June 2023

The Rise of Text Analysis: Using Machine Learning to Explain the Variation in Going Concern Accuracy

Yimei Zhang
University of South Florida

The Rise of Text Analysis: Using Machine Learning to Explain the Variation in Going Concern Accuracy

by

Yimei Zhang

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
with a concentration in Accounting
Lynn Pippenger School of Accountancy
Muma College of Business
University of South Florida

Major Professor: Thomas Smith, Ph.D.
Uday Murthy, Ph.D.
Jong Chool Park, Ph.D.
Eun Sook Kim, Ph.D.

Date of Approval:
June 5, 2023

Keywords: Machine Learning, Natural Language Processing, Risk Disclosure, Management Discussion and Analysis, Going Concern Opinion, Type I Error, Type II Error

DEDICATION

This dissertation is dedicated to all the individuals who have supported me throughout my entire life.

ACKNOWLEDGMENTS

I would like to express my heartfelt gratitude to the following individuals who have played a pivotal role in the completion of my dissertation:

First and foremost, I extend my deepest appreciation to my dissertation chair, Dr. Tom Smith. His guidance, expertise, and unwavering support have been instrumental throughout my research journey. His invaluable insights, constructive feedback, and scholarly mentorship have shaped my academic growth and significantly enhanced the quality of this dissertation.

I am also grateful to the members of my dissertation committee, Dr. Uday Murthy, Dr. Jong Chool Park, and Dr. Eun Sook Kim, for their valuable input and critical evaluation of my work. Their expertise and thoughtful suggestions have undoubtedly contributed to the refinement of my research.

I am immensely grateful to the Muma College of Business at the University of South Florida for providing a conducive learning environment and ample opportunities for personal and academic development. To the esteemed faculty members, thank you for your guidance, mentorship, and profound knowledge. To my fellow Ph.D. colleagues, thank you for being a source of inspiration and a collaborative community.

I would like to thank my loving parents for their unwavering support, patience, and understanding during this challenging academic pursuit. Their encouragement and belief in my abilities have been the driving force behind my success. I am forever indebted to my parents for their unconditional love, constant encouragement, and endless motivation. I am also deeply grateful to the other members of my family for their steadfast support. Their constant presence, understanding, and love have given me the strength and motivation to overcome challenges and reach for my goals.

In addition, I extend my gratitude to all my friends who have provided me with a strong support system throughout the whole process. Their encouragement, intellectual discussions, and mental support have been invaluable in keeping me motivated. I would also like to extend a special thanks to my beloved feline friends, who provided endless emotional support and companionship during the ups and downs of this journey.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Auditors are required to issue modified audit opinions if they have sufficient doubts about the client's ability to continue as a going concern. These going concern opinions represent an important information resource for financial statement users to evaluate client performance, and are associated with a number of negative capital market outcomes (e.g. negative returns, increased cost of capital, etc.). Despite being used by capital market participants, going concern opinions are commonly plagued with Type I errors (false positive) and Type II errors (false negative), making them a particularly noisy measure. The purpose of this study is to determine whether machine learning can be leveraged to reduce this noise by (1) identifying disclosure patterns where going concern accuracy is likely lower (higher) and (2) developing measures from these disclosures that can help predict variation in going concern accuracy. Specifically, I use a machine learning technique (Top2Vec) to identify differences in disclosure topics among financially distressed clients' Risk disclosures (Item 1A) and Management Discussion and Analysis disclosures (Item 7) conditioned on the accuracy of the going concern opinion (accurate, Type I error or Type II error). I find significant differences in the topics that are discussed among Type I error/Type II error clients compared to clients receiving accurate going concern opinions/evaluations. Accurate going concern *opinions* are the situation that clients receive going concern opinions in the current year and file for bankruptcy protection in the subsequent year. Accurate going concern *evaluations* not only include the situations of accurate going concern opinions but also include the situations that clients do not receive going concern

opinions in the current year and do not file for bankruptcy protection in the subsequent year (e.g., an accurate omission of a going concern opinion). In the Type I error settings (ignoring Type II errors in this analysis), the probability of accurate going concern ***opinion*** is higher if clients disclose human capital and supply chain risks, or if clients disclose tax related factors. The probability of accurate going concern ***evaluation*** is higher (lower) if clients disclose human capital, dispersion, legal, and macro-economic risks (funding, financial condition, debt, operational, attestation, and stock market risks), and the probability is lower if clients disclose the facts regarding growing potentials, stocks, and political contributions. In the Type II error settings (ignoring Type I errors in this analysis), the probability of accurate going concern ***opinion*** is higher (lower) if clients disclose bankruptcy and operational risks (development, supply chain, and environmental risks), or if clients disclose the facts regarding bankruptcy, performance changes, and costs (operational performance and tax). The probability of accurate going concern ***evaluation*** is higher (lower) if clients disclose macro-economic, intellectual property, and investment risks (development and oil/gas risks), or if clients disclose the facts regarding human capital (loan and operational performance). After providing evidence of which disclosure topics are associated with going concern accuracy, I then examine whether machine learning can be used to create measures (based on the textual information disclosed in Item 1A and Item 7) to improve models attempting to determine whether an observed going concern opinion is accurate. My findings support the validity and effectiveness of these machine learning developed proxies in predicting accurate going concerns, identifying Type I errors, and identifying Type II errors. I further demonstrate their superiority over other common text-based measures that do not utilize machine learning. The findings of this study have important implications for auditors, regulators, and academia.

CHAPTER ONE: INTRODUCTION

According to PCAOB audit standards (AS 2415), the external auditor is required, as a part of a standard audit, to evaluate a client's ability to continue as a going concern for at least one year after the end of the fiscal period. This evaluation is often performed using analytical procedures where indicators of deteriorating financial conditions (e.g., negative cash flows from operations) or other events (e.g. potential legal liabilities) can be uncovered that may lead the auditor to have sufficient doubts about the client's ability to continue as a going concern. If sufficient doubt exists and cannot be alleviated after considering the mitigation plans from managers, the auditor is required to modify their audit report to disclose this potential operational discontinuation and note the explanation that led them to conclude that the client may not be able to continue as a going concern.[1]

Multiple accounting studies document significant negative market outcomes for clients receiving going concern opinions, which include increased cost of equity, lower stock prices, increased risks of bankruptcy surprises, lower IPO valuations, and reduced willingness for investors to rely on financial statements when making decisions (Amin, Krishnan, and Yang 2014; O'Reilly 2010; Chen and Church 1996; Willenborg and McKeown 2000; Blay, Geiger, and North 2011). On the surface, these associations may not seem all that surprising, but it is important to keep in mind that these clients are often experiencing financial hardship well in

[1] I refer to these modified opinions as "going concern opinions" throughout my study.

advance of receiving a going concern opinion. The reason why the opinion is incrementally informative for users of financial information is that auditors have access to more information and are therefore in a unique position to make such an evaluation. In spite of this access to better information, studies have raised concerns about the inaccuracy of going concern opinions. They classify these inaccuracies into Type I error (false positive) and Type II error (false negative) going concern opinions. Type I errors occur when the auditor expresses sufficient doubt about the client's ability to continue as a going concern by evaluating the conditions and events in the current period, but the client does not file for bankruptcy protection in the subsequent year. On the other hand, Type II errors occur when the auditor does not express concern about the client's ability to continue as a going concern, but the client does file for bankruptcy protection in the subsequent year. Accurate going concern *opinions* occur when the auditor issues a going concern opinion in the current year and the client files for bankruptcy protection in the subsequent year. This is the way that prior research has traditionally defined accurate going concern opinions. However, there is another situation where the auditor provides an accurate assessment of a client's ability to continue as a going concern, in which they decide not to issue a going concern opinion and the client does not declare bankruptcy in the following year. I consider these to also be accurate going concern assessments, and combine them with the traditional instance of an accurate going concern and refer to these collectively as accurate going concern *evaluations*.[2]

Blay, Moon, Paterson (2018) and Carson, Fargher, Geiger, Lennox, Raghunandan, and Willekens (2013) provide evidence to show that more than 80 percent of the going concern opinions have Type I errors and more than 40 percent of the going concern opinions have Type

[2] I use this term to identify these two situations, because in both instances the auditor is making the correct going concern *evaluation*.

II errors. Going concern opinion errors are concerning to auditors, as they are associated with larger audit offices' market share reduction and higher dismissal rates (Xu and Kalelkar 2020). Due to the significant economic impact of going concern opinions and the prevalence of going concern opinion errors, the Center of Audit Quality (CAQ) has called for more research to help practitioners reduce going concern opinion errors (CAQ 2012).

Several researchers have responded to the call of CAQ (2012). One stream of studies investigates this question from the auditors' perspective. Blay et al. (2018) build on Geiger and Rama (2006) and find Big 4 auditors have lower Type I and Type II going concern opinion errors relative to non-Big 4 auditors, suggesting that larger audit firms, with more resources, are in a better position to make accurate going concern assessments. In addition, Geiger, Basioudis, and Delange (2022) document a negative association between non-audit fees and Type I/Type II errors, suggesting that client-specific information obtained from non-audit services can improve going concern opinion accuracy. Another stream of studies examines going concern accuracy from management's perspective. Berglund, Herrmann, and Lawson (2018) and Budisantoso, Rahmawati, Bandi, and Probohudono (2017) find that clients with higher management abilities and higher audit committee independence are associated with more accurate going concern opinions. A third stream of research has begun to explore whether variation in going concern accuracy can be explained by using machine learning algorithms. Vasarhelyi, Zhang, and Gu (2022) show that machine learning can improve the explanatory power of going concern accuracy prediction models. An important limitation of this study is that they only test the advantages of the machine learning algorithm per se (compared to linear regressions), and limit the inputs considered by the algorithm to financial ratios (e.g., return on assets and leverage).

Although these recent studies offer some insights on what types of conditions appear to influence the prevalence of Type I and Type II errors, most of them focus on external factors, such as the expertise of certain human capitals (e.g., Big 4 or non-audit service effects), which requires significant structural changes by the auditor and/or management to meaningfully improve accuracy. These studies also do not incorporate the ways in which the auditor evaluates conditions and events (e.g., negative cash flows, supply chain risks, and legal liabilities) of the client into their design, despite the fact that auditors must consider them when assessing doubt about the client's ability to continue as a going concern (PCAOB 2015; AICPA 2015; FASB 2014). My study intends to fill this gap in the literature by offering a lower-cost approach, which specifically utilizes machine learning to improve going concern accuracy. Specifically, I leverage machine learning technology to determine the conditions and events from the textual information in 10-K disclosures that can explain the variation of going concern opinion accuracy and can therefore be used to reduce Type I or Type II going concern errors. Second, I utilize machine learning technology to create textual-based proxies and demonstrate their validity and effectiveness in predicting going concern accuracy.

To obtain the conditions and events that cause auditors to issue accurate and inaccurate going concern opinions, I first use machine learning algorithms to analyze the textual information disclosed by financially distressed clients in their 10-K disclosures. Top2Vec is the machine learning algorithm I use, because it can summarize the common topics presented in multiple textual documents and assign the specific topic that each document pertains to. I import the Risk Factor Section (Item 1A) and the Management Discussion and Analysis Section (Item 7) disclosures of financially distressed clients into Top2Vec algorithms and have them summarize the common topics for Item 1A and Item 7 separately. Each paragraph within the disclosure is defined as having one topic and is imported as an individual input into the machine learning algorithms. After generating the common topics among all the observations in each going concern setting (Type I error and Type II error), the machine learning algorithms then assign a specific topic to each paragraph. After aggregating the topics shown in all the paragraphs within each client, the content of Item 1A and Item 7 disclosures can be explicitly identified as whether Topic_N is present in a given client's disclosure.[3] Each topic is described using fifty keywords.

After using the machine learning algorithms to generate the topics disclosed by financially distressed clients, I empirically examine whether certain topics are associated with different going concern opinion outcomes. There are four possible outcomes to compare - accurate going concern *opinions* (GC_t=1 and $Bankrupt_{t+1}$=1), accurate going concern *evaluations* (GC_t=1 and $Bankrupt_{t+1}$=1 or GC_t=0 and $Bankrupt_{t+1}$=0), Type I errors (GC_t=1 and $Bankrupt_{t+1}$=0), Type II errors (GC_t=0 and $Bankrupt_{t+1}$=1). I conduct my analysis by comparing Type I error outcomes to both accurate going concern *opinions* and accurate going concern *evaluations*. I then repeat this analysis comparing Type II error outcomes to both accurate going concern *opinions* and accurate going concern *evaluations*.

With respect to Type I errors, I find that the probability of receiving accurate going concern *opinions* (and avoiding Type I errors) is higher if clients disclose human capital risks and supply chain risks in Item 1A or tax related information in Item 7 section.[4] I further find that the probability of accurate going concern *evaluations* (and avoiding Type I errors) is higher if clients disclosed human capital, dispersion, legal, and macro-economic risks and is lower if

[3] For example, one observation in the sample is assigned as variables item1a_topic0 – item1a_topic16 =1, variables item1a_topic17=0, variable item1a_topic18– item1a_topic26= 1, and variables item1a_topic27– item1a_topic29=0.
[4] The topics with significant coefficients are labeled based on interpreting the keywords.

clients disclose funding, financial condition, debt, operational, attestation, and stock market risks in item 1A. For the Item 7 disclosures, clients that explain the facts regarding growing potentials, stocks, and political contributions are less likely to receive accurate going concern *evaluations* (and avoiding Type I errors).

With respect to Type II errors, I find that the probability of accurate going concern *opinions* (and avoiding Type II errors) is higher if clients disclose bankruptcy and operational risks, and lower if clients disclose development, supply chain, and environmental risks in Item 1A. For Item 7 sections, clients are more likely to receive accurate going concern *opinions* (and avoiding Type II errors) if they disclose facts regarding bankruptcy, performance changes, and costs, but are less likely to receive those Type II if they disclose facts regarding operational performance and taxes. I further find that clients with more macro-economic risk, intellectual property risk, and investment risk disclosures have a higher probability of receiving accurate going concern *evaluations* (and avoiding Type II errors), while clients with more development risk and oil/gas risk disclosures have a lower probability of receiving accurate going concern *evaluations* (and avoiding Type II errors). Regarding the topics disclosed in Item 7, clients that disclose information about human capital are more likely to have accurate going concern evaluations (and avoiding Type II errors). Clients that disclose more facts of loan and operational performance are less likely to have accurate going concern *evaluations* (and avoiding Type II errors).

After using machine learning to find disclosure topics within the business Risk and MD&A sections of the 10K that make clients more or less likely to suffer Type I and Type II errors, I next examine whether machine learning can help develop textual-based proxies that can improve our understanding of going concern accuracy. To conduct this analysis, I use multiple machine learning algorithms (Naïve Bayes Classification, Supporting Vector Machine, and Random Forest Classification) to create three proxies that are then used to predict the probability of identifying an accurate going concern opinion, Type I error, or Type II error. These proxies are developed using the textual information disclosed in Item 1A and Item 7 sections. All three algorithms are trained based on 70% of the observations (training sample), and the empirical tests are then conducted based on 30% of the observations (testing sample). I then evaluate the validity and effectiveness of each proxy by observing whether they are associated with the occurrence of accurate, Type I, and Type II going concern opinions and whether the addition of this new variable significantly increases the explanatory power of the model. My results provide robust support for the textual-based proxies generated from machine learning to predict going concern accuracy variation. More importantly, the explanatory power after adding each of the machine learning proxies is significantly higher than what is attained when using traditional textual analysis attributes (e.g, readability, specificity, hard information, and tone). Therefore, this study provides compelling evidence as to the effectiveness of using machine learning to improve going concern accuracy prediction models.

The findings of this study have multiple contributions. First, this study provides evidence that the presence (or absence) of key topics disclosed within the Item 1A or Item 7 section of financial statements are associated with Type I and Type II errors surrounding going concern opinions. These disclosure topics may represent the conditions and events that cause auditors to be more (less) likely to issue accurate going concern *opinions*, or more (less) likely to make accurate going concern *evaluations*. If auditors wish to improve their going concern reporting accuracy, my findings suggest that auditors should consider the presence (or absence) of the topics identified in this study among their client's disclosures. A strength of this machine learning approach is that it would be relatively easy to modify/refine (focus on specific industries, expand this to include even more disclosures) for a given audit firm. This tool can provide another point of data that auditors can consider when opining on their clients' ability to continue as an ongoing concern.

Second, this study demonstrates the validity and effectiveness of using machine learning in predicting going concern accuracy. The proxies generated from Naïve Bayes Classification, Supporting Vector Machine, and Random Forest Classification are all able to predict going concern accuracy. These proxies also perform better than other common text-based measures that are not generated by machine learning algorithms. From an investor perspective, the ability to determine whether a going concern opinion (or lack thereof in the case of a Type II error) is appropriate can be very value relevant. My findings suggest that the use of machine learning can greatly improve the ability of an investor to make this important assessment.

Third, the findings of this study should be of interest to the regulators as it responds to their call for research to improve the evaluation processes for going concern opinions. They note that many clients faced unprecedented challenges in their continuous operations after the Covid-19 pandemic, so the going concern evaluation processes should be refreshed in order to incorporate and more accurately estimate the impact of events and conditions on going concern opinions (Deloitte 2020; CAQ 2020). The qualitative disclosure topics relating to emerging risks and forward-looking projections that are found in Item 1A and 7 are likely to contain information on the very events and conditions that the CAQ suggests auditors should consider when making going concern assessments. The findings of my study therefore act to provide evidence in support of the CAQ recommendation.

Finally, this study expands the going concern literature by showing how risks and firm facts disclosed in financial statements are associated with the probability of going concern accuracy variation. The literature shows that the Big 4 effect, audit committee independence, non-audit services, and managerial abilities of clients can reduce going concern opinion errors due to certain auditors' or managers' expertise. However, no study investigates what specific conditions and events mislead or contribute to auditors' judgment on going concern evaluation processes. This study represents an important first step to show how one can leverage machine learning techniques to analyze the textual information of Risk Factor and MD&A disclosures, in order to reveal the conditions and events that are associated with going concern accuracy.

This study proceeds as follows: section two summarizes the literature on going concern opinion accuracy, dictionary-based textual analysis, and machine-learning-based textual analysis; section three presents the machine learning design in this study; section four provides the empirical evidence to support the validity of machine learning outputs; and section five mention the conclusion of this study.

CHAPTER TWO: LITERATURE REVIEW

Going-Concern Accuracy

AS 2415 (PCAOB 2015) and AU-C Section 9570 (AICPA 2015) state that auditors have the responsibility to assess whether there is substantial doubt regarding the client's ability to continue as a going concern for a reasonable time, which is defined as the subsequent 12 months after the financial statement date. When analyzing financial reports with unmodified audit opinions, investors operate under the maintained assumption that the entity has the ability to continue operations for at least the following year. This is the reason why auditors are required to modify their audit opinions if they have substantial doubt about the company's ability to continue as a going concern. When evaluating a client's ability to continue as a going concern, auditors identify the client-specific conditions and events that may portend future bankruptcy and then collect evidence to assess whether those conditions or events are present for their client. The PCAOB organizes the kinds of conditions and events that could lead to bankruptcy into four broad groups: negative trends, other indications of possible financial difficulties, internal matters, and external matters that have occurred (PCAOB 2015). Negative trends include certain financial indicators or ratios (e.g., negative operational cash flows and recurring operating losses) that are associated with negative operational performance. Other indicators of possible financial difficulties include conditions that refer to financial difficulties, such as trade credit declines and debt restructurings. Internal matters include significant changes inside the client, such as labor difficulties. External matters include events from external parties that have significant impacts on clients' operations, such as legal liabilities and unexpected disasters.

If sufficient doubt exists regarding the ability of the client to continue as a going concern based on the evaluation of certain conditions and events, auditors are then required to ask management to provide mitigation plans (e.g. disposal of certain assets or plans to raise additional capital) to evaluate whether the plans can sufficiently reduce the going concern risks that have been identified. If, after considering any and all mitigation plans, auditors still have sufficient doubt about the clients' ability to continue as going concern, auditors are then required to modify their audit opinions to reflect the uncertainty of continuous operation. An explanatory paragraph is included in the audit report to inform financial statement users what conditions and events lead auditors to issue modified audit opinions and what the impacts are regarding those events. In addition, auditors need to appropriately document their assessment and communicate their going concern evaluation to the audit committee.

Studies have shown that auditors often incorrectly evaluate the extent and timing of the negative impact of conditions and events on a client's ability to continue as a going concern, resulting in two types of errors. The first type of going concern opinion error occurs when a client does not file for bankruptcy protection and continues to operate one year after the auditor issues a going concern opinion (Type I error). The other type of going concern error occurs when a client files for bankruptcy protection within a year of receiving an auditor's opinion that does not include the going concern modification (Type II error). Prior research has documented the prevalence of Type I and Type II errors of going concern opinions. The prevalence of these two types of errors is nontrivial. Carson et al. (2013) report that 40 to 50 percent of bankrupt companies do not receive going concern options in the year prior (Type II errors). In addition, they note that 80 to 90 percent of going concern companies do not file for bankruptcy protection in the subsequent year (Type I errors).

Prior research has investigated the determinants of going concern accuracy. Geiger and Rama (2006) find that Big 4 auditors have fewer Type I and Type II errors in issuing going concern opinions than non-Big 4 auditors. Based on this finding, other studies further test the factors that lead to an accuracy difference between Big 4 and non-Big 4 auditors. Blay et al. (2018) also point out that the going concern issuance rate of auditors' home states is associated with the probability of auditors issuing going concern opinions. Their results indicate that non-Big 4 auditors are more likely to issue going concern opinions if they are located in the high-going-concern-issuance-rate states. More importantly, they point out that this higher propensity of issuing going concern opinions increases both Type I and Type II errors. Their findings suggest that non-Big 4 auditors are more likely to issue false positive or false negative going concern opinions if they are based on the states that have high going concern issuance rates. Moreover, Hardies, Vandenhaute, and Breesch (2018) focus on going concern accuracy of the auditors who service private firms, failing to find any differences in going concern accuracy between Big 4 and non-Big 4. Similarly, Yang, Simnett, and Carson (2022) do not find significant differences in going concern accuracy between Big 4 and non-Big 4 firms when auditing charities. These findings suggest that the increased litigation and reputation risks that public firm auditors face may be contributing to the higher error rates (particularly type 1 errors).

Prior research has also examined going concern accuracy through the lens of managerial ability, audit committee independence, non-audit services, and the global financial crisis. Berglund, Herrmann, and Lawson (2018) find that high managerial ability is associated with lower Type I going concern errors. Their findings suggest that auditors issue less going concern opinions for those clients that continuously operate in the subsequent year. Interestingly, the authors also find that higher managerial ability is associated with higher Type II errors, in which auditors mistakenly issue more “clean” (or unmodified) opinions for those clients that are under financial distresses in the subsequent year. Budisantoso et al. (2017) investigate the moderating effect of going concern opinions. They find that auditors are less likely to be downward switching if there are audit committees and if audit committees are independent when going concern opinions are accurate. Geiger et al. (2022) investigate and find that non-audit fees are positively associated with going concern accuracy due to the client-specific knowledge obtained from non-audit services. However, the authors fail to find evidence that industry experts can improve going concern accuracy. Sanoran (2018) provides evidence that the major economic financial crisis shock impacts the accuracy of auditors’ going concern opinions. Specifically, she finds that Type I errors are lower for the financial crisis period, but Type II errors do not change significantly by the financial crisis, suggesting an asymmetric impact of macroeconomic shocks on going concern accuracy. Bakke, Kubick, and Wilkins (2020) investigate and find that deferred tax valuation allowances are positively associated with going concern accuracy because of the signaling effects for financial distress.

A more recent stream of studies investigates whether advanced statistical analysis methods can improve going concern accuracy. Jan (2021) compares the power of machine learning algorithms to predict going concern opinions. By including financial and non-financial variables in the algorithms, the author finds that the classification and regression tree algorithm outperformance the rest with an accuracy rate of 95 %. Vasarhelyi et al. (2022) also compare the performance of different algorithms in predicting going-concern accuracy. They confirm that advanced machine learning algorithms can better reduce the error, noise, and bias in going concern opinions than the traditional linear regression algorithm. Gutierrez, Krupa, Minutti-Meza, and Vulcheva (2020) find that going concern opinions significantly increase the explanatory power of the default risk prediction model when being added as an indicator, which results in no increases on going concern Type I and Type II errors. However, those studies still focus exclusively on using financial ratios as the input for machine learning algorithms to test their research questions regarding going concern accuracy. My study is different because it elaborates on the advantages of machine learning algorithms in investigating natural language processes and provides insights into how machine learning can benefit the public by analyzing textual information in financial statements. My approach is particularly relevant in light of the recent guidance by the PCAOB as both Risk and MD&A disclosures are likely to provide a rich environment to capture idiosyncratic variation in specific events and conditions that may signal going concern doubts.

In addition to investigating how to improve going concern opinions' accuracy, researchers have tested the impacts of going concern errors. Xu and Kalelkar (2020) focus on the office-level consequences of issuing incorrect going concern opinions. They find audit offices with higher going concern errors are more likely to experience market share reduction and increased dismissal rates. These negative consequences of going concern errors are mainly driven by Type I errors. Ahn and Jensen (2017) also test the consequences of going concern opinion errors on auditor conservatism, which is measured by abnormal accruals, the propensity of issuing going concern opinions, and subsequent going concern opinion issuance. They provide evidence of a negative association between Type I errors and auditor conservatism and a positive association between Type II errors and auditor conservation. These studies provide empirical economic support for why auditors should care about improving going concern opinion accuracy.

Dictionary-Based Textual Analyses

Dictionary-based textual analyses have been used in accounting literature for decades. Some of the dictionary-based measurements focus on the linguistic attributes of disclosures, such as readability, specificity, stickiness, hardness, boilerplate, and conciseness. Other dictionary-based measurements create their keyword lists (bag of words) to measure firm characteristics, such as the level of competition, different types of strategies, and the level of collaboration. Moreover, some other studies in textual analysis literature investigate the impact of positive/negative tones. Those dictionaries of linguistic attributes, firm characteristics, and the tone of disclosures are publicly available (Guay, Samuels, and Taylor 2016; Henry and Leone 2016; Loughran and McDonald 2011; Loughran and McDonald 2015; Henry 2006; Lewis and Young 2019). My study compares the effectiveness of traditional textual analysis attributes (dictionary-based attributes) with machine-learning-based textual proxies. Consequently, I provide the following review of how textual analysis in the accounting domain has evolved over time.

Linguistic attribute

Readability is one of the more common linguistic attributes investigated in the accounting and finance disciplines. Miller (2010) measures it by the Fog index and the length of documents.[5] This measurement provides a score to proxy for how difficult a textual document is for humans to read. Studies in accounting and financial literature have investigated how the

[5] Fog Index Score is calculated by 0.4 [(words/sentences) + 100 (complex words/words)] and the length of documents is measured by the logarithm of the total words in a document.

readability of financial disclosures can impact a variety of factors. Bonsall and Miller (2017) focus on the impact of disclosure readability on the bond market, documenting a positive association with favorable ratings, bond rating agency disagreement, and cost of debt. Kim, Wang, and Zhang (2019) document a negative association between the level of readability and the likelihood of crash risks. Further, their results suggest that managers hide some negative information through hard-to-read annual reports, which results in a high probability of future stock price crashes.

Researchers have also examined the effect of the readability of financial disclosures on external factors, such as the financial analyst environment and the economic growth of countries. Bozanic and Thevenot (2015) show that disclosures with high levels of readability and similarity, and low levels of uniqueness are associated with less information uncertainty. Lu, Qiao, Tan, and Yao (2021) investigate whether the transparency level of the clients in a country is associated with the economic growth of that area. They measure transparency using the readability, specificity, boilerplate, and conciseness of the clients in a specific country, and find a positive association between the local clients' transparency and economic growth.

Readability can be measured using the full text of various documents or based on a specific section of a filing, often yielding different results. For example, Campbell, Chen, Dhaliwal, Lu, and Steele (2013) analyze the content of risk disclosure sections in 10-K filings, counting the keywords in the Risk and MD&A sections to document that high-risk firms disclose more information in annual reports. Gan and Qiu (2020) measure the size of 10-K filings by counting the total number of words in the entire document as well as counting the total number of words for specific sections in 10-K filings. They provide evidence to support that the word count of risk factor sections in 10-K filings is not associated with firms' future stock returns, but the overall word count of 10-K filings is. Li, No, and Wang (2018) focus on testing the impact of cybersecurity disclosure on cybersecurity risks. They illustrate that the presence of cybersecurity disclosure and the length of cybersecurity disclosures are positively associated with the likelihood of the disclosed firms' reporting a breach in the subsequent year. Clarkson, Ponn, Richardson, Rudzicz, Tsang, and Wang (2020) analyze the corporate social responsibility report and find that the total number of words and the total number of sentences can predict corporate social responsibility performance, with an average accuracy rate of 81%. Melloni, Caglio, and Perego (2017) document a negative association between firms' financial performance and the length/readability of their disclosures. They also find a negative association between firms' social responsibility and the readability of their disclosures.

The level of hard information (numeric information), comparability, boilerplate, and Plain English has also been measured and explored by studies in the textual analysis literature. Lang and Lawrence (2015) examine the textual disclosures of non-U.S. firms. They find positive associations between the numeric information of disclosures/the comparability of disclosures and firms' liquidity, financial analyst coverage, and institutional ownership. Campbell, Zheng, and Zhou (2021) measure the proportion of numerical disclosure in conference calls and find that stock returns are higher when conference calls include more numerical information. The authors also find this positive association is stronger when the information environment is poor and when performance uncertainty is high. Loughran and McDonald (2014) find evidence to support that the SEC "Plain English" policy significantly impacts linguistic features of the disclosures, which firms issue more plain textual information in their disclosures before the publishment of the rule than before.

Bag of words

The previous section summarizes the studies that examine how linguistic attributes of textual disclosures are associated with various capital market-based variables. A separate stream of textual analysis literature creates dictionaries to measure firm characteristics. These "bag of words" dictionaries measure various financial attributes in disclosures by searching documents (or specific disclosures within a document) for pre-specified keywords.

Creating a dictionary to measure firms' financial performance is a popular output from the "bag of words" studies. Bodnaruk, Loughran, and Mcdonald (2015) create a dictionary to proxy the financial constraints of public firms by analyzing their financial disclosures. They show that this new proxy is better than the previous-used measurements because of its advancement in predicting future liquidity events. Banker, Huang, Li, and Yan (2021) develop a separate "bag of words" to measure three different types of strategies - product leadership, customer intimacy, and operational excellence. They then calculate the extent that firms have implemented these three strategies based on Item 1 sections in 10-K filings and find that cost rigidity is higher for product-leader firms than the firms that use the other two strategies. Balakrishnan and Darendeli (2020) use the bag of words approach to create a new measurement for localized competition by using the geographical operation-related information disclosed in 10-K filings. They also verify the measurement by showing an association between their measure and profitability reversion. Cazier and Pfeiffer (2017) create a proxy to measure the repetition level of 10-K disclosures based on N-grams. They find that managers strategically repeat textual information in financial disclosures that is not informative for investors. Chen, Francis, Hasan, and Wu (2021) generate a bag of word technique to build a dictionary, which includes collaboration-related keywords, to measure the collaboration level of firms by their 10-K disclosures. They then investigate and find that auditors indeed incorporate the effect of collaboration culture into their audit fees. Fassas, Bellos, and Kladakis (2021) implement textual analysis to identify the keyword differences between and after the Covid-19 pandemic, which measures the factors raised due to the pandemic. They find that firms pay attention to the negative effect of this shock on supply chain disruption, liquidity risks, and economic recessions.

Bauer and Klassen (2017) also use the bag of words approach to develop a proxy related to the uncertain tax benefit liability. Their proxy calculates the similarity between a specific textual disclosure and a list of unfavorable-outcome tax-related keywords. The authors find that investors react negatively to unfavorable outcomes calculated by their dictionary. Allen, O'Leary, Qu, and Swenson (2021) propose a new dictionary for tax-related textual research. They provide a tax-specific dictionary that is being consulted with experts in the tax domain. By counting the total words captured by this tax-specific dictionary, the authors conclude that this new bag of words can capture more tax-specific information than the previous one.

Cyber security is another characteristic where the bag of words approach has been used. Jeyaraj, Zadeh, and Sethi (2020) create measurements to identify different types of cyber threats and cyber responses based on the 10-K filings. They group cyber threats into physical threats, personnel threats, communication and data threats, and operational threats; and they group cyber responses into general responses, technical responses, and non-technical responses. Berkman, Jona, Lee, and Soderstrom (2018) focus on cybersecurity disclosure in 10-K filings. They develop a new proxy called the cybersecurity awareness measurement, which provides a numeric measurement for the extent of cybersecurity disclosures in 10-K filings. The authors also point out that investors have positive reactions or firms with higher cybersecurity awareness.

Although most of the “bag of words” dictionaries are measured based on the textual information in 10-K filings, some “bag of words” dictionaries are created by other text sources. For example, Brochet, Loumioti, and Serafeim (2015) create a new proxy for short-termism by capturing time-related keywords that are present in conference calls. Their results suggest that the time horizon in disclosure is relevant for capital market pressures and the financial incentives of top managers, as well as earnings management. Hu, Shohfi, and Wang (2020) investigate the textual information in merger & acquisition conference calls, creating a dictionary to proxy the financial and strategic motivations of the merger. The empirical findings of this study show that the firms that disclose more financial-related keywords in their merger & acquisition conference calls exhibit better capital market reactions than the firms that disclose more strategic-related keywords.

Tone analysis

Tone analysis is a specific type of dictionary-based textual analysis, which counts the positive/negative words and identifies the positive/negative/neutral tone in textual documents. In accounting literature, the dictionaries from Loughran and McDonald (2015) and the Diction software developed by Roderick Hart are widely implemented.

Some of the studies investigate the tone in SEC filings. Cho, Roberts, and Patten (2010) focus on the tone of firms’ environmental disclosures, finding that firms with bad environmental performance express more optimism in environmental disclosure than firms with good environmental performance. Similarly, Du and Yu (2021) provide evidence of the predictive power of tone in corporate social responsibility reports (CSR reports), revealing that the positive tone of current-year CSR reports indicates better CSR performance in the subsequent years. Huang, Krishnan, and Lin (2018) examine the correlation between the tone in press releases and firms' earnings management activities and conclude that firms strategically manage their tone in earnings press releases to be more positive when they have higher levels of discretionary accruals or their earnings just meet or just analyst forecasts. Wang (2021) discusses the effect of tone on debt markets, testing whether the negative tone in MD&A disclosures signals risks and uncertainty to investors and whether investors react to those negative tones. The paper documents a positive association between negative tone in MD&A sections and short-window credit default swaps spread. Jiang, Pittman, and Saffar (2019) test the impact of political uncertainty on the tone of 10-K disclosures, finding that negative tone is higher when the firm experiences more political uncertainty. Finally, Katsafados, Androutsopoulos, Chalkidis, Fergadiotis, Leledakis, and Yrgiotakis (2021) investigate the impact of 10-K disclosure tone on the likelihood of becoming a bidder in the bank industry, and find that banks with positive tones are more likely to become bidders in merger and acquisition activities.

Other studies explore tone using a variety of sources to develop their measures. For example, Huang, Roberts, and Tan (2018) explore the tone of media coverage. They calculate the tone of the articles in the Wall Street Journal, the Washington Post, the New York Times, and USA Today and find that negative tone is associated with mitigated CEO power. Amoozegar, Berger, Cao, and Pukthuanthong (2019) focus on whether different types of institutional investors impact the tone in conference calls. They illustrate that short-term (long-term) institutional investors are positively associated with positive (negative) tones in conference calls. Breuer and Ghufran (2021) investigate the effect of disclosure tone on merger and acquisition performance. By measuring the positive and negative tone of acquirers' MD&A, the authors find that the positive (positive) tone from acquirers is related to positive (negative) post-takeover performance. In addition, Koelbl (2020) tests the predictive power of tone regarding future performance in the US Real Estate Investment Trust, finding more optimistic tone in the MD&A sections of 10-K filings increases the probability that the firms experience better future performance. Moreover, Mayew, Sethuraman, and Venkatachalam (2015) explore the predictive power of tone in an auditing setting, specifically how the tone of MD&A sections predicts bankruptcy incidents in the subsequent period. They find that a positive tone in the MD&A is negatively associated with the probability that the firms will file for bankruptcy protection in the next period and vice versa. Cheng, Smith, and Tanyi (2018) explore how the leadership structure affects the tone of leadership justification disclosures. They find that CEO duality structured disclosure has higher informativeness compared to split structured disclosure.

There are some additional studies that investigate the determination of tone in disclosures. Brochet, Miller, Naranjo, and Yu (2019) investigate the effect of managers' culture on the tone in conference calls and whether financial analysts differentially interpret information based on distinct cultural backgrounds. The authors find that executives who have individualistic cultural backgrounds express a more optimistic tone in conference calls. In addition, financial analysts capture the optimistic tone in their forecasts if the cultural backgrounds are highlighted to analysts. In addition, Bassyouny, Abdelfattah, and Tao (2020) explore the determination of using different tones (positive/negative tone) in disclosures. Specifically, they show that a positive (negative) tone is positively (negatively) associated with narcissistic CEOs (age, gender, financial expertise level, and the independence of audit committees). Liu and Nguyen (2020) also investigate the area of CEOs and their tones in 10-K filings and CEO letters, finding that while CEO letters have more positive words than MD&A sections for the same company, the tone used in CEO letters cannot predict future performance. Craig and Amernic (2018) find that hubris CEOs are more likely to have high levels of realism in their letters to shareholders. Swift, Colon, and Davis (2020) test how cyber incidents impact the tone of 10-K filings. However, they do not find significant results to support that the tone of MD&A disclosures changes after a cyber incident. Cazier and Pfeiffer (2016) examine the determinations of 10-K length, finding that operational complexity and disclosure redundancy are the key factors driving different lengths of 10-K filings. Finally, Elsayed and Elshadidy (2021) investigate the effect of internal control quality on managers' risk disclosure choices, finding that managers with high-quality internal controls disclose more risk factors compared to those with low-quality internal controls.

Another stream of studies in tone analysis literature investigates the determinants or impacts of abnormal tone from textual disclosures. Davis, Ge, Matsumoto, and Zhang (2015) document the existence of a residual part of tone in conference calls that cannot be explained by firm performance, and is associated with managers' characteristics, such as the length of their tenure. Huang, Teoh, and Zhang (2014) investigate whether firms strategically manage their tone in earnings press releases. They point out that an abnormal positive (negative) tone is positively (negatively) associated with upward (downward) perception management, and that this abnormal tone is related to negative future performance. D'augusta and Deangelis (2020) study the impact of accounting conservatism on tone management, finding that managers with higher levels of accounting conservatism are less likely to use upward tone management in the MD&A section of 10-K filings. Osma, Grande-Herrera, and Saorin (2018) document a positive association between CEOs' positive tone and higher abnormal returns, more future debt, and more capital investments as well as more dividends of their firms.

Moreover, other studies focus on tone dispersion and tone changes over time in textual disclosures. Allee and DeAngelis (2015) investigate the factors that affect the tone dispersion in companies' disclosures and find that those factors include firms' current/future performance, expectation management, and financial reporting choices. After demonstrating the determination of tone dispersion, the authors further test how analysts and investors react to tone dispersion in firm disclosures. They observe more positive questions during conference calls if the positive tone is more dispersed and vice versa. Also, investors are affected by the tone of the questions that analysts ask during the calls. Bochkay, Chychyla, and Nanda (2019) examine how the tone of CEOs changes over time; specifically, whether the level of forward-looking discussion and optimism/ pessimism in quarterly conference calls are different among their tenure. They find that CEOs discuss less forward-looking information and become less optimistic as their tenure increases. They then show that non-CEO executives do not follow this pattern. Feldman, Govindaraj, Livnat, and Segal (2010) also investigate the tone changes in MD&A sections of 10-K and 10-Q filings and support the notion that the capital market reacts significantly to the tone changes of filings after control firms' operational performance in a short window. Efretuei (2021) documents a positive association between disclosure readability and negative tone. This finding suggests that firms are more likely to make their disclosures harder to read if those disclosures have negative tones. Besides the positive and negative tone, some studies in the textual analysis literature investigate the impact of strong and weak tones on some financial and audit attributes. Finally, Liu and Moffitt (2016) find that the likelihood of restatements is positively associated with how strong the tone is in the firms' SEC comment letters.

Machine Learning in Textual Analysis

As previously discussed, dictionary-based textual analysis represents the predominate method used by prior research to analyze qualitative information in financial disclosures, requiring researchers to search and count keywords in disclosures based on published dictionaries in the literature (Loughran and Mcdonald 2011; Loughran and Mcdonald 2015;

Loughran and Mcdonald 2016). However, recent studies have pointed out the limitation of this approach, which publicly available dictionaries may not include sufficient amounts of keywords regarding a specific concept (Guo, Shi, and Tu 2016; Frankel, Jennings, and Lee 2021).[6] This has led some more recent researchers to introduce innovative machine learning approaches to analyze the textual information in SEC filings, news, conference call, etc. These approaches include but are not limited to Naïve Bayes Classification, Latent Dirichlet Allocation, and Supporting Vector Machine. Previous studies provide evidence that machine learning approaches have higher explanatory power than dictionary-based textual analysis. These studies also verify the reliability and feasibility of machine learning approaches (Guo et al. 2016; Frankel et al. 2021; Lewis and Young 2019). I summarize the machine learning in textual analysis literature by categorizing the studies based on the machine learning method each study uses.

Naïve bayes classification

Guo et al. (2016) and Lewis and Young (2019) summarize four machine learning techniques relevant to accounting and financial research. First, they introduce Naïve Bayes Classification, where the authors download news from the News Analytics database which provides the text of news articles with sentiment scores to indicate their level of positivity (negativity). The authors then randomly choose 3000 positive news articles as the training sample to "teach" the machine learning algorithm how to identify positive news, and then select

[6] Besides textual analyses, machine learning approaches are also used to improve model performance. This stream of studies focuses on testing whether machine learning algorithms can outperform linear regressions in analyzing variouscapital market or audit outcomes. The findings show that machine learning algorithm has superior performance in predicting future earnings, abnormal returns and discrete outcomes by using traditional financial ratios (Chen, Cho, Dou, and Lev 2022; Hunt, Myers, and Myers 2022; Krupa and Minutti-Meza 2021; Amel-Zadeh, Calliess, Kaiser, and Roberts 2020). Other studies provide evidence of machine learning algorithms' outperformance in predicting auditor switches, misstatements, lending decisions, insurance payments, asset pricing, and audit quality (Hunt, Rosser, and Rowe 2021; Bertomeu, Cheynel, Floyd, and Pan 2021; Liu 2022; Ding, Lev, Pend, Sun, and Vasarhelyi 2020; Gu, Kelly, and Xiu 2020; Hu, Sun, Vasarhelyi, and Zhang 2022).

500 news articles at random (positive and negative) to verify the accuracy of the algorithm. Compared to the dictionary-based textual analyses, Naïve Bayes Classification outperforms because 1) it assigns weight to each keyword based on the frequency of a single word appearing in a training sample; and 2) it recognizes the positive/negative word based on context (an "*increase*" of costs is recognized as negative in Naïve Bayes Classification but recognized as positive in dictionary-based textual analyses). Therefore, this technique is suitable for situations where there is no existing dictionary.

Li (2010) uses Naïve Bayes Classification to measure the tone of forward-looking sentences in the MD&A sections. She documents that a positive tone in the current year is associated with better future earnings. Huang, Zang, and Zheng (2014) measure the sentiment positivity/negativity in analyst reports, and find that the earnings growth rates are affected positively when more positive sentiments are present in analyst reports, and this effect persists for multiple years. Buehlmaier and Whited (2014) create a proxy for financial constraints based on textual information in annual reports and find that firms with less financial flexibility often experience higher stock returns.

Latent dirichlet allocation (lda)

Guo et al. (2016) and Lewis and Young (2019) provide illustrations of LDA. LDA is a popular unsupervised machine learning approach to reveal the underlying latent topics of documents by using word and topic distributions. It begins with extracting the words of all the documents and transferring a text corpus into word-frequency matrices. Then, LDA reduces the textual data dimension and weights the terms of each topic. In the final stage, this technique produces the word matrices for all topics and the researchers "label" topics based on the high-frequency words in each latent topic.

My study utilizes LDA to summarize the latent topics appearing in the qualitative 10-K disclosures of firms receiving inaccurate going concern opinions. I assume that those latent topics can be used to represent the conditions and events that impact going concern accuracy.

Huang, Lehavy, Zang, and Zheng (2018) utilize the Latent Dirichlet Allocation approach to identify the topics from conference calls and analyst reports separately. By isolating the topics from analyst reports, the authors demonstrate that financial analysts offer additional and meaningful information in their analyst reports, and those insights are not specified in the conference calls. Dyer, Lang, and Lawrence (2017) investigate the evolution of 10-k disclosure length, which specifically tests how the topics change over time. They use LDA to identify the overall topics in 10-K disclosures and find that new FASB and SEC requirements are the key events to trigger the increase of 10-K length over time. Specifically, the authors point out that fair value, internal controls, and risk factors account for the majority of the increase in disclosure length in 10-Ks. They also reveal boilerplate, stickiness, and redundancy are increased, and specificity, readability, and hardness have decreased over their sample period.

Other studies use the LDA method to classify the common topics disclosed in the risk factor of 10-K filings. Agarwal, Gupta, and Israelsen (2017) examine the effect of the Jumpstart Our Business Startups (JOBS) Act on the accounting information and textual disclosures of IPO firms. JOBS Act releases the mandatory disclosure requirements regarding hard accounting information by IPO firms. By using this Act as a natural experiment, the authors find that firms are more likely to reduce accounting information disclosures and are more likely to disclose more risk-related textual information regarding payouts after JOBS Act. However, the authors highlight that only the textual disclosure changes are associated with IPO underpricing. Cheong, Yoon, Cho, and No (2021) provide a classification mechanism for cybersecurity risk disclosures.

By utilizing LDA, they separate cybersecurity disclosure into nine topics, which include incident control and risk mitigation, operational risk, customer-related risk, etc. The authors also reveal that cybersecurity incidents and internal control weaknesses affect how companies decide to disclose the topics regarding incident control and risk mitigation, operational risk, business continuity, and third-party software providers.

Bao and Datta (2014) propose a new LDA approach for textual analyses in risk disclosures. The new method, which is called Sent-LDA, takes into account the impact of sentence structure in assigning topics among contexts. Specifically, sent-LDA assumes that there is only one risk topic in each sentence that is disclosed in the risk section. By using this innovative unsupervised machine learning method, the authors label 30 topics that are commonly disclosed by firms in item 1A of 10-K filings, such as cost risks, debt risks, credit risks, and human resource risks. Then, the authors utilize the results from sent-LDA and test how different risk topics can be recognized differently by investors. They document that only two-thirds of the risk topics are informative for investors which affects the stock return volatility. Among the informative topics, only the systematic and liquidity risks enhance investors' risk perception.

Supporting vector machine

Guo et al. (2016) highlight the advantage of using Supporting Vector Regression in accounting and finance research. Compared to other approaches, Supporting Vector Machine can handle an unbalanced dataset and will not introduce excess noise to the classification results. In addition, this method can map data into three dimensions and process non-linear data points. Specifically, the authors separate firm news into positive and negative groups, then they create a Supporting Vector Machine algorithm to create word vectors from the sentences in the two categories. Supporting Vector Machine technique uses the training sample to "learn" what the distinguishing features between positive news and negative news are. The program then generates a *hyper-plane* for separating observations, where the *hyper-plane* maximizes the classification margin between the two data dots that are closest to the *hyper-plane* in each classification group. More importantly, the Supporting Vector Machine approach can generate a proxy to recognize the out-of-sample news (observations that are not used in the training sample) into either positive ones or negative ones. Donovan, Jennings, Koharki, and Lee (2021) implement three machine learning approaches (Random Forecast Tree, Support Vector Regression, and supervised Latent Dirichlet Allocation) to create a new proxy to measure credit risk. They first use the three techniques to identify the quantitative information in the 10-K disclosures that captures credit risk in the training sample. They also verify the accuracy of this quantitative-information-based credit risk proxy and document the increased explanatory power of this new proxy compared to the popular Z-score. They find that their proxy has a better ability to measure the within-firm variation of credit risk. Frankel et al. (2016) use a support vector machine to proxy accruals based on the textual information disclosure in each 10-K filing section. They point out that the textual-based accruals proxy generated by the information in MD&A sections has the highest predictive power compared to the proxies generated by the texts from other sections. Manela and Moreira (2017) use a Support Vector Machine to create a new proxy (News Implied Volatility) for uncertainty based on the news printed in The Wall Street Journal. They find that high levels of News Implied Volatility are followed by abnormal stock returns.

Other machine learning approaches

Lewis and Young (2019) explain the mechanism of Cosine Similarity, which converts documents into word distribution vectors and then calculates the cosine of the angle between two documents. The range of Cosine Similarity is between 0 and 1. Therefore, if two documents are similar, their Cosine Similarity is close to 1 and vice versa. Florackis, Louca, Michaely, and Weber (2022) use the Cosine Similarity approach to proxy cybersecurity risks based on the context in Item 1A. They use the Item 1A disclosures of breached companies as the “targets”, then calculate the similarity score of Item 1As between those targets and non-breached firms. Those non-breached firms are then proxied as having high cybersecurity risks if Item 1A disclosures of those non-breached firms are mathematically “similar” to that of the breached firms. The authors also verify their cybersecurity risk proxy by demonstrating high-cyber-risk firms have more cyber disclosures in their 10-K filings.

Random Forest Classification is another machine learning approach that can be used to create proxies by imported textual information. This approach utilizes decision tree functions to create a “forest” by repeatedly selecting input and creating fitted trees to predict the target variable. Donovan et al. (2021) build a textual-based credit risk proxy by Random Forecast Tree as one of the three machine learning approaches being used in the study. They demonstrate the outperformance of using Random Forest Tree in measuring the credit risk compared to using the traditional credit risk measurement.

CHAPTER THREE: MACHINE LEARNING DESIGN

Data Preparation

I extract 149,450 10-K filings between 2006 to 2022 from the SEC EDGAR dataset and delete 1,190 duplicate observations based on CIK and fiscal year.[7] When merge with Compustat and Audit Analytics, 71,015 observations are excluded from the sample due to missing unique Identifiers. I then limit the sample to only include observations that are experiencing financial distress in the current year, which are defined as clients with negative operating cash flows or negative income before extraordinary items (Blay et al. 2016; Berglund et al. 2018; Gutierrez et al. 2020). This step provides me with a sample of clients experiencing similar financial difficulties and thus at a heightened risk of bankruptcy. This also allows me to eliminate bankruptcies that are driven by unique situations, such as lawsuits or labor negotiations (Geiger, Raghunandan, Riccardi 2014). This step results in the elimination of 44,101 non-financial distress clients from the sample. I also delete 4,573 observations that belong to financial industries. The remaining observations are then merged with the dataset of bankruptcy obtained from Audit Analytics, in which the bankruptcy filing date needs to be within the current fiscal year end and subsequent fiscal year ends (Casterella, Desir, Stallings, and Wainberg 2020).

My sample focuses on “first-time” going concern opinions, for which I delete 6,024 observations that receive consecutive going concern opinions in both the current year (t) and the prior year (t-1). An additional 2,920 observations are also eliminated because they have missing

[7] The SEC start to require registered firms to disclose risk factors in Item1A sections at 2006.

prior year going concern opinions (Geiger and Rama 2003; Geiger and Rama 2006; Blay et al. 2016; Berglund et al. 2018). The reason for deleting those consecutive going concern opinions is that the "first-time" going concern opinions are issued by auditors with more risks and difficulties and following a separate decision model than the subsequent going concern opinions (Mutchler and Williams 1990; Geiger and Rama 2003; Geiger and Rama 2006). Type I error firm-year observations are clients that receive going concern opinions but do not file for bankruptcy protection within 12 months after the current fiscal year-end dates. Type II error firm-year observations are clients that do not receive going concern opinions but file for bankruptcy protection within 12 months after the current fiscal year end dates (Gutierrez et al. 2020; Berglund et al. 2018). Accurate Going Concern firm-year observations are clients that receive going concern opinions and file for bankruptcy protection within 12 months after the current fiscal year end dates. There is another type of clients that do not receive going concern opinions in the current year and do not file bankruptcy protection within 12 months after the fiscal year end dates.

In order to analyze the textual information of Item 1A (Risk Factor) and Item 7 (MD&A) by machine learning algorithms, I follow Dyer et al. (2017) to extract those two sections from the 10-K filings and clean the texts to make sure that they are ready to be imported into the LDA algorithm. First, I download the 10-Ks filed with EDGAR from 2006 to 2022 using a Python algorithm which includes the links of those filings for all clients. Then, I use the Python algorithm to create a meta-dataset that shows each firm-year 10-K filing as a row and includes the CIK, Company Name, Type of Filing, Date of Filing, HTML Link, Fiscal Period End, SIC Code, State of the Company, Fiscal Year, and Filename. This meta-dataset is created to link the textual information with their identifiers. Next, I download the raw 10-K filings from EDGAR based on the links in the meta-dataset. Then, I sparse Item 1A section and Item 7 sections from the individual raw filing; this procedure results in multiple JSON files that every file includes the footnotes of Item 1A and Item 7 as well as their unique identifiers from the meta-datasets (each type of information is quoted by a quotation mark and separated by a comma). This step also includes a cleaning procedure for the text, which removes all the graphics, tables, page numbers, line breaks, etc. All HTML tags are removed as well. I further set a threshold of the length of the textual information, which removes any sentences that have less than 20 characters or 15 alphanumeric characters and removes paragraphs that have less than 120 characters or have more than 50 percent of non-alphabetic characters. Deleting those sentences and paragraphs can eliminate section headers and any paragraphs that only deliver boilerplate language, such as “As a smaller reporting company we are not required to provide any information under this item”. Last, I load all the individual JSON files into a Stata file which has each firm-year observation as a row, and the textual information from Item 1A and Item 7 sections as two columns. Other financial ratios and audit-related information (obtained from Compustat and Audit Analytics) are also included as separate columns for each observation.

Table 1 summarizes the data preparation for machine learning algorithms in this study. Table 2 Panel A presents the full sample distribution matrix and the subsamples that are used in machine learning algorithms.[8] Top2Vec only uses a subsample at each test when it generates the common topics in order to examine the impact of those topics on going concern accuracy variations. Table 2 Panel B and C provide details about the subsamples used by Top2Vec in each setting. Table 3 Panel D presents the training sample distributions used in generating Naïve

[8] No going concern and no subsequent bankruptcy observations are in the upper left cell; no going concern but has subsequent bankruptcy observations are in the lower left cell; receive going concern but no subsequent bankruptcy observations are in the upper right cell; and receive going concern opinions and has subsequent bankruptcy observations are in the lower right cell.

Bayes Classification, Supporting Vector Machine, and Random Forest Classification based proxies, which require the full sample to be split into a training sample (70%) and a testing sample (30%). The training samples are used for fitting the machine learning models and the testing samples are used for empirical tests.

Using Machine Learning to Summarize the Common Topics in Financial Statements

My study intends to identify the risks and firm characteristics that impact going concern accuracy. If auditors misevaluate the impact of those risks and firm characteristics on clients' likelihood to continuously operate, it can result in false positive (Type I error) going concern opinions or false negative (Type II error) going concern opinions. I use a machine learning algorithm called Top2Vec, to summarize the topics in the observations of the samples, which will provide some "keyword" tables to describe the topics that are prevalently in my sample clients' financial statements. My sample of clients are all financially distressed but are ***not*** only limited to Going Concern Error clients.[9] Then, the algorithm analyzes and marks the topics that are included in the financial statement of each client.[10] I then estimate regressions (details provided in the next section) to test which specific topics are associated with incorrect going concern opinions.

Top2Vec is one of the machine learning approaches that classifies textual disclosures into latent topics by utilizing the *joint semantic* embedding of documents and words (Angelov 2020).

[9] For example, the observations used in Top2Vec for generating the topics in order to examine the effect of those topics on the probability of accurate going concern opinions (Model [1]) are either Type I Error firms or Accurate Going Concern Opinion firms. The observations used in Top2Vec for generating the topics in order to examine the effect of those topics on the probability of accurate going concern evaluations (Model [3]) are Type I Error firms, Accurate Going Concern Opinion firms, and firms that do not receive going concern opinions in the current period and do not file for bankruptcy protection in the subsequent year.
[10] I keep the paragraph separator in the financial statement to allow the machine learning algorithm to recognize each paragraph as an individual input for each firm.

Angelov (2020) proposed a distributed representations topic model to generate topic vectors. The distributed representation in neural networks can automatically learn concepts based on the involved neurons. It learns the concepts based on the notion that "words with similar meanings are used in similar contexts." (Angelov, 2020, p2) Top2Vec is firstly jointly embedding documents and words. In this step, the documents themselves and the words in each document are converted to an embedded multi-dimensional vector in which the words and documents are represented by numeric indicators. Toc2Vec utilizes either Word2Vec or Doc2Vec embedding methods for this process. In my study, I use Doc2Vec approach due to its outperformance in analyzing large datasets.

The embedding process requires calculating the distance between document vectors and word vectors. This distance represents the semantic similarity between different objects so that words/documents with small distances (high semantic similarity) are clustered closely. Any words that are far away from the others are considered outliers and are removed from the output. A hierarchical density-based clustering algorithm called HDBSCAN is included in Top2Vec for the cluster function. There are several documents and words in each cluster (semantic space) which have similar semantic meanings.

The topic assignment is based on the results of jointly embedded word/document vectors (clusters). Each cluster is considered to be an individual topic and the important keywords that represent the topic are recognized by the n-closest words to the centroid of the vector (cluster). Because the number of clusters equals the number of topics, Top2Vec algorithms can automatically provide the optimal number of topics regarding the input. Each input is by default considered to have one topic. Therefore, I separate the Item1a and Item7 contexts into paragraphs and import all the individual paragraphs into the machine learning algorithm. Then, Top2Vec conducts the joint embedding procedure, generates common topics based on all the inputs, and assigns the represented topic to each paragraph.[11] Top2Vec also incorporates a function called hierarchical topic reduction, in which programmers can set the number of topics that they would like the algorithm to generate. This hierarchical reduction is based on merging the smallest topic into the topic that has the highest semantic similarity. This procedure is repeated several times until reaching the number of topics that the programmer has specified.

Previous studies in accounting and finance disciplines use LDA as the topic modeling approach to generate latent topics in financial statements. However, Angelov (2020) and Egger and Yu (2022) point out that Top2Vec is a better algorithm compared to LDA in topic modeling for a variety of reasons. First, Top2Vec focuses on summarizing topics based on the document/word semantics, which provides more informative topics compared to LDA. LDA, on the other hand, focuses on the probability of word occurrences in documents, which does not analyze the semantic meaning and does not generate semantic topics. Second, Top2Vec can automatically generate the optimal number of topics for the input and has the function to reduce it hierarchically. LDA needs programmers separately to use a measurement called perplexity to identify how many topics should be created to describe the corpus effectively. Third, Top2Vec does not need programmers to manually create a stop-word list and eliminate the stop words based on the list, because it is unlikely to have a document that is semantically similar to any of the stop words. Finally, Stemming and Lemmatization are not required in Top2Vec. However, they are used prevalently in LDA.

[11] Appendix B has an example of topic assignment to each paragraph.

To utilize the Top2Vec for analyzing the topics in Item1A and Item7 textual information, I import the clean dataset into the algorithm and identify each paragraph using a paragraph separator, allowing each paragraph of the Item1A/Item7 disclosures to be assigned as an individual input. Therefore, the final corpus used by Top2Vec includes multiple clients (CIK and Fyear), and each client has multiple rows to represent its disclosed paragraphs individually. Then I run the Top2Vec modeling algorithm to generate the joint embedding of documents and words, as well as cluster the vectors. Due to the limitation of computing capacity and the degree of freedom issue in the empirical models, I cannot use the optimal number of topics generated from the algorithm automatically for the empirical tests, so I use the hierarchical reduction function and the number of topics for Item 1A disclosures (Item 7) disclosures is set to be 30 (21) based on previous studies (Bao and Datta 2014; Brown, Hinson, and Tucker 2021).

The algorithm then summarizes the common topics and provides the 50 most-important keywords in each topic for interpretation. In the end, the machine learning algorithms assign the corresponding topic to each paragraph (each row) and then I create an algorithm to aggregate those individual rows of a specific client into different columns with indicator variables (Item1a_Topic_N), to show which topics are presented in that client's financial statement. Those indicator variables are used for empirical tests in the next section.

Using Machine Learning to Create Going Concern Accuracy Proxy

This study also intends to create proxies by machine learning algorithms to measure going concern accuracy. I create three proxies to measure accurate going concern opinions, Type I going concern errors, and Type II going concern errors separately. To create those three proxies, I choose to use three different machine learning algorithms, which are Naïve Bayes Classification, Supporting Vector Machine, and Random Forest Classification. Those three algorithms are commonly used in classifying two groups of observations and creating probabilistic proxies.

Naïve Bayes Classification has been used in accounting and finance literature for an extended period (Li 2010; Guo et al. 2016; Lewis and Young 2019). This classification is created based on the Bayes Theorem, which calculates the conditional probability that an event occurs. When used in natural language processing studies, the machine is trained based on textual information, which in my study are the Item1A and Item7 disclosures. Each observation in the sample has a label of going concern accuracy (e.g., Accurate Going Concern = 1 or Accurate Going Concern = 0) which the Naïve Bayes Classification algorithms calculate the likelihood of each word occurring in each condition. Then the machine learning algorithm creates a proxy and calculates the probability of any given observations in the testing sample in those conditions.

Supporting Vector Machine is another popular machine learning algorithm to generate proxies for predictions (Hearst, Dumais, Osuna, Platt, and Scholkopf 1998; Noble 2006). Different from Naïve Bayes Classification, Supporting Vector Machine creates a hyperplane to best separate different groups and then extract the corresponding keyword on each side of the hyperplane. It requires a training sample as well with clear labels to identify the groups they belong to (e.g., Accurate Going Concern = 1 or Accurate Going Concern = 0). The machine learning algorithm loads the textual information in Item1A and Item7 disclosures and relies on this label in the training sample to build the hyperplane. The keywords within each side of the hyperplane represent the corresponding keywords for the specific group. After being trained, the machine learning algorithm can generate a proxy by using the keywords and their weights in each group for predicting the probability of an observation belonging to each group in the testing sample.

Random Forest Classification focuses on using ensembled decision trees to make predictions. Donovan et al. (2021) use this machine learning approach to create credit risk proxies based on clients' textual disclosures. The mechanism of Random Forest Classification is that the machine randomly chooses some observations and then creates decision trees for those observations based on some features (keywords). After repeating this procedure extensively, the machine uses ensemble learning to combine all the decision trees and then create a forest. That is Random Forest Classification and the trained forest should be able to make predictions for the observations in the testing sample. The prediction is generated by a majority voting of all the individual prediction results from each decision tree in the forest. Specifically in my study, Random Forest Classification loads the embedded keywords from Item1a and Item7 in the training sample and selects only a subset of the training sample to create a decision tree. When the first decision tree is completed, the machine learning algorithm creates the second decision tree by randomly selecting another subset in the training sample. This repeating process ends when the most efficient results are met. Then the algorithm conducts ensemble learning to combine the individual decision trees into a consolidated forest. The forest is then used for predicting the classification in the testing sample. Each observation in the testing sample has the tokenized words in their Item1A and Item7 disclosures and those "words" are processed by all the individual decision trees in the forest, which will generate multiple prediction results. Random Forest Classification summarizes the results and creates an aggregate prediction as the final prediction.

Before being imported into the machine learning algorithms, the textual information needs to be further cleaned for these three machine learning algorithms for generating proxies. The textual information of Item1A and Item7 is appended, and the consolidated context is loaded into the machine learning algorithms. The algorithms then remove any blank rows and convert the text into lowercase. Next, the algorithms remove all the high frequency stop words, such as "a", "the", and "or." This step improves the quality of our machine learning output, which makes sure that the output words are "meaningful" in order to represent the classification. Then, the algorithms conduct lemmatization to convert the words into their base roots.[12] After the textual information is cleaned, I split the sample by 70%/30%, of which 70% of the sample is used as the training sample and 30% of the sample is used as the testing sample. Both the training and testing sample needs to be vectorized into numeric presentations so the machine learning algorithms can analyze the data. Due to the imbalance classification in my training sample, I imply oversampling technique to randomly duplicate observations that belong to the underrepresented group.[13] This procedure can increase the validity and accuracy of the proxy since the features (keywords) in the underrepresented group might be treated as noises and excluded from the proxy model. The machine learning algorithm fits the balanced training sample to the three different classification approaches (Naïve Bayes Classification, Supporting Vector Machine, and Random Forest Classification) and creates the proxy models with important keywords and their weights. Once the proxy models are trained, the algorithm loads the testing sample and calculates the prediction probability based on the proxy models. These predicted probabilities are used for the empirical studies in the next section.

[12] For example, "reports" is lemmatized as "report".

[13] The samples are imbalanced, which the majority of the observations belong to the group that does not receive going concern opinions and does not file for bankruptcy protection. Table 2 presents the sample distribution matrix. For example, to create the accurate going concern opinion proxy, oversampling technique randomly choose some observations in the accurate going concern opinion =1 group at the training sample and duplicate those observations. Therefore, the sample becomes balanced which the accurate going concern opinion =1 group should have the same number of observations as the accurate going concern opinion = 0 group.

Table 1 Data Preparation for Machine Learning Algorithms

10-K filings from SEC EDGAR between 2006 to 2022:	149,450
Less: duplications	(1,190)
Less: observations that their CIK and Fiscal Year values are missing in Compustat	(71,015)
Less: non-financial distressed observations	(44,101)
Less: observations that belong to the financial industry	(4,573)
Less: observations that receive consecutive going concern opinions in both year t and year t-1 or miss going concern opinions in year t-1	(8,944)
10-K filings that can be loaded into the machine learning algorithms	19,627

Table 1 presents the data preparation process for the machine learning algorithms. Eventually, there are 19,627 observations, as the full sample, that are cleaned to be used in the machine learning algorithms.

Table 2 Sample Distribution Matrix for Machine Learning Algorithms

Panel A Full Sample Distribution

	Going Concern$_t$=0	Going Concern$_t$=1
Bankruptcy$_{t+1}$=0	18,074	1,225
Bankruptcy$_{t+1}$=1	131	197

Panel B Subsample used in Machine learning algorithms for Topics in Model [1] -Model [4]:

Accurate Going Concern Opinion in the Type I Error Setting

	Going Concern$_t$=0	Going Concern$_t$=1
Bankruptcy$_{t+1}$=0		1,225
Bankruptcy$_{t+1}$=1		197

Accurate Going Concern Evaluation in the Type I Error Setting

	Going Concern$_t$=0	Going Concern$_t$=1
Bankruptcy$_{t+1}$=0	2,611	1,225
Bankruptcy$_{t+1}$=1		32

Panel C Subsample used in Machine learning algorithms for Topics in Model [5] -Model [8]:

Accurate Going Concern Opinion in the Type II Error Setting

	Going Concern$_t$=0	Going Concern$_t$=1
Bankruptcy$_{t+1}$=0		
Bankruptcy$_{t+1}$=1	131	197

Table 2 (Continued**)**

Accurate Going Concern Evaluation in the Type II Error Setting

	Going $Concern_t$=0	Going $Concern_t$=1
$Bankruptcy_{t+1}$=0	360	
$Bankruptcy_{t+1}$=1	131	7

Panel D Training Sample in Creating Proxies for Model [9] – Model [11]:

Accurate Going Concern Opinion Proxy

	Going $Concern_t$=0	Going $Concern_t$=1
$Bankruptcy_{t+1}$=0	12,657	852
$Bankruptcy_{t+1}$=1	91	138

Type I Error Proxy

	Going $Concern_t$=0	Going $Concern_t$=1
$Bankruptcy_{t+1}$=0	12,641	862
$Bankruptcy_{t+1}$=1	87	148

Type II Error Proxy

	Going $Concern_t$=0	Going $Concern_t$=1
$Bankruptcy_{t+1}$=0	12,670	847
$Bankruptcy_{t+1}$=1	90	131

Table 2 panel A presents the sample distribution matrix of the cleaned disclosures for machine learning algorithms, which the total number of observations is 19,627. Panel B provides the sample distribution matrix for the subsamples used in Top2Vec in order to test the impact of the topics on accurate going concern opinion/evaluation in the Type I error setting. Panel C provides the sample distribution matrix for the subsamples used in Top2Vec in order to test the impact of the topics on accurate going concern opinion/evaluation in the Type II error setting. Panel D shows the sample matrix for the original training sample that is slipped as 70 % of the observations. This is the sample distribution before implying the oversampling approach.

CHAPTER FOUR: EMPIRICAL STUDY

In this section, I investigate which topics generated by the machine learning algorithms are empirically associated with going concern accuracy variation. This empirical evidence enhances the economic contributions of those topics by documenting the correlations between the topics and the probability of accurate going concern opinions/evaluations. In addition, I test whether the going concern accuracy proxies generated by the machine learning algorithms are associated with the likelihood of accurate going concern opinion, Type I going concern error, and Type I going concern errors. These tests prove the validity and effectiveness of machine-learning-based going concern accuracy proxies.

Research Questions

The textual disclosures in Item 1A (Risk Factor) and Item 7 (MD&A) sections of 10-K filings present the information regarding firms' contingencies and performance to the public (Kravet and Muslu, 2011). Auditors are required to evaluate the conditions and events that can bring sufficient doubts about clients' abilities to continue as a going concern, and the topics generated by machine learning in clients' Item 1A and Item 7 sections in their 10-K filings can be used to represent the conditions and events (e.g. risks and performance) that auditors evaluate during the auditing procedures for issuing going concern opinions (Mayew et al. 2015). However, which specific topics generated from the machine learning algorithms will be systematically different between accurate going concern opinions/evaluations and Type I/II errors remains an empirical question. The purpose of this study is to explore the nature of the disclosure topics that precede these different outcomes so that auditors can use this information to help reduce Type I and Type II error rates.

This study also creates proxies to measure the probability of accurate going concern opinions, Type I going concern errors, and Type II going concern errors based on the current-year textual disclosures in Item 1A and Item 7. If effective, these proxies could help investors and other 10K users to make more informed decisions when assessing bankruptcy risks of public companies. However, whether these proxies can effectively improve current going concern accuracy prediction models, and do so better than other text-based measures already used in the literature (e.g. readability, tone, etc.) is also an empirical question.

Research Design

Data description for empirical testing

I begin with the previously discussed sample which has 19,627 observations between 2006 to 2022 for the empirical tests. All of the clients in my sample are experiencing financial distress, which is measured by negative operating cash flows or negative income before extraordinary items. I use the full sample of financially distressed clients with no missing going concern opinions and no missing future bankruptcy data to generate my machine learning algorithm-based proxies. I do not exclude observations that have missing control variables to calculate these proxies. This decision was made because these observations can still contribute to the training process in my machine learning algorithms so long as I am able to correctly label their going concern accuracy. The performance of the training process is also enhanced by importing more textual information into the algorithms.[14]

[14] The observations with missing control variables are eliminated from my regression analyses.

Model

To test whether the topics in Item 1A and Item 7 textual disclosures are associated with going concern accuracy variation, I follow the research design of Bakke et al. (2020), Bao and Datta (2014), and Geiger and Rama (2006) to create the regression models. Model [1] to Model [4] are used for examining the going concern accuracy variation when *Type I* error clients are introduced as inaccurate going concern opinions. In Model [1] and Model [2], the dependent variable is the probability of accurate going concern *opinions*, which equals 1 if client i receives a going concern opinion in year t and files for bankruptcy protection in year t+1; and the dependent variable equals 0 if client i does not receive going concern opinion in year t but file for bankruptcy in year t+1.[15] The independent variable of interest is Item 1A topics or Item 7 topics that are generated from machine learning. In Model [3] and Model [4], the dependent variable is the probability of accurate going concern *evaluation*, which equals 1 if 1) client i receives a going concern opinion in year t and files for bankruptcy in year t+1 or 2) client i does not receive a going concern opinion in year t and does not file for bankruptcy in year t+1. The probability of accurate going concern evaluation equals 0 if client i receives going concern opinion in year t but does not file for bankruptcy in year t+1. The independent variable of interest is the topic generated from Item1a or Item7 disclosures by machine learning.

Due to computer processing limitations, Model [3] and [4] are estimated using a matched sample design. My main sample only includes financially distressed clients, which already eliminates clients that would not have a reasonable likelihood of receiving a going concern opinion and acts as an initial level of matching based on financial health. However, to provide a

[15] Table 2 Panel B and C have the sample distribution used in Model [1] to Model [8]

second level of matching, I further restrict my sample of Type I error clients with accurate going concern *evaluation*s (GC_t=1 and $Bankruptcy_{t+1}$=1 and GC_t=0 and $Bankruptcy_{t+1}$=0) to those from the same 2-digit SIC and of similar size as my Type I error clients. My similar size restriction allows up to the three accurate *evaluation* clients to be matched to an error client based on absolute size differences. Any duplicate accurate going concern *evaluation* firm year observations (this would occur when an accurate going concern *evaluation* client is matched to more than one Type I error client) are eliminated.

$$\mathbf{Pr}(\boldsymbol{Accurate\ Going\ Concern\ Opinion_{it}}) = \alpha_0 + \alpha_1 LogSale_{it} + \alpha_2 Zscore_{it} + \boldsymbol{\alpha_3} NYSE_{it} + \alpha_4 DFT_{it} + \alpha_5 BIG4_{it} + \alpha_6 Leverage_{it} + \boldsymbol{\alpha_T item1a\_Topic} + Industry\ Fixed\ Effects + Year\ Fixed\ Effects + \varepsilon \qquad [1]$$

$$\mathbf{Pr}(\boldsymbol{Accurate\ Going\ Concern\ Opinion_{it}}) = \alpha_0 + \alpha_1 LogSale_{it} + \alpha_2 Zscore_{it} + \alpha_3 NYSE_{it} + \alpha_4 DFT_{it} + \alpha_5 BIG4_{it} + \alpha_6 Leverage_{it} + \boldsymbol{\alpha_T item7\_Topic} + Industry\ Fixed\ Effects + Year\ Fixed\ Effects + \varepsilon \qquad [2]$$

$$\mathbf{Pr}(\boldsymbol{Accurate\ Going\ Concern\ Evaluation_{it}}) = \alpha_0 + \alpha_1 LogSale_{it} + \alpha_2 Zscore_{it} + \alpha_3 NYSE_{it} + \alpha_4 DFT_{it} + \alpha_5 BIG4_{it} + \alpha_6 Leverage_{it} + \boldsymbol{\alpha_T item1a\_Topic} + Industry\ Fixed\ Effects + Year\ Fixed\ Effects + \varepsilon \qquad [3]$$

$$\mathbf{Pr}(\boldsymbol{Accurate\ Going\ Concern\ Evaluation_{it}}) = \alpha_0 + \alpha_1 LogSale_{it} + \alpha_2 Zscore_{it} + \alpha_3 NYSE_{it} + \alpha_4 DFT_{it} + \alpha_5 BIG4_{it} + \alpha_6 Leverage_{it} +$$

$$\boldsymbol{\alpha_T item7\_Topic} + Industry\ Fixed\ Effects +$$

$$Year\ Fixed\ Effects + \ \varepsilon \qquad [4]$$

Model [5] to Model [8] are used for examining the going concern accuracy variation when *Type II error* clients are defined as inaccurate going concern opinions. Similar to Model [1] and Model [2], the dependent variable in Model [5] and Model [6] is the probability of accurate going concern *opinions*. The independent variable of interest is the topic in Item 1a or Item 7 disclosures. In Model [7] and Model [8], the dependent variable is the probability of accurate going concern *evaluations*. The independent variable of interest is the topic generated from Item 1A or Item7 disclosures. Model [7] to Model [8] also use a matched sample created by the same process as Model [3] and [4] but use Type II error clients in the matching procedure.

$$\mathbf{Pr}(\boldsymbol{Accurate\ Going\ Concern\ Opinion_{it}}) = \alpha_0 + \ \alpha_1 LogSale_{it} + \ \alpha_2 Zscore_{it} + \boldsymbol{\alpha_3} NYSE_{it} + \alpha_4 DFT_{it} + \ \alpha_5 BIG4_{it} + \alpha_6 Leverage_{it} + \alpha_7 banklag + \alpha_8 reportlag + \boldsymbol{\alpha_T item1a\_Topic} + Industry\ Fixed\ Effects + \ Year\ Fixed\ Effects + \ \varepsilon \qquad [5]$$

$$\mathbf{Pr}(\boldsymbol{Accurate\ Going\ Concern\ Opinion_{it}}) = \alpha_0 + \ \alpha_1 LogSale_{it} + \ \alpha_2 Zscore_{it} + \alpha_3 NYSE_{it} + \alpha_4 DFT_{it} + \ \alpha_5 BIG4_{it} + \alpha_6 Leverage_{it} + \alpha_7 banklag + \alpha_8 reportlag + \boldsymbol{\alpha_T item7\_Topic} + Industry\ Fixed\ Effects + \ Year\ Fixed\ Effects + \ \varepsilon \qquad [6]$$

$$\mathbf{Pr}(\boldsymbol{Accurate\ Going\ Concern\ Evaluation_{it}}) = \alpha_0 + \ \alpha_1 LogSale_{it} + \ \alpha_2 Zscore_{it} + \alpha_3 NYSE_{it} + \alpha_4 DFT_{it} + \ \alpha_5 BIG4_{it} + \alpha_6 Leverage_{it} + \alpha_7 banklag + \alpha_8 reportlag + \boldsymbol{\alpha_T item1a\_Topic} + Industry\ Fixed\ Effects + \ Year\ Fixed\ Effects + \ \varepsilon \qquad [7]$$

$$\mathbf{Pr}(\boldsymbol{Accurate\ Going\ Concern\ Evaluation_{it}}) = \alpha_0 + \alpha_1 LogSale_{it} + \alpha_2 Zscore_{it} + \alpha_3 NYSE_{it} + \alpha_4 DFT_{it} + \alpha_5 BIG4_{it} + \alpha_6 Leverage_{it} + \alpha_7 banklag + \alpha_8 reportlag + \boldsymbol{\alpha_T item7\_Topic} + Industry\ Fixed\ Effects + Year\ Fixed\ Effects + \varepsilon \quad [8]$$

To eliminate the possibility that the effect of the topics on going concern accuracy variation is driven by other factors, I include the following control variables: $Zscore_{it}$ is the bankruptcy scores for client i in year t calculated by Altman (1968); $LogSale_{it}$ is the natural log of sales for client i in year t; $NYSE_{it}$ = 1 if client i is listed in New York Stock Exchange in year t, and 0 otherwise; $Reportlag_{it}$ is the lag between the date of fiscal-year end and the date of auditor signature for client i in year t. *BIG4* =1 if client i in year t is audited by Deloitte, PwC, E&Y or KPMG. *DFT* =1 if client i in year t is under payment or technical default. *Leverage* is calculated as total liabilities divided by total assets. *Banklag* is calculated as the total days between audit report dates and bankruptcy date.

To test the effectiveness of machine learning methods in generating proxies for going concern accuracy variation, I estimate the following parsimonious model below to predict going concern accuracy, Type I errors, and Type II errors. I then separately add machine-learning-based proxies and other textual attributes that are used frequently in textual analysis studies to compare the significance of the coefficient estimates and the change in explanatory power of the models after adding each variable. In Model [9], the dependent variable is the probability of accurate going concern opinion. It equals 1 if client i receives going concern opinion in year t and files for bankruptcy protection in year t+1, and equals 0 otherwise.[16] The independent

[16] The probability of accurate going concern opinion equals 0 if firm i 1) does not receive going concern opinion in year t and does not file for bankruptcy protection in year t+1, 2) receives a going concern opinion in year t but does

variable of interest is the probability of accurate going concern opinion that is predicted by the machine learning algorithms. In Model [10] and Model [11], the dependent variable is replaced by the probability of Type I (Type II) going concern error, which equals 1 if client i receives (does not receive) a going concern opinion in year t but does not (does) file for bankruptcy protection in year t+1 and 0 otherwise. The independent variable of interest is Type I (Type II) going concern error prediction by machine learning algorithms. The machine learning algorithms use 70% of the sample to train the algorithms and then generate the predictions for the rest 30% of the sample. Then this testing sample is used for the empirical testing. The control variables are the same as the previous model.[17]

$$\begin{aligned}\mathbf{Pr}(\boldsymbol{Accurate\ Going\ Concern\ Opinion_{it}}) = {} & \alpha_0 + \boldsymbol{\alpha_1 Accuracy\_Score_{it}} + \\ & \alpha_2 LogSale_{it} + \alpha_3 Zscore_{it} + \alpha_4 NYSE_{it} + \alpha_5 DFT_{it} + \\ & \alpha_6 BIG4_{it} + \alpha_7 Leverage_{it} + \alpha_8 banklag + \alpha_9 reportlag + \\ & Industry\ Fixed\ Effects + Year\ Fixed\ Effects + \varepsilon \qquad [9]\end{aligned}$$

$$\begin{aligned}\mathbf{Pr}(\boldsymbol{Type\_I\_Error_{it}}) = {} & \alpha_0 + \boldsymbol{\alpha_1 Type\_I\_Score_{it}} + \alpha_2 LogSale_{it} + \alpha_3 Zscore_{it} + \\ & \alpha_4 NYSE_{it} + \alpha_5 DFT_{it} + \alpha_6 BIG4_{it} + \alpha_7 Leverage_{it} + \\ & \alpha_8 banklag + \alpha_9 reportlag + Industry\ Fixed\ Effects + \\ & Year\ Fixed\ Effects + \varepsilon \qquad [10]\end{aligned}$$

$$\begin{aligned}\mathbf{Pr}(\boldsymbol{Type\_II\_Error_{it}}) = {} & \alpha_0 + \boldsymbol{\alpha_1 Type\_II\_Score_{it}} + \alpha_2 LogSale_{it} + \alpha_3 Zscore_{it} + \\ & \alpha_4 NYSE_{it} + \alpha_5 DFT_{it} + \alpha_6 BIG4_{it} + \alpha_7 Leverage_{it} +\end{aligned}$$

not file for bankruptcy protection in year t+1, 3) does not receive a going concern opinion in year t but files for bankruptcy protection in year t+1.

[17] Appendix A presents the list of variable definitions.

$$\alpha_8 banklag + \alpha_9 reportlag + Industry\ Fixed\ Effects + Year\ Fixed\ Effects + \varepsilon \quad [11]$$

Empirical Results

To estimate the regressions for examining which topics in Item 1A and Item 7 disclosure are associated with going concern accuracy variation, I first use the machine learning algorithms mentioned in section 3.2. The topics generated from the two disclosures examined (Item 1A and Item 7) and relating to the four different outcomes (Type I error versus accurate *opinion* and accurate *evaluation*, Type II error versus accurate *opinion* and accurate *evaluation*) are presented in Figure 1 to Figure 8. There are 50 keywords in each topic that can be used for interpretation. The size of the keyword is related to the weight (how close each keyword is to the centroid of the topic cluster). In addition, to estimating the regressions, I follow the machine learning procedures in section 3.3 to generate the prediction regarding the probabilities of accurate going concern opinions, Type I errors, and Type II errors in the testing sample.

Descriptive statistics and regression results for accurate going concern opinion test in type i going concern error settings

Figure 1 and Figure 2 present the topics generated from the machine learning algorithms to be used in Model [1] and Model [2]. The textual data analyzed by the algorithms is from Item 1A and Item 7 disclosures of Type I error clients and accurate going concern *opinion* clients.

Table 3 provides the industry distribution of Model [1] and Model [2], in which the sample includes Type I error clients and accurate going concern *opinion* clients. The Manufacturing industry is the domain industry with 474 observations, and Services industry has

132 observations which is the second largest industry in the sample. The remaining industries are Mining (84 observations), Transportation, Communications, Electric, Gas and Sanitary service (62 observations), Retail (38 observations), Wholesale (8 observations), and Construction (3 observations).

Table 4 Panel A provides the value of the mean, median, standard deviation, first quantile, and third quantile for Model [1] and Model [2]. The majority of the Item 1A/Item 7 topics have means larger than 0.5, which can be interpreted as those topics are presented in over 50 percent of the clients' Item 1A/Item 7 disclosures. The control variables have similar means and medians compared to the previous studies in the literature. *Accurate Going Concern Opinion* has a mean of 0.172, which means that the majority of the observations in Model [1] and Model [2] are Type I error clients which *Accurate Going Concern Opinion* equals 0 and is consistent with the conclusion that 80% of going concern opinions has Type I errors (Blay et al. 2018; Carson et al. 2013).

Table 4 Panel B provides the descriptive statistics for the Type I error clients (Accurate Going Concern Opinion= 0) and the accurate going concern opinion clients (Accurate Going Concern Opinion = 1) separately. The t-tests show that the majority of Item 1A and Item 7 topics are significantly different between the two groups, which provides univariate support that there are differences regarding the types of disclosure topics that precede various going concern opinion related outcomes.

Table 5 reports the probit model results for testing which topics of Item 1A and topics of Item 7 impact the likelihood of going concern opinion accuracy in a Type I error setting. The first column shows the results of Item 1A topics and the second column shows the results of Item 7 topics. In the first column, *Item1a_Topic_12* and *Item1a_Topic_14* are positively and significantly associated with the probability of accurate going concern opinions, which suggests that those two topics are correlated with a lower probability of Type I going concern errors.[18] In Figure 1, *Item1a_Topic_12* has the keywords "key, personnel, attract, retain, and train" which appear to be related to human capital risks.[19] *Item1a_Topic_14* has the keywords "supply, manufacture, suppliers, and manufacturing" which relate to supply chain risks. This suggests that the probability of accurate going concern *opinions* (and specifically avoiding Type I errors) is higher if clients disclose human capital risk and supply chain risk. In the second column, *Item7_Topic_17* is positively and significantly associated with the probability of accurate going concern *opinions*, which means that this topic is correlated with a lower probability of Type I going concern errors. In Figure 2*, Item7_Topic_17* has the keywords "taxable, tax, allowance, deferred, and valuation", which can be defined as tax-related client information. This suggests that the probability of accurate going concern opinions (and avoiding Type I errors) is higher if a client discloses tax related information.

Descriptive statistics and regression results for accurate going concern evaluation test in type i going concern error settings

Figure 3 and Figure 4 show the common topics generated by Top2Vec for analyzing the impact of Item 1A and Item 7 disclosures on the probability of accurate going concern *evaluations*. Compared to accurate going concern opinion clients, clients with accurate going concern *evaluations* include both observations where clients 1) do not receive going concern opinions in the current year and do not file for bankruptcy protection in the subsequent year, as well as those that 2) receive going concern opinions in the current year and file for bankruptcy

[18] *Item1a_Topic_14* is positively and significantly associated with the probability of accurate going concern opinions at p=0.1 level.
[19] All the topics that have significant results in regressions are bold in the figures.

protection in the subsequent year (the accurate going concern *opinion* observations). As previously mentioned, both types of going concern related outcomes could be considered ‘accurate’ because auditors accurately evaluate those clients’ abilities to continue as going concern. Due to the limitation of the computer capacities for running machine learning algorithms, I cannot incorporate all the accurate clients’ Item 1A and Item 7 disclosures. Therefore, I use the smaller matched sample to conduct this analysis. Recall, the matched sample is created by pairing the Type I error clients with accurate going concern *evaluation* clients based on industry and size. Figure 3 exhibits the topics generated from Item 1A disclosures, and Figure 4 exhibits the topics generated from Item 7 disclosures.

Table 6 shows the industry distribution of the observations that are used in Model [3] and Model [4]. Similar to the previous test using the main sample, the Manufacturing industry (1,444 observations) comprises the most observations and is followed by Services industry (421 observations). The Mining industry (213 observations) and Transportation, Communications, Electric, Gas and Sanitary service industry (152 observations) are in the third and fourth place. Retail (96 observations), Wholesale (38 observations), and Construction industry (10 observations) are also represented in the matched sample.

Table 7 Panel A displays the descriptive statistics for Model [3] and Model [4]. The topics in Item 1A disclosures have means in a range between 0.31 to 0.925, and the topics in Item 7 disclosures have means in a range between 0.3 to 0.985. *Accurate Going Concern Evaluation* has a mean of 0.721, which means that most of the observations used in Model [3] and Model [4] are clients with accurate going concern *evaluations* that do not have Type I going concern errors.

Table 7 Panel B provides the descriptive statistical differences between accurate going concern evaluation clients and Type I error clients. There are some significant differences between the Item 1A topics and Item 7 topics disclosed by the two groups, demonstrated by the t-test results.

Table 8 presents the probit model results for testing which topics of Item 1A and topics of Item 7 impact the likelihood of accurate going concern *evaluations* in a Type I error setting. Regarding Item 1A disclosures, there are positive and significant associations between *Item1a_Topic_0*, *Item1a_Topic_1*, *Item1a_Topic_12,* and *Item1a_Topic_15* and the probability of accurate going concern *evaluation*, which indicates a negative association between the topics and probability of Type I going concern errors.[20] Figure 3 provides evidence regarding the interpretation of the topics, in which *Item1a_Topic_0* has the keywords "motivate, personnel, integrate, key, and skilled" which appears to indicate human capital risk. *Item1a_Topic_1* has the keywords "fluctuations, factors, analysts, expectation, decline" which seems to indicate dispersion risk. In addition, *Item1a_Topic_12* has the keywords "penalties, kickback, criminal, civil, and fines" which could represent legal risks. *Item1a_Topic_15* has the keywords "recession, downturn, economy economic, slowdown" which represent macro-economic risk. The results show that the probability of accurate going concern *evaluation* (and avoiding Type I errors) is higher if clients disclose human capital risk, dispersion risk, legal risk, and macro-economic risk.

In addition, *Item1a_Topic_16, Item1a_Topic_17, Item1a_Topic_20, Item1a_Topic_21, Item1a_Topic_27, and Item1a_Topic_29* are negatively and significantly associated with the

[20] Item1a_Topic_15 has a positive and significant coefficient at p=0.1 level.

likelihood of going concern evaluation accuracy. I consider *Item1a_Topic_16* to represent disclosure of a funding risk because of the keywords "inception, achieve, profitable, generate, and profitability". *Item1a_Topic_17* seems to represent financial condition risk because of the keywords "financing, financings, raise, debt, and funds". Moreover, *Item1a_Topic_20* refers to debt risk based on "indebtedness, secured, and lenders". *Item1a_Topic_21* refers to operational risk based on the keywords "net, income, table, carryovers, and forwards". *Item1a_Topic_27* refers to attestation risk based on the keywords "sarbanes, oxley, weaknesses, and attestation". *Item1a_Topic_29* refers to stock market risk based on the keywords "delist, nasdaq, delisted, delisting, and quotation". The regression results show negative and significant coefficients of those topics in Item 1A disclosures, which indicate that funding risk, financial condition risk, debt risk, operational risk, attestation risk, and stock market risk can reduce the likelihood of auditors providing accurate going concern *evaluation* (and avoiding Type I errors).

Table 8 also presents the regression results for testing which topics of Item 7 disclosures impact the likelihood of going concern evaluation accuracy in a Type I error setting. *Item7_Topic_3, Item7_Topic_5, and Item7_Topic_16* are negatively and significantly associated with the likelihood of going concern evaluation accuracy.[21] *Item7_Topic_3* appears to represent disclosures related to a firms growth potential (growing, grow, channels, network, and expanding). *Item7_Topic_5* appears to indicate the firm characteristic regarding contributed capital (shares, holder, preferred, and common). *Item7_Topic_16* is considered to indicate the firm characteristic regarding political contributions (candidates, obtain, raise, relinquish, and financings). This result shows that the likelihood of accurate going concern *evaluation* (and avoiding Type I errors) is lower if a client discloses information regarding growth potential,

[21] *Item7_Topic_3* has a significant and negative coefficient at p=0.1 level.

contributed capital, and political contribution. Collectively, these findings suggest that auditors appear to have difficulty correctly *evaluating* clients' ability to continue as going concern (and are more likely to commit a Type I error) if the client discusses funding risk, financial condition risk, debt risk, operational risk, attestation risk, and stock market risk, growth potential, contributed capital, and political contributions in its financial statement. Interestingly, these topics are not the same as the ones that cause auditors to make inaccurate going concern *opinions*, and suggest that it is important to distinguish between *opinion* and *evaluation* outcomes.

Descriptive statistics and regression results for accurate going concern opinion test in type ii going concern error settings

Figure 5 and Figure 6 contain the topics generated from Top2Vec algorithms which are used in Model [5] and Model [6], in order to test the impact of Item 1A and Item 7 disclosures on accurate going concern *opinion* in Type II error settings.

Table 9 presents the industry distribution of the observations for testing the impact of Item 1A/Item 7 topics on the likelihood of issuing an accurate going concern *opinion* and avoiding a Type II error. Manufacturing is the most dominant industry, with 90 observations. The second and third domain industries are the Mining and Retail industries, with 61 observations and 34 observations respectively. Transportation, Communication, Electric, Gas, and Sanitary services industry, Service industry, and Wholesale trade industry represent the remaining with 27, 16, and 6 observations.

Table 10 Panel A reports the descriptive statistics of Model [5] and Model [6], which test the association between Item 1A/Item 7 topics and the likelihood of accurate going concern *opinion* relative to committing a Type II error. Similar to the descriptive results in Model [1] and Model [2], most of the Item 1A topics have means that are larger than 0.5. *Item1a_Topic_5* has the largest mean of 0.923 and *Item1a_Topic_27* has the smallest mean of 0.265. Regarding the descriptive statistics in Item 7 topics, the means of most of the topics are larger than 0.5, with ranging from a mean of 0.316 (*Item7_Topic_16*) to a mean of 0.953 (*Item7_Topic_6/Item7_Topic_11*). *Accurate Going Concern Opinion* has a mean of 0.59 which means that the number of observations with accurate going concern opinions is larger than the number of observations with Type II going concern errors. This is consistent with previous studies that document that 'only' 40% of going concern opinions suffer from Type II errors (Blay et al. 2018; Carson et al. 2013).

Table 10 Panel B shows the descriptive statistics for the Type II going concern error clients and accurate going concern opinion clients respectively. The t-test suggests that *Item1a_Topic_8, Item1a_Topic_28, Item7_Topic_18, and Item1a_Topic_19* are significantly different in the error group and the accurate going concern opinion group.

Table 11 reports the results of the probit model regarding the likelihood of accurate going concern opinions in Type II error settings. For Topics in Item 1A disclosures, *Item1a_Topic_8* and *Item1a_Topic_11* have positive and significant coefficients, which means those two topics are positively and significantly associated with the likelihood of accurate going concern opinions. In Figure 5, *Item1a_Topic_8* includes the keywords "bankruptcy, chapter, debtors, reorganize, and confirmation", which refers to the bankruptcy risk. *Item1a_Topic_11* has the keywords "statements, reports, weakness, and explanatory", which refers to operational risk. Therefore, the results demonstrate that the likelihood of accurate going concern *opinion* (and avoiding a Type II error) is higher if clients disclose risks regarding bankruptcy and operation.

Meanwhile, the coefficients of *Item1a_Topic_0*, *Item1a_Topic_5, Item1a_Topic_29* are negative and significant, which indicates that the probability of accurate going concern *opinion* (and avoiding a Type II error) is lower for clients that disclose those topics in Item 1A disclosures.[22] Figure 5 provides the keywords for each significant topic, in which *Item1a_Topic_0* indicates development risk (candidates, product, trials, and clinical), *Item1a_Topic_5* indicate supply chain risk (customer, suppliers, reputation, vendors, and relationships), and *Item1a_Topic_29* indicates environmental risk (emissions, greenhouse, climate). Thus, it appears that auditors are more likely to accurately evaluate the effect of bankruptcy and operational risks, but less likely to accurately evaluate the effect of development, supply chain, and environmental risks on clients' ability to continue as going concern and avoid committing Type II errors.

Table 11 also provides evidence of the Item 7 topics that can impact the likelihood of accurate going concern *opinions*, in which *Item7_Topic_1*, *Item7_Topic_8,* and *Item7_Topic_19* have positive and significant coefficients.[23] Specifically, *Item7_Topic_1* includes keywords "decreased, compared, lower, increase, and primarily" which refers to performance changes. *Item7_Topic_8* has the keywords "maintenance, equipment, increased, decreased, and costs" which refers to capital expenditure related costs. *Item7_Topic_19* has the keywords "bankruptcy, debtors, court, creditors, and chapter" which refers to bankruptcy. The likelihood of accurate going concern opinion (and avoiding Type II errors) is higher if those firm characteristics are disclosed in Item 7.

[22] *Item1a_Topic_5* has a significant coefficient at p=0.1 level.
[23] *Item7_Topic_1* has a significant coefficient at p=0.1 level.

On the other hand, *Item7_Topic_6* and *Item7_Topic_15* are negatively associated with the likelihood of accurate going concern opinion in Type II error settings.[24] *Item7_Topic_6* appear to refer to operational performance based on the keywords "table, gaap, thousands, reconciliation, ebitda". *Item7_Topic_15* appear to relate to tax information on the keywords "taxable, tax, and allowance". Thus, auditors are more likely to accurately issue going concern *opinions* (and avoiding Type II errors) when clients disclose information about performance changes, capital expenditure related costs, and bankruptcy, but less likely to accurately issue going concern *opinions* (and avoiding Type II errors) based on the firm fact regarding operational performance and tax.

Descriptive statistics and regression results for accurate going concern evaluation test in type ii going concern error settings

Figure 7 and Figure 8 show the topics that are generated from Top2vec that are used in Model [7] and Model [8]. Similar to the sample match procedure in testing the likelihood of accurate going concern evaluation in the Type I error settings, clients with accurate going concern *evaluations* in this test include those 1) that do not receive going concern opinions in the current year and do not file for bankruptcy protection in the subsequent year, and 2) that receive going concern opinions in the current year and file for bankruptcy protection in the subsequent year. The matched sample is also used and is similarly created by matching Type II error clients with accurate going concern *evaluation* clients by industry and size.

The industry distribution for Model [7] and Model [8] is shown in Table 12. Manufacturing and Mining are the top 2 industries with 140 observations and 112 observations

[24] *Item7_Topic_6* has a significant coefficient at p=0.1 level.

respectively. Then Retail industry (78 observations), Transportation, Communications, Electric, Gas and Sanitary service industry (28 observations), Wholesale industry (20 observations), and Services industry (12 observations) are incorporated in the matched sample as well.

Table 13 panel A shows the descriptive statistics of Model [7] and Model [8], which tests the effect of Item 1A/Item 7 topics on the likelihood of accurate going concern *evaluation* in Type II error settings. *Item1a_Topic_8* has the smallest mean of 0.2 and *Item1a_Topic_11* has the largest mean of 0.9 among Item 1A topics. For Item 7 topics, *Item7_Topic_5* has the smallest mean of 0.315 and *Item7_Topic_0* has the largest mean of 0.956. The mean of *Accurate Going Concern Evaluation* is 0.754, which indicates the sample has more observations with accurate going concern *evaluations* than Type II going concern errors.

Table 13 panel B shows the descriptive statistics by two groups, which are accurate going concern *evaluation*=0 (clients with a Type II error) group and accurate going concern *evaluation*=1 group. The t-test results illustrate that there are some significant differences between those two groups regarding *Item1a_Topic_10, Item7_Topic_1, Item7_Topic_3,* and *Item7_Topic_4*.

Table 14 presents the probit model results of testing how the topics disclosed in Item 1A and Item 7 sections can impact the accuracy of going concern evaluations. In Item 1A disclosures, *Item1a_Topic_4, Item1a_Topic_6,* and *Item1a_Topic_24* are positively and significantly correlated with the likelihood of accurate going concern evaluations.[25] *Item1a_Topic_4* contains the keywords "spending, economic, prices, demand, and downturn", which represent macro-economic risk. *Item1a_Topic_6* has the keywords "intellectual, patents,

[25] *Item1a_Topic_6* has a significant coefficient at p=0.1 level.

infringe, and infringement", which I assume relate to intellectual property risk. *Item1a_Topic_24* has keywords "derivatives, derivative, hedging, and swaps", which I consider to be related to investment risk. Meanwhile, *Item1a_Topic_8* and *Item1a_Topic_9* have negative and significant coefficients, which the former one can be interpreted as development risk (commercialize, candidates, candidate, trials, indications) and the latter one can be interpreted as oil and gas risk (reserves, proved, estimated, undeveloped, properties, geological, and drill).[26] Therefore, the results suggest that auditors are more likely to make accurate going concern *evaluations* (and commit less Type II errors) when their clients disclose macro-economic risk, intellectual property risk, and investment risk. However, they are less likely to have accurate going concern *evaluations* (and more likely to commit Type II errors) when their clients disclose development risk and oil/gas risk.

For Item 7 disclosures, *Item7_Topic_2* is positively and significantly associated with the likelihood of accurate going concern *evaluation* (and avoiding Type II errors), but *Item7_Topic_1* and *Item7_Topic_4* are negatively and significantly associated with the likelihood of accurate going concern *evaluation*s.[27] *Item7_Topic_2* appears to relate to human capital (expense, administrative, headcount, costs, primarily, and salaries), *Item7_Topic_1* is related to loans (covenants, secured, loan, libor, and credit), *Item7_Topic_4* is related to operational performance (following, reconciliation, table, presents, and gaap). These associations illustrate that auditors are more likely to have accurate going concern *evaluations* (and avoiding Type II errors) when clients disclose human capital related information, but less likely to have

[26] *Item1a_Topic_9* has a significant coefficient at p=0.1 level.
[27] *Item7_Topic_2* has a significant coefficient at p=0.1 level.

accurate going concern *evaluations* (and avoiding Type II errors) when clients disclose loan and operational performance related information.

Descriptive statistics and regression results for accurate going concern opinion proxy

In this section, I estimate multiple versions of my probit model (Model [9]) to examine the validity of the machine learning generated accurate going concern opinion proxy based on the textual information disclosed in Item 1A and Item7. The proxy captures the probability of accurate going concern opinion by three different machine learning algorithms - Naïve Bayes Classification, Supporting Vector Machine, and Random Forest Classification. The change in explanatory power after adding each of these proxies is then compared with some commonly used textual analysis attributes (Readability, Specificity, Hardinformaiton, and Tone) in order to evaluate the relative performance of using machine learning to predict accurate going concern opinions.

Table 15 Panel A exhibits the descriptive statistics of Model [9]. The means of Readability, Specificity, HardInformation, and Tone are 10.071, 0.071, 0.003, and -0.018. For the textual-based accurate going concern opinion proxies generated by the three machine learning approaches, the means are 0.235 for *Accurate_Score_NB*, 0.399 for *Accurate_Score_SVM*, and *0.02 for Accurate_Score_RFC*. They represent the average probabilities that the machine learning algorithm provided for the accurate going concern opinion equals 1.

Table 15 Panel B exhibits the descriptive statistics for either the group of accurate going concern opinions or the group of inaccurate going concern opinions.[28] The t-test results show

[28] The inaccurate going concern opinion group (Accurate Going Concern Opinion =0) includes firms with no going concern and no bankruptcy, firms with Type I errors, and firms with Type II errors.

significant differences for *Accurate_Score_NB, Accurate_Score_SVM,* and *Accurate_Score_RFC* between those two groups. In addition, *Tone* has significant differences as well between the accurate and inaccurate going concern opinion groups.

Table 16 provides evidence regarding the validity of the accurate going concern opinion proxies created by machine learning algorithms. All three proxies, *Accurate_Score_NB*, *Accurate_Score_SVM,* and *Accurate_Score_RFC* are positively and significantly associated with the likelihood of accurate going concern opinions. For the traditional textual analysis attributes, only *Tone* has a significant coefficient, and is negative which suggests that more positive tones are correlated with less going concern opinion accuracy. The baseline Pseudo R-squared is 0.178 for the model with no text based proxies included. The addition of Tone increases the Pseudo R-squared to 0.21 which represents a 18 percent improvement in explanatory power. However, the explanatory power of *Accurate_Score_NB*, *Accurate_Score_SVM,* and *Accurate_Score_RFC* (Pseudo R2 =0.238, 0.375, 0.41 respectively) are much higher than *Tone*, which represent 34%, 111%, and 130% increase in explanatory power compared to the baseline model. This provides compelling evidence that textual-based machine learning proxies are better at predicting accurate going concern opinions compared to the traditional textual analysis attributes. Among those three proxies, the one created by Random Forest Classification has the best-predicted performance.

Descriptive statistics and regression results for type i going concern error proxy

In addition to creating accurate going concern proxies, the three machine learning algorithms can also be used to predict Type I going concern errors. Table 17 Panel A presents the descriptive statistics of the Type I going concern proxy test, in which the means of three textual-based Type I error proxies (*Type_I_Score_NB*, *Type_I_Score_SVM*, and *Type_I_Score_RFC*) are 0.349, 0.443, and 0.09 respectively. The means of the traditional textual analysis attributes

(*Readability, Specificity, Hardinformaiton,* and *Tone*) are 10.082, 0.071, 0.003, and -0.018. The mean of *Type I error* is 0.047, which indicates the majority of the observation used in Model [10] do not have Type I going concern errors.

Table 17 Panel B displays the descriptive statistics of Model [10] by groups. The three proxies (*Type_I_Score_NB, Type_I_Score_SVM,* and *Type_I_Score_RFC*) that predict the probability of Type I going concern errors have systematic differences, which is demonstrated by the t-test results. None of the traditional textual analysis attributes (*Readability, Specificity, Hardinformaiton,* and *Tone*) have systematic differences between the two groups.

Table 18 reports the probit model results of Model [10], which again proves the validity of the Type I error proxies generated by the machine learning algorithms. All three proxies (*Type_I_Score_NB, Type_I_Score_SVM,* and *Type_I_Score_RFC*) are positively and significantly associated with the likelihood of Type I going concern errors. *Tone*, as the traditional textual analysis attribute, once again has a negative and significant coefficient. When comparing the increase in explanatory power of adding each of the text based variables, the textual-based machine learning proxies (Pseudo R2 of *Type_I_Score_NB* = 0.158, Pseudo R2 of *Type_I_Score_SVM* = 0.384, Pseudo R2 of *Type_I_Score_RFC* = 0.339) once again outperform the traditional textual analysis attribute (Pseudo R2 of *Tone* = 0.145), which confirm the validity of machine learning algorithms and highlight the effectiveness of them in predicting Type I going concern errors. Interestingly, in this setting the Supporting Vector Machine creates the proxy that has the highest explanatory power compared to the other two.

Descriptive statistics and regression results for type ii going concern error proxy

I also create proxies to predict Type II going concern errors by the machine learning algorithms and the descriptive statistics are shown in Table 19. Panel A presents the descriptive statistics of Model [11], which tests the validity of the textual-based Type II going concern error proxy. The three proxies generated by machine learning algorithms (*Type_II_Score_NB, Type_II_Score_SVM,* and *Type_II_Score_RFC)* have means of 0.278, 0.423, and 0.022 respectively. The means of the traditional textual analysis attributes (*Readability, Specificity, Hardinformaiton,* and *Tone*) are 10.082, 0.071, 0.003, and -0.018. The mean of *Type_II_Error* is 0.006, which indicates the majority of the observations in this testing sample do not have Type II going concern errors.

Table 19 Panel B presents the descriptive statistics of Model [11] by two groups. *Type_II_Score_NB, Type_II_Score_SVM,* and *Type_II_Score_RFC* are significantly different between the Type II error groups and the non-Type II error groups. Yet, *Readability, Specificity, Hardinformaiton,* and *Tone* are not significantly different between the two groups.

Table 20 reports the probit regression results of Model [11], which examines the validity of Type II going concern error proxies generated by machine learning algorithms. *Type_II_Score_NB, Type_II_Score_SVM,* and *Type_II_Score_RFC* have positive and significant coefficients which confirm their validity in predicting Type II going concern errors. On the other hand, none of the traditional textual analysis attributes (*Readability, Specificity, Hardinformaiton,* or *Tone*) are significantly associated with the likelihood of Type II going concern errors. Among the three machine learning proxies, *Type_II_Score_NB* has the highest Pseudo R2 compared to *Type_II_Score_SVM* and *Type_II_Score_RFC,* which suggests that

Naïve Bayes Classification provides superior improvement in explanatory power relative to Supporting Vector Machine and Random Forest Classification in creating Type II error proxies.

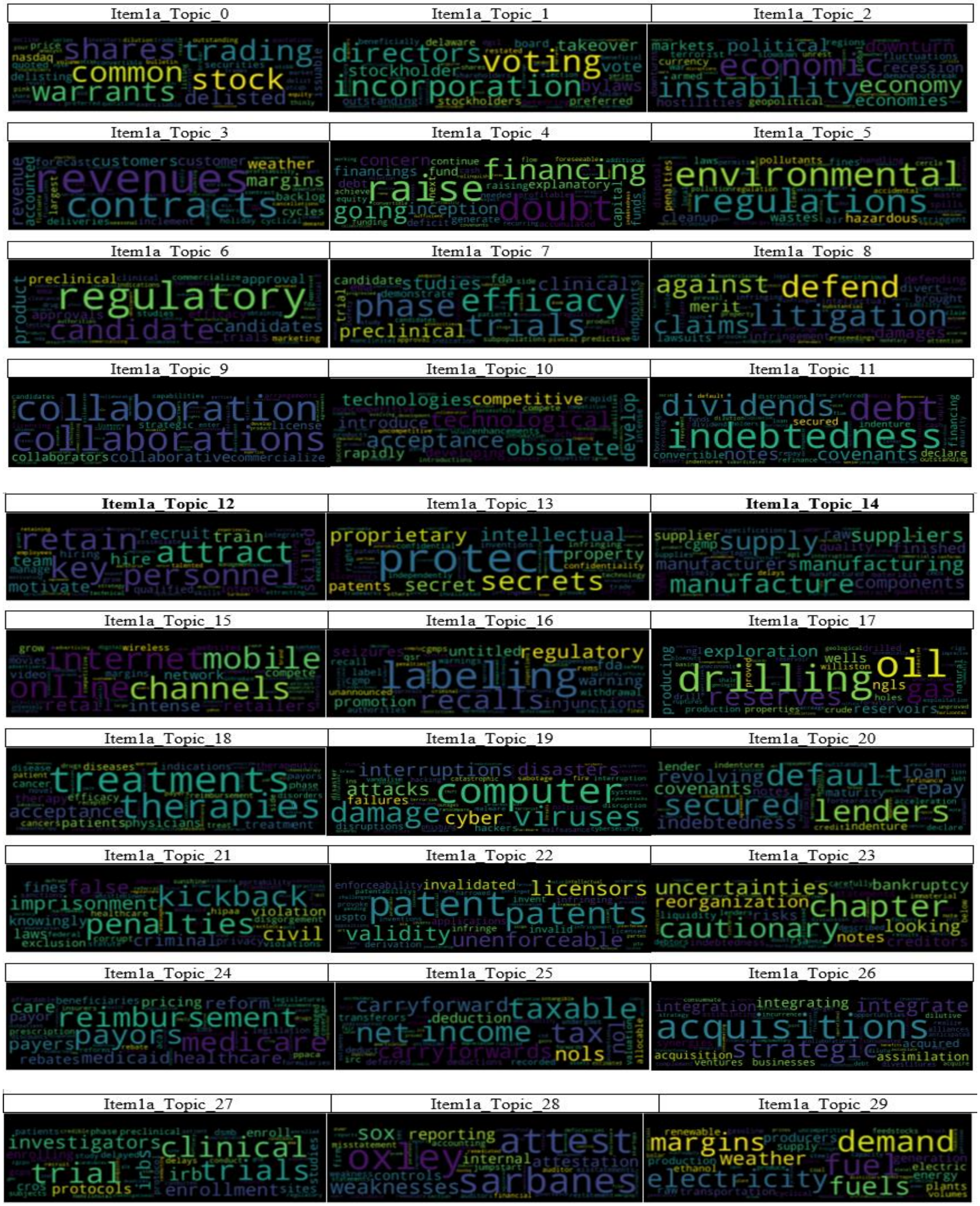

Figure 1. Item 1a Topic for Model [1]

Figure 2. Item7 Topic for Model [2]

Figure 3. Item1a Topic for Model [3]

Figure 4. Item7 Topic for Model [4]

Figure 5. Item1a Topic for Model [5]

Figure 6. Item7 Topic for Model [6]

Figure 7. Item1a Topic for Model [7]

Figure 8. Item7 Topic for Model [8]

Table 3 Industry Distribution of Model [1] and Model [2]

SIC Codes	Division	No. of Observations
0100-0999	Agriculture, Forestry, and Fishing	0
1000-1499	Mining	84
1500-1799	Construction	3
2000-3999	Manufacturing	474
4000-4999	Transportation, Communications, Electric, Gas and Sanitary service	62
5000-5199	Wholesale Trade	8
5200-5999	Retail Trade	38
6000-6799	Finance, Insurance and Real Estate	0
7000-8999	Services	132
9100-9729	Public Administration	0
9900-9999	Non-Classifiable	0
Total No. of Observations		802

Table 3 present the results of industry distribution of Model [1] and Model [2], which tests what topics in Item 1A and Item 7 are associated with the likelihood of accurate going concern opinion in Type I error settings. The sample include accurate going concern opinion clients and Type I going concern error clients. Observations with missing control variables are eliminated for the regressions.

Table 4 Descriptive Statistics for Model [1] and Model [2]

Panel A Descriptive Statistics for Accurate Going Concern Opinion Models (Type I Error) (N=802)

	Mean	**Median**	**Sd**	**p25**	**p75**
Item1a_Topic_0	0.868	1	0.339	1	1
Item1a_Topic_1	0.782	1	0.413	1	1
Item1a_Topic_2	0.858	1	0.349	1	1
Item1a_Topic_3	0.845	1	0.362	1	1
Item1a_Topic_4	0.890	1	0.313	1	1
Item1a_Topic_5	0.818	1	0.386	1	1
Item1a_Topic_6	0.436	0	0.496	0	1
Item1a_Topic_7	0.330	0	0.471	0	1
Item1a_Topic_8	0.863	1	0.344	1	1
Item1a_Topic_9	0.504	1	0.500	0	1
Item1a_Topic_10	0.805	1	0.396	1	1
Item1a_Topic_11	0.824	1	0.381	1	1

Table 4 (Continued).

Item1a_Topic_12	0.905	1	0.293	1	1
Item1a_Topic_13	0.677	1	0.468	0	1
Item1a_Topic_14	0.622	1	0.485	0	1
Item1a_Topic_15	0.687	1	0.464	0	1
Item1a_Topic_16	0.490	0	0.500	0	1
Item1a_Topic_17	0.224	0	0.417	0	0
Item1a_Topic_18	0.394	0	0.489	0	1
Item1a_Topic_19	0.756	1	0.430	1	1
Item1a_Topic_20	0.712	1	0.453	0	1
Item1a_Topic_21	0.522	1	0.500	0	1
Item1a_Topic_22	0.493	0	0.500	0	1
Item1a_Topic_23	0.877	1	0.329	1	1
Item1a_Topic_24	0.481	0	0.500	0	1
Item1a_Topic_25	0.746	1	0.436	0	1
Item1a_Topic_26	0.732	1	0.443	0	1

Table 4 (Continued).

Item1a_Topic_27	0.320	0	0.467	0	1
Item1a_Topic_28	0.650	1	0.477	0	1
Item1a_Topic_29	0.451	0	0.498	0	1
Item7_Topic_0	0.802	1	0.399	1	1
Item7_Topic_1	0.986	1	0.116	1	1
Item7_Topic_2	0.865	1	0.342	1	1
Item7_Topic_3	0.960	1	0.196	1	1
Item7_Topic_4	0.786	1	0.411	1	1
Item7_Topic_5	0.960	1	0.196	1	1
Item7_Topic_6	0.966	1	0.180	1	1
Item7_Topic_7	0.510	1	0.500	0	1
Item7_Topic_8	0.764	1	0.425	1	1
Item7_Topic_9	0.761	1	0.427	1	1
Item7_Topic_10	0.784	1	0.412	1	1
Item7_Topic_11	0.879	1	0.326	1	1

Table 4 (Continued).

Item7_Topic_12	0.769	1	0.422	1	1
Item7_Topic_13	0.897	1	0.305	1	1
Item7_Topic_14	0.716	1	0.451	0	1
Item7_Topic_15	0.440	0	0.497	0	1
Item7_Topic_16	0.774	1	0.418	1	1
Item7_Topic_17	0.782	1	0.413	1	1
Item7_Topic_18	0.191	0	0.393	0	0
Item7_Topic_19	0.663	1	0.473	0	1
Item7_Topic_20	0.758	1	0.428	1	1
Logsale	3.336	3.478	2.653	1.734	5.068
Zscore	0.809	0.521	0.980	0.101	1.284
EXCHCD	0.151	0	0.416	0	0
DFT	0.110	0	0.313	0	0
Big4	0.468	0	0.499	0	1
Leverage	0.441	0.297	0.516	0.084	0.620

Table 4 (Continued).

Accurate Going Concern Opinion	0.172	0	0.378	0	0

Panel B Descriptive Statics for Accurate Going Concern Opinion Models (Type I error) by Groups

	Accurate Going Concern = 0 (N=664)					**Accurate Going Concern = 1 (N=138)**					
	Mean	**Median**	**Sd**	**p25**	**p75**	**Mean**	**Median**	**Sd**	**p25**	**p75**	**t test**
Item1a_Topic_0	0.88	1	0.33	1	1	0.812	1	0.392	1	1	0.0679*
Item1a_Topic_1	0.801	1	0.4	1	1	0.688	1	0.465	0	1	0.113**
Item1a_Topic_2	0.843	1	0.36	1	1	0.928	1	0.26	1	1	-0.0842**
Item1a_Topic_3	0.834	1	0.37	1	1	0.899	1	0.303	1	1	-0.0642
Item1a_Topic_4	0.899	1	0.3	1	1	0.848	1	0.36	1	1	0.0513
Item1a_Topic_5	0.809	1	0.39	1	1	0.862	1	0.346	1	1	-0.0536
Item1a_Topic_6	0.479	0	0.5	0	1	0.232	0	0.424	0	0	0.247***
Item1a_Topic_7	0.37	0	0.48	0	1	0.138	0	0.346	0	0	0.233***
Item1a_Topic_8	0.875	1	0.33	1	1	0.804	1	0.398	1	1	0.0707*
Item1a_Topic_9	0.541	1	0.5	0	1	0.326	0	0.47	0	1	0.215***

Table 4 (Continued).

Item1a_Topic_10	0.827	1	0.38	1	1	0.703	1	0.459	0	1	0.124***
Item1a_Topic_11	0.821	1	0.38	1	1	0.841	1	0.367	1	1	-0.0198
Item1a_Topic_12	0.901	1	0.3	1	1	0.928	1	0.26	1	1	-0.0269
Item1a_Topic_13	0.723	1	0.45	0	1	0.457	0	0.5	0	1	0.266***
Item1a_Topic_14	0.636	1	0.48	0	1	0.558	1	0.498	0	1	0.0776
Item1a_Topic_15	0.678	1	0.47	0	1	0.732	1	0.445	0	1	-0.0542
Item1a_Topic_16	0.53	1	0.5	0	1	0.297	0	0.459	0	1	0.233***
Item1a_Topic_17	0.191	0	0.39	0	0	0.384	0	0.488	0	1	-0.193***
Item1a_Topic_18	0.438	0	0.5	0	1	0.181	0	0.387	0	0	0.257***
Item1a_Topic_19	0.741	1	0.44	0	1	0.826	1	0.38	1	1	-0.0851*
Item1a_Topic_20	0.679	1	0.47	0	1	0.87	1	0.338	1	1	-0.190***
Item1a_Topic_21	0.547	1	0.5	0	1	0.406	0	0.493	0	1	0.141**
Item1a_Topic_22	0.538	1	0.5	0	1	0.275	0	0.448	0	1	0.262***
Item1a_Topic_23	0.864	1	0.34	1	1	0.935	1	0.248	1	1	-0.0703*
Item1a_Topic_24	0.514	1	0.5	0	1	0.326	0	0.47	0	1	0.187***

Table 4 (Continued).

Item1a_Topic_25	0.729	1	0.44	0	1	0.826	1	0.38	1	1	-0.0972*
Item1a_Topic_26	0.724	1	0.45	0	1	0.768	1	0.424	1	1	-0.0437
Item1a_Topic_27	0.352	0	0.48	0	1	0.167	0	0.374	0	0	0.186***
Item1a_Topic_28	0.669	1	0.47	0	1	0.558	1	0.498	0	1	0.111*
Item1a_Topic_29	0.413	0	0.49	0	1	0.638	1	0.482	0	1	-0.225***
Item7_Topic_0	0.774	1	0.418	1	1	0.920	1	0.272	1	1	-0.146***
Item7_Topic_1	0.988	1	0.109	1	1	0.971	1	0.168	1	1	0.0169
Item7_Topic_2	0.852	1	0.355	1	1	0.928	1	0.260	1	1	-0.0751*
Item7_Topic_3	0.964	1	0.187	1	1	0.942	1	0.235	1	1	0.0218
Item7_Topic_4	0.782	1	0.413	1	1	0.797	1	0.404	1	1	-0.0155
Item7_Topic_5	0.965	1	0.183	1	1	0.935	1	0.248	1	1	0.0306
Item7_Topic_6	0.968	1	0.175	1	1	0.957	1	0.205	1	1	0.0119
Item7_Topic_7	0.556	1	0.497	0	1	0.290	0	0.455	0	1	0.266***
Item7_Topic_8	0.750	1	0.433	0	1	0.833	1	0.374	1	1	-0.0833*
Item7_Topic_9	0.742	1	0.438	0	1	0.848	1	0.360	1	1	-0.105**

Table 4 (Continued).

Item7_Topic_10	0.797	1	0.403	1	1	0.717	1	0.452	0	1	0.0793*
Item7_Topic_11	0.892	1	0.311	1	1	0.819	1	0.387	1	1	0.0727*
Item7_Topic_12	0.745	1	0.436	0	1	0.877	1	0.330	1	1	-0.131***
Item7_Topic_13	0.887	1	0.317	1	1	0.920	1	0.272	1	1	-0.0332
Item7_Topic_14	0.684	1	0.465	0	1	0.870	1	0.338	1	1	-0.186***
Item7_Topic_15	0.489	0	0.500	0	1	0.203	0	0.404	0	0	0.287***
Item7_Topic_16	0.791	1	0.407	1	1	0.688	1	0.465	0	1	0.102**
Item7_Topic_17	0.753	1	0.432	1	1	0.913	1	0.283	1	1	-0.160***
Item7_Topic_18	0.158	0	0.365	0	0	0.348	0	0.478	0	1	-0.190***
Item7_Topic_19	0.667	1	0.472	0	1	0.645	1	0.480	0	1	0.0222
Item7_Topic_20	0.735	1	0.442	0	1	0.862	1	0.346	1	1	-0.127**
Logsale	2.900	3.084	2.541	1.452	4.433	5.433	5.713	2.132	4.352	6.724	-2.533***
Zscore	0.765	0.485	0.966	0.062	1.234	1.019	0.691	1.023	0.288	1.487	-0.254**
Big4	0.434	0	0.496	0	1	0.630	1	0.484	0	1	-0.197***
EXCHCD	0.107	0	0.387	0	0	0.362	0	0.482	0	1	-0.255***

Table 4 (Continued).

DFT	0.099	0	0.299	0	0	0.159	0	0.367	0	0	-0.20
Leverage	0.368	0.240	0.450	0.057	0.524	0.793	0.671	0.652	0.431	1.011	-0.425***

Table 4 Panel A presents the descriptive statistics Model [1] and Model [2], which examines which topics in Item 1A and Item 7 are associated with the likelihood of accurate going concern opinion in the Type I error setting. $Item1A\_Topic\_N$ equals 1 if a client disclosure topic N in its Item1A disclosure and 0 otherwise. $Item7\_TopicN$ equals 1 if a client disclosure topic N in its Item7 disclosure and 0 otherwise. $LogSale_{it}$ is the natural log of sales. $Zscore_{it}$ is the bankruptcy score calculated by Altman (1968). *Big4* equals 1 if a client is audited by big 4 auditors and 0 otherwise. *EXCHCD* equals 1 if a client is listed in New York Stock Exchange and 0 otherwise. *DFT* equals 1 if a client is under technical or payment default and 0 otherwise. *Leverage* is calculated as total liabilities divided by total assets. Panel B represents the descriptive statistics for Type I error clients and accurate going concern clients separately. The Accurate Going Concern = 0 group includes Type I error clients and Accurate Going Concern = 1 group includes accurate going concern opinion clients.T-test are shown in the last column. *** p<0.01, ** p<0.05, * p<0.1.

Table 5 Regression Results of Model [1] and Model [2]

VARIABLES	1 PROB(*Accurate Going Concern Opinion*)	2 PROB(*Accurate Going Concern Opinion*)
Item1a_Topic_12	**0.710****	
	(2.092)	
Item1a_Topic_14	**0.275***	
	(1.672)	
Item1a_Topic_0	-0.204	
	(-0.871)	
Item1a_Topic_1	-0.136	
	(-0.713)	
Item1a_Topic_2	0.00972	
	(0.0341)	
Item1a_Topic_3	-0.123	
	(-0.478)	
Item1a_Topic_4	-0.116	
	(-0.481)	
Item1a_Topic_5	-0.199	
	(-0.895)	
Item1a_Topic_6	-0.171	
	(-0.800)	
Item1a_Topic_7	-0.0895	
	(-0.286)	
Item1a_Topic_8	-0.254	
	(-1.151)	
Item1a_Topic_9	-0.0886	
	(-0.517)	

Table 5 (Continued).

Item1a_Topic_10	-0.0646
	(-0.364)
Item1a_Topic_11	0.0194
	(0.0855)
Item1a_Topic_13	-0.0308
	(-0.163)
Item1a_Topic_15	0.0331
	(0.181)
Item1a_Topic_16	-0.183
	(-1.012)
Item1a_Topic_17	0.21
	(1.141)
Item1a_Topic_18	0.0371
	(0.144)
Item1a_Topic_19	0.136
	(0.689)
Item1a_Topic_20	-0.0131
	(-0.0657)
Item1a_Topic_21	-0.179
	(-1.122)
Item1a_Topic_22	0.0853
	(0.452)
Item1a_Topic_23	0.212
	(0.69)
Item1a_Topic_24	-0.0574
	(-0.330)
Item1a_Topic_25	0.0326
	(0.178)

Table 5 (Continued).

Item1a_Topic_26	0.163	
	(0.977)	
Item1a_Topic_27	0.323	
	(1.288)	
Item1a_Topic_28	-0.0515	
	(-0.355)	
Item1a_Topic_29	0.0324	
	(0.197)	
Item7_Topic_17		**0.561****
		(2.22)
Item7_Topic_0		0.123
		(0.512)
Item7_Topic_1		-0.600
		(-0.800)
Item7_Topic_2		0.0973
		(0.367)
Item7_Topic_3		0.176
		(0.329)
Item7_Topic_4		-0.0803
		(-0.410)
Item7_Topic_5		-0.287
		(-0.648)
Item7_Topic_6		0.219
		(0.428)
Item7_Topic_7		-0.112
		(-0.674)
Item7_Topic_8		-0.0174
		(-0.0813)

Table 5 (Continued).

Item7_Topic_9		0.147
		(0.809)
Item7_Topic_10		-0.127
		(-0.813)
Item7_Topic_11		-0.221
		(-0.965)
Item7_Topic_12		0.152
		(0.686)
Item7_Topic_13		-0.271
		(-0.928)
Item7_Topic_14		-0.034
		(-0.168)
Item7_Topic_15		-0.246
		(-1.400)
Item7_Topic_16		-0.0447
		(-0.288)
Item7_Topic_18		-0.0862
		(-0.476)
Item7_Topic_19		-0.158
		(-1.099)
Item7_Topic_20		0.275
		(1.494)
Logsale	0.236***	0.189***
	(5.336)	(4.633)
Zscore	-0.0961	-0.0781
	(-1.193)	(-0.994)
EXCHCD	0.0404	0.0521
	(0.264)	(0.35)

Table 5 (Continued).

DFT	0.0383	-0.0726
	(0.204)	(-0.373)
Big4	0.0209	0.149
	(0.136)	(1.01)
Leverage	0.727***	0.732***
	(5.978)	(6.18)
Constant	22.44	17.47
	(0.701)	(0.594)
Observations	802	802
Industry Fixed Effect	Yes	Yes
Year Fixed Effect	Yes	Yes
Pseudo R2	0.275	0.273

Table 5 presents the regression results of Model [1] and Model [2], which tests which topics disclosed in either Item 1A or Item 7 sections are associated with the likelihood of accurate going concern opinions in the Type I error setting. *Accurate Going Concern Opinion* equals 1 if a client receives a going concern opinion in the current year and files for bankruptcy protection in the subsequent year. *Accurate Going Concern Opinion* equals 0 if a client receives a going concern opinion in the current year and does not file for bankruptcy protection in the subsequent year (Type I error). $Item1A\_Topic\_N$ equals 1 if a client disclosure topic N in its Item1A disclosures and 0 otherwise. $Item7\_TopicN$ equals 1 if a client disclosure topic N in its Item7 disclosures and 0 otherwise. $LogSale_{it}$ is the natural log of sales. $Zscore_{it}$ is the Bankruptcy score calculated by Altman (1968). *Big4* equals 1 if a client is audited by big 4 auditors and 0 otherwise. *EXCHCD* equals 1 if a client is listed in New York Stock Exchange and 0 otherwise. *DFT* equals 1 if a client is under technical or payment default and 0 otherwise. *Leverage* is calculated as total liabilities divided by total assets. Z-scores are shown in the last column. *** p<0.01, ** p<0.05, * p<0.1.

Table 6 Industry Distribution of Model [3] and Model [4]

SIC Codes	Division	No. of Observations
0100-0999	Agriculture, Forestry, and Fishing	0
1000-1499	Mining	213
1500-1799	Construction	10
2000-3999	Manufacturing	1,444
4000-4999	Transportation, Communications, Electric, Gas and Sanitary service	152
5000-5199	Wholesale Trade	38
5200-5999	Retail Trade	96
6000-6799	Finance, Insurance and Real Estate	0
7000-8999	Services	421
9100-9729	Public Administration	0
9900-9999	Non-Classifiable	0
Total No. of Observations		2,376

Table 3 presents the results of industry distribution of Model [3] and Model [4], which tests what topics in Item 1A and Item 7 are associated with the likelihood of accurate going concern evaluation in Type I error settings. The sample include accurate going concern evaluation clients and Type I going concern error clients. Observations with missing control variables are eliminated.

Table 7 Descriptive Statistics for Type I Error Test (Model [3] -Model [4])

Panel A Descriptive Statistics for Accurate Going Concern Evaluation Models (Type I Error) (N=2,376)

	Mean	**Median**	**Sd**	**p25**	**p75**
Item1a_Topic_0	0.915	1	0.28	1	1
Item1a_Topic_1	0.925	1	0.263	1	1
Item1a_Topic_2	0.749	1	0.434	0	1
Item1a_Topic_3	0.435	0	0.496	0	1
Item1a_Topic_4	0.678	1	0.467	0	1
Item1a_Topic_5	0.392	0	0.488	0	1
Item1a_Topic_6	0.324	0	0.468	0	1
Item1a_Topic_7	0.376	0	0.485	0	1
Item1a_Topic_8	0.826	1	0.379	1	1
Item1a_Topic_9	0.67	1	0.47	0	1
Item1a_Topic_10	0.718	1	0.45	0	1
Item1a_Topic_11	0.793	1	0.405	1	1

Table 7 (Continued).

Item1a_Topic_12	0.585	1	0.493	0	1
Item1a_Topic_13	0.774	1	0.419	1	1
Item1a_Topic_14	0.417	0	0.493	0	1
Item1a_Topic_15	0.789	1	0.408	1	1
Item1a_Topic_16	0.734	1	0.442	0	1
Item1a_Topic_17	0.828	1	0.378	1	1
Item1a_Topic_18	0.464	0	0.499	0	1
Item1a_Topic_19	0.678	1	0.467	0	1
Item1a_Topic_20	0.604	1	0.489	0	1
Item1a_Topic_21	0.727	1	0.445	0	1
Item1a_Topic_22	0.561	1	0.496	0	1
Item1a_Topic_23	0.701	1	0.458	0	1
Item1a_Topic_24	0.466	0	0.499	0	1
Item1a_Topic_25	0.848	1	0.359	1	1

Table 7 (Continued).

Item1a_Topic_26	0.313	0	0.464	0	1
Item1a_Topic_27	0.637	1	0.481	0	1
Item1a_Topic_28	0.575	1	0.494	0	1
Item1a_Topic_29	0.653	1	0.476	0	1
Item7_Topic_0	0.799	1	0.401	1	1
Item7_Topic_1	0.985	1	0.122	1	1
Item7_Topic_2	0.957	1	0.204	1	1
Item7_Topic_3	0.785	1	0.411	1	1
Item7_Topic_4	0.737	1	0.44	0	1
Item7_Topic_5	0.791	1	0.407	1	1
Item7_Topic_6	0.862	1	0.345	1	1
Item7_Topic_7	0.895	1	0.307	1	1
Item7_Topic_8	0.454	0	0.498	0	1
Item7_Topic_9	0.854	1	0.353	1	1

Table 7 (Continued).

Item7_Topic_10	0.779	1	0.415	1	1
Item7_Topic_11	0.909	1	0.288	1	1
Item7_Topic_12	0.878	1	0.327	1	1
Item7_Topic_13	0.3	0	0.458	0	1
Item7_Topic_14	0.732	1	0.443	0	1
Item7_Topic_15	0.803	1	0.397	1	1
Item7_Topic_16	0.623	1	0.485	0	1
Item7_Topic_17	0.394	0	0.489	0	1
Item7_Topic_18	0.687	1	0.464	0	1
Item7_Topic_19	0.934	1	0.248	1	1
Item7_Topic_20	0.742	1	0.437	0	1
Logsale	3.355	3.508	2.334	2.145	4.758
Zscore	0.883	0.691	0.889	0.218	1.301
EXCHCD	0.108	0	0.356	0	0

Table 7 (Continued).

DFT	0.074	0	0.261	0	0
Big4	0.418	0	0.493	0	1
Leverage	0.256	0.142	0.365	0.004	0.366
Accurate Going Concern	0.721	1	0.449	0	1

Panel B Descriptive Statics for Accurate Going Concern Evaluation Models by Groups (Type I Error)

	Accurate Going Concern Evaluation = 0 (N=664)					**Accurate Going Concern Evaluation= 1 (N=1,712)**					
	Mean	**Median**	**Sd**	**p25**	**p75**	**Mean**	**Median**	**Sd**	**p25**	**p75**	**t test**
Item1a_Topic_0	0.919	1	0.274	1	1	0.913	1	0.282	1	1	0.00571
Item1a_Topic_1	0.923	1	0.266	1	1	0.926	1	0.262	1	1	-0.0026
Item1a_Topic_2	0.791	1	0.407	1	1	0.733	1	0.442	0	1	0.0576**
Item1a_Topic_3	0.486	0	0.5	0	1	0.415	0	0.493	0	1	0.0717**
Item1a_Topic_4	0.675	1	0.469	0	1	0.679	1	0.467	0	1	-0.004
Item1a_Topic_5	0.39	0	0.488	0	1	0.393	0	0.489	0	1	-0.0031

Table 7 (Continued).

Item1a_Topic_6	0.396	0	0.489	0	1	0.296	0	0.456	0	1	0.101***
Item1a_Topic_7	0.438	0	0.497	0	1	0.352	0	0.478	0	1	0.0860***
Item1a_Topic_8	0.863	1	0.344	1	1	0.811	1	0.391	1	1	0.0516**
Item1a_Topic_9	0.667	1	0.472	0	1	0.671	1	0.47	0	1	-0.004
Item1a_Topic_10	0.795	1	0.404	1	1	0.689	1	0.463	0	1	0.107***
Item1a_Topic_11	0.821	1	0.384	1	1	0.783	1	0.413	1	1	0.0381*
Item1a_Topic_12	0.608	1	0.488	0	1	0.575	1	0.494	0	1	0.0331
Item1a_Topic_13	0.779	1	0.415	1	1	0.772	1	0.42	1	1	0.007
Item1a_Topic_14	0.473	0	0.5	0	1	0.395	0	0.489	0	1	0.0780***
Item1a_Topic_15	0.771	1	0.42	1	1	0.796	1	0.403	1	1	-0.0245
Item1a_Topic_16	0.815	1	0.389	1	1	0.703	1	0.457	0	1	0.112***
Item1a_Topic_17	0.907	1	0.291	1	1	0.797	1	0.402	1	1	0.109***
Item1a_Topic_18	0.536	1	0.499	0	1	0.436	0	0.496	0	1	0.0998***
Item1a_Topic_19	0.717	1	0.451	0	1	0.662	1	0.473	0	1	0.0545*
Item1a_Topic_20	0.706	1	0.456	0	1	0.564	1	0.496	0	1	0.142***

Table 7 (Continued).

Item1a_Topic_21	0.78	1	0.414	1	1	0.707	1	0.455	0	1	0.0733***
Item1a_Topic_22	0.634	1	0.482	0	1	0.532	1	0.499	0	1	0.102***
Item1a_Topic_23	0.667	1	0.472	0	1	0.714	1	0.452	0	1	-0.0472*
Item1a_Topic_24	0.527	1	0.5	0	1	0.443	0	0.497	0	1	0.0844***
Item1a_Topic_25	0.866	1	0.341	1	1	0.841	1	0.366	1	1	0.0254
Item1a_Topic_26	0.37	0	0.483	0	1	0.291	0	0.454	0	1	0.0796***
Item1a_Topic_27	0.729	1	0.445	0	1	0.602	1	0.49	0	1	0.127***
Item1a_Topic_28	0.599	1	0.49	0	1	0.566	1	0.496	0	1	0.0334
Item1a_Topic_29	0.762	1	0.426	1	1	0.61	1	0.488	0	1	0.152***
Item7_Topic_0	0.764	1	0.425	1	1	0.813	1	0.39	1	1	-0.0495**
Item7_Topic_1	0.988	1	0.109	1	1	0.984	1	0.127	1	1	0.00431
Item7_Topic_2	0.968	1	0.175	1	1	0.952	1	0.214	1	1	0.0163
Item7_Topic_3	0.788	1	0.409	1	1	0.783	1	0.412	1	1	0.00436
Item7_Topic_4	0.762	1	0.426	1	1	0.727	1	0.446	0	1	0.0348
Item7_Topic_5	0.86	1	0.347	1	1	0.765	1	0.424	1	1	0.0953***

Table 7 (Continued).

Item7_Topic_6	0.861	1	0.346	1	1	0.863	1	0.344	1	1	-0.00129
Item7_Topic_7	0.91	1	0.287	1	1	0.889	1	0.314	1	1	0.0206
Item7_Topic_8	0.532	1	0.499	0	1	0.423	0	0.494	0	1	0.108***
Item7_Topic_9	0.878	1	0.328	1	1	0.845	1	0.362	1	1	0.0328*
Item7_Topic_10	0.761	1	0.427	1	1	0.786	1	0.411	1	1	-0.0251
Item7_Topic_11	0.928	1	0.259	1	1	0.901	1	0.298	1	1	0.0264*
Item7_Topic_12	0.898	1	0.303	1	1	0.871	1	0.335	1	1	0.0267
Item7_Topic_13	0.288	0	0.453	0	1	0.304	0	0.46	0	1	-0.0167
Item7_Topic_14	0.694	1	0.461	0	1	0.747	1	0.435	0	1	-0.0528**
Item7_Topic_15	0.756	1	0.43	1	1	0.822	1	0.383	1	1	-0.0658***
Item7_Topic_16	0.735	1	0.442	0	1	0.579	1	0.494	0	1	0.156***
Item7_Topic_17	0.452	0	0.498	0	1	0.371	0	0.483	0	1	0.0809***
Item7_Topic_18	0.693	1	0.462	0	1	0.685	1	0.465	0	1	0.00761
Item7_Topic_19	0.944	1	0.23	1	1	0.93	1	0.254	1	1	0.0138
Item7_Topic_20	0.758	1	0.429	1	1	0.737	1	0.441	0	1	0.0210

Table 7 (Continued).

Logsale	2.9	3.084	2.541	1.452	4.433	5.433	5.713	2.132	4.352	6.724	-0.631***
Zscore	0.765	0.485	0.966	0.062	1.234	1.019	0.691	1.023	0.288	1.487	-0.164***
EXCHCD	0.107	0	0.387	0	0	0.362	0	0.482	0	1	-0.00172
DFT	0.064	0	0.244	0	0	0.064	0	0.244	0	0	0.0357**
Big4	0.434	0	0.496	0	1	0.63	1	0.484	0	1	0.0214
Leverage	0.368	0.24	0.45	0.057	0.524	0.793	0.671	0.652	0.431	1.011	0.155***

Table 7 Panel A presents the descriptive statistics Model [3] and Model [4], which examines which topics in Item 1A and Item 7 are associated with the likelihood of accurate going concern evaluation in the Type I error setting. $Item1A\_Topic\_N$ equals 1 if a client disclosure topic N in its Item1A disclosures and 0 otherwise. $Item7\_TopicN$ equals 1 if a client disclosure topic N in its Item7 disclosures and 0 otherwise. $LogSale_{it}$ is the natural log of sales. $Zscore_{it}$ is the Bankruptcy score calculated by Altman (1968). *Big4* equals 1 if a client is audited by big 4 auditors and 0 otherwise. *EXCHCD* equals 1 if a client is listed in New York Stock Exchange and 0 otherwise. *DFT* equals 1 if a client is under technical or payment default and 0 otherwise. *Leverage* is calculated as total liabilities divided by total assets. Panel B represents the descriptive statistics for Type I error clients and accurate going evaluation clients separately. The Accurate Going Concern Evaluation = 0 group includes Type I error clients and Accurate Going Concern Evaluation = 1 group includes 1) accurte going concern opinion clients and 2) clients with no going concern opinions and do not file for bankruptcy in the next year. This table is produced based on a matched sample, which the Type I error clients are matched with accurate going concern evaluation clients by SIC and size (the three closest sizes). T-test are shown in the last column. *** p<0.01, ** p<0.05, * p<0.1.

Table 8 Regression Results of Model [3] and Model [4]

VARIABLES	**1** **PROB(*Accurate Going Concern Evaluation*)**	**2** **PROB(*Accurate Going Concern Evaluation*)**
Item1a_Topic_0	**0.509*****	
	(2.969)	
Item1a_Topic_1	**0.808*****	
	(4.776)	
Item1a_Topic_12	**0.171****	
	(2.316)	
Item1a_Topic_15	**0.163***	
	(1.847)	
Item1a_Topic_16	**-0.240*****	
	(-2.656)	
Item1a_Topic_17	**-0.434*****	
	(-3.529)	
Item1a_Topic_20	**-0.359*****	
	(-4.863)	
Item1a_Topic_21	**-0.244*****	
	(-2.961)	
Item1a_Topic_27	**-0.258*****	
	(-3.596)	
Item1a_Topic_29	**-0.292*****	
	(-3.864)	
Item1a_Topic_2	0.0929	
	(1.047)	
Item1a_Topic_3	0.0948	
	(1.007)	
Item1a_Topic_4	-0.108	

Table 8 (Continued).

	(-1.358)
Item1a_Topic_5	-0.126
	(-1.638)
Item1a_Topic_6	-0.183
	(-1.622)
Item1a_Topic_7	0.00122
	(0.0113)
Item1a_Topic_8	-0.0927
	(-0.841)
Item1a_Topic_9	0.0798
	(0.989)
Item1a_Topic_10	-0.0233
	(-0.260)
Item1a_Topic_11	-0.0942
	(-1.020)
Item1a_Topic_13	0.127
	(1.423)
Item1a_Topic_14	-0.0363
	(-0.405)
Item1a_Topic_18	-0.0986
	(-1.108)
Item1a_Topic_19	0.017
	(0.182)
Item1a_Topic_22	-0.0389
	(-0.518)
Item1a_Topic_23	-0.095
	(-1.110)
Item1a_Topic_24	0.053

Table 8 (Continued).

	(0.64)	
Item1a_Topic_25	0.0566	
	(0.521)	
Item1a_Topic_26	0.0389	
	(0.403)	
Item1a_Topic_28	0.025	
	(0.354)	
Item7_Topic_3		**-0.144***
		(-1.739)
Item7_Topic_5		**-0.217*****
		(-2.682)
Item7_Topic_16		**-0.301*****
		(-4.309)
Item7_Topic_0		-0.0103
		(-0.112)
Item7_Topic_1		0.496
		(1.467)
Item7_Topic_2		-0.0624
		(-0.325)
Item7_Topic_4		-0.0972
		(-1.297)
Item7_Topic_6		0.0619
		(0.641)
Item7_Topic_7		-0.0762
		(-0.704)
Item7_Topic_8		-0.131
		(-1.639)
Item7_Topic_9		-0.0542

Table 8 (Continued).

		(-0.595)
Item7_Topic_10		-0.0461
		(-0.570)
Item7_Topic_11		0.0791
		(0.685)
Item7_Topic_12		-0.139
		(-1.398)
Item7_Topic_13		-0.0365
		(-0.502)
Item7_Topic_14		0.0393
		(0.53)
Item7_Topic_15		0.123
		(1.581)
Item7_Topic_17		0.0967
		(1.195)
Item7_Topic_18		0.0148
		(0.228)
Item7_Topic_19		0.0326
		(0.237)
Item7_Topic_20		0.0871
		(1.199)
Logsale	0.122***	0.115***
	(6.106)	(6.003)
Zscore	-0.0799*	-0.0685
	(-1.790)	(-1.582)
EXCHCD	-0.163*	-0.161*
	(-1.810)	(-1.822)
DFT	-0.159	-0.176*

Table 8 (Continued).

	(-1.476)	(-1.661)
Big4	-0.194***	-0.112
	(-2.730)	(-1.630)
Leverage	-0.610***	-0.658***
	(-7.521)	(-8.242)
Constant	37.99**	42.00***
	(2.564)	(3.153)
Observations	2,376	2,376
Industry Fixed Effect	Yes	Yes
Year Fixed Effect	Yes	Yes
Pseudo R2	0.111	0.0759

Table 8 presents the regression results of Model [3] and Model [4], which test which topics disclosed in either Item 1A or Item 7 sections are associated with the likelihood of accurate going concern evaluation in a Type I error setting. *Accurate Going Concern Evaluation* equals 1 if a client 1) receives a going concern opinion in the current year and files for bankruptcy protection in the subsequent year, or 2) does not receive a going concern opinion in the current period and does not file for bankruptcy protection in the subsequent year. *Accurate Going Concern Evaluation* equals 0 if a client receives a going concern opinion in the current year and does not file for bankruptcy protection in the subsequent year (Type I error). $Item1A\_Topic\_N$ equals 1 if a client disclosure topic N in its Item1A disclosures and 0 otherwise. $Item7\_TopicN$ equals 1 if a client disclosure topic N in its Item7 disclosures and 0 otherwise. $LogSale_{it}$ is the natural log of sales. $Zscore_{it}$ is the Bankruptcy score calculated by Altman (1968). *Big4* equals 1 if a client is audited by big 4 auditors and 0 otherwise. *EXCHCD* equals 1 if a client is listed in New York Stock Exchange and 0 otherwise. *DFT* equals 1 if a client is under technical or payment default and 0 otherwise. *Leverage* is calculated as total liabilities divided by total assets. This regression is produced based on a matched sample, which the Type I error clients are matched with accurate going concern evaluation clients by SIC and size (the three closest sizes). Z-scores are shown in the last column. *** p<0.01, ** p<0.05, * p<0.1.

Table 9 Industry Distribution of Model [5] and Model [6]

SIC Codes	Division	No. of Observations
0100-0999	Agriculture, Forestry, and Fishing	0
1000-1499	Mining	61
1500-1799	Construction	0
2000-3999	Manufacturing	90
4000-4999	Transportation, Communications, Electric, Gas and Sanitary Service	27
5000-5199	Wholesale Trade	6
5200-5999	Retail Trade	34
6000-6799	Finance, Insurance and Real Estate	0
7000-8999	Services	16
9100-9729	Public Administration	0
9900-9999	Non-Classifiable	0
Total No. of Observations		234

Table 9 presents the results of industry distribution of Model [5] and Model [6], which tests what topics in Item 1A and Item 7 are associated with the likelihood of accurate going concern opinion in Type II error settings. The sample includes accurate going concern opinion clients and Type II going concern error clients. Observations with missing control variables are eliminated.

Table 10 Descriptive Statistics for Model[5] and Model[6]

Panel A Descriptive Statistics for Accurate Going Concern Opinion Models (Type II Error) (N=234)

	Mean	**Median**	**Sd**	**p25**	**p75**
Item1a_Topic_0	0.299	0	0.459	0	1
Item1a_Topic_1	0.744	1	0.438	0	1
Item1a_Topic_2	0.859	1	0.349	1	1
Item1a_Topic_3	0.919	1	0.274	1	1
Item1a_Topic_4	0.722	1	0.449	0	1
Item1a_Topic_5	0.923	1	0.267	1	1
Item1a_Topic_6	0.85	1	0.357	1	1
Item1a_Topic_7	0.838	1	0.37	1	1
Item1a_Topic_8	0.432	0	0.496	0	1
Item1a_Topic_9	0.692	1	0.463	0	1
Item1a_Topic_10	0.491	0	0.501	0	1

Table 10 (Continued).

Item1a_Topic_11	0.897	1	0.304	1	1
Item1a_Topic_12	0.765	1	0.425	1	1
Item1a_Topic_13	0.778	1	0.417	1	1
Item1a_Topic_14	0.855	1	0.353	1	1
Item1a_Topic_15	0.765	1	0.425	1	1
Item1a_Topic_16	0.701	1	0.459	0	1
Item1a_Topic_17	0.355	0	0.479	0	1
Item1a_Topic_18	0.85	1	0.357	1	1
Item1a_Topic_19	0.641	1	0.481	0	1
Item1a_Topic_20	0.632	1	0.483	0	1
Item1a_Topic_21	0.85	1	0.357	1	1
Item1a_Topic_22	0.551	1	0.498	0	1
Item1a_Topic_23	0.671	1	0.471	0	1
Item1a_Topic_24	0.782	1	0.414	1	1
Item1a_Topic_25	0.474	0	0.5	0	1

Table 10 (Continued).

Item1a_Topic_26	0.893	1	0.31	1	1
Item1a_Topic_27	0.265	0	0.442	0	1
Item1a_Topic_28	0.671	1	0.471	0	1
Item1a_Topic_29	0.47	0	0.5	0	1
Item7_Topic_0	0.906	1	0.292	1	1
Item7_Topic_1	0.812	1	0.392	1	1
Item7_Topic_2	0.923	1	0.267	1	1
Item7_Topic_3	0.97	1	0.171	1	1
Item7_Topic_4	0.949	1	0.221	1	1
Item7_Topic_5	0.842	1	0.366	1	1
Item7_Topic_6	0.953	1	0.212	1	1
Item7_Topic_7	0.85	1	0.357	1	1
Item7_Topic_8	0.786	1	0.411	1	1
Item7_Topic_9	0.944	1	0.23	1	1
Item7_Topic_10	0.791	1	0.408	1	1

Table 10 (Continued).

Item7_Topic_11	0.953	1	0.212	1	1
Item7_Topic_12	0.791	1	0.408	1	1
Item7_Topic_13	0.795	1	0.405	1	1
Item7_Topic_14	0.932	1	0.253	1	1
Item7_Topic_15	0.915	1	0.28	1	1
Item7_Topic_16	0.316	0	0.466	0	1
Item7_Topic_17	0.855	1	0.353	1	1
Item7_Topic_18	0.466	0	0.5	0	1
Item7_Topic_19	0.547	1	0.499	0	1
Item7_Topic_20	0.419	0	0.494	0	1
Logsale	5.649	5.94	2.155	4.47	7.031
Zscore	1.047	0.719	1	0.315	1.565
EXCHCD	0.393	0	0.489	0	1
DFT	0.158	0	0.366	0	0
Big4	0.641	1	0.481	0	1

Table 10 (Continued).

Leverage	0.702	0.605	0.582	0.368	0.886
Banklag	144.821	85	146.832	71	105
Reportlag	86.701	75	55.183	67	90
Accurate Going Concern Opinion	0.59	1	0.493	0	1

Panel B Descriptive Statics for Accurate Going Concern Opinion Models (Type II Error) by Groups (Model [5] -Model [6])

	Accurate Going Concern Opinion = 0 (N=96)					**Accurate Going Concern Opinion= 1 (N=138)**					
	Mean	**Median**	**Sd**	**p25**	**p75**	**Mean**	**Median**	**Sd**	**p25**	**p75**	**t test**
Item1a_Topic_0	0.344	0	0.477	0	1	0.268	0	0.445	0	1	0.0756
Item1a_Topic_1	0.719	1	0.452	0	1	0.761	1	0.428	1	1	-0.0421
Item1a_Topic_2	0.844	1	0.365	1	1	0.87	1	0.338	1	1	-0.0258
Item1a_Topic_3	0.917	1	0.278	1	1	0.92	1	0.272	1	1	-0.0036

Table 10 (Continued).

Item1a_Topic_4	0.75	1	0.435	0	1	0.703	1	0.459	0	1	0.0471
Item1a_Topic_5	0.927	1	0.261	1	1	0.92	1	0.272	1	1	0.00679
Item1a_Topic_6	0.854	1	0.355	1	1	0.848	1	0.36	1	1	0.00634
Item1a_Topic_7	0.823	1	0.384	1	1	0.848	1	0.36	1	1	-0.0249
Item1a_Topic_8	0.229	0	0.423	0	0	0.572	1	0.497	0	1	-0.343***
Item1a_Topic_9	0.677	1	0.47	0	1	0.703	1	0.459	0	1	-0.0258
Item1a_Topic_10	0.479	0	0.502	0	1	0.5	1	0.502	0	1	-0.0208
Item1a_Topic_11	0.865	1	0.344	1	1	0.92	1	0.272	1	1	-0.0557
Item1a_Topic_12	0.781	1	0.416	1	1	0.754	1	0.432	1	1	0.0276
Item1a_Topic_13	0.781	1	0.416	1	1	0.775	1	0.419	1	1	0.00589
Item1a_Topic_14	0.844	1	0.365	1	1	0.862	1	0.346	1	1	-0.0186

Table 10 (Continued).

Item1a_Topic_15	0.729	1	0.447	0	1	0.79	1	0.409	1	1	-0.0607
Item1a_Topic_16	0.719	1	0.452	0	1	0.688	1	0.465	0	1	0.0303
Item1a_Topic_17	0.365	0	0.484	0	1	0.348	0	0.478	0	1	0.0168
Item1a_Topic_18	0.823	1	0.384	1	1	0.87	1	0.338	1	1	-0.0466
Item1a_Topic_19	0.677	1	0.47	0	1	0.616	1	0.488	0	1	0.0611
Item1a_Topic_20	0.625	1	0.487	0	1	0.638	1	0.482	0	1	-0.0127
Item1a_Topic_21	0.823	1	0.384	1	1	0.87	1	0.338	1	1	-0.0466
Item1a_Topic_22	0.583	1	0.496	0	1	0.529	1	0.501	0	1	0.0543
Item1a_Topic_23	0.698	1	0.462	0	1	0.652	1	0.478	0	1	0.0457
Item1a_Topic_24	0.76	1	0.429	1	1	0.797	1	0.404	1	1	-0.0367
Item1a_Topic_25	0.542	1	0.501	0	1	0.428	0	0.497	0	1	0.114

Table 10 (Continued).

Item1a_Topic_26	0.885	1	0.32	1	1	0.899	1	0.303	1	1	-0.0131
Item1a_Topic_27	0.271	0	0.447	0	1	0.261	0	0.441	0	1	0.00996
Item1a_Topic_28	0.76	1	0.429	1	1	0.609	1	0.49	0	1	0.152*
Item1a_Topic_29	0.531	1	0.502	0	1	0.428	0	0.497	0	1	0.104
Item7_Topic_0	0.927	1	0.261	1	1	0.891	1	0.312	1	1	0.0358
Item7_Topic_1	0.792	1	0.408	1	1	0.826	1	0.38	1	1	-0.0344
Item7_Topic_2	0.938	1	0.243	1	1	0.913	1	0.283	1	1	0.0245
Item7_Topic_3	0.969	1	0.175	1	1	0.971	1	0.168	1	1	-0.0023
Item7_Topic_4	0.948	1	0.223	1	1	0.949	1	0.22	1	1	-0.0014
Item7_Topic_5	0.813	1	0.392	1	1	0.862	1	0.346	1	1	-0.0498
Item7_Topic_6	0.969	1	0.175	1	1	0.942	1	0.235	1	1	0.0267

Table 10 (Continued).

Item7_Topic_7	0.844	1	0.365	1	1	0.855	1	0.353	1	1	-0.0113
Item7_Topic_8	0.74	1	0.441	0	1	0.819	1	0.387	1	1	-0.0793
Item7_Topic_9	0.948	1	0.223	1	1	0.942	1	0.235	1	1	0.00589
Item7_Topic_10	0.802	1	0.401	1	1	0.783	1	0.414	1	1	0.0195
Item7_Topic_11	0.958	1	0.201	1	1	0.949	1	0.22	1	1	0.00906
Item7_Topic_12	0.781	1	0.416	1	1	0.797	1	0.404	1	1	-0.0159
Item7_Topic_13	0.771	1	0.423	1	1	0.812	1	0.392	1	1	-0.0408
Item7_Topic_14	0.927	1	0.261	1	1	0.935	1	0.248	1	1	-0.0077
Item7_Topic_15	0.948	1	0.223	1	1	0.891	1	0.312	1	1	0.0566
Item7_Topic_16	0.333	0	0.474	0	1	0.304	0	0.462	0	1	0.029
Item7_Topic_17	0.875	1	0.332	1	1	0.841	1	0.367	1	1	0.0344

Table 10 (Continued).

Item7_Topic_18	0.385	0	0.489	0	1	0.522	1	0.501	0	1	-0.136*
Item7_Topic_19	0.365	0	0.484	0	1	0.674	1	0.47	0	1	-0.309***
Item7_Topic_20	0.438	0	0.499	0	1	0.406	0	0.493	0	1	0.0317
Logsale	5.96	6.295	2.161	4.997	7.169	5.433	5.713	2.132	4.352	6.724	0.527
Zscore	1.088	0.771	0.971	0.328	1.733	1.019	0.691	1.023	0.288	1.487	0.0693
Big4	0.438	0	0.499	0	1	0.362	0	0.482	0	1	0.0752
EXCHCD	0.156	0	0.365	0	0	0.159	0	0.367	0	0	-0.0032
DFT	0.656	1	0.477	0	1	0.63	1	0.484	0	1	0.0258
Leverage	0.571	0.52	0.433	0.3	0.721	0.793	0.671	0.652	0.431	1.011	-0.222**
Banklag	146.802	72	151.261	59.5	101	143.442	90	144.211	75	105	3.36
Reportlag	70	68	18.114	58	75	98.319	89	67.971	75	101	-28.32***

Table 10 Panel A presents the descriptive statistics Model [5] and Model [6], which examines which topics in Item 1A and Item 7 are associated with the

likelihood of accurate going concern opinion in the Type II error setting. $Item1A\_Topic\_N$ equals 1 if a client disclosure topic N in its Item1A disclosures and 0 otherwise. $Item7\_TopicN$ equals 1 if a client disclosure topic N in its Item7 disclosures and 0 otherwise. $LogSale_{it}$ is the natural log of sales. $Zscore_{it}$ is the Bankruptcy score calculated by Altman (1968). *Big4* equals 1 if a client is audited by big 4 auditors and 0 otherwise. *EXCHCD* equals 1 if a client is listed in New York Stock Exchange and 0 otherwise. *DFT* equals 1 if a client is under technical or payment default and 0 otherwise. *Leverage* is calculated as total liabilities divided by total assets. *Banklag* equals the difference between audit report dates and bankruptcy dates. *Reportlag* equals the difference between fiscal year end dates and audit report dates. Panel B represents the descriptive statistics for Type II error clients and accurate going concern clients separately. The Accurate Going Concern = 0 group includes Type II error clients and Accurate Going Concern = 1 group includes accurate going concern opinion clients. T-test are shown in the last column. *** p<0.01, ** p<0.05, * p<0.1.

Table 11 Regression Results of Model [5] and Model [6]

VARIABLES	**1** **PROB(*Accurate Going Concern Opinion*)**	**2** **PROB(*Accurate Going Concern Opinion*)**
Item1a_Topic_0	**-0.763****	
	(-2.471)	
Item1a_Topic_5	**-1.077***	
	(-1.689)	
Item1a_Topic_8	**1.341*****	
	(4.759)	
Item1a_Topic_11	**0.991****	
	(1.999)	
Item1a_Topic_29	**-0.951*****	
	(-2.819)	
Item1a_Topic_1	0.401	
	(1.124)	
Item1a_Topic_2	0.558	
	(1.234)	
Item1a_Topic_3	0.415	
	(0.616)	
Item1a_Topic_4	-0.605	
	(-1.579)	
Item1a_Topic_6	-0.801*	
	(-1.934)	
Item1a_Topic_7	0.265	
	(0.671)	
Item1a_Topic_9	0.184	
	(0.551)	
Item1a_Topic_10	0.0227	

Table 11 (Continued).

	(0.0773)
Item1a_Topic_12	-0.0568
	(-0.160)
Item1a_Topic_13	0.235
	(0.698)
Item1a_Topic_14	0.249
	(0.608)
Item1a_Topic_15	-0.484
	(-1.460)
Item1a_Topic_16	-0.0744
	(-0.251)
Item1a_Topic_17	-0.246
	(-0.661)
Item1a_Topic_18	-0.0223
	(-0.0559)
Item1a_Topic_19	-0.0132
	(-0.0452)
Item1a_Topic_20	0.275
	(0.864)
Item1a_Topic_21	0.292
	(0.755)
Item1a_Topic_22	-0.314
	(-1.019)
Item1a_Topic_23	0.17
	(0.505)
Item1a_Topic_24	0.0757
	(0.217)
Item1a_Topic_25	-0.166

Table 11 (Continued).

	(-0.486)	
Item1a_Topic_26	0.333	
	(0.691)	
Item1a_Topic_27	-0.158	
	(-0.503)	
Item1a_Topic_28	-0.425	
	(-1.414)	
Item7_Topic_1		**0.767***
		(1.93)
Item7_Topic_6		**-1.668***
		(-1.712)
Item7_Topic_8		**0.740****
		(2.499)
Item7_Topic_15		**-1.804****
		(-2.000)
Item7_Topic_19		**0.747*****
		(3.082)
Item7_Topic_0		-0.308
		(-0.568)
Item7_Topic_2		-0.529
		(-0.929)
Item7_Topic_3		0.886
		(0.668)
Item7_Topic_4		-0.352
		(-0.442)
Item7_Topic_5		0.556
		(1.482)
Item7_Topic_7		-0.0499

Table 11 (Continued).

		(-0.130)
Item7_Topic_9		0.503
		(0.676)
Item7_Topic_10		-0.114
		(-0.406)
Item7_Topic_11		0.0478
		(0.0458)
Item7_Topic_12		0.395
		(1.18)
Item7_Topic_13		0.101
		(0.316)
Item7_Topic_14		0.993
		(1.327)
Item7_Topic_16		0.17
		(0.533)
Item7_Topic_17		-0.457
		(-1.090)
Item7_Topic_18		0.302
		(1.312)
Item7_Topic_20		0.00379
		(0.0149)
Logsale	-0.148*	-0.0517
	(-1.648)	(-0.703)
Zscore	-0.0371	0.0495
	(-0.207)	(0.334)
EXCHCD	-0.199	-0.0511
	(-0.702)	(-0.189)
DFT	0.361	0.145

Table 11 (Continued).

	(1.103)	(0.455)
Big4	0.434	0.425
	(1.453)	(1.645)
Leverage	0.326	0.471**
	(1.252)	(2.102)
Banklag	-0.00213**	-0.000439
	(-2.118)	(-0.494)
Reportlag	0.0454***	0.0409***
	(5.602)	(5.778)
Constant	11.37	-40.01
	(0.171)	(-0.757)
Observations	234	234
Industry Fixed Effect	Yes	Yes
Year Fixed Effect	Yes	Yes
Pseudo R2	0.403	0.34

Table 11 presents the regression results of Model [5] and Model [6], which tests which topics disclosed in either Item 1A or Item 7 sections are associated with the likelihood of accurate going concern opinions in the Type II error setting. *Accurate Going Concern Opinion* equals 1 if a client receives a going concern opinion in the current year and files for bankruptcy protection in the subsequent year. *Accurate Going Concern Opinion* equals 0 if a client does not receive a going concern opinion in the current year and file for bankruptcy protection in the subsequent year (Type II error). $Item1A\_Topic\_N$ equals 1 if a client disclosure topic N in its Item1A disclosures and 0 otherwise. $Item7\_TopicN$ equals 1 if a client disclosure topic N in its Item7 disclosures and 0 otherwise. $LogSale_{it}$ is the natural log of sales. $Zscore_{it}$ is the Bankruptcy score calculated by Altman (1968). *Big4* equals 1 if a client is audited by big 4 auditors and 0 otherwise. *EXCHCD* equals 1 if a client is listed in New York Stock Exchange and 0 otherwise. *DFT* equals 1 if a client is under technical or payment default and 0 otherwise. *Leverage* is calculated as total liabilities divided by total assets. *Banklag* equals to the difference between audit report dates and bankruptcy dates. *Reportlag* equals to the difference between fiscal year end dates and audit report dates. Z-scores are shown in the last column. *** p<0.01, ** p<0.05, * p<0.1.

Table 12 Industry Distribution of Model [7] and Model [8]

SIC Codes	**Division**	**No. of Observations**
0100-0999	Agriculture, Forestry, and Fishing	0
1000-1499	Mining	112
1500-1799	Construction	0
2000-3999	Manufacturing	140
4000-4999	Transportation, Communications, Electric, Gas and Sanitary Service	28
5000-5199	Wholesale Trade	20
5200-5999	Retail Trade	78
6000-6799	Finance, Insurance and Real Estate	0
7000-8999	Services	12
9100-9729	Public Administration	0
9900-9999	Non-Classifiable	0
Total No. of Observations		390

Table 12 presents the results of industry distribution of Model [7] and Model [8], which tests what topics in Item 1A and Item 7 are associated with the likelihood of accurate going concern evaluation on in Type II error settings. The sample includes accurate going concern evaluation clients and Type II going concern error clients. Observations with missing control variables are eliminated.

Table 13 Descriptive Statistics for Model [7] and Model [8]

Panel A Descriptive Statistics for Accurate Going Concern Evaluation Models (Type II Error) (N=390)

	Mean	**Median**	**Sd**	**p25**	**p75**
Item1a_Topic_0	0.762	1	0.427	1	1
Item1a_Topic_1	0.874	1	0.332	1	1
Item1a_Topic_2	0.815	1	0.388	1	1
Item1a_Topic_3	0.618	1	0.487	0	1
Item1a_Topic_4	0.882	1	0.323	1	1
Item1a_Topic_5	0.226	0	0.419	0	0
Item1a_Topic_6	0.554	1	0.498	0	1
Item1a_Topic_7	0.759	1	0.428	1	1
Item1a_Topic_8	0.2	0	0.401	0	0
Item1a_Topic_9	0.344	0	0.476	0	1
Item1a_Topic_10	0.7	1	0.459	0	1
Item1a_Topic_11	0.91	1	0.286	1	1

Table 13 (Continued).

Item1a_Topic_12	0.733	1	0.443	0	1
Item1a_Topic_13	0.797	1	0.402	1	1
Item1a_Topic_14	0.718	1	0.451	0	1
Item1a_Topic_15	0.726	1	0.447	0	1
Item1a_Topic_16	0.592	1	0.492	0	1
Item1a_Topic_17	0.687	1	0.464	0	1
Item1a_Topic_18	0.595	1	0.492	0	1
Item1a_Topic_19	0.867	1	0.34	1	1
Item1a_Topic_20	0.49	0	0.501	0	1
Item1a_Topic_21	0.592	1	0.492	0	1
Item1a_Topic_22	0.767	1	0.423	1	1
Item1a_Topic_23	0.723	1	0.448	0	1
Item1a_Topic_24	0.597	1	0.491	0	1
Item1a_Topic_25	0.685	1	0.465	0	1
Item1a_Topic_26	0.654	1	0.476	0	1

Table 13 (Continued).

Item1a_Topic_27	0.867	1	0.34	1	1
Item1a_Topic_28	0.549	1	0.498	0	1
Item1a_Topic_29	0.559	1	0.497	0	1
Item7_Topic_0	0.956	1	0.204	1	1
Item7_Topic_1	0.859	1	0.348	1	1
Item7_Topic_2	0.928	1	0.258	1	1
Item7_Topic_3	0.879	1	0.326	1	1
Item7_Topic_4	0.908	1	0.29	1	1
Item7_Topic_5	0.315	0	0.465	0	1
Item7_Topic_6	0.795	1	0.404	1	1
Item7_Topic_7	0.826	1	0.38	1	1
Item7_Topic_8	0.844	1	0.364	1	1
Item7_Topic_9	0.813	1	0.391	1	1
Item7_Topic_10	0.577	1	0.495	0	1
Item7_Topic_11	0.908	1	0.29	1	1

Table 13 (Continued).

Item7_Topic_12	0.91	1	0.286	1	1
Item7_Topic_13	0.718	1	0.451	0	1
Item7_Topic_14	0.936	1	0.245	1	1
Item7_Topic_15	0.746	1	0.436	0	1
Item7_Topic_16	0.826	1	0.38	1	1
Item7_Topic_17	0.49	0	0.501	0	1
Item7_Topic_18	0.821	1	0.384	1	1
Item7_Topic_19	0.872	1	0.335	1	1
Item7_Topic_20	0.354	0	0.479	0	1
Logsale	6.043	6.337	2.005	5.003	7.186
Zscore	1.064	0.837	0.922	0.349	1.61
EXCHCD	0.456	0	0.499	0	1
DFT	0.108	0	0.31	0	0
Big4	0.726	1	0.447	0	1
Leverage	0.385	0.345	0.332	0.153	0.534

Table 13 (Continued).

Banklag	144.177	70	151.181	59	91
Reportlag	66.246	64	24.329	57	75
Accurate Going Concern Evaluation	0.754	1	0.431	1	1

Panel B Descriptive Statics for Accurate Going Concern Evaluation Models by Groups (Model [7] -Model [8])

	Accurate Going Concern Evaluation = 0 (N=96)					**Accurate Going Concern Evaluation= 1 (N=294)**					
	Mean	**Median**	**Sd**	**p25**	**p75**	**Mean**	**Median**	**Sd**	**p25**	**p75**	**t test**
Item1a_Topic_0	0.792	1	0.408	1	1	0.752	1	0.433	1	1	0.04
Item1a_Topic_1	0.927	1	0.261	1	1	0.857	1	0.351	1	1	0.0699
Item1a_Topic_2	0.823	1	0.384	1	1	0.813	1	0.391	1	1	0.00999
Item1a_Topic_3	0.635	1	0.484	0	1	0.612	1	0.488	0	1	0.0232
Item1a_Topic_4	0.885	1	0.32	1	1	0.881	1	0.324	1	1	0.00446
Item1a_Topic_5	0.26	0	0.441	0	1	0.214	0	0.411	0	0	0.0461
Item1a_Topic_6	0.531	1	0.502	0	1	0.561	1	0.497	0	1	-0.03
Item1a_Topic_7	0.781	1	0.416	1	1	0.752	1	0.433	1	1	0.0295

Table 13 (Continued).

Item1a_Topic_8	0.25	0	0.435	0	1	0.184	0	0.388	0	0	0.0663
Item1a_Topic_9	0.385	0	0.489	0	1	0.33	0	0.471	0	1	0.0555
Item1a_Topic_10	0.823	1	0.384	1	1	0.66	1	0.475	0	1	0.163**
Item1a_Topic_11	0.927	1	0.261	1	1	0.905	1	0.294	1	1	0.0223
Item1a_Topic_12	0.75	1	0.435	0	1	0.728	1	0.446	0	1	0.0221
Item1a_Topic_13	0.813	1	0.392	1	1	0.793	1	0.406	1	1	0.02
Item1a_Topic_14	0.729	1	0.447	0	1	0.714	1	0.453	0	1	0.0149
Item1a_Topic_15	0.698	1	0.462	0	1	0.735	1	0.442	0	1	-0.0368
Item1a_Topic_16	0.615	1	0.489	0	1	0.585	1	0.494	0	1	0.0295
Item1a_Topic_17	0.76	1	0.429	1	1	0.663	1	0.473	0	1	0.0972
Item1a_Topic_18	0.656	1	0.477	0	1	0.575	1	0.495	0	1	0.0814
Item1a_Topic_19	0.896	1	0.307	1	1	0.857	1	0.351	1	1	0.0387
Item1a_Topic_20	0.469	0	0.502	0	1	0.497	0	0.501	0	1	-0.0278
Item1a_Topic_21	0.625	1	0.487	0	1	0.582	1	0.494	0	1	0.0434
Item1a_Topic_22	0.771	1	0.423	1	1	0.765	1	0.425	1	1	0.00553

Table 13 (Continued).

Item1a_Topic_23	0.74	1	0.441	0	1	0.718	1	0.451	0	1	0.0219
Item1a_Topic_24	0.583	1	0.496	0	1	0.602	1	0.49	0	1	-0.0187
Item1a_Topic_25	0.656	1	0.477	0	1	0.694	1	0.462	0	1	-0.0376
Item1a_Topic_26	0.635	1	0.484	0	1	0.66	1	0.475	0	1	-0.0244
Item1a_Topic_27	0.896	1	0.307	1	1	0.857	1	0.351	1	1	0.0387
Item1a_Topic_28	0.573	1	0.497	0	1	0.541	1	0.499	0	1	0.0321
Item1a_Topic_29	0.615	1	0.489	0	1	0.541	1	0.499	0	1	0.0738
Item7_Topic_0	0.969	1	0.175	1	1	0.952	1	0.213	1	1	0.0164
Item7_Topic_1	0.927	1	0.261	1	1	0.837	1	0.37	1	1	0.0903*
Item7_Topic_2	0.948	1	0.223	1	1	0.922	1	0.269	1	1	0.0261
Item7_Topic_3	0.938	1	0.243	1	1	0.861	1	0.347	1	1	0.0770*
Item7_Topic_4	0.969	1	0.175	1	1	0.888	1	0.316	1	1	0.0810*
Item7_Topic_5	0.344	0	0.477	0	1	0.306	0	0.462	0	1	0.0376
Item7_Topic_6	0.823	1	0.384	1	1	0.786	1	0.411	1	1	0.0372
Item7_Topic_7	0.865	1	0.344	1	1	0.813	1	0.391	1	1	0.0517

Table 13 (Continued).

Item7_Topic_8	0.865	1	0.344	1	1	0.837	1	0.37	1	1	0.0278
Item7_Topic_9	0.823	1	0.384	1	1	0.81	1	0.393	1	1	0.0134
Item7_Topic_10	0.542	1	0.501	0	1	0.588	1	0.493	0	1	-0.0468
Item7_Topic_11	0.948	1	0.223	1	1	0.895	1	0.308	1	1	0.0534
Item7_Topic_12	0.938	1	0.243	1	1	0.901	1	0.299	1	1	0.0361
Item7_Topic_13	0.792	1	0.408	1	1	0.694	1	0.462	0	1	0.0978
Item7_Topic_14	0.958	1	0.201	1	1	0.929	1	0.258	1	1	0.0298
Item7_Topic_15	0.802	1	0.401	1	1	0.728	1	0.446	0	1	0.0742
Item7_Topic_16	0.813	1	0.392	1	1	0.83	1	0.376	1	1	-0.0174
Item7_Topic_17	0.51	1	0.503	0	1	0.483	0	0.501	0	1	0.0274
Item7_Topic_18	0.833	1	0.375	1	1	0.816	1	0.388	1	1	0.017
Item7_Topic_19	0.875	1	0.332	1	1	0.871	1	0.336	1	1	0.00425
Item7_Topic_20	0.333	0	0.474	0	1	0.361	0	0.481	0	1	-0.0272
Logsale	5.96	6.295	2.161	4.997	7.169	6.07	6.35	1.954	5.003	7.186	-0.109
Zscore	1.088	0.771	0.971	0.328	1.733	1.056	0.871	0.908	0.358	1.558	0.0323

Table 13 (Continued).

EXCHCD	0.438	0	0.499	0	1	0.463	0	0.499	0	1	-0.0251
DFT	0.156	0	0.365	0	0	0.092	0	0.289	0	0	0.0644
Big4	0.656	1	0.477	0	1	0.748	1	0.435	0	1	-0.092
Leverage	0.571	0.52	0.433	0.3	0.721	0.324	0.296	0.265	0.109	0.467	0.247***
Banklag	146.802	72	151.261	59.5	101	143.32	69	151.403	58	90	3.482
Reportlag	70	68	18.114	58	75	65.02	62	25.949	56	74	4.98

Table 13 Panel A presents the descriptive statistics Model [7] and Model [8], which examines which topics in Item 1A and Item 7 are associated with the likelihood of accurate going concern evaluation in the Type II error setting. $Item1A\_Topic\_N$ equals 1 if a client disclosure topic N in its Item1A disclosures and 0 otherwise. $Item7\_TopicN$ equals 1 if a client disclosure topic N in its Item7 disclosures and 0 otherwise. $LogSale_{it}$ is the natural log of sales. $Zscore_{it}$ is the Bankruptcy score calculated by Altman (1968). *Big4* equals 1 if a client is audited by big 4 auditors and 0 otherwise. *EXCHCD* equals 1 if a client is listed in New York Stock Exchange and 0 otherwise. *DFT* equals 1 if a client is under technical or payment default and 0 otherwise. *Leverage* is calculated as total liabilities divided by total assets. *Banklag* equals to the difference between audit report dates and bankruptcy dates. *Reportlag* equals the difference between fiscal year end dates and audit report date. Panel B represents the descriptive statistics for Type II error clients and accurate going concern evaluation clients separately. The Accurate Going Concern Evaluation= 0 group includes Type II error clients and Accurate Going Concern Evaluation = 1 group includes accurate going concern evaluation clients that 1) receive a going concern opinion in the current year and file for bankruptcy protection in the subsequent year, or 2) does not receive a going concern opinion in the current period and does not file for bankruptcy protection in the subsequent year. This table is produced based on a matched sample, which the Type II error clients are matched with accurate going concern evaluation clients by SIC and size (the three closest sizes). T-test are shown in the last column. *** p<0.01, ** p<0.05, * p<0.1

Table 14 Regression Results of Model [7] and Model [8]

VARIABLES	**1** **PROB(*Accurate Going Concern Evaluation*)**	**2** **PROB(*Accurate Going Concern Evaluation*)**
Item1a_Topic_4	**0.759****	
	(2.104)	
Item1a_Topic_6	**0.380***	
	(1.774)	
Item1a_Topic_8	**-0.638****	
	(-2.537)	
Item1a_Topic_9	**-0.472***	
	(-1.879)	
Item1a_Topic_24	**0.449****	
	(2.081)	
Item1a_Topic_0	0.134	
	(0.513)	
Item1a_Topic_1	-0.628	
	(-1.581)	
Item1a_Topic_2	-0.181	
	(-0.609)	
Item1a_Topic_3	0.0206	
	(0.0946)	
Item1a_Topic_5	-0.0163	
	(-0.0723)	
Item1a_Topic_7	-0.0266	
	(-0.111)	
Item1a_Topic_10	-0.413	
	(-1.564)	

Table 14 (Continued).

Item1a_Topic_11	-0.377
	(-0.880)
Item1a_Topic_12	0.165
	(0.73)
Item1a_Topic_13	0.182
	(0.73)
Item1a_Topic_14	0.0786
	(0.357)
Item1a_Topic_15	0.319
	(1.297)
Item1a_Topic_16	0.0641
	(0.321)
Item1a_Topic_17	0.0153
	(0.0657)
Item1a_Topic_18	-0.26
	(-1.180)
Item1a_Topic_19	-0.331
	(-0.972)
Item1a_Topic_20	0.0618
	(0.282)
Item1a_Topic_21	-0.148
	(-0.719)
Item1a_Topic_22	-0.157
	(-0.686)
Item1a_Topic_23	0.051
	(0.235)
Item1a_Topic_25	0.351
	(1.491)

Table 14 (Continued).

Item1a_Topic_26	0.0351	
	(0.169)	
Item1a_Topic_27	-0.167	
	(-0.531)	
Item1a_Topic_28	0.0567	
	(0.278)	
Item1a_Topic_29	0.0612	
	(0.313)	
Item7_Topic_1		**-0.946****
		(-2.274)
Item7_Topic_2		**1.100***
		(1.939)
Item7_Topic_4		**-1.501*****
		(-2.946)
Item7_Topic_0		0.547
		(0.744)
Item7_Topic_3		-0.373
		(-1.117)
Item7_Topic_5		-0.0286
		(-0.104)
Item7_Topic_6		-0.289
		(-1.084)
Item7_Topic_7		-0.248
		(-0.805)
Item7_Topic_8		-0.134
		(-0.479)
Item7_Topic_9		0.183
		(0.724)

Table 14 (Continued).

Item7_Topic_10		0.357
		(1.638)
Item7_Topic_11		-0.184
		(-0.371)
Item7_Topic_12		-0.0868
		(-0.217)
Item7_Topic_13		0.0253
		(0.112)
Item7_Topic_14		0.658
		(1.262)
Item7_Topic_15		-0.115
		(-0.498)
Item7_Topic_16		0.352
		(1.388)
Item7_Topic_17		0.0153
		(0.0841)
Item7_Topic_18		-0.0705
		(-0.257)
Item7_Topic_19		0.048
		(0.147)
Item7_Topic_20		-0.0824
		(-0.441)
Logsale	-0.0803	0.0508
	(-1.175)	(0.781)
Zscore	-0.271**	-0.332***
	(-2.286)	(-2.744)
EXCHCD	0.0921	0.0365
	(0.413)	(0.166)

Table 14 (Continued).

DFT	-0.275	-0.209
	(-1.095)	(-0.824)
Big4	0.35	0.333
	(1.612)	(1.607)
Leverage	-1.761***	-1.827***
	(-5.660)	(-5.994)
Banklag	-0.000136	-0.000487
	(-0.203)	(-0.764)
Reportlag	-0.0146**	-0.0133**
	(-2.241)	(-2.150)
Constant	82.70*	52.92
	(1.696)	(1.292)
Observations	390	390
Industry Fixed Effect	Yes	Yes
Year Fixed Effect	Yes	Yes
Pseudo R2	0.203	0.203

Table 14 presents the regression results of Model [7] and Model [8], which tests which topics disclosed in either Item 1A or Item 7 sections are associated with the likelihood of accurate going concern evaluation in a Type II error setting. *Accurate Going Concern Evaluation* equals 1 if a client 1) receives a going concern opinion in the current year and files for bankruptcy protection in the subsequent year, or 2) does not receive a going concern opinion in the current period and does not file for bankruptcy protection in the subsequent year. *Accurate Going Concern Evaluation* equals 0 if a client does not receive a going concern opinion in the current year and files for bankruptcy protection in the subsequent year (Type II error). $Item1A\_Topic\_N$ equals 1 if a client disclosure topic N in its Item1A disclosures and 0 otherwise. $Item7\_TopicN$ equals 1 if a client disclosure topic N in its Item7 disclosures and 0 otherwise. $LogSale_{it}$ is the natural log of sales. $Zscore_{it}$ is the Bankruptcy score calculated by Altman (1968). *Big4* equals 1 if a client is audited by big 4 auditors and 0 otherwise. *EXCHCD* equals 1 if a client is listed in New York Stock Exchange and 0 otherwise. *DFT* equals 1 if a client is under technical or payment default and 0 otherwise. *Leverage* is calculated as total liabilities divided by total assets. *Banklag* equals the difference between audit report dates and bankruptcy dates. *Reportlag* equals the difference between fiscal year end dates and audit report dates. This regression is produced based on a matched sample, which the Type II error clients are matched with accurate going concern evaluation clients by SIC and size (the three closest sizes). Z-scores are shown in the last column. *** p<0.01, ** p<0.05, * p<0.1.

Table 15 Descriptive Statistics of Model [9]

Panel A Descriptive Statistics for Accurate Going Concern Proxies

	Mean	**Median**	**sd**	**p25**	**p75**
Readability	10.071	10.3	1.939	9.745	10.92
Specificity	0.071	0.071	0.017	0.062	0.08
HardInformation	0.003	0.003	0.002	0.002	0.004
Tone	-0.018	-0.018	0.005	-0.022	-0.015
Accurate_Score_NB	0.235	0.098	0.283	0.027	0.353
Accurate_Score_SVM	0.399	0.392	0.07	0.343	0.448
Accurate_Score_RFC	0.02	0	0.107	0	0.002
Logsale	4.925	5.041	2.378	3.588	6.481
Zscore	0.91	0.734	0.847	0.367	1.214
EXCHCD	0.242	0	0.438	0	0
DFT	0.089	0	0.285	0	0
Big4	0.622	1	0.485	0	1

Table 15 (Continued).

Leverage	0.264	0.195	0.321	0.013	0.406
Reportlag	70.302	69	27.313	58	77
Banklag	114.908	72	120.474	60	88
Accurate Going Concern Opinion	0.01	0	0.101	0	0

Panel B Descriptive Statistics of Accurate Going Concern Proxies by groups

	Accurate Going Concern Opinion = 0 (N=4217)					**Accurate Going Concern Opinion= 1 (N=44)**					
	Mean	**Median**	**Sd**	**p25**	**p75**	**Mean**	**Median**	**Sd**	**p25**	**p75**	**t test**
Readability	10.071	10.3	1.941	9.745	10.92	10.043	10.21	1.741	9.518	10.88	0.03
Specificity	0.071	0.071	0.017	0.062	0.08	0.071	0.069	0.016	0.064	0.079	0.00
HardInformation	0.003	0.003	0.002	0.002	0.004	0.003	0.003	0.002	0.002	0.004	0.00
Tone	-0.018	-0.018	0.005	-0.022	-0.015	-0.022	-0.022	0.006	-0.026	-0.019	0.004***
Accurate_Score_NB	0.23	0.096	0.279	0.027	0.346	0.636	0.71	0.347	0.352	0.979	-0.405***
Accurate_Score_SVM	0.398	0.391	0.068	0.342	0.447	0.561	0.562	0.092	0.493	0.622	-0.163***

Table 15 (Continued).

Accurate_Score_RFC	0.015	0	0.088	0	0.001	0.45	0.356	0.43	0.015	0.976	-0.435***
Logsale	4.916	5.027	2.378	3.579	6.463	5.782	5.878	2.215	4.447	7.435	-0.866*
Zscore	0.908	0.733	0.846	0.367	1.213	1.085	0.84	0.884	0.43	1.759	-0.18
EXCHCD	0.24	0	0.437	0	0	0.409	0	0.497	0	1	-0.169*
DFT	0.088	0	0.283	0	0	0.182	0	0.39	0	0	-0.094*
Big4	0.622	1	0.485	0	1	0.659	1	0.479	0	1	-0.04
Leverage	0.259	0.192	0.31	0.012	0.399	0.77	0.603	0.756	0.398	0.949	-0.511***
Reportlag	69.904	69	24.968	58	76	108.5	90	106.165	75	95.5	-38.60***
Banklag	114.436	72	119.943	60	88	160.205	90	158.926	75.5	144	-45.77*

Table 15 Panel A presents the descriptive statistics Model [9], which examines the validity and effectiveness of the textual-based accurate going concern proxies generated by machine learning algorithms. *Accurate_Score_NB* is the probabilities of accurate going concern opinion that are calculated by Naives Bayes Classification. *Accurate_Score_SVM* is the probabilities of accurate going concern opinion that are calculated by Supporting Vector Machine. *Accurate_Score_RFC* is the probabilities of accurate going concern opinion that are calculated by Random Forest Classification. All three proxies are calculated based on combinations of Item 1A and Item 7 disclosures. *Readability* is the average of readabilities in Item 1A and Item 7 disclosures that are calculated by the Gunning (1952) Fog index. *Specificity* is the average of specificities in Item 1A and Item 7 disclosures that are calculated as the total number of entities (e.g., locations, people, and organizations) explicitly identified in the disclosure divided by the total number of words. *HandInformation* is the average quantitative information disclosed in Item 1A and Item 7 divided by the total number of words. *Tone* is the average of tone in Item 1A and Item 7 disclosures that are calculated as positive words minus negative words (Loughran and McDonald 2015) and divided by total number of words. $LogSale_{it}$ is the natural log of sales. $Zscore_{it}$ is the Bankruptcy score calculated by Altman (1968). *Big4* equals 1 if a client is audited by big 4 auditors and 0 otherwise. *EXCHCD* equals 1 if a client is listed in New York Stock Exchange and 0 otherwise. *DFT* equals 1 if a client is under technical or payment default and 0 otherwise. *Leverage* is calculated as total liabilities divided by total assets. *Banklag* equals the difference between audit report dates and bankruptcy dates. *Reportlag* equals the difference between

fiscal year end dates and audit report dates. Panel B represents the descriptive statistics for inaccurate going concern opinion clients and accurate going concern opinion clients separately. The Accurate Going Concern Opinion= 1 group includes accurate going concern opinion clients and Accurate Going Concern Opinion = 0 group includes clients that 1) do not receive going concern opinions in the current year and do not file for bankruptcy protection in the subsequent year, 2) have Type I going concern errors, 3) have Type II going concern errors. The results are based on the testing sample. T-test are shown in the last column. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 16 Empirical Testing for Model [9]

	1	2	3	4	5	6	7	8
VARIABLES	**PROB(Accurate Going Concern Opinion)**							
Readability		-0.00191						
		(-0.0580)						
Specificity			-0.954					
			(-0.263)					
HardInformation				-4.475				
				(-0.154)				
Tone					**-51.90*** **			
					(-3.766)			
Accurate_Score_NB						**1.214*** **		
						(5.221)		
Accurate_Score_SVM							**2.336*** **	
							(10.17)	
Accurate_Score_RFC								**11.70*** **
								(8.867)
Logsale	0.0657*	0.0657*	0.0658*	0.0656*	0.0394	0.00295	0.0318	-0.0719
	(1.684)	(1.683)	(1.685)	(1.678)	(0.954)	(0.0706)	(0.723)	(-1.496)
Zscore	0.0487	0.0487	0.049	0.0491	0.035	0.0528	0.0668	0.0995
	(0.76)	(0.76)	(0.764)	(0.765)	(0.508)	(0.91)	(0.971)	(1.574)
EXCHCD	0.149	0.149	0.148	0.149	0.137	0.0298	0.174	-0.133
	(0.97)	(0.97)	(0.962)	(0.971)	(0.873)	(0.178)	(0.981)	(-0.669)
DFT	0.301*	0.301*	0.299*	0.300*	0.343**	0.195	0.279	0.0495
	(1.734)	(1.732)	(1.724)	(1.728)	(1.967)	(1.042)	(1.367)	(0.221)
Big4	-0.0404	-0.0403	-0.0399	-0.0391	-0.0125	0.0801	-0.0299	0.138
	(-0.260)	(-0.259)	(-0.256)	(-0.251)	(-0.0781)	-0.481	(-0.169)	-0.727

Table 16 (Continued).

Leverage	0.648***	0.648***	0.650***	0.648***	0.680***	0.591***	0.508***	0.272*
	(6.232)	(6.231)	(6.232)	(6.233)	(6.433)	(5.301)	(4.221)	(1.706)
Reportlag	0.00626***	0.00627***	0.00627***	0.00626***	0.00549***	0.00602***	0.00484***	0.00298*
	(4.625)	(4.622)	(4.631)	(4.621)	(3.969)	(4.151)	(3.364)	(1.709)
Banklag	0.000509	0.000508	0.000505	0.000507	0.000603	0.000504	0.000668	0.00113**
	(1.007)	(1.003)	(0.997)	(1.002)	(1.179)	(0.951)	(1.2150	(1.961)
Constant	7.939	7.764	7.897	7.924	16.41	-4.264	2.619	-15.78
	(0.275)	(0.267)	(0.273)	(0.274)	(0.553)	(-0.138)	(0.078)	(-0.437)
Observations	4,261	4,261	4,261	4,261	4,261	4,261	4,261	4,261
Industry Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R2	0.178	0.178	0.178	0.178	0.21	0.238	0.375	0.41

Table 16 presents the regression results of Model [9], which examines the validity and effectiveness of the textual-based accurate going concern proxies generated by machine learning algorithms. *Accurate_Score_NB* is the probabilities of accurate going concern opinion that are calculated by Naives Bayes Classification. *Accurate_Score_SVM* is the probabilities of accurate going concern opinion that are calculated by Supporting Vector Machine. *Accurate_Score_RFC* is the probabilities of accurate going concern opinion that are calculated by Random Forest Classification. All three proxies are calculated based on combinations of Item 1A and Item 7 disclosures. *Readability* is the average of readabilities in Item 1A and Item 7 disclosures that are calculated by the Gunning (1952) Fog index. *Specificity* is the average of Specificities in Item 1A and Item 7 disclosures that are calculated as the total number of entities (e.g., locations, people, and organizations) explicitly identified in the disclosure divided by the total number of words. *HandInformation* is the average quantitative information disclosed in Item 1A and Item 7 divided by the total number of words. *Tone* is the average tone in Item 1A and Item 7 disclosures that are calculated as positive words minus negative words (Loughran and McDonald 2015) and divided by the total number of words. $LogSale_{it}$ is the natural log of sales. $Zscore_{it}$ is the Bankruptcy score calculated by Altman (1968). *Big4* equals 1 if a client is audited by big 4 auditors and 0 otherwise. *EXCHCD* equals 1 if a client is listed in New York Stock Exchange and 0 otherwise. *DFT* equals 1 if a client is under technical or payment default and 0 otherwise. *Leverage* is calculated as total liabilities divided by total assets. *Banklag* equals the difference between audit report dates and bankruptcy dates. *Reportlag* equals the difference between fiscal year end dates and audit report dates. The results are based on the testing sample, which has 200 Type I error observations, 26 Type II error observations, 44 accurate going concern observations, and 3,991 observations that have no going concern opinions and no subsequent bankruptcy. Z scores are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 17 Descriptive Statistics for Model [10]

Panel A Descriptive Statics for Type I Error Proxy Test (Model [10])

	Mean	**Median**	**sd**	**p25**	**p75**
Readability	10.082	10.335	1.963	9.775	10.945
Specificity	0.071	0.071	0.017	0.063	0.08
HardInformation	0.003	0.003	0.002	0.002	0.004
Tone	-0.018	-0.019	0.006	-0.022	-0.015
Type_I_Score_NB	0.349	0.247	0.257	0.146	0.523
Type_I_Score_SVM	0.443	0.431	0.058	0.399	0.481
Type_I_Score_RFC	0.09	0.004	0.221	0.001	0.034
Logsale	4.868	5.016	2.368	3.54	6.418
Zscore	0.911	0.727	0.831	0.355	1.214
EXCHCD	0.236	0	0.438	0	0
DFT	0.084	0	0.277	0	0
Big4	0.61	1	0.488	0	1

Table 17 (Continued).

Leverage	0.253	0.18	0.301	0.01	0.386
Type_I_Error	0.047	0	0.211	0	0

Panel B Descriptive Statics for Type I Error Proxy Testing (Model [10])

	Type I Error = 0 (N=4,075)					**Type I Error = 1 (N=199)**					
	Mean	**Median**	**sd**	**p25**	**p75**	**Mean**	**Median**	**sd**	**p25**	**p75**	**t-test**
Readability	10.082	10.335	1.964	9.775	10.940	10.083	10.340	1.947	9.675	10.980	-0.002
Specificity	0.071	0.071	0.017	0.063	0.080	0.071	0.072	0.018	0.063	0.081	0.000
HardInformation	0.003	0.003	0.002	0.002	0.004	0.003	0.003	0.002	0.002	0.004	0.000
Tone	-0.018	-0.019	0.006	-0.022	-0.015	-0.019	-0.019	0.006	-0.021	-0.016	0.001
Type_I_Score_NB	0.338	0.236	0.253	0.143	0.500	0.577	0.603	0.247	0.358	0.824	-0.240***
Type_I_Score_SVM	0.439	0.428	0.054	0.397	0.474	0.538	0.541	0.062	0.492	0.588	-0.0993***
Type_I_Score_RFC	0.064	0.004	0.174	0.001	0.025	0.616	0.739	0.369	0.254	0.967	-0.552***

Table 17 (Continued).

Logsale	4.961	5.095	2.320	3.648	6.489	2.957	3.195	2.537	1.565	4.412	2.004***
Zscore	0.916	0.732	0.816	0.368	1.210	0.815	0.504	1.083	0.061	1.423	0.100
EXCHCD	0.242	0	0.439	0	0	0.106	0	0.394	0	0	0.137***
DFT	0.082	0	0.274	0	0	0.121	0	0.326	0	0	-0.039
Big4	0.618	1	0.486	0	1	0.462	0	0.500	0	1	0.155***
Leverage	0.246	0.177	0.282	0.009	0.381	0.411	0.226	0.547	0.036	0.598	-0.165***

Table 17 Panel A presents the descriptive statistics Model [10], which examines the validity and effectiveness of the textual-based Type I going concern error proxies generated by machine learning algorithms. *Type_I_Score_NB* is the probability of Type I going concern error that is calculated by Naives Bayes Classification. *Type_I_Score _SVM* is the probabilities of Type I going concern error that is calculated by Supporting Vector Machine. *Type_I_Score _RFC* is the probabilities of Type I going concern error that is calculated by Random Forest Classification. All three proxies are calculated based on the combinations of Item 1A and Item 7 disclosures. *Readability* is the average of readabilities in Item 1A and Item 7 disclosures that are calculated by the Gunning (1952) Fog index. *Specificity* is the average of Specificities in Item 1A and Item 7 disclosures that are calculated as the total number of entities (e.g., locations, people, and organizations) explicitly identified in the disclosure divided by the total number of words. *HandInformation* is the average quantitative information disclosed in Item 1A and Item 7 divided by the total number of words. *Tone* is the average tone in Item 1A and Item 7 disclosures that are calculated as positive words minus negative words (Loughran and McDonald 2015) and divided by the total number of words. $LogSale_{it}$ is the natural log of sales. $Zscore_{it}$ is the Bankruptcy score calculated by Altman (1968). *Big4* equals 1 if a client is audited by big 4 auditors and 0 otherwise. *EXCHCD* equals 1 if a client is listed in New York Stock Exchange and 0 otherwise. *DFT* equals 1 if a client is under technical or payment default and 0 otherwise. *Leverage* is calculated as total liabilities divided by total assets. Panel B represents the descriptive statistics for non Type I error clients and Type I error clients separately. The Type_I_Error = 1 group includes Type I going concern error clients and Type_I_error = 0 group includes clients that 1) do not receive going concern opinions in the current year and do not file for bankruptcy protection in the subsequent year, 2) receive accurate going concern opinions, 3) have Type II going concern errors. The results are based on the testing sample. T-test are shown in the last column. *** p<0.01, ** p<0.05, * p<0.1.

Table 18 Empirical Testing for Model [10]

	1	2	3	4	5	6	7	8
VARIABLES	**PROB(Type_I_Error)**							
Readability		-0.00392						
		(-0.216)						
Specificity			-0.571					
			(-0.282)					
HardInformation				11.39				
				(-0.757)				
Tone					**-37.41*** **			
					(-5.052)			
Type_I_Score_NB						**1.325*** **		
						(6.826)		
Type_I_Score_RFC							**15.91*** **	
							(16.24)	
Type_I_Score_SVM								**2.413*** **
								(19.31)
Logsale	-0.197***	-0.197***	-0.197***	-0.197***	-0.221***	-0.113***	0.0452*	-0.102***
	(-10.68)	(-10.68)	(-10.69)	(-10.69)	(-11.42)	(-5.129)	-1.777	(-4.551)
Zscore	0.179***	0.179***	0.179***	0.178***	0.164***	0.210***	0.206***	0.158***
	(3.911)	(3.907)	(3.922)	(3.893)	(3.467)	(4.677)	(4.094)	(2.983)
EXCHCD	0.0459	0.0451	0.046	0.0468	0.0276	0.142	0.265**	0.105
	(0.430)	(0.422)	(0.431)	(0.439)	(0.255)	(1.289)	(2.121)	(0.871)
DFT	0.208*	0.208*	0.208*	0.210*	0.210*	0.189	0.089	0.142
	(1.739)	(1.742)	(1.744)	(1.760)	(1.746)	(1.540)	(0.634)	(1.013)

Table 18 (Continued).

Big4	-0.00372	-0.003	-0.00359	-0.00558	0.0173	-0.13	-0.0505	0.218**
	(-0.0477)	(-0.0384)	(-0.0460)	(-0.0714)	-0.219	(-1.584)	(-0.553)	-2.242
Leverage	0.719***	0.719***	0.719***	0.719***	0.747***	0.659***	0.484***	0.373***
	(7.858)	(7.861)	(7.861)	(7.858)	(8.008)	(7.160)	(4.894)	(3.640)
Constant	3.855	3.596	3.739	4.531	9.657	16.42	8.171	9.776
	(0.243)	(0.226)	(0.236)	(0.286)	(0.601)	(1.010)	(0.445)	(0.513)
Observations	4,274	4,274	4,274	4,274	4,274	4,274	4,274	4,274
Industry Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R2	0.128	0.128	0.128	0.128	0.145	0.158	0.339	0.384

Table 18 presents the regression results of Model [10], which examines the validity and effectiveness of the textual-based Type I going concern error proxies generated by machine learning algorithms. *Type_I_Score_NB* is the probabilities of Type I going concern error that is calculated by Naives Bayes Classification. *Type_I_Score_SVM* is the probability of Type I going concern error that is calculated by Supporting Vector Machine. *Type_I_Score_RFC* is the probabilities of Type I going concern error that is calculated by Random Forest Classification. Type IType IType I All three proxies are calculated based on combinations of Item 1A and Item 7 disclosures. *Readability* is the average of readabilities in Item 1A and Item 7 disclosures that are calculated by the Gunning (1952) Fog index. *Specificity* is the average of Specificities in Item 1A and Item 7 disclosures that are calculated as the total number of entities (e.g., locations, people, and organizations) explicitly identified in the disclosure divided by the total number of words. *HandInformation* is the average quantitative information disclosed in Item 1A and Item 7 divided by the total number of words. *Tone* is the average tone in Item 1A and Item 7 disclosures that are calculated as positive words minus negative words (Loughran and McDonald 2015) and divided by the total number of words. $LogSale_{it}$ is the natural log of sales. $Zscore_{it}$ is the Bankruptcy score calculated by Altman (1968). *Big4* equals 1 if a client is audited by big 4 auditors and 0 otherwise. *EXCHCD* equals 1 if a client is listed in New York Stock Exchange and 0 otherwise. *DFT* equals 1 if a client is under technical or payment default and 0 otherwise. *Leverage* is calculated as total liabilities divided by total assets. The results are based on the testing sample, which has 199 Type I error observations, 35 Type II error observations, 37 accurate going concern observations, and 4,003 observations that have no going concern opinions and no subsequent bankruptcy. Z scores are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 19 Descriptive Statistics for Model [11]

Panel A Descriptive Staticstics for Type II Error Proxy Test

	Mean	**Median**	**sd**	**p25**	**p75**
Readability	10.082	10.315	1.95	9.75	10.955
Specificity	0.071	0.071	0.017	0.063	0.08
HardInformation	0.003	0.002	0.002	0.001	0.004
Tone	-0.018	-0.019	0.006	-0.022	-0.015
Type_II_Score_NB	0.278	0.128	0.309	0.029	0.488
Type_II_Score_SVM	0.423	0.417	0.092	0.345	0.499
Type_II_Score_RFC	0.022	0	0.111	0	0.002
Logsale	4.857	5.005	2.394	3.467	6.481
Zscore	0.888	0.706	0.788	0.346	1.195
EXCHCD	0.236	0	0.429	0	0
DFT	0.086	0	0.28	0	0
Big4	0.619	1	0.486	0	1
Leverage	0.266	0.191	0.32	0.011	0.41

Table 19 (Continued).

Reportlag	70.105	69	27.276	58	76
Banklag	113.716	72	119.537	60	88
Type_II_Error	0.006	0	0.078	0	0

Panel B Descriptive Staticstics for Type II Error Proxy Test by groups

	Type II Error=0 (N=4218)					Type II Error=1 (N=26)					
	Mean	**Median**	**sd**	**p25**	**p75**	**Mean**	**Median**	**sd**	**p25**	**p75**	**t-test**
Readability	10.081	10.315	1.954	9.75	10.955	10.169	10.003	1.24	9.73	10.915	-0.088
Specificity	0.071	0.071	0.017	0.063	0.08	0.074	0.073	0.015	0.067	0.081	-0.003
HardInformation	0.003	0.002	0.002	0.001	0.004	0.003	0.003	0.003	0.002	0.004	0
Tone	-0.018	-0.019	0.006	-0.022	-0.015	-0.02	-0.021	0.005	-0.023	-0.018	0.002
Type_II_Score_NB	0.276	0.126	0.307	0.028	0.485	0.646	0.863	0.379	0.263	0.96	-0.371***
Type_II_Score_SVM	0.422	0.416	0.091	0.345	0.499	0.528	0.554	0.084	0.484	0.596	-0.106***

Table 19 (Continued).

Type_II_Score_RFC	0.021	0	0.109	0	0.002	0.135	0.013	0.264	0.001	0.104	-0.114***
Logsale	4.849	4.992	2.396	3.461	6.474	6.221	6.217	1.604	5.822	6.722	-1.372**
Zscore	0.886	0.707	0.786	0.348	1.194	1.107	0.663	1.025	0.323	1.733	-0.221
EXCHCD	0.235	0	0.429	0	0	0.423	0	0.504	0	1	-0.188*
DFT	0.086	0	0.28	0	0	0.154	0	0.368	0	0	-0.068
Big4	0.618	1	0.486	0	1	0.654	1	0.485	0	1	-0.036
Leverage	0.264	0.19	0.319	0.011	0.406	0.541	0.542	0.362	0.279	0.727	-0.277***
Reportlag	70.098	69	27.308	58	76	71.346	63.5	21.77	57	75	-1.248
Banklag	113.628	72	119.44	60	88	127.885	68	136.321	58	90	-14.26

Table 19 Panel A presents the descriptive statistics Model [11], which examines the validity and effectiveness of the textual-based Type II going concern opinion error proxies generated by machine learning algorithms. *Type_II_Score_NB* is the probabilities of Type II going concern error opinion that are calculated by Naives Bayes Classification. *Type_II_Score_SVM* is the probabilities of Type II going concern opinion error that are calculated by Supporting Vector Machine. *Type_II_Score_RFC* is the probabilities of Type II going concern opinion error that are calculated by Random Forest Classification. All three proxies are calculated based on combinations of Item 1A and Item 7 disclosures. *Readability* is the average of readabilities in Item 1A and Item 7 disclosures that are calculated by the Gunning (1952) Fog index. *Specificity* is the average of Specificities in Item 1A and Item 7 disclosures that are calculated as the total number of entities (e.g., locations, people, and organizations) explicitly identified in the disclosure divided by the total number of words. *HandInformation* is the average quantitative information disclosed in Item 1A and Item 7 divided by the total number of words. *Tone* is the average of tone in Item 1A and Item 7 disclosures that are calculated as positive words minus negative words (Loughran and McDonald 2015) and divided by the total number of words. $LogSale_{it}$ is the natural log of sales. $Zscore_{it}$ is the Bankruptcy score calculated by Altman (1968). *Big4* equals 1 if a client is audited by big 4 auditors and 0 otherwise. *EXCHCD* equals 1 if a client is listed in New York Stock Exchange and 0 otherwise. *DFT* equals 1 if a client is under technical or payment default and 0 otherwise. *Leverage* is calculated as total liabilities divided by total assets. *Banklag* equals the difference between audit report dates and bankruptcy dates. *Reportlag* equals the difference between fiscal year end dates and audit report dates. Panel B represents the descriptive statistics for inaccurate going concern

opinion clients and accurate going concern opinion clients separately. The Type_II_Error = 1 group includes Type II going concern error clients and Type_II_Error = 0 group includes clients that 1) do not receive going concern opinions in the current year and do not file for bankruptcy protection in the subsequent year, 2) have accurate going concern opinions, 3) have Type I going concern errors. The results are based on the testing sample. T-test are shown in the last column. *** p<0.01, ** p<0.05, * p<0.1.

Table 20 Empirical Testing for Model [11]

	1	2	3	4	5	6	7	8
VARIABLES	**PROB(Type_II_Error)**							
Readability		0.0124						
		(0.28)						
Specificity			3.594					
			(0.757)					
HardInformation				13.74				
				(0.464)				
Tone					-16.97			
					(-1.118)			
Type_II_Score_RFC						**0.934*** **		
						(2.769)		
Type_II_Score _SVM							**0.874*** **	
							(3.221)	
Type_II_Score _NB								**4.472*** **
								(3.61)
Logsale	0.0976*	0.0985*	0.0974*	0.0978*	0.0911*	0.0870*	0.0657	0.0172
	(1.904)	(1.918)	(1.901)	(1.906)	(1.744)	(1.696)	(1.206)	(0.300)
Zscore	0.152	0.151	0.15	0.152	0.143	0.148	0.0972	0.122
	(1.623)	(1.620)	(1.592)	(1.621)	(1.503)	(1.577)	(1.005)	(1.299)
EXCHCD	-0.035	-0.0382	-0.0321	-0.0343	-0.0468	-0.0276	-0.178	-0.17
	(-0.177)	(-0.193)	(-0.163)	(-0.174)	(-0.236)	(-0.139)	(-0.866)	(-0.835)
DFT	0.129	0.13	0.129	0.125	0.132	0.129	0.0752	0.109
	(0.590)	(0.595)	(0.593)	(0.574)	(0.605)	(0.580)	(0.332)	(0.483)
Big4	-0.0608	-0.0624	-0.0619	-0.0633	-0.0673	-0.0777	-0.0236	0.0136
	(-0.330)	(-0.338)	(-0.335)	(-0.344)	(-0.363)	(-0.416)	(-0.123)	-0.0701
Leverage	0.437***	0.439***	0.435***	0.436***	0.443***	0.429***	0.451***	0.400***

Table 20 (Continued).

	(3.211)	(3.221)	(3.203)	(3.199)	(3.255)	(3.121)	(3.143)	(2.640)
Reportlag	0.00153	0.00152	0.00149	0.00156	0.00116	0.00169	0.00205	0.00131
	(3.211)	(3.221)	(3.203)	(3.199)	(3.255)	(3.121)	(3.143)	(2.640)
Banklag	0.000641	0.000642	0.00065	0.000637	0.000656	0.000471	0.000222	0.000378
	(1.076)	(1.075)	(1.084)	(1.065)	(1.099)	(0.763)	(0.346)	(0.598)
Constant	-61.16*	-59.94*	-59.46	-61.01*	-56.28	-63.07*	-69.71*	-84.08**
	(-1.688)	(-1.648)	(-1.644)	(-1.685)	(-1.538)	(-1.713)	(-1.852)	(-2.178)
Observations	4,244	4,244	4,244	4,244	4,244	4,244	4,244	4,244
Industry Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R2	0.113	0.113	0.115	0.114	0.117	0.133	0.147	0.16

Table 20 presents the regression results of Model [11], which examines the validity and effectiveness of the textual-based Type II going concern error proxies generated by machine learning algorithms. . *Type_II_Score_NB* is the probabilities of Type II going concern error opinion that are calculated by Naives Bayes Classification. *Type_II_Score_SVM* is the probabilities of Type II going concern opinion error that are calculated by Supporting Vector Machine. *Type_II_Score_RFC* is the probabilities of Type II going concern opinion error that are calculated by Random Forest Classification. All three proxies are calculated based on combinations of Item 1A and Item 7 disclosures. *Readability* is the average of readabilities in Item 1A and Item 7 disclosures that are calculated by the Gunning (1952) Fog index. *Specificity* is the average of Specificities in Item 1A and Item 7 disclosures that are calculated as the total number of entities (e.g., locations, people, and organizations) explicitly identified in the disclosure divided by the total number of words. *HandInformation* is the average quantitative information disclosed in Item 1A and Item 7 divided by the total number of words. *Tone* is the average tone in Item 1A and Item 7 disclosures that are calculated as positive words minus negative words (Loughran and McDonald 2015) and divided by total number of words. $LogSale_{it}$ is the natural log of sales. $Zscore_{it}$ is the Bankruptcy score calculated by Altman (1968). *Big4* equals 1 if a client is audited by big 4 auditors and 0 otherwise. *EXCHCD* equals 1 if a client is listed in New York Stock Exchange and 0 otherwise. *DFT* equals 1 if a client is under technical or payment default and 0 otherwise. *Leverage* is calculated as total liabilities divided by total assets. *Banklag* equals the difference between audit report dates and bankruptcy dates. *Reportlag* equals the difference between fiscal year end dates and audit report dates. The results are based on the testing sample, which has 121 Type I error observations, 26 Type II error observations, 52 accurate going concern observations, and 3,954 observations that have no going concern opinions and no subsequent bankruptcy. Z scores are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

CHAPTER FIVE: CONCLUSION

This study utilizes machine learning techniques to conduct textual analysis, in order to determine the disclosure topics that lead auditors to issue inaccurate going concern opinions. In the Type I setting, I find that the probability of observing an accurate going concern *opinion* is higher when clients disclose human capital and supply chain risks in Item 1A and when clients disclose tax-related information in Item 7. The probability of accurate going concern *evaluation* is higher (lower) when human capital, fluctuation, legal, and macro-economic risks (funding, financial condition, debt, operational, attestation, and stock market risks) are disclosed in Item 1A. This probability is lower when clients disclose growing potential, stocks, and political contribution related information in Item 7.

In the Type II error settings, the probability of accurate going concern *opinion* is higher when clients disclose bankruptcy and operational risks (development, supply chain, and environmental risks) in Item 1A. For Item 7, the probability is higher when clients disclose bankruptcy, performance changes, and costs (operational performance and tax) related information. The probability of accurate going concern *evaluation* is higher when clients disclose macro-economic, intellectual property, and investment risks (development and oil/gas risks) in Item 1A, or if clients disclose human capital (loan and operational performance) related information in Item 7.

In addition, I create new proxies for going concern opinion accuracy, Type I going concern errors, and Type II going concern errors by utilizing various machine learning techniques, and then conduct empirical tests to test the validity and effectiveness of those proxies. The regression results confirm the validity and effectiveness of the proxies in predicting going concern accuracy variations. More importantly, the textual-based proxies generated by machine learning algorithms are demonstrated to have higher explanatory power than the traditional textual attributes. Interestingly, while I find that all machine learning-based proxies outperform the traditional text-based proxies, there is not one machine learning technique that consistently outperforms the others. This suggests that the optimal machine learning technique is highly dependent on the type of outcome under scrutiny (e.g. evaluating the accuracy of a going concern opinion, or determining whether a Type I error or Type II error is likely).

My study is not free of limitations. One limitation of using machine learning techniques for textual analysis is that those latent topics need human judgment in the labeling procedure. It may introduce biases in which the latent topics may not represent the subjects that human beings assigned them to be. Despite this limitation, my study provides an economically significant use case to demonstrate the benefits of machine learning for textual analysis. Future research could test whether the disclosure topics elicit different capital market reactions to the going concern opinions. In addition, future studies can also utilize machine learning algorithms to generate proxies for other audit or financial outcomes, based on the textual information disclosed in 10-K filings.

REFERENCES

Agarwal, S., Gupta, S., & Israelsen, R. D. (2017). Public and private information: Firm disclosure, sec letters, and the jobs act. *Georgetown McDonough School of Business Research Paper*, (2891089), 17-4.

Ahn, J., & Jensen, K. L. (2017). Quality Control in Audit Firms: Do Auditors Learn from Going Concern Errors?. *Available at SSRN 3153078*.

AICPA. 2015. The Auditor's Consideration of an Entity's Ability to Continue as a Going Concern: Auditing Interpretations of Section 570. *Available at: https://us.aicpa.org/content/dam/aicpa/research/standards/auditattest/downloadabledocuments/au-c-09570.pdf*

Allee, K. D., & DeAngelis, M. D. (2015). The structure of voluntary disclosure narratives: Evidence from tone dispersion. *Journal of Accounting Research*, 53(2), 241-274.

Allen, E., O'Leary, D. E., Qu, H., & Swenson, C. W. (2021). Tax Specific versus Generic Accounting-Based Textual Analysis and the Relationship with Effective Tax Rates: Building Context. *Journal of Information Systems*, 35(2), 115-147.

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The journal of finance*, 23(4), 589-609.

Amel-Zadeh, A., Calliess, J. P., Kaiser, D., & Roberts, S. (2020). Machine learning-based financial statement analysis. *Available at SSRN 3520684*.

Amin, K., Krishnan, J., & Yang, J. S. (2014). Going concern opinion and cost of equity. *Auditing: A Journal of Practice & Theory*, 33(4), 1-39.

Amoozegar, A., Berger, D., Cao, X., & Pukthuanthong, K. (2020). Earnings conference calls and institutional monitoring: Evidence from textual analysis. *Journal of Financial Research*, 43(1), 5-36.

Angelov, D. (2020). Top2vec: Distributed representations of topics. arXiv preprint arXiv:2008.09470.

Balakrishnan, K., & Darendeli, A. (2020). Do Firms Respond Differently to Local Competition? Evidence from Textual Analysis of 10-K Filings. *Evidence from Textual Analysis of*.

Banker, R. D., Huang, R., Li, X., & Yan, Y. (2021). Strategy Typology and Cost Structure: A Textual Analysis Approach. *Fox School of Business Research Paper Forthcoming*.

Bao, Y., & Datta, A. (2014). Simultaneously discovering and quantifying risk types from textual risk disclosures. *Management Science*, 60(6), 1371-1391.

Bassyouny, H., Abdelfattah, T., & Tao, L. (2020). Beyond narrative disclosure tone: The upper echelons theory perspective. *International Review of Financial Analysis*, 70, 101499.

Bauer, A. M., & Klassen, K. J. (2017). Assessing the market reaction to unfavorable tax settlements: Using textual analysis to categorize ambiguous tabulated disclosures. *Available at SSRN 2379666*.

Berglund, N. R., Herrmann, D. R., & Lawson, B. P. (2018). Managerial ability and the accuracy of the going concern opinion. *Accounting and the Public Interest*, 18(1), 29-52.

Berkman, H., Jona, J., Lee, G., & Soderstrom, N. (2018). Cybersecurity awareness and market valuations. *Journal of Accounting and Public Policy*, 37(6), 508-526.

Bertomeu, J. (2020). Machine learning improves accounting: discussion, implementation and research opportunities. *Review of Accounting Studies*, 25(3), 1135-1155.

Bertomeu, J., Cheynel, E., Floyd, E., & Pan, W. (2021). Using machine learning to detect misstatements. *Review of Accounting Studies*, 26(2), 468-519.

Blay, A. D., Geiger, M. A., & North, D. S. (2011). The auditor's going-concern opinion as a communication of risk. *Auditing: A Journal of Practice & Theory*, 30(2), 77-102.

Blay, A. D., Moon Jr, J. R., & Paterson, J. S. (2016). There's no place like home: The influence of home-state going-concern reporting rates on going-concern opinion propensity and accuracy. *Auditing: A Journal of Practice & Theory*, 35(2), 23-51.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan), 993-1022.

Bochkay, K., Chychyla, R., & Nanda, D. (2019). Dynamics of CEO disclosure style. *The Accounting Review*, 94(4), 103-140.

Bodnaruk, A., Loughran, T., & McDonald, B. (2015). Using 10-K text to gauge financial constraints. *Journal of Financial and Quantitative Analysis*, 50(4), 623-646.

Bonsall, S. B., & Miller, B. P. (2017). The impact of narrative disclosure readability on bond ratings and the cost of debt. *Review of Accounting Studies*, 22(2), 608-643.

Bozanic, Z., & Thevenot, M. (2015). Qualitative disclosure and changes in Sell-Side financial analysts' information environment. *Contemporary Accounting Research*, 32(4), 1595-1616.

Breuer, W., & Ghufran, B. (2020). The Predictive Power of Managerial Tone: A Text-Based Analysis of Takeover Performance. *Available at SSRN 3763744*.

Brochet, F., Loumioti, M., & Serafeim, G. (2015). Speaking of the short-term: Disclosure horizon and managerial myopia. *Review of Accounting Studies*, 20(3), 1122-1163.

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

Brown, S. V., Hinson, L. A., & Tucker, J. W. (2021). Financial Statement Adequacy and Firms' MD&A Disclosures. *Available at SSRN 3891572*.

Budisantoso, T., Rahmawati, R., Bandi, B., & Probohudono, A. N. (2017). Determinant of Downward Auditor Switching. *Jurnal Akuntansi Multiparadigma*, 8(3), 444-457.

Buehlmaier, M. M., & Whited, T. M. (2018). Are financial constraints priced? Evidence from textual analysis. *The Review of Financial Studies*, 31(7), 2693-2728.

Burks, J. J., Cuny, C., Gerakos, J., & Granja, J. (2018). Competition and voluntary disclosure: Evidence from deregulation in the banking industry. *Review of Accounting Studies*, 23(4), 1471-1511.

Campbell, J. L., Chen, H., Dhaliwal, D. S., Lu, H. M., & Steele, L. B. (2014). The information content of mandatory risk factor disclosures in corporate filings. *Review of Accounting Studies*, 19(1), 396-455.

Carson, E., Fargher, N. L., Geiger, M. A., Lennox, C. S., Raghunandan, K., & Willekens, M. (2013). Audit reporting for going-concern uncertainty: A research synthesis. *Auditing: A Journal of Practice & Theory*, 32(Supplement 1), 353-384.

Campbell, J. L., Zheng, X., & Zhou, D. (2021). Number of Numbers: Does Quantitative Textual Disclosure Reduce Information Risk?. *Available at SSRN 3775905*.

Elsayed r, R. A., & Pfeiffer, R. J. (2016). Why are 10-K filings so long?. *Accounting Horizons*, 30(1), 1-21.

Cazier, R. A., & Pfeiffer, R. J. (2017). 10-K disclosure repetition and managerial reporting incentives. *Journal of Financial Reporting*, 2(1), 107-131.

Casterella, J. R., Desir, R., Stallings, M. A., & Wainberg, J. S. (2020). Information transfer of bankruptcy announcements: Examining the impact of auditor opinions. *Accounting Horizons*, 34(1), 45-66.

Center for Audit Quality (CAQ). 2012. Request for Proposals for Academic Research in Auditing. (December). Washington, DC: CAQ.

Center for Audit Quality (CAQ). 2020. Profession in focus: Going concern and covid-19: The Center for Audit Quality. The Center for Audit Quality |. Retrieved from https://www.thecaq.org/profession-in-focus-going-concern-and-covid-19/

Chang, J., Gerrish, S., Wang, C., Boyd-Graber, J., & Blei, D. (2009). Reading tea leaves: How humans interpret topic models. *Advances in neural information processing systems*, 22.

Chen, H., Francis, B. B., Hasan, T., & Wu, Q. (2022). Does corporate culture impact audit pricing? Evidence from textual analysis. *Journal of Business Finance & Accounting*, 49(5-6), 778-806.

Chen, K. C., & Church, B. K. (1996). Going concern opinions and the market's reaction to bankruptcy filings. *Accounting Review*, 117-128.

Chen, X., Cho, Y. H., Dou, Y., & Lev, B. (2022). Predicting Future Earnings Changes Using Machine Learning and Detailed Financial Data. *Journal of Accounting Research*, 60(2), 467-515.

Cheng, X., Smith, D., & Tanyi, P. (2018). An analysis of proxy statement leadership structure justification disclosures. *Review of Quantitative Finance and Accounting*, 51(4), 1071-1106.

Cheong, A., Yoon, K., Cho, S., & No, W. G. (2021). Classifying the contents of cybersecurity risk disclosure through textual analysis and factor analysis. *Journal of information Systems*, 35(2), 179-194.

Cho, C. H., Roberts, R. W., & Patten, D. M. (2010). The language of US corporate environmental disclosure. *Accounting, Organizations and Society*, 35(4), 431-443.

Clarkson, P. M., Ponn, J., Richardson, G. D., Rudzicz, F., Tsang, A., & Wang, J. (2020). A textual analysis of US corporate social responsibility reports. Abacus, 56(1), 3-34.

Craig, R., & Amernic, J. (2018). Are there language markers of hubris in CEO letters to shareholders?. *Journal of business ethics*, 149(4), 973-986.

D'Augusta, C., & DeAngelis, M. D. (2020). Does accounting conservatism discipline qualitative disclosure? Evidence from tone management in the MD&A. *Contemporary Accounting Research*, 37(4), 2287-2318.

Davis, A. K., Ge, W., Matsumoto, D., & Zhang, J. L. (2015). The effect of manager-specific optimism on the tone of earnings conference calls. *Review of Accounting Studies*, 20(2), 639-673.

Deloitte 2020. Going concern - key considerations related to performing a comprehensive assessment (July 8, 2020). DART. (n.d.). Retrieved from https://dart.deloitte.com/USDART/home/publications/deloitte/accounting-spotlight/going-concern-assessment

Ding, K., Lev, B., Peng, X., Sun, T., & Vasarhelyi, M. A. (2020). Machine learning improves accounting estimates: Evidence from insurance payments. *Review of Accounting Studies*, 25(3), 1098-1134.

Ding, K., Peng, X., & Wang, Y. (2019). A machine learning-based peer selection method with financial ratios. *Accounting Horizons*, 33(3), 75-87.

Dong, M., Jondeau, E., & Rockinger, M. (2019). Textual Analysis of Banks' Pillar 3 Documents. *Available at SSRN 3365005*.

Donovan, J., Jennings, J., Koharki, K., & Lee, J. (2021). Measuring credit risk using qualitative disclosure. *Review of Accounting Studies*, 26(2), 815-863.

Du, S., & Yu, K. (2021). Do corporate social responsibility reports convey value relevant information? Evidence from report readability and tone. *Journal of business ethics*, 172(2), 253-274.

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

Efretuei, E. (2021). Year and industry-level accounting narrative analysis: readability and tone variation. *Journal of Emerging Technologies in Accounting*, 18(2), 53-76.

Egger, R., & Yu, J. (2022). A topic modeling comparison between lda, nmf, top2vec, and bertopic to demystify twitter posts. *Frontiers in sociology*, 7.

Elsayed, M., & Elshandidy, T. (2021). Internal control effectiveness, textual risk disclosure, and their usefulness: US evidence. *Advances in accounting*, 53, 100531.

Fassas, A., Bellos, S., & Kladakis, G. (2021). Corporate liquidity, supply chain and cost issues awareness within the Covid-19 context: evidence from us management reports' textual analysis. *Corporate Governance: The International Journal of Business in Society*, 21(6), 1155-1171.

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

Financial Accounting Standards Board (FASB). 2014. No. 2014-15, Presentation of Financial Statements—Going Concern(Subtopic 205-40): Disclosure of Uncertainties about an Entity's Ability to Continue as a Going Concern. Available at: https://www.fasb.org/page/ShowPdf?path=FASB+in+Focus--Going+Concern+8.27.14.pdf&title=FIF%20(August%202014)%20Accounting%20Standards%20Update%E2%80%94Presentation%20of%20FinancialStatements%E2%80%94Going%20Concern%20(Subtopic%20205-40):%20Disclosure%20of%20Uncertainties%20about%20an%20Entity%27s%20Ability%20to%20Continue%20as%20a%20Going%20Concern

Florackis, C., Louca, C., Michaely, R., & Weber, M. (2020). Cybersecurity risk (No. w28196). *National Bureau of Economic Research*.

Frankel, R., Jennings, J., & Lee, J. (2016). Using unstructured and qualitative disclosures to explain accruals. *Journal of Accounting and Economics*, 62(2-3), 209-227.

Frankel, R., Jennings, J., & Lee, J. (2022). Disclosure sentiment: machine learning vs. dictionary methods. *Management Science*, 68(7), 5514-5532.

Gan, Q., & Qiu, B. (2021). The information content of 10-K file size change. *International Review of Finance*, 21(4), 1251-1285.

Geiger, M. A., Basioudis, I. G., & DeLange, P. (2022). The effect of non-audit fees and industry specialization on the prevalence and accuracy of auditor's going-concern reporting decisions. *Journal of International Accounting, Auditing and Taxation*, 100473.

Geiger, M. A., Raghunandan, K., & Riccardi, W. (2014). The global financial crisis: US bankruptcies and going-concern audit opinions. *Accounting Horizons*, 28(1), 59-75.

Geiger, M. A., & Rama, D. V. (2003). Audit fees, nonaudit fees, and auditor reporting on stressed companies. *Auditing: A journal of practice & theory*, 22(2), 53-69.

Geiger, M. A., & Rama, D. V. (2006). Audit firm size and going-concern reporting accuracy. *Accounting horizons*, 20(1), 1-17.

Gogas, P., & Papadimitriou, T. (2021). Machine learning in economics and finance. *Computational Economics*, 57(1), 1-4.

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies*, 33(5), 2223-2273.

Guay, W., Samuels, D., & Taylor, D. (2016). Guiding through the fog: Financial statement complexity and voluntary disclosure. *Journal of Accounting and Economics*, 62(2-3), 234-269.

Gutierrez, E., Krupa, J., Minutti-Meza, M., & Vulcheva, M. (2020). Do going concern opinions provide incremental information to predict corporate defaults?. *Review of Accounting Studies*, 25(4), 1344-1381.

Guo, L., Shi, F., & Tu, J. (2016). Textual analysis and machine leaning: Crack unstructured data in finance and accounting. *The Journal of Finance and Data Science*, 2(3), 153-170.

Hardies, K., Vandenhaute, M. L., & Breesch, D. (2018). An analysis of Auditors' going-concern reporting accuracy in private firms. *Accounting Horizons*, 32(4), 117-132.

He, J., & Plumlee, M. A. (2020). Measuring disclosure using 8-K filings. *Review of Accounting Studies*, 25(3), 903-962.

Hearst, M. A., Dumais, S. T., Osuna, E., Platt, J., & Scholkopf, B. (1998). Support vector machines. *IEEE Intelligent Systems and their applications*, 13(4), 18-28.

Henry, E. (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.

Henry, E., & Leone, A. J. (2016). Measuring qualitative information in capital markets research: Comparison of alternative methodologies to measure disclosure tone. *The Accounting Review*, 91(1), 153-178.

Hoepner, A. G., McMillan, D., Vivian, A., & Wese Simen, C. (2021). Significance, relevance and explainability in the machine learning age: an econometrics and financial data science perspective. *The European Journal of Finance*, 27(1-2), 1-7.

Hrazdil, K., Novak, J., Rogo, R., Wiedman, C., & Zhang, R. (2020). Measuring executive personality using machine-learning algorithms: A new approach and audit fee-based validation tests. *Journal of Business Finance & Accounting*, 47(3-4), 519-544.

Hu, H., Sun, T., Vasarhelyi, M. A., & Zhang, M. (2020). A Machine Learning Approach of Measuring Audit Quality: Evidence From China. *Available at SSRN 3732563*.

Hu, W., Shohfi, T., & Wang, R. (2021). What’s really in a deal? Evidence from textual analysis of M&A conference calls. *Review of Financial Economics*, 39(4), 500-521.

Huang, A. H., Lehavy, R., Zang, A. Y., & Zheng, R. (2018). Analyst information discovery and interpretation roles: A topic modeling approach. *Management science*, 64(6), 2833-2855.

Huang, A. H., Zang, A. Y., & Zheng, R. (2014). Evidence on the information content of text in analyst reports. *The Accounting Review*, 89(6), 2151-2180.

Huang, J., Roberts, H., & Tan, E. K. (2018). Media Tone and CEO Power. *Available at SSRN 3220885.*

Huang, J., Zhang, X., & Tan, L. (2016, November). Detecting sensitive data disclosure via bi-directional text correlation analysis. In Proceedings of the 2016 24th ACM SIGSOFT International Symposium on Foundations of Software Engineering (pp. 169-180).

Huang, K. W., & Li, Z. (2011). A multilabel text classification algorithm for labeling risk factors in SEC form 10-K. *ACM Transactions on Management Information Systems (TMIS)*, 2(3), 1-19.

Huang, X., Krishnan, S., & Lin, P. (2018). Tone Analysis and Earnings Management. *Journal of Accounting & Finance* (2158-3625), 18(8).

Huang, X., Teoh, S. H., & Zhang, Y. (2014). Tone management. *The Accounting Review*, 89(3), 1083-1113.

Hunt, J. O., Myers, J. N., & Myers, L. A. (2022). Improving earnings predictions and abnormal returns with machine learning. *Accounting Horizons*, 36(1), 131-149.

Hunt, J. O., Rosser, D. M., & Rowe, S. P. (2021). Using machine learning to predict auditor switches: How the likelihood of switching affects audit quality among non-switching clients. *Journal of Accounting and Public Policy*, 40(5), 106785.

Jan, C. L. (2021). Using deep learning algorithms for CPAs' going concern prediction. *Information*, 12(2), 73.

Jeyaraj, A., Zadeh, A., & Sethi, V. (2021). Cybersecurity threats and organisational response: textual analysis and panel regression. *Journal of Business Analytics*, 4(1), 26-39.

Jiang, L., Pittman, J. A., & Saffar, W. (2017). Policy uncertainty and textual disclosure. *Accounting Horizons*.

Jiang, J., Srinivasan, K. MoreThanSentiments: A text analysis package. *Software Impacts*, 100456 (2022). https://doi.org/10.1016/J.SIMPA.2022.100456

Katsafados, A. G., Androutsopoulos, I., Chalkidis, I., Fergadiotis, E., Leledakis, G. N., & Pyrgiotakis, E. G. (2021). Using textual analysis to identify merger participants: Evidence from the US banking industry. *Finance Research Letters*, 42, 101949.

Kim, C., Wang, K., & Zhang, L. (2019). Readability of 10-K reports and stock price crash risk. *Contemporary accounting research*, 36(2), 1184-1216.

Koelbl, M. (2020). Is the MD&A of US REITs informative? A textual sentiment study. *Journal of Property Investment & Finance*, 38(3), 181-201.

Kravet, T., & Muslu, V. (2013). Textual risk disclosures and investors' risk perceptions. *Review of Accounting Studies*, 18(4), 1088-1122.

Krupa, J., & Minutti-Meza, M. (2021). Regression and Machine Learning Methods to Predict Discrete Outcomes in Accounting Research. *Journal of Financial Reporting*.

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

Lee, J., & Park, J. (2019). The impact of audit committee financial expertise on management discussion and analysis (MD&A) tone. *European Accounting Review*, 28(1), 129-150.

Leung, E., & Veenman, D. (2018). Non-GAAP earnings disclosure in loss firms. *Journal of Accounting Research*, 56(4), 1083-1137.

Lewis, C., & Young, S. (2019). Fad or future? Automated analysis of financial text and its implications for corporate reporting. *Accounting and Business Research*, 49(5), 587-615.

Li, F. (2010). The information content of forward-looking statements in corporate filings—A naïve Bayesian machine learning approach. *Journal of Accounting Research*, 48(5), 1049-1102.

Li, H., No, W. G., & Wang, T. (2018). SEC's cybersecurity disclosure guidance and disclosed cybersecurity risk factors. *International Journal of Accounting Information Systems*, 30, 40-55.

Liu, M. (2022). Assessing human information processing in lending decisions: A machine learning approach. *Journal of Accounting Research*, 60(2), 607-651.

Liu, P., & Nguyen, H. T. (2020). CEO characteristics and tone at the top inconsistency. *Journal of Economics and Business*, 108, 105887.

Liu, Y., & Moffitt, K. C. (2016). Text mining to uncover the intensity of SEC comment letters and its association with the probability of 10-K restatement. *Journal of Emerging Technologies in Accounting*, 13(1), 85-94.

Loukas, L., Fergadiotis, M., Androutsopoulos, I., & Malakasiotis, P. (2021). EDGAR-CORPUS: Billions of tokens make the world go round. arXiv preprint arXiv:2109.14394.

Lopez-Lira, A. (2021). Why do managers disclose risks accurately? Textual analysis, disclosures, and risk exposures. *Economics Letters*, 204, 109896.

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

Loughran, T., & McDonald, B. (2014). Measuring readability in financial disclosures. the *Journal of Finance*, 69(4), 1643-1671.

Loughran, T., & McDonald, B. (2014). Regulation and financial disclosure: The impact of plain English. *Journal of Regulatory Economics*, 45(1), 94-113.

Loughran, T., & McDonald, B. (2015). The use of word lists in textual analysis. *Journal of Behavioral Finance*, 16(1), 1-11.

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

Lu, M., Qiao, Z., Tan, H., & Yao, L. (2022). Corporate Textual Transparency and Economic Growth. *Available at SSRN 3549968*.

Manela, A., & Moreira, A. (2017). News implied volatility and disaster concerns. *Journal of Financial Economics*, 123(1), 137-162.

Marquez-Illescas, G., Zebedee, A. A., & Zhou, L. (2019). Hear me write: does CEO narcissism affect disclosure?. *Journal of business ethics*, 159(2), 401-417.

Mayew, W. J., Sethuraman, M., & Tanyi, M. (2015). MD&A Disclosure and the Firm's Ability to Continue as a Going Concern. *The Accounting Review*, 90(4), 1621-1651.

McCallum, Andrew Kachites. "MALLET: A Machine Learning for Language Toolkit." http://mallet.cs.umass.edu. 2002.

Melloni, G., Caglio, A., & Perego, P. (2017). Saying more with less? Disclosure conciseness, completeness and balance in Integrated Reports. Jou*rnal of Accounting and Public Policy*, 36(3), 220-238.

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

Mousa, G. A., Elamir, E. A., & Hussainey, K. (2022). Using machine learning methods to predict financial performance: Does disclosure tone matter?. *International Journal of Disclosure and Governance*, 19(1), 93-112.

Mutchler, J. F., & Williams, D. D. (1990). The relationship between audit technology, client risk profiles, and the going-concern opinion decision. *AUDITING-A JOURNAL OF PRACTICE & THEORY*, 9(3), 39-54.

Noble, W. S. (2006). What is a support vector machine?. *Nature biotechnology*, 24(12), 1565-1567.

O'Reilly, D. M. (2010). Do investors perceive the going-concern opinion as useful for pricing stocks?. *Managerial Auditing Journal*.

Osma, B. G., Grande-Herrera, C., & Saorin, E. G. (2018). Optimistic disclosure tone and CEO career concerns. *SSRN Electron. J*.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. the Journal of machine Learning research, 12, 2825-2830.

Public Company Accounting Oversight Board (PCAOB). 2015. AS 2415: Consideration of an Entity's Ability to Continue as a Going Concern. Available at: https://pcaobus.org/oversight/standards/auditing-standards/details/AS2415

Qu, Y., Quan, P., Lei, M., & Shi, Y. (2019). Review of bankruptcy prediction using machine learning and deep learning techniques. *Procedia Computer Science*, 162, 895-899.

Sanoran, K. L. (2018). Auditors' going concern reporting accuracy during and after the global financial crisis. Journal of Contemporary Accounting & Economics, 14(2), 164-178.

Siano, F., & Wysocki, P. (2021). Transfer learning and textual analysis of accounting disclosures: Applying big data methods to small (er) datasets. *Accounting Horizons*, 35(3), 217-244.

Swift, O., Colon, R., & Davis, K. (2020). The impact of cyber breaches on the content of cybersecurity disclosures. *Journal of Forensic and Investigative Accounting*, 12(2), 197-212.

Van der Heijden, H. (2022). Predicting industry sectors from financial statements: An illustration of machine learning in accounting research. *The British Accounting Review*, 101096.

Vasarhelyi, M., Gu, Y., & Zhang, C. A. (2021). Error, Noise, and Bias of Auditors' Going Concern Opinions and the Role of Machine Learning. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3984462

Wang, K. (2021). Is the tone of risk disclosures in MD&As relevant to debt markets? Evidence from the pricing of credit default swaps. *Contemporary Accounting Research*, 38(2), 1465-1501.

Warin, T., & Stojkov, A. (2021). Machine Learning in Finance: A Metadata-Based Systematic Review of the Literature. *Journal of Risk and Financial Management*, 14(7), 302.

Willenborg, M., & McKeown, J. C. (2000). Going-concern initial public offerings. *Journal of Accounting and Economics*, 30(3), 279-313.

Xu, Q., & Kalelkar, R. (2020). Consequences of going-concern opinion inaccuracy at the audit office level. *Auditing: A Journal of Practice & Theory*, 39(3), 185-208.

Yang, R., Yu, Y., Liu, M., & Wu, K. (2018). Corporate risk disclosure and audit fee: A text mining approach. *European Accounting Review*, 27(3), 583-594.

Yang, Y., Simnett, R., & Carson, E. (2022). Auditors' propensity and accuracy in issuing going-concern modified audit opinions for charities. *Accounting & Finance*, 62, 1273-1306.

APPENDIX A: VARIABLE DEFINITION

Variable	Definition
Item1a_Topic_N	equals 1 if a client disclosure topic N in its Item1A disclosures and 0 otherwise
Item7_Topic_N	equals 1 if a client disclosure topic N in its Item7 disclosures and 0 otherwise
Accurate(Type I/Type II)_Score_NB	probabilities of accurate (Type I/Type II) going concern opinion that are calculated by Naives Bayes Classification
Accurate(Type I/Type II)_Score_SVM	probabilities of accurate (Type I/Type II) going concern opinion that are calculated by Supporting Vector Machine
Accurate(Type I/Type II)_Score_RFC	probabilities of accurate (Type I/Type II) going concern opinion that are calculated by Random Forest Classification
Accurate Going Concern Opinion	equals 1 if clients receive going concern opinions in the current year and file for bankruptcy protection in the subsequent year, and 0 with definitions depending on the setting
Accurate Going Concern Evaluation	equals 1 if clients 1) receive going concern opinions in the current year and file for bankruptcy protection in the subsequent year and 2) do not receive going concern opinions in the current year and do not file for bankruptcy protection in the subsequent year, and 0 with definitions depending on the setting
Type I Error	equals 1 if clients receive going concern opinions in the current year and do not file for bankruptcy protection in the subsequent year, and 0 with definitions depending on the setting

Type II Error	equals 1 if clients do not receive going concern opinions in the current year and file for bankruptcy protection in the subsequent year, and 0 with definitions depending on the setting
Readability	average of readabilities in Item 1A and Item 7 disclosures that are calculated by the Gunning (1952) Fog index
Specificity	average of Specificities in Item 1A and Item 7 disclosures that are calculated as the total number of entities (e.g., locations, people, and organizations) explicitly identified in the disclosure divided by the total number of words
HandInformation	average quantitative information disclosed in Item 1A and Item 7 divided by the total number of words
Tone	average tone in Item 1A and Item 7 disclosures that are calculated as positive words minus negative words
LogSale	natural log of sales
Zscore	Bankruptcy score calculated by Altman (1968)
Big4	equals 1 if a client is audited by big 4 auditors and 0 otherwise
EXCHCD	equals 1 if a client is listed in New York Stock Exchange and 0 otherwise
DFT	equals 1 if a client is under technical or payment default and 0 otherwise
Leverage	total liabilities divided by total assets
Banklag	the difference between audit report dates and bankruptcy dates
Reportlag	the difference between fiscal year end dates and audit report dates

APPENDIX B: EXAMPLES OF TOPIC ASSIGNMENT

Topic Assigned	Company Index	Paragraph
13	0	We cannot guarantee that the patents issued to us will be broad enough to provide any meaningful protection of our proprietary technologies.
8	0	We may incur substantial costs as a result of litigation or other proceedings relating to patent and other intellectual property rights.
1	0	While we currently have 56,814,833 shares of common stock outstanding after implementing the 5 to 1 reverse split in 2018, we are authorized to issue up to 250,000,000 shares of common stock. In the event we elect to issue additional shares of common stock in connection with any financing, acquisition or otherwise, current shareholders could find their holdings substantially diluted, which means they will own a smaller percentage of our company. There are also 5 million shares of preferred stock that the board can issue under any terms it wants and without any shareholder approval. Shareholders approved the Company's proposal to increase the authorized capital and/or a reverse split, the risk described above will is heightened even more.
14	1	specified supplier. The qualification of a new supplier could delay our development and marketing efforts. If for any reason we are unable to obtain sufficient quantities of any of the raw materials or components required to produce and package our products, we may not be able to manufacture our products as planned, which could have a material adverse effect on our business, financial condition and results of operations.
3	1	In 2008 and 2007, our major sales were through the three large wholesale drug distributors noted below. These three large wholesale drug distributors account for a large portion of our gross sales, revenues and accounts receivable in all our business segments except for contract services.
16	1	The FDA conducted another inspection of the Decatur facility from July 23, 2007 to August 17, 2007. The FDA investigators identified a number of observations representing potential violations of the cGMP regulations. We submitted comprehensive responses to these observations on September 28, 2007. Subsequently, we were notified by the FDA on December 20, 2007, that all cGMP issues had been satisfactorily resolved resulting in removal of the Warning Letter's potential restrictions on new product approvals; approval of the

lyophilization and filling operations of the Decatur facility; and approval of the site transfer for manufacture of IC Green to the Decatur facility. Since then, we have received FDA approval of several ANDAs and NDAs for manufacture of product at the Decatur facility. We were prompted to initiate one product recall of our Cyanide Antidote Kit during 2008, due to the third quarter recall notification by Becton, Dickinson and Company of their 60ml syringe. This syringe is included as part of a packaged kit along with drug components manufactured and sourced by us, to support the Cyanide Antidote Kit. Our recall of the Cyanide Antidote Kit was necessitated by the BD recall, and has resulted in no patient impact and no shortage of product supply to the marketplace. We recorded a $440,000 additional provision to sales returns in 2008 to recognize the impact of this recall. Our supporting efforts were reviewed by the FDA, as part of our due diligence in apprising the Agency of our reaction to the BD recall. There were no product recalls during 2007 or 2006. We also manufacture and distribute several controlled-drug substances, the distribution and handling of which are regulated by the DEA. Failure to comply with DEA regulations can result in fines or seizure of product. We do not anticipate any material adverse effect from compliance with federal, state and local provisions that have been enacted or adopted regulating the discharge of materials into the environment, or otherwise relating to the protection of the environment. During 2008, 2007 and 2006, approximately $1,384,000, $1,320,000 and $1,104,000, respectively, of our revenues were from customers located in foreign countries. Most of our business segments do not experience significant seasonality other than Td vaccines in the spring through fall seasons and flu vaccine products which are typically sold in the August through November period.