Audit Effort on Tone Ambiguity in 10-K Filings

Martin M. Kim
Saint Joseph's University

Hambisa Belina
Morgan State University

This study examines the association between ambiguous words (uncertain and weak modal words) in 10-K filings and audit fees. We find a positive association between ambiguous words in 10-K filings and audit fees. Results show that auditors put more time and effort into analyzing clients' reports containing more ambiguous words. We conduct a propensity score matching technique to address potential endogeneity in client characteristics, which presents further empirical support for our main results. We re-estimate the main analysis for robustness tests by excluding firms with going concern opinions, firms audited by Big 4, and firms that report a net loss. We find the main results still hold in all robustness tests. The results of this study can be in the interest of multiple stakeholders, such as issuers, investors, auditors, and regulators, in that the 10-K report is mandatory, but the discretion of the management determines the degree of clarity of the text in the 10-K filings. Therefore, the interpretation and decisions of users of these ambiguous reports are non-trivial.

Keywords: tone ambiguity, uncertain and weak modal words, 10-K filings, audit fees, audit effort

INTRODUCTION

The U.S. Securities and Exchange Commission (SEC) has raised concern that firms may intentionally use ambiguous words in their 10-K filings to protect themselves against claims and litigations. Thus, SEC has demanded more clarity and less ambiguity in 10-K filings (SEC, 2007). The 10-K report is mandatory, but the degree of clarity of the text in the 10-K filings is determined by the intent of the management. Therefore, interpretation and decision made by users of these ambiguous reports cannot be easily overlooked (Kanagaretnam, Mawani, Shi, and Zhou 2020).

In this study, we examine the association between ambiguity in financial disclosures and audit pricing. Specifically, we examine the association between ambiguous words (i.e., uncertain and weak modal words) in 10-K filings and audit fees. Prior accounting and finance research examines the tone (e.g., Davis, Piger, and Sedor 2012; Demers and Vega 2014; Frankel, Mayew, and Sun 2010) and readability (e.g., Li 2008; Biddle, Hilary, and Verdi 2009; Miller 2010; Lehavy, Li, and Merkley 2011; Lawrence 2013; Blanco, Coram, Dhole, and Kent 2021) of financial disclosures and finds that market participants make critical financial decisions based on the choice of tone and the level of readability in their written communication (such as 10-Ks, investor message boards, and earnings press releases) and verbal discussions at conference calls. Similarly, the audit report provides valuable information for market participants in their decision

process. According to DeFond and Zhang (2014), auditors play a critical role in providing independent assurance on the credibility of their client's financial information, which improves resource allocation and contracting efficiency.

Prior studies (e.g., Davis et al. 2012; Demers and Vega 2014; Frankel et al. 2010) examine readability and tone (optimistic and pessimistic) of financial disclosures. In addition to the tone and readability analysis of financial statements, Loughran and McDonald (2016) suggest investigating uncertainty, litigious, strong, and weak modal words to examine additional means of parsing sentiment. Our focus on sentiment words in this study is uncertain and weak modal words. Uncertain words, such as *approximate, assume, contingent, depend,* and *indefinite,* express imprecision. *W*eak modal words such as "*might," "could," "maybe," "depending,"* and *"possible,"* express lack of confidence (Loughran and McDonald 2011). In this paper, we examine the association between uncertain and weak modal words in financial statements and audit fees and provide an incremental explanation of audit pricing.

Ambiguous worded 10-K reports may mislead the users of the financial statements and disclosures, decreasing the users' ability to understand and evaluate risks in investment, financing, and valuation (Kanagaretnam et al. 2020). Prior studies find that ambiguous words in financial disclosure may increase information risk and decrease the ability to comprehend and assess the valuation of firms. Loughran and McDonald (2011) report that ambiguous text in 10-K filings increases the stock return volatility in the year after 10-K filings. Loughran and McDonald (2013) find that the ambiguous tone of financial disclosures increases the valuation uncertainty of the firms. Ertugrul, Lei, Qiu, and Wan (2017) find a positive relation between ambiguous words in a client's 10-K filings and perceived information risk evaluated by creditors.

Ambiguous 10-K filings increase information risk for the auditors and decrease auditors' ability to comprehend financial reports. A study by Blanco et al. (2021) finds that a lower readability score[1] in 10-K filings is associated with longer audit delays and higher audit fees for U.S. auditors, indicating the auditor needs more time and effort when the readability of the 10-K is low. The ambiguity of tone in 10-K filings makes the disclosures harder to understand and, therefore, more difficult to interpret preparers' intentions. Consequently, we predict that more uncertain and weak modal words increase audit fees as these ambiguous words decrease auditors' comprehension of the disclosure leading to more audit effort.

The main results of this study show more ambiguous text in 10-K filings increases audit fees because the ambiguous tone in financial disclosures requires more time and analysis to assess a firm's risk characteristics and value-relevant information. To control potential endogeneity, we use propensity score matching, and we find the main results of this study hold with this technique. For the robustness tests, we exclude sample firms with going concern opinions, sample firms audited by Big 4 audit firms, and sample firms who report a net loss for their income statements and rerun the main regression analysis with the above changes. The coefficients of both *uncertain* and *weak modal words* for all the robustness tests remain positive and significant, which indicates the main results of this study still hold in all robustness tests.

This study contributes to the literature on the ambiguity of financial disclosures and audit fees. A recent study by Blanco et al. (2021) predicts and finds lower readability of 10-K filings is associated with higher audit fees. Our study differs from Blanco et al. (2021) because we use the ambiguous tone of financial statements instead of readability and find that uncertain and weak modal words positively relate to audit fees. Auditors can charge higher costs because of higher audit quality, monopoly pricing, or simply due to high-risk factors associated with clients (DeFond and Zhang 2014). Our study indicates that uncertain and weak modal words have significant incremental explanatory power in studying audit pricing research. This study provides some insights for 10-K preparers and auditors to take a strategic approach to audit costs and consider audits' efficiency and effectiveness. The result of this study can be in the interest of multiple stakeholders, such as issuers, auditors, investors, and regulators.

BACKGROUND AND HYPOTHESIS DEVELOPMENT

Ambiguous Words in 10-K Filings

Prior studies have examined the tone and sentiment of the 10-K filing (e.g., Loughran and McDonald, 2011) and find that market players consider management tone in financial disclosures to make important

decisions. Ambiguous words such as uncertain and weak modal words in 10-K filings affect efficient and effective communication about value-relevant information for investors (Loughran and McDonald, 2014). Loughran and McDonald (2011) create word lists to reflect ambiguous words in the financial context. For example, the list of uncertain words includes *"approximate," "assume," "contingent," "depend,"* and *"indefinite"* – words that express imprecision; the list of weak modal words contains *"might," "could," "maybe," "depending,"* and *"possible"* – words that indicate lack of confidence.

Audit Effort and Fees on Ambiguous Words in Financial Disclosures

Auditors play an essential role in providing independent assurance of the credibility of clients' financial statements enabling better resource allocation and client contracting efficiency (DeFond and Zhang, 2014). However, a less readable financial report may decrease audit efficiency and influence investors' investment behavior and public understanding of the firm. As a result, auditors spend more time and effort compensating for less readable 10-K filings, increasing audit costs due to decreased efficiency (Blanco et al. 2021). Prior studies analyze the effectiveness and efficiency of understanding ambiguous financial disclosures. Loughran and McDonald (2011) report that uncertain and weak modal words in 10-K filings are positively associated with the stock return volatility the following year after 10-K filings. Findings in Loughran and McDonald (2013) show that the ambiguous tone of financial disclosures is positively related to the valuation uncertainty of the firms. Ertugrul et al. (2017) argue that ambiguous text in 10-K filings may cause information risk or interfere with understanding the company's report for investors. They find a positive relation between the level of ambiguous words in 10-K filings and the difficulties in assessing a firm's risk factors and value, leading creditors to increase a client firm's perceived information risk. Overall, their results provide considerable evidence that the ambiguous language of 10-K filings is associated with a firm's information-concealing behavior, increasing its information risk and cost of capital. Thus, ambiguous language in annual reports can be a source of firm risk because it may increase informational risk and decrease auditors' ability to comprehend financial statements. Simunic (1980) finds that the auditor's effort and the expected losses from litigation can drive the increase in the firm's audit fees. Auditors may compensate for a greater ambiguity of 10-K filings by increasing audit time and effort, leading to higher audit fees. Thus, we predict the following in an alternative hypothesis form.

H1: *There is a positive relation between uncertain and weak modal words in 10-K filing and audit fees.*

RESEARCH DESIGN

We use the following model for H1 that predicts the relation between uncertain and weak modal words in 10-K filings and audit fees.

$$Auditfees_{it} = \beta_0 + \beta_1(Uncertain_{it}\ or\ Wmodal_{it}) + \beta_2 Size_{it} + \beta_3 Leverage_{it} + \beta_4 BTM_{it} + \beta_5 Loss_{it} + \beta_6 AltmanZ_{it} + \beta_7 Big4_{it} + \beta_8 M\&A_{it} + \beta_9 Foreign_{it} + \beta_{10} AssetTurn_{it} + \beta_{11} Current_{it} + \beta_{12} Quick_{it} + \beta_{13} ROA_{it} + \beta_{14} NAS_{it} + \beta_{15} OCF_{it} + \beta_{16} Extraord_{it} + \beta_{17} GCOpinion_{it} + \beta_{18} BusyFYE_{it} + \varepsilon_{it} \quad (1)$$

In Equation (1) above, the natural logarithm of Audit fees (*Auditfees*) is the dependent variable to proxy for audit effort. Two independent variables - *Uncertain* and *Wmodal* are in this model. Each independent variable will be tested separately because these two variables are highly correlated. Following prior literature, we include other control variables; firms size (*Size*); the total debt divided by total assets (*Leverage*); firm's book-to-market ratio (*BTM*); a dummy variable for net loss in the current year (*Loss*); a proxy for financial distress based on Altman (1968) (*AltmanZ*); a dummy variable for Big 4 auditors (*Big4*); a dummy variable for a firm who had an acquisition that contributed to sales (*M&A*); a dummy variable for a firm who reported foreign taxes (*Foreign*); asset turn over (*AssetTurn*), a ratio of current assets to total assets (*Current*); the ratio of current assets less inventory to current liabilities (*Quick*); return on assets (*ROA*); the ratio of non-audit serves to audit fees (*NAS*); operating cash flow divided by total assets (*OCF*);

a dummy variable for extraordinary items (*Extraord*); a dummy variable for going concern opinion issued by the auditor (*GCOpinion*); and a dummy variable for indicating busy season of the audit period (*BusyFYE*). See variable definitions in Appendix. We estimate Equation (2) below for propensity score matching analysis to address potential endogeneity.

$$DUncertain_{it}|DWmodal_{it} = \beta_0 + \beta_1 Size_{it} + \beta_2 Leverage_{it} + \beta_3 BTM_{it} + \beta_4 Loss_{it} + \beta_5 AltmanZ_{it} + \beta_6 Big4_{it} + \beta_7 M\&A_{it} + \beta_8 Foreign_{it} + \beta_9 AssetTurn_{it} + \beta_{10} Current_{it} + \beta_{11} Quick_{it} + \beta_{12} ROA_{it} + \beta_{13} NAS_{it} + \beta_{14} OCF_{it} + \beta_{15} Extraord_{it} + \beta_{16} GCOpinion_{it} + \beta_{17} BusyFYE_{it} + \varepsilon_{it} \quad (2)$$

For the first stage estimation of this analysis, like the above Equation (2), *DUncertian* (*or DWmodal*) is a dummy variable equal to 1 if the firm has a proportion of uncertain words (or weak modal words) equal to 75% percentile and above, otherwise zero. Control variables are the same as in Equation 1. To construct a propensity score matched sample, we match the sample without replacement using 0.01 caliper distance for the first step of the analysis. For the second stage estimation of this analysis, we re-estimate Equation (1) with a propensity score matched sample to examine if the main result of this study holds with this analysis.

For the robustness tests, first, we exclude firms with going concern opinions and rerun our main model. Second, we exclude firm samples audited by the Big 4 auditors because Big 4 auditors are fundamentally different from non-Big 4 auditors based on audit quality, reputation, client demographic, audit fee level, and audit risk (DeAngelo 1981; Francis and Wang 2005; DeFond and Zhang 2014). Lastly, according to descriptive statistics (Table 1), about 42.8% of firms in the sample report net loss; thus, we have excluded firms with losses to investigate if firms with losses may lead to the main results of this study.

DATA AND SAMPLE

We obtain uncertain and weak modal words from the word list on Professor McDonald's website[2]. We get audit fee data from Audit Analytics. Compustat North America Fundamentals Annual is a source of financial data. Our sample period is from 1999 to 2018[3]. We begin with 160,166 firm-year observations for uncertain and weak modal words from Professor McDonald's website. Following previous studies, we exclude financial firms, SIC 6000 – 6999, for 54,499 firm observations. This yields 105,667 firm-year observations. We merge 105,667 firm-year observations with Compustat financial data and audit fees from Audit Analytics and drop 49,143 with missing audit and financial data. This yields the final sample of 56,524 firm-year observations.

RESULTS

Descriptive Statistics

Table 1 provides descriptive statistics of variables used for correlation tests and regression analyses. *Auditfees* indicates audit fees in a million dollars, and the mean is about 1.5 million dollars. Uncertain and weak modal words are about 1.2% and 0.5% of the total count of words in 10-K filings, accordingly. We winsorize all continuous variables at the 1st and 99th percentiles. We use the natural logarithm of audit fees for the analyses in Tables 2 – 7. We use the robust standard error to control for the possibility that the error terms not having constant variance (i.e., heteroscedasticity). All other variables' definitions are in Appendix.

Table 2 presents Pearson Pairwise Correlation coefficients. Consistent with our expectations, we find a positive and significant relation ($p < .01$) between *Uncertain* and *Auditfees* (coefficient = 0.1218). We also find a positive and significant relation ($p < .01$) between *Wmodal* and *Auditfees* (coefficient = 0.0305). These results show preliminary support for our expectations.

TABLE 1
DESCRIPTIVE STATISTICS
(n = 56,524)

Variable	Mean	Std. Dev.	25th Percentile	Median	75th Percentile
Auditfees ($ million)	1.527	3.176	0.188	0.578	1.600
Uncertain	0.012	0.003	0.010	0.012	0.015
Wmodal	0.005	0.002	0.004	0.005	0.007
Size	5.425	2.663	3.883	5.586	7.213
Leverage	0.360	0.915	0.009	0.182	0.372
BTM	0.035	0.114	0.002	0.009	0.031
Loss	0.428	0.495	0.002	0.002	1.000
AltmanZ	-3.002	44.331	0.950	2.790	5.150
Big4	0.660	0.474	0.000	1.000	1.000
M&A	0.034	0.182	0.000	0.000	0.000
Foreign	0.437	0.496	0.000	0.000	1.000
AssetTurn	1.054	0.892	0.437	0.856	1.429
Current	0.512	0.268	0.296	0.511	0.727
Quick	2.390	3.124	0.842	1.411	2.587
ROA	-0.455	2.194	-0.146	0.018	0.066
NAS	0.355	0.582	0.036	0.157	0.401
OCF	-0.129	0.819	-0.038	0.064	0.121
Extraord	0.039	0.193	0.000	0.000	0.000
GCOpinion	0.111	0.314	0.000	0.000	0.000
BusyFYE	0.720	0.449	0.000	1.000	1.000

TABLE 2
CORRELATIONS

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1 *Auditfees*	1																			
2 *Uncertain*	0.1218*	1																		
3 *Wmodal*	0.0305*	0.7643*	1																	
4 *Size*	0.8603*	0.0289*	-0.0676*	1																
5 *Leverage*	-0.2136*	-0.0947*	-0.0581*	-0.3551*	1															
6 *BTM*	-0.1810*	-0.0801*	-0.1345*	-0.0923*	-0.1729*	1														
7 *Loss*	-0.3350*	0.0550*	0.1970*	-0.4434*	0.1675*	-0.0414*	1													
8 *AltmanZ*	0.2588*	0.0727*	0.0363*	0.4526*	-0.7197*	0.1235*	-0.1951*	1												
9 *Big4*	0.5767*	0.0313*	0.0091	0.5880*	-0.1826*	-0.1130*	-0.2448*	0.2164*	1											
10 *M&A*	0.1306*	0.1047*	0.0713*	0.1086*	-0.0142*	-0.0240*	-0.0521*	0.0255*	0.0362*	1										
11 *Foreign*	0.5104*	0.0926*	-0.0297*	0.4206*	-0.1392*	-0.0689*	-0.2374*	0.1377*	0.2996*	0.0876*	1									
12 *AssetTurn*	-0.0035	-0.1288*	-0.2021*	-0.0348*	-0.0150*	0.0909*	-0.1732*	0.0209*	-0.0327*	-0.0231*	0.0344*	1								
13 *Current*	-0.2689*	0.1258*	0.2023*	-0.4033*	-0.0204*	0.0657*	0.1754*	-0.0776*	-0.1403*	-0.0925*	-0.0067	0.2033*	1							
14 *Quick*	-0.1412*	0.1782*	0.2690*	-0.0894*	-0.1814*	0.0505*	0.1022*	0.2094*	0.007	-0.0409*	-0.0856*	-0.2812*	0.3814*	1						
15 *ROA*	0.3372*	0.0481*	-0.0149*	0.5226*	-0.6439*	0.1193*	-0.2800*	0.7950*	0.2514*	0.0383*	0.1794*	0.0814*	-0.1336*	0.1005*	1					
16 *NAS*	-0.0711*	-0.1712*	-0.1529*	0.1136*	-0.0576*	0.0079	-0.0439*	0.0674*	0.1121*	-0.0350*	0.0443*	-0.0059	-0.0313*	0.0135*	0.0626*	1				
17 *OCF*	0.3618*	0.0263*	-0.0630*	0.5610*	-0.6095*	0.1176*	-0.3212*	0.7491*	0.2685*	0.0423*	0.2138*	0.1304*	-0.1970*	0.0582*	0.8482*	0.0726*	1			
18 *Extraord*	0.0274*	-0.1344*	-0.1353*	0.0921*	-0.0074	0.0017	0.0136*	0.0218*	0.0729*	-0.0373*	-0.0027	-0.0015	-0.0746*	-0.0410*	0.0291*	0.1897*	0.0389*	1		
19 *GCOpinion*	-0.4118*	-0.0579*	0.0134*	-0.5404*	0.4231*	-0.1610*	0.3597*	-0.4511*	-0.3528*	-0.0516*	-0.2515*	-0.0826*	0.0523*	-0.1348*	-0.5016*	-0.0914*	-0.5313*	-0.0346*	1	
20 *BusyFYE*	0.1165*	0.0371*	0.1104*	0.1078*	0.0018	-0.0807*	0.0304*	0.0181*	0.0887*	0.0186*	-0.0368*	-0.0783*	-0.0882*	0.0193*	0.0290*	-0.0120*	0.0202*	0.0278*	-0.0329*	1

This table presents the Pearson pairwise correlations between the dependent variable, ambiguous independent variables, and other control variables.
* denotes significant at p-value < 0.01

Main Results

Table 3 shows the main results, Column 1 represents results with uncertain words, and Column 2 states results with weak modal words in 10-K filings. Column 1 (Column 2) regresses *Auditfees* on *Uncertain* (*Wmodal*) and control variables. Coefficients on *Uncertain* and *Wmodal* are positive and significant (coefficient = 4.350; t-value = 4.67 and coefficient =6.166; t-value = 3.81, respectively), meaning more ambiguity of 10-K filings increases audit fees. As expected, the following control variables are positively associated with audit fees (*Size*, *Loss*, *Big4*, *M&A*, *Foreign*, *AssetTurn*, *Current*, *Extraord*, *GCOpinion*, and *BusyFYE*) and negatively associated with audit fees (*BTM*, *AltmanZ*, *Quick*, *NAS*, and *OCF*). The above results show that more ambiguous words in 10-K filings increase audit fees because the ambiguous tone in financial disclosures requires more time and analysis to assess a firm's risk characteristics and value-relevant information. Overall, the signs on the coefficients of control variables are consistent with prior literature.

TABLE 3
AUDIT FEES AND UNCERTAIN AND WEAK MODAL WORDS IN 10-K FILINGS

	Dependent Variable: Audit Fees	
Variable	**Key IV: *Uncertain***	**Key IV: *Wmodal***
Uncertain	4.350***	
	(4.67)	
Wmodal		6.166***
		(3.81)
Size	0.499***	0.499***
	(257.36)	(257.23)
Leverage	0.000	0.000
	(-0.06)	(-0.08)
BTM	-0.384***	-0.380***
	(-17.25)	(-17.02)
Loss	0.188***	0.187***
	(34.31)	(33.87)
AltmanZ	-0.002***	-0.002***
	(-16.01)	(-16.04)
Big4	0.366***	0.365***
	(55.29)	(54.81)
M&A	0.091***	0.092***
	(7.89)	(8.02)
Foreign	0.309***	0.310***
	(51.41)	(51.67)
AssetTurn	0.101***	0.101***
	(23.77)	(23.77)
Current	0.323***	0.320***
	(20.93)	(20.62)
Quick	-0.027***	-0.028***
	(-25.18)	(-25.23)
ROA	0.000	0.000
	(0.00)	(-0.01)
NAS	-0.218***	-0.218***
	(-41.92)	(-41.97)
OCF	-0.150***	-0.149***
	(-16.92)	(-16.87)

Extraord	0.119***	0.119***
	(9.04)	(9.05)
GCOpinion	0.050***	0.050***
	(4.26)	(4.23)
BusyFYE	0.104***	0.102***
	(18.38)	(18.13)
Constant	9.29***	9.313***
	(102.00)	(101.94)
Industry/Year FE	Yes	Yes
Robust SE	Yes	Yes
Observations	56,524	56,524
R-squared	0.8769	0.8769

*, **, *** Indicate significance at the 0.10, 0.05, and 0.01 levels, respectively (based on two-tailed tests). Models are estimated using OLS regression with industry and year fixed effects. We winsorize all continuous variables at the 1st and 99th percentiles. Standard errors are robust to heteroscedasticity. T-statistics are in parentheses below the coefficients.

Results From Additional Tests

Propensity Score Matching to Address Potential Endogeneity

Firms may not randomly choose the level of uncertain and weak modal words in their 10-K filings, which can cause bias in non-randomized and observational research. We use the propensity score matching technique to address potential endogeneity and document the results of a two-stage estimation in Table 4. The first and third column in Table 4 presents the result of the first stage estimation. We find a positive relation between *Loss, M&A, Current, Quick, BusyFYE* and uncertain and weak modal words. Firms with higher ambiguous words in 10-K filings are more likely to make a net loss and include sales from Merger and Acquisition. Moreover, we find a negative relation between *Leverage*, *BTM*, *AssetTurn*, *NAS*, *Extraord*, and *GCOpinion* and both *Uncertain* and *Wmodal*. The second steps for both uncertain and weak modal words present a positive relation between ambiguous words in 10-K filings and audit fees, and coefficients on other variables are constant with those for Table 3, meaning that the main result of this study holds with propensity score matching analysis.

TABLE 4
PROPENSITY SCORE MATCHED REPORT

	Dependent Variable: Auditfees			
Variable	**First Step *Uncertain***	**Second Step *Uncertain***	**First Step *Wmodal***	**Second Step *Wmodal***
Uncertain		3.543***		
		(2.77)		
Wmodal				4.571**
				(2.05)
Size	-0.011	0.489***	-0.034***	0.476***
	(-1.62)	(175.04)	(-4.60)	(162.11)
Leverage	-0.155***	-0.016**	-0.124***	-0.013*
	(-7.94)	(-2.08)	(-7.19)	(-1.95)
BTM	-1.386***	-0.435***	-1.985***	-0.436***
	(-12.50)	(-10.82)	(-17.09)	(-10.92)
Loss	0.266***	0.169***	0.529***	0.162***
	(11.28)	(22.05)	(21.90)	(20.46)
AltmanZ	0.001	-0.002***	0.001**	-0.002***
	(1.35)	(-10.83)	(2.57)	(-11.98)

Big4	0.000	0.413***	0.117***	0.437***
	(0.01)	(44.77)	(4.22)	(45.92)
M&A	0.857***	0.100***	0.792***	0.118***
	(17.64)	(7.76)	(15.56)	(8.12)
Foreign	0.353***	0.286***	-0.043*	0.290***
	(15.03)	(33.40)	(-1.75)	(32.89)
AssetTurn	-0.307***	0.111***	-0.432***	0.118***
	(-20.05)	(16.65)	(-26.85)	(17.47)
Current	0.931***	0.303***	1.468***	0.281***
	(18.32)	(14.56)	(28.46)	(13.2)
Quick	0.038***	-0.026***	0.057***	-0.026***
	(10.07)	(-19.94)	(14.47)	(-19.89)
ROA	0.001	0.005	0.021**	0.004
	(0.05)	(0.90)	(2.06)	(0.81)
NAS	-0.509***	-0.264***	-0.525***	-0.252***
	(-21.36)	(-28.17)	(-21.59)	(-26.60)
OCF	0.025	-0.182***	-0.041	-0.161***
	(0.88)	(-13.80)	(-1.58)	(-14.28)
Extraord	-1.221***	0.153***	-1.442***	0.103***
	(-14.02)	(4.73)	(-14.94)	(2.92)
GCOpinion	-0.281***	-0.002	-0.252***	-0.004
	(-6.38)	(-0.10)	(-5.83)	(-0.23)
BusyFYE	0.137***	0.089***	0.426***	0.107***
	(5.92)	(11.03)	(17.19)	(12.12)
Constant	-1.422***	9.615***	-1.759***	9.527***
	(-24.45)	(234.39)	(-29.22)	(92.69)
Industry/Year FE	Yes	Yes	Yes	Yes
Robust SE	Yes	Yes	Yes	Yes
Observations	56,524	28,028	56,524	25,326
R-squared	0.0640	0.8719	0.1183	0.8706

*, **, *** Indicate significance at the 0.10, 0.05, and 0.01 levels, respectively (based on two-tailed tests). We estimate First Step models using logistic regression models and Second Step models using OLS regression models. We winsorize all continuous variables at the 1st and 99th percentiles. Standard errors are robust to heteroscedasticity. T-statistics are shown in parentheses below the coefficients.

Test Without Firms Reporting Going Concern Opinions

We exclude firms with going concern opinions and rerun our main model because previous research (Abernathy, Guo, Kubick, and Masli 2019 and Blanco et al. 2021) includes a test with going concern opinion variables, and these variables may lead to the result of our study. The results in Table 5 show that the coefficients of both uncertain and weak modal words are positive and significant (column 1: coefficient = 2.346 and t-value = 2.47; column 2: coefficient = 2.728 and t-value = 1.65). These results show that the main results of this study are consistent with this robustness test without going concern firms.

TABLE 5
ANALYSIS WITHOUT GOING CONCERN FIRMS

	Dependent Variable: *Auditfees*	
Variable	**Key IV: *Uncertain***	**Key IV: *Wmodal***
Uncertain	2.346**	
	(2.47)	
Wmodal		2.728*
		(1.65)
Size	0.509***	0.509***
	(253.30)	(252.98)
Leverage	0.002	0.001
	(0.14)	(0.07)
BTM	-0.381***	-0.38***
	(-16.02)	(-15.91)
Loss	0.178***	0.177***
	(29.87)	(29.72)
AltmanZ	-0.003***	-0.003***
	(-8.29)	(-8.30)
Big4	0.339***	0.339***
	(49.93)	(49.52)
M&A	0.075***	0.076***
	(6.58)	(6.66)
Foreign	0.294***	0.294***
	(48.17)	(48.32)
AssetTurn	0.112***	0.112***
	(23.29)	(23.25)
Current	0.310***	0.309***
	(18.37)	(18.19)
Quick	-0.025***	-0.025***
	(-20.2)	(-20.21)
ROA	0.010	0.010
	(0.99)	(0.98)
NAS	-0.210***	-0.211***
	(-40.39)	(-40.43)
OCF	-0.277***	-0.276***
	(-13.83)	(-13.8)
Extraord	0.121***	0.121***
	(9.07)	(9.07)
GCOpinion		
BusyFYE	0.100***	0.099***
	(17.36)	(17.22)
Constant	9.196***	9.211***
	(128.20)	(128.41)
Industry/Year FE	Yes	Yes
Robust SE	Yes	Yes
Observations	50,263	50,263
R-squared	0.8647	0.8647

*, **, *** Indicate significance at the 0.10, 0.05, and 0.01 levels, respectively (based on two-tailed tests). Models are estimated using OLS regression with industry and year fixed effects. We winsorize all continuous variables at the 1st and 99th percentiles. Standard errors are robust to heteroscedasticity. T-statistics are shown in parentheses below the coefficients.

The Effect of Big-4 Auditors on Ambiguous Tone in 10-K Filings and Audit Fees

Previous studies document that clients are willing to pay a higher price to the Big N audit firms for the job because they expect a higher quality audit from the Big N auditors (DeFond and Zhang 2014). We exclude firm samples audited by Big 4 because Big 4 auditors are fundamentally different from non-Big 4 auditors based on audit quality, reputation, client demographic, audit fee level, and audit risk (DeAngelo 1981; Francis and Wang 2005; DeFond and Zhang 2014). Therefore, sample firms audited by Big-4 auditors may drive the result. We present the result in Table 6. The results without Big 4 auditor's engagement show a positive and significant result (Column 1: coefficient = 9.792 and t-value = 5.77; Column 2: coefficient = 15.269 and t-value = 5.60). These result indicates that Big 4 auditors do not drive our main findings.

TABLE 6
ANALYSIS WITHOUT BIG 4 AUDIT FIRMS

	Dependent Variable: *Auditfees*	
Variable	**Key IV: *Uncertain***	**Key IV: *Wmodal***
Uncertain	9.792***	
	(5.77)	
Wmodal		15.269***
		(5.60)
Size	0.479***	0.479***
	(128.7)	(128.49)
Leverage	-0.003	-0.003
	(-0.66)	(-0.64)
BTM	-0.462***	-0.453***
	(-15.62)	(-15.26)
Loss	0.168***	0.165***
	(17.04)	(16.64)
AltmanZ	-0.002***	-0.002***
	(-16.20)	(-16.25)
Big4		
M&A	0.093***	0.096***
	(3.68)	(3.80)
Foreign	0.338***	0.341***
	(28.88)	(29.18)
AssetTurn	0.089***	0.090***
	(15.23)	(15.32)
Current	0.234***	0.231***
	(10.02)	(9.89)
Quick	-0.023***	-0.024***
	(-15.02)	(-15.11)
ROA	-0.006*	-0.006*
	(-1.73)	(-1.72)
NAS	-0.261***	-0.261***
	(-22.84)	(-22.85)
OCF	-0.127***	-0.126***
	(-13.62)	(-13.49)
Extraord	0.109***	0.109***
	(3.27)	(3.27)
GCOpinion	-0.015	-0.016
	(-0.98)	(-1.06)

BusyFYE	0.092***	0.088***
	(9.19)	(8.86)
Constant	9.494***	9.531***
	(53.03)	(52.78)
Industry/Year FE	Yes	Yes
Robust SE	Yes	Yes
Observations	19,203	19,203
R-squared	0.7817	0.7817

*, **, *** Indicate significance at the 0.10, 0.05, and 0.01 levels, respectively (based on two-tailed tests). Models are estimated using OLS regression with industry and year fixed effects. We winsorize all continuous variables at the 1st and 99th percentiles. Standard errors are robust to heteroscedasticity. T-statistics are shown in parentheses below the coefficients.

The Effect of Loss Firms on Ambiguous Tone in 10-K Filings and Audit Fees

Interestingly, descriptive statistics of Table 1 show that 42.8% of sample firms report a net loss. Thus, these loss firms may lead to the main results of this study. We exclude sample firms that report a net loss for their financial statement and present a report in Table 7. We find positive and significant coefficients (Column 1: coefficient = 2.711 and t-value = 2.27; Column 2: coefficient = 3.716 and t-value = 1.73) on both *Uncertain* and *Wmodal* indicating that loss firms in our sample do not lead to the main result of this study.

TABLE 7
ANALYSIS WITHOUT LOSS FIRMS

	Dependent Variable: *Auditfees*	
Variable	**Key IV: *Uncertain***	**Key IV: *Wmodal***
Uncertain	2.711**	
	(2.27)	
Wmodal		3.716*
		(1.73)
Size	0.514***	0.514***
	(204.85)	(204.50)
Leverage	0.050***	0.049***
	(3.41)	(3.38)
BTM	-0.421***	-0.419***
	(-11.65)	(-11.55)
Loss		
AltmanZ	-0.002***	-0.002***
	(-5.42)	(-5.44)
Big4	0.327***	0.327***
	(36.37)	(36.21)
M&A	0.058***	0.058***
	(4.54)	(4.60)
Foreign	0.288***	0.288***
	(37.38)	(37.52)
AssetTurn	0.108***	0.108***
	(16.63)	(16.62)
Current	0.310***	0.309***
	(13.10)	(13.01)

Quick	-0.025***	-0.026***
	(-12.54)	(-12.57)
ROA	-0.075	-0.075
	(-1.38)	(-1.39)
NAS	-0.187***	-0.187***
	(-28.70)	(-28.73)
OCF	-0.253***	-0.252***
	(-7.57)	(-7.56)
Extraord	0.132***	0.132***
	(7.83)	(7.84)
GCOpinion	0.008	0.007
	(0.18)	(0.16)
BusyFYE	0.106***	0.106***
	(15.03)	(14.91)
Constant	9.184***	9.199***
	(109.02)	(108.66)
Industry/Year FE	Yes	Yes
Robust SE	Yes	Yes
Observations	32,316	32,316
R-squared	0.8733	0.8733

*, **, *** Indicate significance at the 0.10, 0.05, and 0.01 levels, respectively (based on two-tailed tests). Models are estimated using OLS regression with industry and year fixed effects. We winsorize all continuous variables at the 1st and 99th percentiles. Standard errors are robust to heteroscedasticity. T-statistics are shown in parentheses below the coefficients.

CONCLUSION

This study examines the effect of ambiguous words in 10-K filings on audit fees. This study finds that more uncertain and weak modal words in 10-K filings increase audit fees as these ambiguous words require more auditors' efforts to analyze clients' financial disclosures. The main results of this study hold when we re-estimate the result with propensity score matching to control endogeneity. Moreover, we re-estimate the main regression test by excluding clients that were issued going concern opinions, audited by Big-4 auditors, and reported a net loss in their income statements. We find the main results still hold in all robustness tests. This study contributes to the literature regarding audit pricing and ambiguous sentiment textual analysis in financial disclosures. This study provides evidence that the auditor increases the engagement risk proxied by audit fees when clients' annual report includes more ambiguous words. The result of this study can be in the interest of multiple stakeholders, such as issuers, investors, auditors, and regulators, as SEC has raised concerns that firms may purposely use ambiguous language in 10-K reports to protect themselves against possible legal claims or poor financial performances (SEC 2007). Even though a 10-K report is mandatory, the degree of clarity of the text in the 10-K filings is determined by the discretion of the management, so interpretation and decisions made by users of these ambiguous reports are non-trivial (Kanagaretnam et al. 2020).

ENDNOTES

1. Blanco et al. (2021) use the Bog Index to measure a readability of firms' 10-K disclosures. The bog index is more accurate than other readability measures such as the FOG index, 10-K file size, and the number of words in the document (Bonsall, Leone, Miller, and Rennekamp, 2017). Please visit Professor Miller's webpage for more detail. https://host.kelley.iu.edu/bpm/activities/bogindex.html
2. You can find a file containing all summary data for all 10-K filings for sentiment word counts (e.g., uncertain, weak modal, litigious words) here. https://sraf.nd.edu/sec-edgar-data/lm_10x_summaries/. You can find

Loughran-McDonald Master Dictionary with Sentiment Word Lists here. https://sraf.nd.edu/loughranmcdonald-master-dictionary/.

3. We start our sample period from 1999 because the earliest year of audit fees we can obtain is 2000. And we stop our sample period in 2018 because of the availability of uncertain and weak modal words from Professor McDonald's webpage.

REFERENCES

Abernathy, J.L., Guo, F., Kubick, T.R., & Masli, A. (2019). Financial statement footnote readability and corporate audit outcomes. *Auditing: A Journal of Practice & Theory*, *38*(2), 1–26.

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

Biddle, G.C., Gilles, H., & Verdi, R.S. (2009). How does financial reporting quality relate to investment efficiency? *Journal of Accounting and Economics*, *48*(2–3), 112–131.

Blanco, B., Coram, P., Dhole, S., & Kent, P. (2021). How do auditors respond to low annual report readability? *Journal of Accounting and Public Policy*, *40*(3), 106769.

Bonsall, S.B., IV, Leone, A.J., Miller, B.P., & Rennekamp, K. (2017). A plain English measure of financial reporting readability. *Journal of Accounting and Economics*, *63*(2–3), 329–357.

Davis, A.K., Piger, J.M., & Sedor, L.M. (2012). Beyond the numbers: Measuring the information content of earnings press release language. *Contemporary Accounting Research*, *29*(3), 845–868.

DeAngelo, L.E. (1981). Auditor size and audit quality. *Journal of Accounting and Economics*, *3*(3), 183–199.

DeFond, M., & Zhang, J. (2014). A review of archival auditing research. *Journal of Accounting and Economics*, *58*(2–3), 275–326.

Demers, E., & Vega, C. (2014). *Understanding the role of managerial optimism and uncertainty in the price formation process: Evidence from the textual content of earnings announcements*. [Working paper, University of Virginia].

Ertugrul, M., Lei, J., Qiu, J., & Wan, C. (2017). Annual Report Readability, Tone Ambiguity, and the Cost of Borrowing. *Journal of Financial and Quantitative Analysis*, *52*(2), 811–836.

Francis, J.R., & Wang, D. (2005). Impact of the SEC's public fee disclosure requirement on subsequent period fees and implications for market efficiency. *Auditing: A Journal of Practice & Theory*, *24*(s-1), 145–160.

Frankel, R., Mayew, W., & Sun, Y. (2010). Do pennies matter? Investor relations consequences of small negative earnings surprises. *Review of Accounting Studies*, *15*, 220–242.

Kanagaretnam, K., Mawani, A., Shi, G., & Zhou, Z. (2020). Impact of social capital on tone ambiguity in banks' 10-K filings. *Journal of Behavioral and Experimental Finance*, *28*, 100411.

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

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

Li, F. (2008). Annual Report Readability, Current Earnings, and Earnings Persistence. *Journal of Accounting and Economics*, *45*, 221–47.

Loughran, T., & McDonald, B. (2011). When Is a Liability Not a Liability? Textual analysis, Dictionaries, and 10-Ks. *Journal of Finance*, *66*, 35–65.

Loughran, T., & McDonald, B. (2013). IPO First-Day Returns, Offer Price Revisions, Volatility, and Form S-1 Language. *Journal of Financial Economics*, *109*, 307–26.

Loughran, T., & McDonald, B. (2014). Measuring Readability in Financial Disclosures. *Journal of Finance*, *69*, 1643–71.

Loughran, T., & McDonald, B. (2016). Textual Analysis in Accounting and Finance: A Survey. *Journal of Accounting Research*, *54*(4), 1187–1230.

Miller, B.P. (2010). The effects of reporting complexity on small and large investor trading. *The Accounting Review*, *85*(6), 2107–2143.
SEC. (2007). *Speech by SEC chairman: Closing remarks to the second annual corporate governance summit.* Retrieved from https://www.sec.gov/news/speech/2007/spch032307cc.htm
Simunic, D.A. (1980). The pricing of audit services: Theory and evidence. *Journal of Accounting Research*, pp. 161–190.

APPENDIX: VARIABLE DEFINITIONS

Variable		Description
Dependent Variable		
Auditfees	=	Natural logarithm of audit fees in a million dollars.
Independent Variables		
Uncertain	=	The proportion of uncertain words to the total words in 10-K filings as defined in Loughran and McDonald (2011).
Wmodal	=	The proportion of weak modal words to the total words in 10-K filings as defined in Loughran and McDonald (2011).
Other Variables		
Size	=	The natural logarithm of total assets.
Leverage	=	The ratio of total debt divided by total assets.
BTM	=	The book value of shareholders' equity divided by market value of equity.
Loss	=	a dummy variable for a net loss in the current year
AltmanZ	=	Measurement of risk of financial distress (Altman 1968). It is calculated from 1.2 × T1 + 1.4 × T2 + 3.3 × T3 + 0.6 × T4 + 0.999 × T5, where T1 stands for working capital divided by total assets; T2 stands for retained earnings divided by total assets; T3 stands for earnings before interest and tax expenses divided by total assets; T4 stands for market value of equity divided by total liabilities; T5 stands for sales divided by total assets.
Big4	=	1 if audited by one of the Big 4 auditors, else 0.
M&A	=	1 if a firm had an acquisition that contributed to sales, and 0 otherwise.
Foreign	=	1 if a firm had foreign exchange income (loss), else 0.
AssetTurn	=	The ratio of sales to total assets.
Current	=	The ratio of current assets to total assets.
Quick	=	The ratio of current assets less inventory to current liabilities.
ROA	=	The ratio of net income to total assets.
NAS	=	The ratio of non-audit services to audit fees.
OCF	=	Cash flow from operating activities is divided by total assets.
Extraord	=	1 if the extraordinary item is reported in the income statement, 0 otherwise
GCOpinion	=	1 if a going concern opinion is issued by the auditor to the firm, and 0 otherwise.
BusyFYE	=	1 if a client's fiscal year ends in either December or January and 0 otherwise.