study shows that ICFR-related risk factor disclosure is informative of the internal control risk from an audit risk perspective. Prior studies (e.g., Chan, Farrell, and Lee 2008; Ashbaugh-Skaife et al. 2009) show that internal control deficiencies are also related to earning management and cost of equity. Future studies may investigate whether ICFR-related risk factor disclosure benefits information users as well as the firm itself. For example, it may be interesting to know if analyst behaviors and market reactions change when firms disclose ICFR-related risk factors.

Chapter 5. Conclusion and Future Research

This dissertation contributes to the auditing literature by showing three examples of how textual analysis can be used to gain information about the likelihood of restatement of annual reports from the SEC comment letters, to identify a new proxy of audit risks from the management qualitative earnings press releases, and to improve the understanding of internal control risks from the risk factor disclosures.

The first essay examines the intensity of SEC comment letter and how it relates to the probability of reviewed 10-K filings. The SEC comment letter is the correspondence between SEC staff and SEC filers about the filers' public information disclosures. The intensity of comment letters in terms of the use of strong/weak modal language can reflect perceived deficiencies in the reviewed filings. This essay uses text mining to examine the intensity of SEC comment letters and develops an intensity measure based on the modality of comment letters. Empirical analysis on a sample of initial comment letters related to 10-K filings shows that the intensity is positively associated with the probability of restatement of the reviewed 10-K filings and that this association is robust using both original Loughran and McDonald (2011) word lists and modified word lists. This essay contributes to the literature by introducing a direct measure of the intensity of SEC comment letters based on the strong/weak modality of the letter.

There are some limitations of this essay. First, comment letters can review several different filings simultaneously but we calculate the intensity measure based on the whole comment letter rather than for the specific parts referring to the 10-K filings. Second, we modify the word lists by simply eliminating irrelevant words and it is possible that some words not in the word lists may suggest higher or lower intensity. In addition, we can

develop an intensity measure based on the sentence structure rather than simply based on the use of words. Future research can further improve the intensity measure in these ways. Furthermore, if we could get some SEC staff to read a random sample of comment letters and evaluate their intensity, then we can use the labeled data to train a machine learning algorithm to create a new list of words and improve the intensity measure.

The second essay proposes that the abnormal tone of management voluntary earnings press releases can be used as a proxy for audit risks to estimate audit fee decisions. Using publicly issued management qualitative earnings press releases, this essay examines the association between abnormal tone and audit fees. We find that abnormally negative earnings press release tone is negatively associated with audit fees, whereas abnormally positive tone is not associated with audit fees except in extreme cases. This demonstrates the association between audit fees and the tone of press releases as a measurement of client business risk if managers are overly pessimistic. On the other hand, we find the association between audit fees and the tone of press releases as a proxy of opportunistic disclosure behavior only if the abnormal tone is extremely positive. In addition, we find that the association between abnormal negative tone and audit fees is moderated by the credibility of the earnings press releases issued in the previous year. These findings generally support the idea that qualitative management disclosures serve as a proxy for a client business and financial reporting risks.

A question relates to the earnings press releases and many other corporate disclosures is that whether these disclosures are composed by the management or simply by a group of lawyers. Many such disclosures contain different levels of platitudes.

Future research can look into this issue and examine how the level of platitudes in the disclosures impact the usefulness of the qualitative disclosures.

The third essay investigates the informativeness of firms' disclosures of risk factors related to internal control over financial reporting (ICFR-related risk factors). This essay finds that the ICFR-related risk factor disclosure incrementally predicts future adverse consequences including annual report restatements and material internal control weaknesses. The results also indicate that audit fees for firms with ICFR-related risk factors are likely to be higher, consistent with the intuition that these firms are having higher internal control risks. In addition, the contents of ICFR-related risk factors can help financial statement users assess the severity of internal control issues. Taken together, these findings suggest that ICFR-related risk factors reflect potential internal control deficiencies.

Since there is no prior study developing a specific dictionary for identifying topics related to internal control, we use a simple way of manually reading 100 randomly extracted risk factors with the keyword "internal control" or "material weakness(es)" to create the keyword list for this study. Future research may use a more systematic way to create the dictionary to identify the ICFR-related risks.

This dissertation uses textual analysis on three types of disclosures to extract useful information. In this big data era, there are much more texts with the potential value to benefit our decision making that could be available, such as social media, product reviews, contracts, emails and messages, and so on. Future research can explore the possibility of using textual analysis to examine different sources of texts and see how it can provide new information that contributes to the auditing field.

Bibliography

Abraham, Santhosh, and Philip J. Shrives. 2014. Improving the relevance of risk factor disclosure in corporate annual reports. The British Accounting Review 46, (1): 91-107.

AICPA. Consideration of fraud in a financial statement audit. 2002. Available at https://www.aicpa.org/Research/Standards/AuditAttest/DownloadableDocuments/AU-00316.pdf.

———. Audit Evidence. In AU Section 326. 2006. Available at: https://www.aicpa.org/content/dam/aicpa/research/standards/auditattest/downloadabledocuments/au-00326-9.pdf

———. The importance of internal control in financial reporting and safeguarding plan assets. 2014. Available at https://www.aicpa.org/content/dam/aicpa/interestareas/employeebenefitplanauditquality/resources/planadvisories/downloadabledocuments/plan-advisoryinternalcontrol-hires.pdf.

Altman, Edward I. 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. The Journal of Finance 23, (4): 589-609.

Arslan-Ayaydin, Özgür, Kris Boudt, and James Thewissen. 2016. Managers set the tone: Equity incentives and the tone of earnings press releases. Journal of Banking & Finance 72: S132-47.

Ashbaugh-Skaife, Hollis, Daniel W. Collins, and Ryan Lafond. 2009. The effect of SOX internal control deficiencies on firm risk and cost of equity. Journal of Accounting Research 47, (1): 1-43.

Bamber, Matthew, and Kevin McMeeking. 2016. An examination of international accounting standard-setting due process and the implications for legitimacy. The British Accounting Review 48, (1): 59-73.

Bao, Yang, and Anindya Datta. 2014. Simultaneously discovering and quantifying risk types from textual risk disclosures. Management Science 60, (6): 1371-91.

Beatty, Anne, Lin Cheng, and Haiwen Zhang. 2018. Are Risk Factor Disclosures Still relevant? Evidence from Market reactions to Risk Factor Disclosures Before and after the Financial Crisis. Contemporary Accounting Research.

Bonini, Stefano and Pavesi, Filippo and Scotti, Massimo, Financial Analysts and Collective Reputation: Theory and Evidence (January 28, 2011). Available at SSRN: https://ssrn.com/abstract=1965451 or http://dx.doi.org/10.2139/ssrn.1965451.

Boone, Jeff P., Cheryl L. Linthicum, and April Poe. 2013. Characteristics of accounting standards and SEC review comments. Accounting Horizons 27, (4): 711-36.

Bowen, Robert M., Angela K. Davis, and Dawn A. Matsumoto. 2005. Emphasis on pro forma versus GAAP earnings in quarterly press releases: Determinants, SEC intervention, and market reactions. The Accounting Review 80, (4): 1011-38.

Bozanic, Zahn and Dietrich, J. Richard and Johnson, Bret A., SEC Comment Letters and Firm Disclosure (June 12, 2017). Journal of Accounting and Public Policy, 2017, Vol. 36(5), 337-357; George Mason University School of Business Research Paper No. 18-6. Available at SSRN: https://ssrn.com/abstract=2989164.

Campbell, John L., Mark Cecchini, Anna M. Cianci, Anne C. Ehinger, and Edward M. Werner. 2018. Tax-related mandatory risk factor disclosures, future profitability, and stock returns. Review of Accounting Studies: 1-45.

Campbell, John L., Hsinchun Chen, Dan S. Dhaliwal, Hsin-min Lu, and Logan B. Steele. 2014. The information content of mandatory risk factor disclosures in corporate filings. Review of Accounting Studies 19, (1): 396-455.

Carcello, Joseph V., Dana R. Hermanson, Terry L. Neal, and Richard A. Riley Jr. 2002. Board characteristics and audit fees. Contemporary Accounting Research 19, (3): 365-84.

Carson, Elizabeth, and Neil Fargher. 2007. Note on audit fee premiums to client size and industry specialization. Accounting & Finance 47, (3): 423-46.

Cassell, Cory A., Lauren M. Dreher, and Linda A. Myers. 2013. Reviewing the SEC's review process: 10-K comment letters and the cost of remediation. The Accounting Review 88, (6): 1875-908.

Colbert, Janet L., Michael S. Luehlfing, and C. Wayne Alderman. "Engagement risk." The CPA Journal 66, no. 3 (1996): 54.

Chan, Kam C., Barbara Farrell, and Picheng Lee. 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): 161-79.

Chen, Reining, Rick Johnston, and Sundaresh Ramnath. 2010. The Effect of Regulator Oversight on Firms' Information Environment: Securities and Exchange Commission Comment Letters. Working paper. Massachusetts Institute of Technology and Purdue University, 2010.

Cheng, Mei, Dan Dhaliwal, and Yuan Zhang. 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.

Chin, Michael V. and Liu, Yue and Moffitt, Kevin, Voluntary Disclosure Through the Ranking of Risk Factors in the 10-K (March 18, 2018). Available at SSRN: https://ssrn.com/abstract=3142990or http://dx.doi.org/10.2139/ssrn.3142990.

Choi, Jong-Hag, Chansog Kim, Jeong-Bon Kim, and Yoonseok Zang. 2010. Audit office size, audit quality, and audit pricing. Auditing: A Journal of Practice & Theory 29, (1): 73-97.

Coleman, Derryck, and Plante, Kristyn. Risk of internal controls failures. 2017. Available at: http://www.auditanalytics.com/blog/the-risks-of-internal-controls/.

Davis, Angela K., Jeremy M. Piger, and Lisa M. Sedor. 2012. Beyond the numbers: Measuring the information content of earnings press release language. Contemporary Accounting Research 29, (3): 845-68.

Davis, Angela K., and Isho Tama-Sweet. 2012. Managers' use of language across alternative disclosure outlets: Earnings press releases versus MD&A. Contemporary Accounting Research 29, (3): 804-37.

Dechow, Patricia M., Weili Ge, Chad R. Larson, and Richard G. Sloan. 2011. Predicting material accounting misstatements. Contemporary Accounting Research 28, (1): 17-82.

Demers, Elizabeth, and Clara Vega. Soft information in earnings announcements: News or noise?. Vol. 951. Board of Governors of the Federal Reserve System, 2008.

Druz, Marina and Petzev, Ivan and Wagner, Alexander F. and Zeckhauser, Richard J., When Managers Change Their Tone, Analysts and Investors Change Their Tune (December 15, 2017). Swiss Finance Institute Research Paper No. 15-02. Available at SSRN: https://ssrn.com/abstract=2559157 or http://dx.doi.org/10.2139/ssrn.2559157.

Doyle, Jeffrey, Weili Ge, and Sarah McVay. 2007. Determinants of weaknesses in internal control over financial reporting. Journal of Accounting and Economics 44, (1-2): 193-223.

Dyer, Travis, Mark Lang, and Lorien Stice-Lawrence. 2017. The evolution of 10-K textual disclosure: Evidence from latent dirichlet allocation. Journal of Accounting and Economics 64, (2-3): 221-45.

Elshandidy, Tamer, Philip J. Shrives, Matt Bamber, and Santhosh Abraham. 2018. Risk reporting: A review of the literature and implications for future research. Journal of Accounting Literature 40, : 54-82.

Feng, Mei, and Chan Li. 2014. Are auditors professionally skeptical? evidence from auditors' going-concern opinions and management earnings forecasts. Journal of Accounting Research 52, (5): 1061-85.

Francis, Jennifer, Donna Philbrick, and Katherine Schipper. 1994. Shareholder litigation and corporate disclosures. Journal of Accounting Research: 137-64.

Francis, Jere R., and Jagan Krishnan. 1999. Accounting accruals and auditor reporting conservatism. Contemporary Accounting Research 16, (1): 135-65.

Francis, Jere R., Kenneth Reichelt, and Dechun Wang. 2005. The pricing of national and city-specific reputations for industry expertise in the US audit market. The Accounting Review 80, (1): 113-36.

Gaulin, Maclean Peter. 2017. Risk fact or fiction: The information content of risk factor disclosures. Dissertation. Rice University.

Ghosh, Aloke, and Robert Pawlewicz. 2009. The impact of regulation on auditor fees: Evidence from the sarbanes-oxley act. Auditing: A Journal of Practice & Theory 28, (2): 171-97.

Gietzmann, Miles B., and Angela K. Pettinicchio. 2014. External auditor reassessment of client business risk following the issuance of a comment letter by the SEC. European Accounting Review 23, (1): 57-85.

Gietzmann, Miles, Antonio Marra, and Angela Pettinicchio. 2016. Comment letter frequency and CFO turnover: A dynamic survival analysis. Journal of Accounting, Auditing & Finance 31, (1): 79-99.

Hackenbrack, Karl E., Nicole Thorne Jenkins, and Mikhail Pevzner. 2014. Relevant but delayed information in negotiated audit fees. Auditing: A Journal of Practice & Theory 33, (4): 95-117.

Hail, Luzi and Muhn, Maximilian and Oesch, David, Do Risk Disclosures Matter When it Counts? Evidence from the Swiss Franc Shock (January 23, 2019). Available at SSRN: https://ssrn.com/abstract=2939935 or http://dx.doi.org/10.2139/ssrn.2939935.

Hay, David C., W. Robert Knechel, and Norman Wong. 2006. Audit fees: A meta-analysis of the effect of supply and demand attributes. Contemporary Accounting Research 23, (1): 141-91.

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

———. 2006. Market reaction to verbal components of earnings press releases: Event study using a predictive algorithm. Journal of Emerging Technologies in Accounting 3, (1): 1-19.

Hogan, Chris E., and Michael S. Wilkins. 2008. Evidence on the audit risk model: Do auditors increase audit fees in the presence of internal control deficiencies? Contemporary Accounting Research 25, (1): 219-42.

Hoitash, Rani, Udi Hoitash, and Jean C. Bedard. 2008. Internal control quality and audit pricing under the sarbanes-oxley act. Auditing: A Journal of Practice & Theory 27, (1): 105-26.

Hope, Ole-Kristian, Danqi Hu, and Hai Lu. 2016. The benefits of specific risk-factor disclosures. Review of Accounting Studies 21, (4): 1005-45.

Humpherys, Sean L., Kevin C. Moffitt, Mary B. Burns, Judee K. Burgoon, and William F. Felix. "Identification of fraudulent financial statements using linguistic credibility analysis." Decision Support Systems 50, no. 3 (2011): 585-594.

Hutton, Amy P., Gregory S. Miller, and Douglas J. Skinner. "The role of supplementary statements with management earnings forecasts." Journal of Accounting Research 41, no. 5 (2003): 867-890.

Huang, Xuan, Siew Hong Teoh, and Yinglei Zhang. 2014. Tone management. The Accounting Review 89, (3): 1083-113.

Judd, J. Scott, Kari Joseph Olsen, and James Stekelberg. 2017. How do auditors respond to CEO narcissism? evidence from external audit fees. Accounting Horizons 31, (4): 33-52.

Kothari, Sabino P., Xu Li, and James E. Short. 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): 1639-70.

Kravet, Todd, and Volkan Muslu. 2013. Textual risk disclosures and investors' risk perceptions. Review of Accounting Studies 18, (4): 1088-122.

Krishnan, Gopal V., Mikhail Pevzner, and Partha Sengupta. 2012. How do auditors view managers' voluntary disclosure strategy? the effect of earnings guidance on audit fees. Journal of Accounting and Public Policy 31, (5): 492-515.

Krishnan, Jayanthi, Heibatollah Sami, and Yinqi Zhang. 2005. Does the provision of nonaudit services affect investor perceptions of auditor independence? Auditing: A Journal of Practice & Theory 24, (2): 111-35.

Lang, Mark, and Lorien Stice-Lawrence. 2015. Textual analysis and international financial reporting: Large sample evidence. Journal of Accounting and Economics 60, (2-3): 110-35.

Levi, Shai. 2008. Voluntary disclosure of accruals in earnings press releases and the pricing of accruals. Review of Accounting Studies 13, (1): 1-21.

Li, Chan and Wang, Qian, Sox 4O4 Assessments and Financial Reporting Errors (August 2006). Available at
SSRN: https://ssrn.com/abstract=926180 or http://dx.doi.org/10.2139/ssrn.926180.

Li, Feng. 2010. The information content of forward-looking statements in corporate filings—A naïve bayesian machine learning approach. Journal of Accounting Research 48, (5): 1049-102.

———. 2008. Annual report readability, current earnings, and earnings persistence. Journal of Accounting and Economics 45, (2-3): 221-47.

———.2006. Do Stock Market Investors Understand the Risk Sentiment of Corporate Annual Reports? Available at
SSRN: https://ssrn.com/abstract=898181 or http://dx.doi.org/10.2139/ssrn.898181.

Lobo, Gerald J., and Yuping Zhao. 2013. Relation between audit effort and financial report misstatements: Evidence from quarterly and annual restatements. The Accounting Review 88, (4): 1385-412.

Loughran, Tim, and Bill McDonald. 2016. Textual analysis in accounting and finance: A survey. Journal of Accounting Research 54, (4): 1187-230.

———. 2011. When is a liability not a liability? textual analysis, dictionaries, and 10-Ks. The Journal of Finance 66, (1): 35-65.

McNichols, Maureen. "Evidence of informational asymmetries from management earnings forecasts and stock returns." Accounting Review (1989): 1-27.

Moon, James and Swanquist, Quinn Thomas, Measuring Misreporting Risk in Firms' 10-K Disclosures and the Auditor's Role in Mitigating Misstatements (September 2018). Available at
SSRN: https://ssrn.com/abstract=2997967 or http://dx.doi.org/10.2139/ssrn.2997967.

Nagy, Albert L. 2010. Section 404 compliance and financial reporting quality. Accounting Horizons 24, (3): 441-54.

Palmiter, Alan R. 2017. Examples & explanations for securities regulation. Seventh Edition. Wolters Kluwer Law & Business.

Palmrose, Zoe-Vonna. 1986. Audit fees and auditor size: Further evidence. Journal of Accounting Research: 97-110.

Patelli, Lorenzo, and Matteo Pedrini. "Is tone at the top associated with financial reporting aggressiveness?." Journal of Business Ethics 126, no. 1 (2015): 3-19.

Petersen, Mitchell A. 2009. "Estimating standard errors in finance panel data sets: Comparing approaches." The Review of Financial Studies 22.1: 435-480.

Public Company Accounting Oversight Board (PCAOB). 2017. AS 2201: An Audit of Internal Control Over Financial Reporting That Is Integrated with An Audit of Financial Statements. Available at https://pcaobus.org/Standards/Auditing/Pages/AS2201.aspx

———. 2001. Financial forecasts and projections. Available at https://pcaobus.org/Standards/Attestation/Pages/AT301.aspx.

Raghunandan, K., and Dasaratha V. Rama. 2006. SOX section 404 material weakness disclosures and audit fees. Auditing: A Journal of Practice & Theory 25, (1): 99-114.

Regulation S-K 503. 2018. Available at https://www.ecfr.gov/cgi-bin/text-idx?node=17:3.0.1.1.11&rgn=div5#se17.3.229_1503.

Rice, Sarah C., and David P. Weber. 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): 811-43.

Riley, Tracey Jean. 2011. An investigation into managers' language use in earnings press releases. Dissertation. University of Massachusetts Amherst.

Rogers, Jonathan L., Andrew Van Buskirk, and Sarah LC Zechman. 2011. Disclosure tone and shareholder litigation. The Accounting Review 86, (6): 2155-83.

Schrand, Catherine M., and John A. Elliott. 1998. Risk and financial reporting: A summary of the discussion at the 1997 AAA/FASB conference. Accounting Horizons 12, (3): 271.

SEC. How to read a 10-K. 2011a. Available at https://www.sec.gov/fast-answers/answersreada10khtm.html.

———. US securities and exchange commission/comment letters. 2011b. Available at https://www.sec.gov/fast-answers/answerscommentlettershtm.html.

———. Interpretation: Commission guidance regarding disclosure related to climate change. 2010. Available at https://www.sec.gov/rules/interp/2010/33-9106.pdf.

———. Securities and exchange commission final rule, release no. 33-8591 (FR-75). 2005. Available at https://www.sec.gov/rules/final/33-8591.pdf.

Seetharaman, Ananth, Ferdinand A. Gul, and Stephen G. Lynn. 2002. Litigation risk and audit fees: Evidence from UK firms cross-listed on US markets. Journal of Accounting and Economics 33, (1): 91-115.

Simunic, Dan A. 1980. The pricing of audit services: Theory and evidence. Journal of Accounting Research: 161-90.

Stanley, Jonathan D. 2011. Is the audit fee disclosure a leading indicator of clients' business risk? Auditing: A Journal of Practice & Theory 30, (3): 157-79.

Stice, James D. 1991. Using financial and market information to identify pre-engagement factors associated with lawsuits against auditors. Accounting Review: 516-33.

Tama-Sweet, Isho. 2009. Do Managers Alter the Tone of their Earnings Announcements Around Stock Option Grants and Exercises?.

Tetlock, Paul C., Maytal Saar-Tsechansky, and Sofus Macskassy. 2008. More than words: Quantifying language to measure firms' fundamentals. The Journal of Finance 63, (3): 1437-67.

Tsileponis, Nikolaos and Stathopoulos, Konstantinos and Walker, Martin, Do Corporate Press Releases Drive Media Coverage? (July 11, 2016). Available at SSRN: https://ssrn.com/abstract=2740045 or http://dx.doi.org/10.2139/ssrn.2740045.

Williams, Patricia A. "The relation between a prior earnings forecast by management and analyst response to a current management forecast." Accounting Review (1996): 103-115.

Yang, Rong, Yang Yu, Manlu Liu, and Kean Wu. 2018. Corporate risk disclosure and audit fee: A text mining approach. European Accounting Review 27, (3): 583-94.

Appendices

Appendix A: An Example of Qualitative Earnings Press Releases

South Jersey Industries Inc., third quarter results of fiscal year 2015 (SJI, 8-K filed on 2015.11.05)

SJI Reports Third Quarter Results, Maintains Guidance

Folsom, NJ - South Jersey Industries today announced third quarter 2015 and year-to-date results. GAAP income and Economic Earnings* for the year to date and for the third quarter of 2015 are presented below, as compared with the same periods in 2014...

Economic Earnings through September totaled $55.8 million, as compared with $72.8 million for the prior year period. Although these results reflect a nearly $11 million year-over-year impact associated with the shutdown and subsequent write-down of our energy facility serving the former Revel property, performance in our core utility and commodity marketing businesses remains strong. Economic Earnings for the third quarter of 2015 reflect a loss of $5 million, as compared with a loss of $3.4 million in the third quarter of 2014. While the third quarter is generally a lower producing period, these results were further hindered by increases in write-offs for uncollectible accounts in our utility, as well as increased post-retirement account expenses. Lower levels of investment tax credits associated with solar project development also contributed to the year-over-year variance.

“Our 2015 EPS guidance of $1.49 to $1.54 remains intact,” said SJI President and CEO Michael J. Renna. “And, more importantly, our core businesses continue to perform very well, providing the foundation that will enable us to achieve our long-term target of $150 million of Economic Earnings by 2020.”

“Significant infrastructure investments and customer growth in our utility, coupled with strong contributions from commodity marketing and new fuel supply contracts, will not only drive growth but will also steadily improve the quality of earnings,” Renna added. “These results, combined with the benefits from the PennEast pipeline investment that we expect to see in the latter part of the decade, reinforce our commitment to and confidence in our longer-term strategic objectives.” …

Appendix B. Variable Descriptions and Data Sources in Chapter 3

Variable	Description	Data Source
LNFEE	logarithm of audit fees	Audit Analytics
TONE	(#positive words - #negative words)/total non-numerical words, where positive and negative words are identified using Loughran and McDonald's word lists, and if a positive word is within three words after a negation word (no, not, none, neither, never, nobody), it is considered as a negative word	SEEKINF
ABTONE	abnormal positive tone, which is the residual of tone model	Tone regression
ABTONE_G	ABTONE if $ABTONE_{t-1} < 0.0041$ (75% percentile of ABTONE), and 0 otherwise	Tone regression
ABTONE_B	ABTONE if $ABTONE_{t-1} >= 0.0041$ (75% percentile of ABTONE), and 0 otherwise	Tone regression
EXPOSABTONE	indicator that equal to 1 if ABTONE > 0.0104 (95% percentile of ABTONE)	
PSCORE	predicted probability of misstatements, calculated as the fitted value of the misstatement detection model of Dechow et al. (2011)	Misstatement regression
EARN	earnings before extraordinary items scaled by lagged total assets	Compustat
RET	buy-and-hold monthly returns for 12 months ending three months after the fiscal year-end, calculated using CRSP dataset	CRSP
SIZE	logarithm of total assets	Compustat
BTM	book-to-market ratio measured at fiscal year-end, calculated by the difference between total assets (AT) and total liabilities (LT) divided by market value of common equity (PRCC_F × CSHO)	Compustat
STDRET	standard deviation of RET over the last 12 months ending three months after the fiscal year-end	CRSP
STDEARN	standard deviation of EARN calculated over the last five years, with at least three years of data required	Compustat
AGE	log (1 + # of years since the first year the firm entered the CRSP dataset	CRSP
BUSSEG	log (1+ # of business segments), or 1 if item is missing from Compustat	Compustat
GEOSEG	log (1+ # of geographic segments), or 1 if item is missing from Compustat;	Compustat
LOSS	1 if EARN is negative, 0 otherwise	Compustat
ΔEARN	change in earnings before extraordinary items/beginning total assets	Compustat
AFE	(I/B/E/S actual EPS - median of most recent analysts' forecasts)/stock price at the fiscal year-end	I/B/E/S
AF	analyst consensus forecast for one-year-ahead EPS/stock price at the fiscal year-end	I/B/E/S
ROAEARNING	earnings, calculated as operating income after depreciation (OIADP) divided by total asset (AT);	Compustat
INVREC	inventory (INVT) plus accounts receivable (RECT) divided by total assets (AT	Compustat
FOREIGN	1 if the firm has foreign operations (TXFO), 0 otherwise	Compustat
MERGE	1 if the firm reported the item related to acquisition and merger (AQP), 0 otherwise	Compustat
SPECIAL	1 if the firm reported special items (SPI), 0 otherwise	Compustat
LEVERAGE	the difference between total liabilities (LT) and current liabilities (LCT) divided by total assets (AT);	Compustat
CURRENTRATIO	current assets (ACT) divided by current liabilities	Compustat
GROWTH	the percentage of change in sales (SALE) from period n-1 to period n	Compustat
BIG4	1 if a successor auditor is one of the Big 4, 0 otherwise	Audit Analytics
RESIGNATION	1 if a predecessor auditor initiated auditor resigns, 0 otherwise	Audit Analytics
DISMISSAL	1 if predecessor auditor initiated auditor dismissals, 0 otherwise	Audit Analytics
INITIAL	1 if RESIGNATION = 1 or DISMISSAL = 1	Audit Analytics
GC	1 if a successor auditor issues a going-concern opinion, 0 otherwise	Audit Analytics
IW	1 if a successor auditor indicates internal control weakness, 0 otherwise	Audit Analytics
RESTATEMENT	1 if the annual report is restated, and 0 otherwise	Audit Analytics
ABRET	the difference between annual buy-and-hold stock return and annual buy-and-hold value-weighted index return	CRSP
LAGABRET	ABRET lagged by 1 year	CRSP
ΔROA	change in IB/AT from year t-1 to t	Compustat
ΔINV	change in INVT/AT from year t-1 to t	Compustat
ΔREC	change in RECT/AT from year t-1 to t	Compustat

ΔCSALE	change in cash sales divided by prior year cash sale, where cash sales CSALE = SALE = ΔREC	Compustat
TOTALACCURAL	change in noncash assets (noncash total assets minus total liabilities and preferred stocks) from year t_1 to year t scaled by average total assets	Compustat
SOFT_ASSETS	soft assets (AT - PPENT - CHE) scaled by total assets	Compustat
ΔEMP	the difference between the percentage change in the number of employees and the percentage change in total assets	Compustat
LEASE	1 if MRCT > 0, 0 otherwise	Compustat
ISSUANCE	1 if DLTIS > 0 or SSTK>0, 0 otherwise	Compustat
ABRET	difference between annual buy-and-hold stock return and annual buy-and-hold value-weighted market return	CRSP
LAGABRET	lag of ABRET	CRSP
TENURE	natural logarithm of (1+number of years that the auditor has been successively auditing the client)	Audit Analytics
INDEXPERT	1 if total audit fees by the auditor in the industry is over 30% of the total audit fees for the industry both city-wide and national wide	Audit Analytics
MFERROR	management forecast error, calculated as the difference between mean value of management forecast and most recent actual EPS, scaled by stock price	I/B/E/S

Appendix C: Variable Descriptions and Data Sources in Chapter 4

Variable	Description	Data Source
IC_RF	1 if the firm disclose at least one ICFR-related risk factor, and 0 otherwise	EDGAR
ONLYICRF	1 if the firm only discloses ICFR-related risk factors, and 0 otherwise	EDGAR
ONLYMW	1 if the firm discloses material weaknesses, and 0 otherwise	EDGAR
ICRF_MW	1 if the firm discloses both ICFR-related risk factors and material weaknesses, and 0 otherwise	EDGAR
RESTATEMENT	1 if the firm restate the 10-K filing for the fiscal year, and 0 otherwise	Audit Analytics
TOTALACCURAL	Change in noncash assets (noncash total assets minus total liabilities and preferred stocks) from year t_1 to year t scaled by average total assets	Compustat
ΔREC	Change in RECT/AT from year t-1 to t	Compustat
ΔINV	Change in INVT/AT from year t-1 to t	Compustat
SOFT_ASSETS	Soft assets (AT - PPENT - CHE) scaled by total assets	
ΔCSALE	Change in cash sales divided by prior year cash sale, where cash sales CSALE = SALE = ΔREC	Compustat
ΔROA	Change in IB/AT from year t-1 to t	Compustat
ISSUANCE	1 if DLTIS > 0 or SSTK>0, 0 otherwise	Compustat
ΔEMP	The difference between the percentage change in the number of employees and the percentage change in total assets	Compustat
LEASE	1 if MRCT > 0, 0 otherwise	Compustat
ABRET	Difference between annual buy-and-hold stock return and annual buy-and-hold value-weighted market return	CRSP
LAGABRET	Lag of ABRET	CRSP
MW	1 if the firm or auditor discloses material weaknesses in internal control, and 0 otherwise	Audit Analytics
LNASSET	Logarithm of total assets	Compustat
INVREC	Inventory (INVT) plus accounts receivable (RECT) divided by total assets (AT	Compustat
BUSSEG	Log (1+ # of business segments), or 1 if the item is missing from Compustat	Compustat
FOREIGN	1 if the firm has foreign operations (TXFO), 0 otherwise	Compustat
MERGE	1 if the firm reported the item related to acquisition and merger (AQP), 0 otherwise	Compustat
SPECIAL	1 if the firm reported special items (SPI), 0 otherwise	
LEVERAGE	The difference between total liabilities (LT) and current liabilities (LCT) divided by total assets (AT);	Compustat
CURRENTRATIO	Current assets (ACT) divided by current liabilities	Compustat
BTM	Book-to-market ratio measured at fiscal year-end, calculated by the difference between total assets (AT) and total liabilities (LT) divided by market value of common equity (PRCC_F × CSHO)	
GROWTH	The percentage of change in sales (SALE) from period n-1 to period n	Compustat
BIG4	1 if a successor auditor is one of the Big 4, 0 otherwise	Audit Analytics
RESIGNATION	1 if a predecessor auditor initiated auditor resigns, 0 otherwise	Audit Analytics
DISMISSAL	1 if predecessor auditor initiated auditor dismissals, 0 otherwise	Audit Analytics
GC	1 if a successor auditor issues a going-concern opinion, 0 otherwise	Audit Analytics
TENURE	Log (1+number of years that the auditor has been successively auditing the client)	Compustat
INDEXPERT	1 if total audit fees by the auditor in the industry are over 30% of the total audit fees for the industry both city-wide and national wide	Compustat
AUDITLAG	Number of days from end of fiscal year (FISCAL_YEAR_END_OP) to issuance of the audit opinion (FILE_DATE).	Audit Analytics
AZSCORE	Altman (1968) z-score, categorized in the same way as Krishnan and Wang (2015): setting to 2 if the zscore is less than 1.81, 1 if the z-score is between 1.81 and 2.99, and 0 if the z-score is greater than 3.	Compustat

MKBK	Ratio of the market value of equity to the book value of equity	Compustat
RESTRUCTURE	Pretax restructuring charges scaled by total assets (RCP/AT), and equal to 0 if missing.	Compustat
AGE	log (1 + # of years since the first year the firm entered the CRSP dataset	CRSP
OVERCONFIDENT	A dummy variable equal to 1 if the residual from regressing the sum of capital expenditures, research and development expense, and acquisitions, less cash received from the sale of property, plant, and equipment) scaled by lagged total assets (CAPX + XRD + AQC - SPPIV)/AT) on GROWTH is in the top quartile for the firm's industry-year, and 0 otherwise.	Compustat
LNFEE	natural logarithm of audit fees	Audit Analytics
PSCORE	Predicted probability of misstatements, calculated as the fitted value of the original misstatement detection model of Dechow et al (2011)	Misstatement regression
ΔREC	Change in RECT/AT from year t-1 to t	Compustat
ΔINV	Change in INVT/AT from year t-1 to t	Compustat
SOFT_ASSETS	Soft assets (AT - PPENT - CHE) scaled by total assets	
ΔCSALE	Change in cash sales divided by prior year cash sale, where cash sales CSALE = SALE = ΔREC	Compustat
ΔROA	Change in IB/AT from year t-1 to t	Compustat
ISSUANCE	1 if DLTIS > 0 or SSTK>0, 0 otherwise	Compustat
ΔEMP	The difference between the percentage change in the number of employees and the percentage change in total assets	Compustat
LEASE	1 if MRCT > 0, 0 otherwise	Compustat
ABRET	Difference between annual buy-and-hold stock return and annual buy-and-hold value-weighted market return	CRSP
LAGABRET	Lag of ABRET	CRSP
IC_RANK	1-original rank of ICFR-related risk factor divided by total number of risks in Item 1A	Calculated
LENGTH	Total number non-stop words in the ICFR-related risk factor	Calculated
SPECIFICITY	Number of specific words in ICFR-related risk factor divided by LENGTH	Calculated
BOILERPERCENT	Number of words in sentences that include at least one of the 4-word-phrases that appear in at least 50% of all ICFR-related risk factors divided by total number of words in the ICFR-related risk factor	Calculated

Variable	Description	Data Source
ONLYICRF	1 if the firm only discloses ICFR-related risk factors, and 0 otherwise	EDGAR
ONLYMW	1 if the firm discloses material weaknesses, and 0 otherwise	EDGAR
ICRF_MW	1 if the firm discloses both ICFR-related risk factors and material weaknesses, and 0 otherwise	EDGAR
RESTATEMENT	1 if the firm restate the 10-K filing for the fiscal year, and 0 otherwise	Audit Analytics
TOTALACCURAL	Change in noncash assets (noncash total assets minus total liabilities and preferred stocks) from year t_1 to year t scaled by average total assets	Compustat
ΔREC	Change in RECT/AT from year t-1 to t	Compustat
ΔINV	Change in INVT/AT from year t-1 to t	Compustat
SOFT_ASSETS	Soft assets (AT - PPENT - CHE) scaled by total assets	
ΔCSALE	Change in cash sales divided by prior year cash sale, where cash sales CSALE = SALE = ΔREC	Compustat
ΔROA	Change in IB/AT from year t-1 to t	Compustat
ISSUANCE	1 if DLTIS > 0 or SSTK>0, 0 otherwise	Compustat
ΔEMP	The difference between the percentage change in the number of employees and the percentage change in total assets	Compustat
LEASE	1 if MRCT > 0, 0 otherwise	Compustat
ABRET	Difference between annual buy-and-hold stock return and annual buy-and-hold value-weighted market return	CRSP
LAGABRET	Lag of ABRET	CRSP

MW	1 if the firm or auditor discloses material weaknesses in internal control, and 0 otherwise	Audit Analytics
LNASSET	Logarithm of total assets	Compustat
INVREC	Inventory (INVT) plus accounts receivable (RECT) divided by total assets (AT	Compustat
BUSSEG	Log (1+ # of business segments), or 1 if the item is missing from Compustat	Compustat
FOREIGN	1 if the firm has foreign operations (TXFO), 0 otherwise	Compustat
MERGE	1 if the firm reported the item related to acquisition and merger (AQP), 0 otherwise	Compustat
SPECIAL	1 if the firm reported special items (SPI), 0 otherwise	
LEVERAGE	The difference between total liabilities (LT) and current liabilities (LCT) divided by total assets (AT);	Compustat
CURRENTRATIO	Current assets (ACT) divided by current liabilities	Compustat
BTM	Book-to-market ratio measured at fiscal year-end, calculated by the difference between total assets (AT) and total liabilities (LT) divided by market value of common equity (PRCC_F × CSHO)	
GROWTH	The percentage of change in sales (SALE) from period n-1 to period n	Compustat
BIG4	1 if a successor auditor is one of the Big 4, 0 otherwise	Audit Analytics
RESIGNATION	1 if a predecessor auditor initiated auditor resigns, 0 otherwise	Audit Analytics
DISMISSAL	1 if predecessor auditor initiated auditor dismissals, 0 otherwise	Audit Analytics
GC	1 if a successor auditor issues a going-concern opinion, 0 otherwise	Audit Analytics
TENURE	Log (1+number of years that the auditor has been successively auditing the client)	Compustat
INDEXPERT	1 if total audit fees by the auditor in the industry are over 30% of the total audit fees for the industry both city-wide and national wide	Compustat
AUDITLAG	Number of days from end of fiscal year (FISCAL_YEAR_END_OP) to issuance of the audit opinion (FILE_DATE).	Audit Analytics
AZSCORE	Altman (1968) z-score, categorized in the same way as Krishnan and Wang (2015): setting to 2 if the zscore is less than 1.81, 1 if the z-score is between 1.81 and 2.99, and 0 if the z-score is greater than 3.	Compustat
MKBK	Ratio of the market value of equity to the book value of equity	Compustat
RESTRUCTURE	Pretax restructuring charges scaled by total assets (RCP/AT), and equal to 0 if missing.	Compustat
AGE	log (1 + # of years since the first year the firm entered the CRSP dataset	CRSP
OVERCONFIDENT	A dummy variable equal to 1 if the residual from regressing the sum of capital expenditures, research and development expense, and acquisitions, less cash received from the sale of property, plant, and equipment) scaled by lagged total assets (CAPX + XRD + AQC - SPPIV)/AT) on GROWTH is in the top quartile for the firm's industry-year, and 0 otherwise.	Compustat
LNFEE	natural logarithm of audit fees	Audit Analytics
PSCORE	Predicted probability of misstatements, calculated as the fitted value of the original misstatement detection model of Dechow et al (2011)	Misstatement regression
ΔREC	Change in RECT/AT from year t-1 to t	Compustat
ΔINV	Change in INVT/AT from year t-1 to t	Compustat
SOFT_ASSETS	Soft assets (AT - PPENT - CHE) scaled by total assets	
ΔCSALE	Change in cash sales divided by prior year cash sale, where cash sales CSALE = SALE = ΔREC	Compustat
ΔROA	Change in IB/AT from year t-1 to t	Compustat
ISSUANCE	1 if DLTIS > 0 or SSTK>0, 0 otherwise	Compustat
ΔEMP	The difference between the percentage change in the number of employees and the percentage change in total assets	Compustat
LEASE	1 if MRCT > 0, 0 otherwise	Compustat
ABRET	Difference between annual buy-and-hold stock return and annual buy-and-hold value-weighted market return	CRSP

LAGABRET	Lag of ABRET	CRSP
IC_RANK	1-original rank of ICFR-related risk factor divided by total number of risks in Item 1A	Calculated
LENGTH	Total number non-stop words in the ICFR-related risk factor	Calculated
SPECIFICITY	Number of specific words in ICFR-related risk factor divided by LENGTH	Calculated
BOILERPERCENT	Number of words in sentences that include at least one of the 4-word-phrases that appear in at least 50% of all ICFR-related risk factors divided by total number of words in the ICFR-related risk factor	Calculated

Appendix D: Examples of ICFR-related risk factors

1. An informative ICFR-related risk factor

We face risks related to our recent accounting restatements. In 2004, we reported a restatement to previously issued financial statements. More recently, in February 2007, we reported that we had discovered accounting errors in previously reported Consolidated Statements of Operations and Comprehensive Earnings. These errors related to the presentation of deferred charge as a non-interest expense amount compared to the restated presentation as a component of income tax expense. We also reported restated amounts in the Consolidated Statements of Cash Flows to eliminate certain non-cash items related to intercompany transactions and the redesignation of loans from held-for-sale to held-for-investment. The restatement of our financial statements could lead to litigation claims and/or regulatory proceedings against us. The defense of any such claims or proceedings may cause the diversion of management's attention and resources, and we may be required to pay damages if any such claims or proceedings are not resolved in our favor. Any litigation or regulatory proceeding, even if resolved in our favor, could cause us to incur significant legal and other expenses. We also may have difficulty raising equity capital or obtaining other financing, such as lines of credit or otherwise. We may not be able to effectuate our current operating strategy. The occurrence of any of the foregoing could harm our business and reputation and cause the price of our securities to decline. (IMPAC MORTGAGE HOLDINGS, INC., 10-K of the fiscal year 2007)

2. A less informative ICFR-related risk factor

Failure of our internal controls over financial reporting could harm our business and financial results. Our management is responsible for establishing and maintaining effective internal control over financial reporting. Internal control over financial reporting is a process to provide reasonable assurance regarding the reliability of financial reporting for external purposes in accordance with accounting principles generally accepted in the United States. Because of its inherent limitations, internal control over financial reporting is not intended to provide absolute assurance that we would prevent or detect a misstatement of our financial statements or fraud. Any failure to maintain an effective system of internal control over financial reporting could limit our ability to report our financial results accurately and timely or to detect and prevent fraud. A significant financial reporting failure or material weakness in internal control over financial reporting could cause a loss of investor confidence and decline in the market price of our stock. (BUFFALO WILD WINGS, INC., 10-K of the fiscal year 2009)