The Information Content of Forward-looking Statements in Corporate Filings—a Naïve Bayesian Machine Learning Approach *

Feng Li

Ross School of Business, University of Michigan

701 Tappan St., Ann Arbor, MI 48109

Phone: (734) 936-2771

E-mail: Feng@umich.edu

First Draft: January 2008
This Draft: November 29, 2009

*I would like to thank David Aboody, Phil Berger, Dave Burgstahler, Judson Caskey, John Core, Ilia Dichev, Weili Ge, Joseph Gerakos, Jack Hughes, Raffi Indjejikian, Mark Lang, Roby Lehavy, Joshua Livnat, Russell Lundholm, Dawn Matsumoto, Venky Nagar, Jim Ohlson, Marlene Plumlee, Jonathan Rogers, Joshua Ronen, Gil Sadka, Haresh Sapra, Cathy Shakespeare, Terry Shevlin, Doug Skinner, Abbie Smith, K. R. Subramanyam, Wayne Thomas, Brett Trueman, Andy Van Buskirk, Hal White, Regina Wittenberg Moerman, Sarah Zechman, and the participants of the 2008 University of Michigan Traverse City Conference, the 2008 North American Summer Meeting of the Econometric Society, the 2009 FARS mid-year conference, the New York University 2009 Summer Camp, and the workshop participants at Chicago, Notre Dame, University of Washington, UCLA, Indiana, and Florida for helpful comments. I also thank Mike Chen, Qian Fang, Qian Li, Chris Liong, Doris Lu, Amy McMahon, Bharat Modi, Qimin Ni, Darin Styles, Sarah Xia, Quan Yuan, Kathy Wang, Tina Wei, and especially Rencheng Wang and Winnie Wei for their excellent research assistance and the Stephen M. Ross School of Business and the Harry Jones Endowment for Earnings Quality at the University of Michigan for financial support.

Abstract

This paper examines the tone and content of the forward-looking statements (FLS) in the Management Discussion and Analysis section (MD&A) of corporate 10-K and 10-Q filings using a Naïve Bayesian machine learning algorithm. I first manually categorize 30,000 sentences of randomly-selected FLS extracted from the MD&As along two dimensions: (1) tone (i.e., positive versus negative tone); and (2) content (i.e., profitability, operations, liquidity etc.). These manually-coded sentences are then used as training data in a Naïve Bayesian machine learning algorithm to classify the tone and content of about 13 million forward-looking statements from more than 140,000 corporate 10-K and 10-Q MD&As between 1994 and 2007.

I find that firms with better current performance, lower accruals, smaller size, lower market-to-book ratio, and less return volatility tend to have more positive forward-looking statements in their MD&As. Furthermore, the average tone of the forward-looking statements in a firm's MD&A is positively associated with future earnings and liquidity, even after controlling for other determinants of future performance. The results also show that there is no systematic change in the information content of MD&As over time. Finally, the information in MD&As seems to mitigate the mis-pricing of accruals: When managers "warn" about the future performance implications of accruals (i.e., MD&A tone is positive (negative) when accruals are negative (positive)), accruals are not associated with future returns; on the other hand, for firms whose managers do not provide such a "warning," accruals are negatively correlated with future returns.

1 Introduction

This paper analyzes the information content of the forward-looking statements in the Management's Discussion and Analysis section of corporate 10-K and 10-Q filings by examining the tone and content of the statements. In 1980, the Securities and Exchange Commission (SEC) mandated that public companies include in their annual reports a section for Management's Discussion and Analysis of Financial Condition and Results of Operations (MD&A). The MD&A is intended to assess an enterprise's liquidity, capital resources, and operations in a way that many investors can understand. One of the SEC's goals in mandating the MD&A was to make public the information about predictable future events and trends that may affect future operations of the business.

Despite this goal, whether MD&A disclosures are truly informative remains an open empirical question. On the one hand, consistent with the SEC's intention that MD&A disclosures provide relevant information for investors, the MD&A is arguably the most read and most important component of the financial section (Tavcar (1998)). Indeed, of all the disclosure items of the annual report, sell-side financial analysts in the U.S. most frequently use the MD&A when preparing their analyst reports (Knutson (1993) and Rogers and Grant (1997)). Furthermore, the safe harbor provisions of the Private Securities Litigation Reform Act of 1995 encourage more forward-looking information (Grundfest and Perino (1997)) and therefore should also make MD&A disclosures more informative.

On the other hand, the MD&A may not be as informative as intended for several reasons. First, companies may not have incentives to disclose in the MD&A because of concerns over proprietary costs (Verrecchia (1983)), uncertainties about the judicial interpretation of safe harbor protection (Grundfest and Perino (1997)), or the lack of a mandatory auditing requirement for MD&As (Hüfner (2007)). Another concern with MD&As is that many include substantial boilerplate disclaimers and disclosures, generic language, and immaterial detail without much information content (SEC (2003)). Consistent with these arguments, Pava and Epstein (1993) show that most of the 25 randomly-selected companies they study accurately describe historical events, but very few provide useful and accurate forecasts in their MD&As.

In this paper, I study the determinants and information content of the forward-looking statements in the MD&A section, as these are likely to be the most informative part of the section. Specifically, I first explore variations in the tone of the forward-looking statements in the MD&As and study their economic determinants. Second, I examine whether the forward-looking statements contain (incremental) information about future profitability and liquidity. I then explore whether the information content of MD&As changes over time, especially after 2003, when the SEC issued new guidelines for preparing MD&As and the Sarbanes-Oxley Act increased disclosure requirements in MD&As (Bainbridge (2007)). Finally, I examine the implications of MD&A tone for the accrual anomaly.

To assess the content and tone of forward-looking statements (FLS) in MD&As, I rely on a Naïve Bayesian learning algorithm instead of a dictionary-based approach (e.g., Kothari, Li, and Short (2008)).[1] First, I manually categorize 30,000 sentences of randomly-selected forward-looking statements extracted from the MD&A section of 10-Q filings along two dimensions: (1) tone (i.e., positive versus negative tone) and (2) content (i.e., profitability operations, liquidity, etc.). These manually-coded sentences are then used as training data in the Naïve learning algorithm to categorize the tone and content of other forward-looking statements in 10-Q and 10-K filings. N-fold cross-validation tests (with N varying from 3 to 50) indicate that the algorithm predicts the tone (out of three possible tones: positive, neutral, and negative) and content (out of twelve possible categories) of forward-looking statements in 10-Q and 10-K filings with a success rate of about 67% and 63%, respectively.

After developing the training data, I then use the Bayesian learning algorithm to categorize the tone and content of about 13 million forward-looking statements from more than 140,000 corporate 10-Q and 10-K filings between 1994 and 2007. The results indicate that firms with better current performance, lower accruals, smaller size, lower market-to-book ratio, and less return volatility tend to have more positive forward-looking statements in

[1] The dictionary approach relies on a "mapping" algorithm and assigns each word (or phrase) in a document into different categories based on some pre-defined dictionaries.

their MD&As.

I find that the average tone of the forward-looking statements in a firm's MD&A is positively correlated with its future earnings and liquidity and has explanatory power in addition to other variables in predicting future performance. For instance, the return on assets in the next year for firms with a perfectly positive MD&A tone is 8 percentage points higher than that for firms with a perfectly negative tone. An inter-quartile change in MD&A tone implies a difference in annual return on assets of 1.5 percentage points. These effects are found after controlling for current earnings, stock returns, accruals, and other factors that may affect future performance. An examination of the information content of MD&As over time shows that, despite the SEC's continuous effort to strengthen MD&A disclosures and the passage of the Sarbanes-Oxley Act, there is no systematic change in the informativeness of MD&As over time.

The tone of MD&As is also related to the cross-sectional association of accruals with future stock returns. While, on average, accruals are negatively associated with future returns (Sloan (1996)), lower (higher) accruals are not associated with higher (lower) future returns if managers "warn" about the future outlook in the MD&A disclosures (i.e., the accruals are negative (positive) and MD&A tone is positive (negative)). This suggests that MD&A disclosures mitigate the mis-pricing of accruals.

Finally, I decompose the tone of the forward-looking statements in the MD&As into three dimensions: profitability-related FLS (i.e., revenue, cost, and operations), liquidity-related FLS (i.e., capital resources, financing, and investment), and other FLS, where the categories are generated by the learning algorithm classification. The empirical results show that both the profitability-related FLS and the liquidity-related FLS have information content about future fundamentals, but the liquidity-related FLS have stronger implications both economically and statistically.

This paper contributes to the literature in several ways. This study is the first to use statistical learning methodology to analyze corporate disclosures. Because the empirical analyses in the paper are joint tests of the machine learning methodology and the economic hypotheses, the results in the paper show that the statistical learning algorithm, which is widely used in other research areas (e.g., Mitchell (2006)), can be successfully applied to corporate financial statement settings and thus could be useful for future research on disclosure.[2]

This paper is also among the first of several large-sample studies on forward-looking statements made by managers in corporate 10-Q and 10-K filings.[3] This study extends the literature on management disclosures of forward-looking information (Patell (1976), Penman (1980), Pownell, Wasley, and Waymire (1993), Skinner (1994), Dietrich, Kachelmeier, Kleinmuntz, and Linsmeier (1997), Miller and Piotroski (2000), Hutton, Miller, and Skinner (2003), and Lang and Lundholm (2003)).

Several prior small-sample studies based on content analysis by human coders also find that MD&As provide information content about future firm performance (Bryan (1997), Barron, Kile, and O'Keefe (1999), and Callahan and Smith (2004)). The innovations of this paper over these studies are severalfold. First, none of the prior studies examine the determinants of MD&A tones. By contrast, in this paper, I test hypotheses about economic factors that may explain MD&A tone variations. Second, none of the prior papers provide any empirical evidence on the informativeness of MD&As over time, even though the SEC has strengthened its requirements on MD&A disclosures significantly in the last ten years through releases and rules. This study provides evidence on the time-series change in the MD&A information content and thus sheds light on the effectiveness of the SEC regulation on MD&As. Third, accruals and MD&A are two important channels through which managers

[2] I also find that the empirical results based on the tone measures calculated using the dictionary approach do not lend strong support to the hypothesis that MD&As contain information content about future performance. This is perhaps because the dictionaries may not work well for the specific domain of corporate filings even though they may work well with other domains (e.g., news). Furthermore, the Bayesian algorithm is used at a sentence-level analysis in this paper, but it may not work well for analyzing passages.

[3] A contemporaneous study (Muslu, Radhakrishnan, Subramanyam, and Lim (2008)) also examines the information content of forward-looking statements in MD&As, but the focus of that paper is on the intensity of the forward-looking information, rather than the tone. Feldman, Govindaraj, Livnat, and Segal (2009) examine MD&A tone and post-earnings announcement drift using the General Inquirer.

communicate to investors. This paper contributes to the literature by being the first to examine whether information from the MD&As affects investors' processing of the accruals information.

The rest of the paper proceeds as follows. Section 2 discusses the nature of MD&A disclosures and hypotheses. Section 3 presents the details of the Naïve Bayesian learning algorithm and Section 4 discusses its empirical implementation, including the validation test results. Section 5 presents the empirical results and Section 6 concludes.

2 Literature review and research questions

2.1 MD&A and forward-looking statements

Item 303 of Regulation S-K ("Reg. S-K") presents the specific SEC rules for the MD&A, and many SEC releases provide more detailed instructions and interpretive guidance. Since 1968, firms have been required to discuss unusual (non-recurring) components of earnings in the MD&A (SEC (1968)). Later firms were also required to analyze certain trends associated with operations (SEC (1974)). Dissatisfied with the lack of informativeness in the disclosures firms were providing, the SEC granted protection under safe harbor rules in 1979, and issued a revised requirement the following year (SEC (1980)). According to the SEC, the MD&A requirements have three principal objectives: (1) to provide investors with a narrative explanation of the financial statements of a company; (2) to increase overall company disclosure and to supply the contextual basis for investors on which they can analyze financial information; and (3) to provide information about the quality and potential variability of the earnings and cash flows of a company (SEC (2003)).

The SEC also encourages direct forward-looking information in circumstances such as known material trends, events, commitments, and uncertainties. However, the requirements in the rules and releases related to such circumstances are not set out in concrete, objectified terms. Specifically, they leave the decision regarding prospective disclosure to the discretion of management with regard to three different aspects. Whether the disclosure of forward-looking information is mandatory depends cumulatively on management's assessment of (1) whether an above-mentioned circumstance is "presently known," (2) whether such a circumstance is "reasonably likely," and (3) whether "material effects" are to be expected. However, neither the term "reasonably likely" nor the term "material" are clearly defined. Therefore, with regard to providing direct forward-looking information in the MD&A, management's discretion is given considerable leeway.

In order to encourage companies to provide forward-looking statements in their MD&As, any predictive information is explicitly covered by the safe-harbor rule for projections (Reg. S-K Item 303 (a) Instr. 7). However, even with the safe harbor provisions of the Reform Act, companies may avoid disseminating forward-looking information because of their uncertainty regarding judicial interpretation of the safe harbor provisions and because of fears regarding state court litigation where, plaintiffs will argue, no such safe harbor is available (Grundfest and Perino (1997)).

Note that it is not mandatory for companies to have their MD&A reports audited. At best, auditors have a professional responsibility, which is supported by the American Institute of Certified Public Accountants, to check a company's MD&A information for material inconsistencies against the respective financial statements of that company. A company may not be held liable under the federal securities laws for projections and other forward-looking statements if the forward-looking statement is accompanied by meaningful cautionary statements identifying important factors that could cause actual results to differ materially from those in the forward-looking statement, or the plaintiff fails to prove that the forward-looking statement was made with actual knowledge that the statement was false or misleading (Hüfner (2007)).

2.2 Literature review

There is extensive research on the economic implications of voluntary corporate disclosure in the accounting literature.[4] Conceptually, there are at least three characteristics of the disclosures that are interesting to researchers: the level ("how much you say"), the meaning or the tone ("what do you mean"), and the transparency ("how you say it").

Many of these studies can be categorized along two dimensions: dependent variables and independent variables (i.e., the disclosure measures), as illustrated in the table in Appendix A1. The majority of the prior studies fit into column (1) in the table. That is, they examine the level or amount of disclosure, i.e., which topics/subjects managers discuss in the disclosure ("how much you say"). With the exception of the Association for Investment Management and Research (AIMR) scores, most studies rely on manual coding of the disclosure level, perhaps due to the difficulty in creating a large-sample measure of disclosure quality. A few papers examine "how you say it" by examining the readability (Li (2008)) and vocal tone of disclosures (Mayew and Venkatachalam (2008)).

This paper fits into the second column in the table (i.e., studying "what you mean" in corporate disclosures). Most papers in this area carry out a content analysis with human coders and have small sample sizes. For example, Bryan (1997) studies 250 MD&As and Callahan and Smith (2004) examine 71 firms and 420 firm-years. While manual coding may be more precise, it has two disadvantages (Core (2001)): small sample size due to cost considerations and difficulty with replication due to subjectivity in the coding process. In particular, the small sample size may limit the scope of the empirical tests. For instance, to study any potential change in the information content of MD&A disclosures over time, researchers need a relatively large panel of data. One approach to solve the small sample problem is to rely on a dictionary-based content analysis to understand corporate disclosures.

[4] There is also a substantial literature on the determinants of corporate disclosures (e.g., Lang and Lundholm (1993) and Miller (2002)), the frequency of voluntary disclosures (e.g., Botosan and Harris (2000)), management forecasts (e.g., Waymire (1984)), and the consequences of mandatory disclosures (e.g., Leuz and Verrecchia (2000)), which are discussed in detail in Healy and Palepu (2001) and Core (2001).

The difference between the dictionary approach and the approach used in this paper is discussed in greater detail in Section 3.

In the disclosure literature, the empirical results regarding the information content of MD&As are mixed. On the one hand, Bryan (1997) finds that discussions of future operations and planned capital expenditures are associated with future short-term performance. Likewise, Callahan and Smith (2004) find that their disclosure index based on a content analysis provides incremental explanatory power in predicting future firm performance and market valuation while controlling for current income and other related factors. On the other hand, Pava and Epstein (1993) examine 25 randomly-selected companies to test the SEC's contention that MD&As were not reaching their intended goal. Their results show that, while most companies accurately describe historical events, very few provide useful and accurate forecasts (i.e., many companies make practically no predictions.). They also find a strong bias in favor of correctly projecting positive trends, while negative trends tend to be either ignored or not fully reported. A contemporaneous study by Muslu, Radhakrishnan, Subramanyam, and Lim (2008) examines the intensity of the forward-looking information in MD&As and finds that a more intense discussion of forward-looking information makes stock returns incorporate future earnings in a more timely fashion and reduce analysts' forecast errors. Feldman, Govindaraj, Livnat, and Segal (2009) measure MD&A tone using a customized dictionary and found that tone changes are significantly associated with stock price reactions and the post-earnings announcement drift.

The second broad stream of literature into which this paper fits is the research in accounting, finance, and other social science fields that analyzes the content of textual documents using computer algorithms. Within this literature, there is extensive research on the information content of corporate earnings releases (Davis, Piger, and Sedor (2005), Rogers, Buskirk, and Zechman (2009)), accounting policy disclosures (Levine and Smith (2006)), audit opinions (Butler, Leone, and Willenborg (2004)), financial news (Tetlock, Saar-Tsechansky, and Macskassy (2007), and Core, Guay, and Larcker (2008)), Internet stock message board (Antweiler and Frank (2004)), multiple sources of financial text (Kothari, Li, and Short

(2008)), presidential election campaigns (Pennebaker and Stone (2001)), and art history (Martindale (1990)). Appendix A2 lists a sample of textual analysis papers in accounting and finance and tabulates them according to their methodology. The majority of the papers use the dictionary-based approach, and hence have a firm-level measure based on the percentage of specific categories of words.

2.3 Hypotheses

2.3.1 Determinants of MD&A tone

I explore several factors to explain the cross-sectional variations in the tone of MD&A FLS:

Current firm performance. While there is substantial theoretical and empirical literature on the amount of disclosure, little theoretical work exists on the direction of this disclosure (i.e., tones). Some arguments on the relation between the level of disclosure and firm performance can apply to the MD&A tone as well. For example, litigation concern may encourage firms with good current performance to be more cautious in discussing future events in their MD&A. Momentum in firm performance also suggests that the MD&A FLS may be more positive for firms with good current performance. On the other hand, earnings are mean-reverting (i.e., there is a negative relation between current earnings innovation and future earnings change), which implies a more cautious and negative tone in forward-looking statements for firms with better current performance.

Accruals. It is well documented that accruals are negatively associated with future firm performance and investors seem to under-react to this information. One explanation of the accrual anomaly is that managers manipulate the accrual component of earnings (Sloan (1996)). If this is true, it implies that managers know the implications of the accruals for future performance, and thus a negative correlation between accruals and the FLS tone is expected. Alternatively, accruals may simply proxy for a firm's economic conditions (e.g., distress) and managers are likely to understand (at least partially) the implications of accruals for future earnings. In either case, a negative relation between accruals and the MD&A tone is expected. On the other hand, if managers are overconfident or fixated on current earnings, then they may not understand the implications of accruals for future performance and hence no relation between accruals and MD&A tone is expected.

Firm size. Size captures many aspects of a firm's operational and business environment. The accounting literature has used firm size as a proxy for a firm's political cost (Watts and Zimmerman (1986)). Holding current performance constant, bigger firms may have more cautious forward-looking statements because of the higher political and legal cost due to their visibility.

Market-to-book ratio. High market-to-book firms are different from low market-to-book firms in many aspects, including the investment opportunity set and growth potential. To the extent that growth firms face more uncertain future economic conditions, a negative relation between market-to-book ratio and MD&A tone is expected.

Volatility of operations. Firms with more volatile business environments may be more cautious in discussing future events because of information uncertainty with regard to future performance. Alternatively, firms with high performance variability are likely to have more severe information asymmetry between managers and investors. Finally, performance variability may be related to MD&A tone because of its effect on a firm's vulnerability to legal action. These factors all posit a negative relation between volatility and MD&A tone.

Reporting quarter. Prior research has shown that the accounting numbers behave differently across different reporting quarters (Das and Shroff (2002)). To examine the potential implications of reporting quarter on MD&A tone, I include three quarterly dummies ($Q2$, $Q3$, and $Q4$), which are dummies for the 2nd, 3rd, and 4th quarter, respectively.

2.3.2 Information content of MD&As

Next, I examine whether the forward-looking statements in corporate MD&A disclosures contain information about future profitability and liquidity, i.e., when managers are more positive in discussing the firm's future outlook, the future does turn out to be "brighter." Based on the arguments developed in the Introduction section, there are many factors that may lead to non-informative FLS disclosures. *Ex ante*, it is not clear whether, on average, the FLS disclosures contain economically meaningful information about future firm performance.[5]

I also examine whether there is a systematic change in the information content of MD&As over time with a particular focus on a pre- and post-2003 comparison. The SEC issued new guidelines on MD&A disclosures in 2003 and encouraged a significant increase in information content and a reduction in boilerplate language (SEC (2003)). The Sarbanes-Oxley Act also enhanced the MD&A disclosure requirements (Bainbridge (2007)). First, the MD&A section is chosen as the vehicle for more complete disclosure of off-balance-sheet transactions. Second, the Sarbanes-Oxley Act requires CEOs and CFOs to certify that their financial statements, including the MD&A section, fairly present the financial conditions and results of operations of the issuer. Consequently, a test of the change in the information content of MD&As over time will shed light on the effectiveness of the SEC regulations.

2.3.3 MD&A tone and the accrual anomaly

Prior literature has shown that accruals contain information content about future performance and can predict future stock returns (Sloan (1996)). Following Sloan (1996), a large literature has emerged examining the possible explanations for the accrual anomaly (Richardson, Tuna, and Wysocki (2009) provide a comprehensive review of this literature). To shed light on this issue, many prior papers empirically examine whether the accrual anomaly is a function of transaction costs and firm characteristics (such as growth). However, none of these papers link the accrual anomaly to the other disclosures made by managers.[6]

Both accruals and MD&A tone can be considered as signals from managers that contain information about future firm performance. While accruals are subject to the financial reporting rules, managers tend to have a higher degree of freedom with the MD&A disclosures

[5] Note that the empirical tests are a joint test of (1) the machine learning algorithm's ability to capture tone; (2) whether managers follow the SEC instructions and provide information about the future in their MD&A disclosures; and (3) whether managers, at least on average, are truthful in their disclosures.

[6] One exception is a recent study by Huddart and Louis (2009), who study the relation between insider trading and the pricing of "managed" accruals.

(Bainbridge (2007)). To the extent that the MD&A tone provides a more direct prediction about future outlook than are accruals, the information in the MD&A tone about future performance is more salient. Given the limited attention of investors, salient information could affect investors' interpretation of other information (Hirshleifer and Teoh (2003)). I therefore examine the accrual anomaly as a function of the MD&A tone. This sheds light on how investors price accruals conditional on the information communicated by managers in their 10-K and 10-Q filings.

Sloan (1996) suggests that investors fixate on reported earnings and, as a result, accruals are negatively related to future stock returns. If more positive (negative) management discussions about the future outlook accompany negative (positive) accruals, investors may be less likely to fixate on accruals to the extent that the discussions made by managers provide more salient information about future performance and help investors understand the implications of accruals better (Hirshleifer and Teoh (2003)). In this scenario, information in the MD&A helps mitigate the accrual anomaly and one should expect the accrual anomaly to be stronger for firms with managers who do not "warn" investors about the future performance implications of the accruals. Alternatively, investors may ignore the information in the MD&A tone because there is substantial boilerplate disclosures in MD&As and the information processing cost could be high. In this scenario, the degree of mis-pricing of the accruals would be unrelated to the MD&A tone. I conduct empirical tests to distinguish between the two scenarios to shed light on how investors respond to the information in accruals depending on other information signaled by managers.

3 Naïve Bayesian algorithm and text classification

There are two general approaches for conducting content analysis: a rule-based approach (i.e., dictionary approach) and a statistical approach. The first approach uses a "mapping" algorithm, in which a computer program reads the text and classifies the words (or phrases) in the text into different categories based on some pre-defined rules or categories (i.e., dictionary). For instance, the General Inquirer, published by Harvard psychologist Philip J. Stone, and the Linguistic Inquiry and Word Count (LIWC) software by University of Texas psychologist James W. Pennebaker, are often used in content analysis.

The second approach, which was pioneered by computer scientists and mathematicians, relies on statistical techniques to infer the content of those documents and classify documents based on statistical inference (e.g., Manning and Schütze (1999) and Mitchell (2006)). For instance, the algorithm may calculate the statistical correlation between the frequency of some keywords and the document type to draw inferences.

In this paper, I use the statistical approach, which offers several advantages. First, there is no readily available dictionary that is built for the setting of corporate filings—as a result, the dictionary-based approach may have low power for corporate filings. For instance, consider the sentence "In addition, the Company has experienced attrition of its Medicare and commercial business in 1998 and 1999 and expects additional attrition." According to the General Inquirer (http://www.webuse.umd.edu:9090/), the sentence has 2 or 10.53% positive words ("expect" and "experience") and no negative words, even though it is obvious that this sentence has a negative tone. Second, the simple dictionary-based approach does not take into consideration the context of a sentence. For instance, if a sentence is about revenue, then "increase" should be treated as a positive word; however, it is likely to be of a negative tone if the topic is "cost."

Third, the rule-based approach generally ignores any prior knowledge that researchers may have about the text. For example, if most of the sentences that appear in MD&A reports are neutral, then, unless there is strong evidence that a sentence is of negative tone, it might be more efficient to classify a random sentence as being of a neutral tone. This point is especially salient when the topic of interest is managerial disclosure, because managers have incentives to disclose strategically.[7] Finally, the statistical approach typically provides a natural way to validate classification efficiency using the training data. The training data are human coded and thus could be used to test the effectiveness of the algorithm.

[7] On the other hand, if the subject of the analysis is news articles, there may not be a strong prior.

In this paper, I rely on a specific type of statistical learning method—the Bayesian algorithm—to conduct content analysis. Under this method of document classification, a given sentence is first reduced to a list of words (*words*) with each word weighted in some fashion (e.g., by frequency in the sentence). The goal is to classify the sentence into a specific category *cat* from a set of all possible categories (*cats*). For instance, one may want to classify a sentence into a category from a set of four possible categories $cats = (positive, negative, neutral, uncertain)$. The Naïve Bayesian algorithm would choose the best category by solving the following problem:

$$cat^* = \underset{cat \in cats}{\operatorname{argmax}} \frac{P(words|cat)P(cat)}{P(words)}.$$

Since $P(words)$ does not change over the range of categories, it can be eliminated. The problem thus becomes:

$$cat^* = \underset{cat \in cats}{\operatorname{argmax}} P(words|cat)P(cat).$$

Finally, if $w_1, w_2, ..., w_n$ are the words in the document and their probability of showing up in a sentence is assumed to be independent, then this expression is equivalent to:

$$cat^* = \underset{cat \in cats}{\operatorname{argmax}} P(w_1|cat) * P(w_2|cat) * ... * P(w_n|cat) * P(cat),$$

which is the formula used in the document categorization algorithm of this paper.

The last step is the only non-rigorous one in the derivation, and this is the "naïve" part of the Naïve Bayesian technique. Specifically, it assumes that the probability of each word appearing in a document is unaffected by the presence or absence of each other word in the document. The independence assumption simplifies the computation and avoids the "curse of dimensionality" problem (Bellman (1961)). Independence is assumed even though it is not true: For example, the word "iodized" is far more likely to appear in a document that contains the word "salt" than it is to appear in a document that contains the word "subroutine"; likewise, in the financial statement setting, the words "adverse effect" are more likely to show up together with the word "material" in a sentence from a corporate annual report. However, empirical results from other fields suggest that making this assumption even if it is not true may have little effect on the results.[8]

4 Empirical implementation

4.1 Data preparation

To apply the Naïve Bayesian algorithm, I first download all the 10-Ks and 10-Qs filed between 1994 and 2007 from the SEC Edgar website and remove any HTML tag. I then extract the MD&A section of the filings.[9] Next, I split the MD&A text into sentences using the Lingua::EN::Sentence module in Perl, which takes acronyms into consideration in the splitting process.[10] The forward-looking statements from the 10-Qs and 10-Ks are then extracted and cleaned and the details of this process are presented in Appendix B.

To construct the training data, I manually classify 30,000 randomly-selected forward-looking statements along the tone and content dimensions. First, every sentence is classified into one of four tones: positive, neutral, negative, and uncertain nature.[11] The "uncertain" tone is added because prior research has shown that managers tend to convey negative information by using words like "risk" and "uncertainty" (Li (2007)). A typical uncertain tone sentence looks like the following: "Significant additional work will be required for

[8]See, for instance, http:///www.cs.washington.edu/homes/pedrod/mlj97.ps.gz.

[9]Based on a random check of 200 filings, the success rate of extracting MD&As is about 95% for 10-Qs and between 85% and 90% for 10-Ks. The details of the MD&A extraction process are available upon request.

[10]For instance, in the string "FASB No. 123," the dot should not be treated as a delimiter for a sentence.

[11]"Positive" has richer meaning than "optimistic"—while "optimistic" means "tending, or conforming, to the opinion that all events are ordered for the best," "positive" can have other meanings (e.g., "fully assured; confident; certain; sometimes overconfident; dogmatic; overbearing"). The commonly used computational linguistic dictionaries also have slightly different definitions for "optimism" versus "positive." For instance, Rogers, Buskirk, and Zechman (2009) calculate their "net optimism" measure as the sum of three positive components (praise, satisfaction, and inspiration) minus the sum of three negative components (blame, hardship, and denial) from Diction. On the other hand, the 2007 version of the LIWC dictionary classifies "positive emotion" words and "negative emotion" words (which include three sub-categories: "Anxiety," "Anger," and "Sadness") as part of the "Affective processes" and do not have an "Optimism" category. In this paper, I standardize the terms to "positive/negative," rather than "optimistic/pessimistic."

the scaling-up of each new product prior to commercialization, and this work may not be completed successfully." I also divide the content of the forward-looking statements from corporate filings into 12 categories, the details of which are shown in Appendix C. The details of the training data preparation are shown in Appendix D.

The descriptive statistics of the training data set are reported in Table 1. Of the 30,000 FLS that are manually classified, 19.59% are coded as being of positive tone, 39.97% neutral, 17.82% negative, and 22.55% uncertain. The distribution of the sentences across the four groups is not even—the largest group is the neutral tone group. Somewhat inconsistent with the findings in Pava and Epstein (1993), there is a fair amount of negative tone in the MD&As, with a percentage close to that of the positive tone group. Prior research shows that managers tend to express their negative views using words like "risk" and "uncertain" (Li (2007)). Therefore, the 22.55% sentences with uncertain tone are also possibly of negative tone. In the empirical analysis, I combine the uncertain tone group with the negative tone group.

Table 1 also shows the percentages of the training data for each content category. The sum of the percentages across the twelve categories is greater than 100% because one sentence can be assigned to multiple categories. For instance, consider the statement "The Company believes these changes will help move existing inventory and reduce its cash flow concerns, but it is unlikely that these changes alone will provide sufficient capital to fund ongoing operations"; this sentence is about both operations and liquidity and, therefore, is classified as both category 4 and category 5.

The first four categories relate to revenue, cost, profitability, and operations; combined, 62.81% of the discussions are about these issues. Category 5 (liquidity issues) is discussed in 11.57% of the sentences, Category 6 (investment issues) 10.79%, and Category 7 (financing issues) 16.45%. As these three categories are related to financing, liquidity, and capital resources, they are combined in later analyses. Other categories represent a very small share of the MD&A discussions—ranging from Category 10 (regulation issues, 4.05%) to Category 9 (employee relations, 1.41%). The fact that most of the FLS are about profitability, liquidity, and capital resources—as evidenced by the combined 101.62% percentages in categories 1 to 7—is consistent with the SEC's intention that the MD&A section provide investors with information primarily about capital resources and liquidity.

4.2 Computation

I use the Algorithm::NaiveBayes module in Perl to conduct the computation. I first convert the vector of words for each sentence after the stemming and stopwording processes into a hash variable in Perl. I then feed the hash variables from the 30,000 manually-coded sentences into the Bayesian classifier module in Perl and run the training process. After this step, the algorithm predicts the tone and category of all the FLS from 10-Ks and 10-Qs, which total to about 13 million sentences.

4.3 Validation of the algorithm

There are three common methods for evaluating the effectiveness of a text classification algorithm: training error, train and test, and N-fold cross-validation.[12] To evaluate the effectiveness of the Naïve Bayesian algorithm, I first validate the effectiveness of the algorithm by calculating the training errors—carrying out both the training and predicting processes

[12] In the case of training error calculation, the classifier is trained on a training data set and evaluated on the same data set. Although this method is obviously biased toward the training data, it can detect underfitting of the data. In the "train and test" method, the data are divided into two parts: training and testing. The split is usually 90% for training and 10% for testing, but sometimes 2/3 of the data is used for training and 1/3 for testing. This is an unbiased evaluation, and can detect underfitting as well as overfitting. The theory of cross-validation was developed by Geisser (1975). It is useful in guarding against testing hypotheses suggested by the data ("Type III error"), especially where further samples are hazardous, costly, or impossible. In the N-fold cross-validation test, the data is randomly partitioned into N equal parts. N experiments are performed, and in each experiment, one part is taken as the testing data, while the remaining $N-1$ parts are used for training. At the end, the results over the N experiments are averaged. This is unbiased testing that gives more statistical significance than the train-and-test method, but it is not applicable for examining the training data during classifier construction.

using all of the 30,000 manually coded sentences. Untabulated results show that training errors are, in general, less than 10%. This literally means that if the algorithm is used to predict the sentences in the training data set (after learning from the same data set), it will correctly classify the categories and tone more than 90% of the time. This suggests that the chances of underfitting the data are small.

Since the "train and test" method can be seen as a special case of the N-fold cross validation test, I next report the empirical validation of the Naïve Bayesian learning algorithm using the N-fold cross-validation method, with N varying from 3 to 50. For instance, when $N = 3$, I carry out the 3-fold cross-validation by randomly dividing the 30,000 sentences of training data into 3 equal parts with each part containing 10,000 sentences. Three tests are then performed with one part (10,000 sentences) used as the learning data to classify the other two parts (20,000 sentences). In the end, the average success rate of the three tests is reported as the 3-fold cross-validation result.

Table 2 reports the average correct classification rate for the different levels of fN. In row (1) of Table 2, the results show that when we conduct the 3-fold cross-validation test (i.e., $N = 3$) to predict the four-category tones (positive, neutral, negative, and uncertain), the learning algorithm classifies the testing sentences correctly 59.15% of the time. This is a higher rate than that obtained with an informed guessing approach, where we calculate the percentage of each of the four tones using the learning data and apply this probability naïvely to the sentences in the testing data. With this method, the correct classification rate is only 32.44%, only half that obtained by the Bayesian learning algorithm. If we combine the negative and uncertain tones into one category (classifying the sentences into 3 categories), the Bayesian algorithm success rate increases to 66.95% in the 3-fold cross-validation test, and the informed guessing rate becomes 40.47%. These results are stable when N increases from 3 to 50—the success classification rate of the four-category tone by the Bayesian algorithm is about 59% and that of the three-category tone (positive, neutral, negative/uncertain) is about 67%; both are much higher than the rate obtained from the informed guessing strategy (about 32% and 40%, respectively).

Classifying every FLS into a content category out of 12 possible categories as defined in Appendix C, the Bayesian algorithm has a success rate of about 63% in the N-fold cross-validation, while that of the informed guessing strategy is only about 15%. If we combine the 12 content categories into 3 groups—profit (categories 1 to 4), liquidity (categories 5 to 7), and other (categories 8 to 12), the Bayesian algorithm achieves a success rate of more than 82%, while the informed guessing rate is about 44%. Overall, the N-fold cross-validation tests show that the Bayesian learning algorithm achieves a good classification rate compared with am informed guessing strategy.[13]

5 Information content of MD&A FLS

5.1 Descriptive statistics for MD&A tone and content

I start with all the 10-K and 10-Q filings from the SEC Edgar website. To be included in the final data set for further analysis, a firm-quarter has to have the following data: (1) a Central Index Key that can be matched with the GVKEY from Compustat and PERMNO from CRSP; (2) quarterly earnings (item 69 in the Quarterly file) and cash flows from operations (item 108) from Compustat; (3) stock returns in CRSP; and (4) at least five sentences of forward-looking statements in the filing. The requirement of at least five sentences of FLS is arbitrary, and the purpose of this requirement is to make sure that the empirical measures derived from the filing are not due to random noise. Varying this requirement (e.g., requiring at least 10 FLS sentences per filing) does not alter any of the empirical results.

For every forward-looking sentence k, I define its tone as the following: $TONE_k = 1$ if the learning algorithm predicts the sentence to be positive (i.e., the predicted probability of the sentence being positive is higher than the probability of being any of the other three categories); $TONE_k = 0$ if the prediction is neutral; and $TONE_k = -1$ if the prediction is

[13] I also feed the 30,000 training sentences into two dictionaries (General Inquirer dictionary and the LIWC software). Unreported results show that the dictionaries yield a classification rate comparable to that of the informed guessing strategy in the N-fold cross-validation tests.

negative or uncertain. I combine the negative and uncertain tone categories, because there are a large number of uncertain statements that carry a negative implication.

For every firm i in quarter j, I then define the tone of a firm's MD&A forward-looking statements as the average tone of all the K forward-looking sentences in its 10-Q or 10-K filing for that quarter as predicted by the learning algorithm:

$$TONE_{ij} = \frac{1}{K} \sum_{k=1}^{K} TONE_{ij,k} \tag{1}$$

By construction, $TONE_{ij}$ is a variable that is between -1 and 1, with 1 being completely positive and -1 being a completely negative tone. The more positive $TONE_{ij}$ is, the more positive the tone of the forward-looking statements made by firm i in quarter j's 10-Q or 10-K filing.

Table 3A shows the descriptive statistics for MD&A tone and other variables for the final sample of 145,479 firm-quarters. On average, the forward-looking statements in MD&A disclosures are negative, as indicated by a mean (median) of $TONE$ of -0.23 (-0.21). The mean of $TONE$ is significantly different from 0 with a p-value of 0.000 in a t-test. Decomposing $TONE$ into three different components—$PROFIT\_TONE$ (the average tone of sentences in categories 1 to 4 as described in Appendix C), $LIQUIDITY\_TONE$ (the average tone of sentences in categories 5 to 7), and $OTHER\_TONE$ (the average tone of sentences in categories 8 to 12)—shows that the negative tone of the MD&A FLS is mainly because of the profitability-related sentences. The mean of $PROFIT\_TONE$ is -0.42, $LIQUIDITY\_TONE$ 0.16, and $OTHER\_TONE$ -0.26, all of which are statistically significantly different from 0.

Table 3A also presents the descriptive statistics for the content of the MD&A forward-looking statements. $PROFIT\_PCT$ is the percentage of content devoted to profitability and operations (categories 1 to 4) as predicted by the learning algorithm, and $LIQUIDITY\_PCT$ is the percentage devoted to liquidity and capital resources. The means (medians) of $PROFIT\_PCT$ and $LIQUIDITY\_PCT$ are 53.65% (55.32%) and 32.99% (29.63%), respectively.

In Table 3B, the Pearson correlations for tone and content are reported. There is a significant negative correlation between $TONE$ and $PROFIT\_PCT$ (-0.47), indicating that when managers devote more discussion to future profitability, the average tone is more negative. This is consistent with the negative mean of $PROFIT\_TONE$ documented in Table 3A. The positive correlation between $TONE$ and $LIQUIDITY\_PCT$ (0.52) shows that managers devote more discussion to liquidity and capital resource issues when the $TONE$ of the MD&A is more positive. Current earnings ($EARN$) are positively correlated with MD&A tone (with a Pearson correlation coefficient of 0.142), and firms with more positive accruals tend to have more negative MD&A statements, with the correlation between ACC and $TONE$ being -0.082.

Figure 1 plots the means for the MD&A tone for firms sorted into quintiles. In quarter 0, all firms are sorted into five quintiles based on $TONE$ for the 10-K or 10-Q filed for that quarter, with quintile 1 firms having the most negative tone and quintile 5 the most positive. The tones of these firms are then tracked for the next 48 quarters with the mean plotted for each quintile. From Figure 1, it can be seen that $TONE$ is mean-reverting—by quarter 10, the differences in tone between the quintiles are dramatically reduced.

5.2 Determinants of MD&A tone and content

Table 4 reports the OLS regression results for $TONE$ when it is regressed on its hypothesized determinants: $EARN$ (current earnings), RET (contemporaneous stock returns), ACC (accruals), $SIZE$ (the logarithm of market value of equity), MTB (market-to-book ratio), $RETVOL$ (return volatility), and three reporting quarter dummies, $Q2$, $Q3$, and $Q4$. Year and 2-digit SIC industry fixed effects are also included in the regression. Since there are likely to be cross-sectional correlations, the standard errors are clustered by year.[14]

The results indicate that the FLS tone is positively related to current performance (the

[14]In all the pooled cross-section regressions, standard errors are calculated based on clustering by year. The empirical results based on industry-level clustering or Fama-MacBeth approach are qualitatively similar and are available upon requests.

coefficient on $EARN$=0.119 with a t-statistic of 9.59 and that on RET is 0.03 with a t-statistic of 4.74), confirming the univariate correlations in Table 3B. This means that when a firm is performing well in a given quarter, managers tend to discuss its future outlook in a more positive tone. Accruals are significantly negatively related to $TONE$ (the coefficient on ACC being -0.06 with t=-5.69)—suggesting that when current accruals are very positive, management's discussion of the firm's future outlook is more negative. Given that accruals are negatively related to future performance, this suggests that managers understand, at least to some extent, the implications of accruals for future performance.

Furthermore, bigger firms are more likely to use a negative tone in their MD&As, as indicated by the negative coefficient on $SIZE$ (-0.005 with a t-statistic of -3.48). This is consistent with the hypothesis that large firms are more cautious in their disclosures due to political and legal concerns. Additionally, firms with high market-to-book ratio have a less positive tone in their MD&As, consistent with the hypothesis that growth firms have more uncertain information environments and are more conservative in discussing future events. Finally, more volatile firms tend to have a less positive tone when discussing their future outlook (the coefficient on $RETVOL$ is -0.393 with a t-value of -9.04).[15] This suggests that firms with more volatile business environments may be more cautious in forward-looking disclosures because of either information uncertainty with regard to future performance or potential legal concerns. All of the three reporting quarter dummies ($Q2$, $Q3$, and $Q4$) show up insignificantly—both statistically and economically. In particular, the variable $Q4$ captures the difference in tone between the 4th quarter MD&A FLS discussions and 1st quarter discussions (coefficient on $Q4$ being 0.001 with a t-value of 0.15). This suggests that even though 10-Ks are in general much longer and more complicated than 10-Qs, their tone does not differ systematically from that in other quarters. The third quarter dummy $Q3$ has a coefficient of -0.005 (t-statistic=-1.02), which is small in economic magnitude. Overall, most of the economic factors are significant in explaining MD&A tone in the hypothesized direction.

[15] This is also consistent with the findings in Dichev and Tang (2007).

5.3 FLS tone and future earnings and liquidity

In this section, I examine the implications of the FLS tone generated by the Naïve Bayesian algorithm for a firm's future performance beyond that contained in the numeric financial information. To do so, I first check the link between $TONE$ and future profitability.

Table 5A shows the regression results for earnings in the next four quarters scaled by the book value of assets at the end of the current quarter ($EARN(t+1)$, $EARN(t+2)$, $EARN(t+3)$, and $EARN(t+4)$) when they are regressed on $TONE$ and control variables, including all the variables examined as the determinants of $TONE$ as well as year and industry-fixed effects. The residuals from the four equations are likely to be correlated with each other. In unreported Seemingly Unrelated Regressions, when these cross-equation correlations are taken care of, the coefficients become even more significant statistically.[16]

In column (1), the coefficient on $TONE$ is 0.010 with a t-statistic of 5.99, suggesting that when managers are more positive in discussing a firm's future outlook in the MD&A, the earnings in the next quarter are significantly higher. The economic magnitude of this effect is quite substantial—the next quarter's earnings scaled by the book value of assets of firms with extremely positive FLS (i.e., $TONE = 1$) is higher than that of firms with extremely negative FLS (i.e., $TONE = -1$) by two percentage points (0.010×2). This translates into an annual difference in ROA of eight percentage points—a substantial difference for a return-on-asset metric.[17] Examining this effect from a more realistic angle, we see that the inter-quartile range of $TONE$ validates this finding. From Table 3A, the 25th percentile and 75th percentile of $TONE$ are -0.41 and -0.03 respectively—thus, an inter-quartile change in $TONE$ implies a difference in annual ROA of 1.5 percentage points, even after controlling for the level of current stock returns and accruals. The same positive relation exists for $TONE$ when it is used to predict earnings in the next four quarters. In columns (2) to (4), the

[16]The coefficients remain the same because all four equations contain the same explanatory variables. The statistical significance in the SUR setting, however, is much bigger. This is likely due to the fact that, in an SUR setting, it is not easy to implement the usual adjustment for within-year correlation in error terms and to control for heteroskedasticity.

[17]Technically, this is not ROA because the numerator is net income, rather than unleveraged income.

coefficients on $TONE$ are 0.009 (t=4.93), 0.009 (t=6.79), and 0.005 (t=3.54), respectively, showing that $TONE$ has predictive power for future earnings for at least four quarters after the current quarter, although this effect becomes smaller economically and statistically over time. Not surprisingly, the coefficients on current earnings and stock returns both positively predict earnings in the next four quarters. Current quarter accruals, on the other hand, have a significantly negative relation with future performance, consistent with prior findings (Sloan (1996)).

In Table 5B, the dependent variable is the change (rather than level) in earnings in the next four quarters. For instance, in column (1), the dependent variable is $DEARN(t+1)$, which is calculated as earnings in the next quarter minus this quarter's earnings scaled by the book value of assets at the end of this quarter. The implications of $TONE$ for the change in future earnings are quite substantial both economically and statistically. For example, the coefficient on $TONE$ in column (1) of Table 5B is 0.010 (t=6.23), the same as that in Table 5A.

Next, I examine whether $TONE$ is systematically related to future liquidity. There are many financial ratios that measure firm liquidity. In this analysis, I focus on the cash flow ratio, defined as operating cash flows divided by current liabilities. Other common ratios include the current ratio and the interest coverage ratio. However, each of these has disadvantages for our analysis. The disadvantage of using the current ratio is that, as a balance sheet measure, it tends to be quite "sticky" and thus is not the ideal measure of what "happens" in a future quarter. The disadvantages of the interest coverage ratio are that fewer firms on Compustat report interest expense and that a small interest expense may lead to a scaling problem. Empirical tests using these two alternative measures of liquidity are almost identical to those using the cash flow measure and are not reported here.

Table 6 presents the empirical results. In column (1), the dependent variable is $CFRATIO(t+1)$, the operating cash flows in the next quarter divided by the current liabilities at the end of the next quarter. The coefficient on $TONE$ is 0.280 with a t-statistic of 10.37. The economic magnitude of the coefficient is substantial—the inter-quartile range of the cash flow ratio is about 0.50 and increasing $TONE$ from the 25th to the 75th percentile implies an increase in the next quarter's cash flow ratio of 0.11, after controlling for other economic factors that may affect future liquidity. In columns (2) to (4) of Table 6, where the cash flow ratios from the next three quarters are the dependent variables, there is still a significantly positive association between $TONE$ and future liquidity, although the size of the coefficient on $TONE$ is decreasing when the forecasting horizon increases.

Note that the control variables in predicting the future cash flow ratio do not include the current cash flow ratio, because current earnings and current accruals are already included in the regression as explanatory variables and including a cash flow ratio would make it difficult to interpret the coefficients. In unreported results, I include the current cash flow ratio instead of accruals as a control variable, and the results remain essentially the same. Of all the control variables, current earnings and stock returns predict future liquidity positively, while current accruals are negatively associated with future liquidity. Larger firms also tend to have better cash flow ratios, and firms with more volatile businesses have lower cash flow ratios.

It is interesting to test whether positive and negative tones have asymmetric implications for future performance. To examine this issue, in untabulated results, $TONE$ is interacted with a dummy variable $PTONE$, which is set to one if $TONE \geq 0$ and zero otherwise. The results indicate that there is no non-linear relation between $TONE$ and future earnings, as the coefficient on $TONE \times PTONE$ is insignificant. This suggests that positive and negative MD&As have about the same implications for a firm's future earnings.

5.4 FLS tone based on the dictionary approach and future earnings

In this section, I examine the relation between future earnings and FLS tone calculated using Diction 6.0, the General Inquirer, and the LIWC dictionary. The goal is to test the hypothesis that MD&As have information content about future performance using the tone measures based on these dictionaries.

I calculate the percentages of positive and negative words using each dictionary. Table 7A shows the summary statistics for the MD&A FLS tone based on the dictionaries. Following Davis, Piger, and Sedor (2005) and Rogers, Buskirk, and Zechman (2009), I use the Diction 6.0 software to estimate the percentage of positive words for each MD&A as the sum of the percentages of praise, satisfaction, and inspiration words: $DICTION\_POS = praise + satisfaction + inspiration$ and the percentage of negative words as the sum of the percentages of blame, hardship, and denial words: $DICTION\_NEG = blame + hardship + denial$.[18] The positive and negative tone based on the General Inquirer are $GI\_POSITIV$ and $GI\_NEGATIV$ respectively, which are the percentages of positive and negative words in each MD&A as classified by the GI. Similarly, $LIWC\_POSEMO$ and $LIWC\_NEGEMO$ are the percentages of positive emotion and negative emotion words respectively based on the LIWC package. I also calculate a summary measure of the MD&A tone based on each dictionary by taking the difference between the positive tone and negative tone percentages: $DICTION\_TONE = DICTION\_POS - DICTION\_NEG$, $GI\_TONE = GI\_POSITIV - GI\_NEGATIV$, and $LIWC\_TONE = LIWC\_POSEMO - LIWC\_NEGEMO$.

Table 7A shows that Diction classified 5.64% of MD&A forward-looking statements as positive and 10.05% negative. In contrast, both GI and LIWC report more positive words (5.81% and 2.00% respectively) than negative words (2.87% and 0.79% respectively) for MD&A forward-looking statements.[19] As a result, the mean of $DICTION\_TONE$ is negative (-4.41) and those of GI_TONE and $LIWC\_TONE$ are positive (2.94 and 1.21 respectively).

Table 7B presents the Pearson correlations between $TONE$ and the dictionary-based tone measures. Not surprisingly, the percentages of positive and negative words classified by the three dictionaries are positively correlated. The Pearson correlation coefficient between $DICTION\_POS$ and $GI\_POSITIV$ is 0.246, that between $DICTION\_POS$ and $LIWC\_POSEMO$ 0.240, and that between $GI\_POSITIV$ and $LIWC\_POSEMO$

[18]The results are based on the Diction 6.0 "corporate financial reports" norm.

[19]The percentages of positive and negative words based on the GI are comparable to those reported in Kothari, Li, and Short (2008).

0.483; The Pearson correlation coefficient between $DICTION\_NEG$ and $GI\_NEGATIV$ is 0.339, that between $DICTION\_NEG$ and $LIWC\_NEGEMO$ 0.440, and that between $GI\_NEGATIV$ and $LIWC\_NEGEMO$ 0.627. Of the three dictionaries, GI and LIWC appear to be closer to each other as the correlations of the tone measures based on them are higher than those with the measures from Diction.

The correlation between GI_TONE and $LIWC\_TONE$ is 0.565, higher than the correlation coefficient between $DICTION\_TONE$ and GI_TONE (0.290) and that between $DICTION\_TONE$ and $LIWC\_TONE$ (0.324). There are also positive correlations between $TONE$ and the MD&A tone measures based on the dictionary approaches. The Pearson correlations between $TONE$ and $DICTION\_TONE$ (GI_TONE, $LIWC\_TONE$) is 0.247 (0.299, 0.254). This showss that the machine learning algorithm and the dictionaries capture MD&A tone similarly, but the correlation between the dictionary-approach measures is much higher.

Table 7C examines the information content of the different dictionary-based measures of MD&A tone for future performance by regressing next quarter's earnings on the tone measures and control variables. Overall, the percentage of negative words is informative about future performance as indicated by the negative coefficients on $DICTION\_NEG$ (-0.000 with a t-statistic of -2.03), $GI\_NEGATIV$ (-0.001 with a t-statistic of -2.57), and $LIWC\_NEGEMO$ (-0.001 with a t-statistic of -1.15). However, the percentages of positive words from the dictionaries ($DICTION\_POS$, $GI\_POSITIV$, and $LIWC\_NEGEMO$) are also negatively and significantly associated with future earnings.

As a result, when the percentages of positive and negative words are combined into a summary measure of tone, the dictionary-based measures do not lend support to the hypothesis that corporate MD&A forward-looking statements are informative about future performance. In column (4), GI_TONE and other control variables are used to explain $EARN(t+1)$ and the coefficient on GI_TONE is -0.001 (t=-2.57), indicating that a more positive MD&A tone based on the General Inquirer dictionary is associated with lower future earnings. Similarly, in column (6), $LIWC\_TONE$ is negatively correlated with future earnings (coefficient -0.001 with a t-value of -4.49).[20] The only positive coefficient comes from $DICTION\_TONE$ (0.000 with a t-statistic of 1.06 in column (2)). This is likely due to the fact that Diction has a "corporate financial reporting" norm.

Overall, the evidence based on the MD&A tone measured using the dictionary approach does not support the hypothesis that there is information content about future performance in MD&A disclosures.

5.5 Information content of MD&A over time

To assess whether the information content of MD&As changes over time, Table 8 shows the regression of future earnings and liquidity on $TONE$ and its interaction with a time dummy, $POST2003$, which is equal to one if the report is filed in or after 2003 and zero otherwise. This test is designed to capture any systematic change in the information content of MD&As after the new SEC guidelines related to MD&As and the passage of the Sarbanes-Oxley Act, which significantly enhanced MD&A disclosure. A significant positive (negative) interaction term indicates that MD&As have become more (less) informative over time.

In this test, six of the eight interaction terms of $POST2003$ with $TONE$ are negative and one of these is statistically significant. The other two interaction terms are positive but insignificant. This result suggests that, despite the continuous effort by the SEC to strengthen the disclosure requirements for MD&As, there is no systematic change in their information content over time. In unreported regressions, replacing $POST2003$ with the calendar year leads to similar findings. Overall, the empirical results show that there is no significant increase in the informativeness of MD&As after 2003.

5.6 Aggregate tone

In this section, I investigate the aggregate tone for all U.S. public filers and examine whether there are any systematic changes in tone over time. Specifically, I calculate the tone of each

[20] If the summary tone measure is constructed as the logarithm difference of the percentages of positive and negative words based on GI or LIWC, the statistical significance of the measure is much weaker.

MD&A in the 10-Ks and 10-Qs filed by firms with Compustat and CRSP coverage for each month and then plot the average tone over time.

Figure 2 plots the equal-weighted aggregate tones of MD&As of U.S. public firms over time. This figure shows substantial variation in the aggregate tone, with three vertical lines indicating three possible dates of interest. The first vertical line indicates March 2000, when the NASDAQ index peaked at 4572.83. Here, the figure shows that, contrary to the dramatic increase in the index in the late 1990s, the average tone in MD&As actually became more negative. The second vertical line in Figure 2 shows that after the September 11, 2001 terrorist attack, there is a downward trend in management tone, although it is not clear whether this reflects a revision in management expectation or is simply a continuation of a previous trend. The third vertical line shows that the passage of the Sarbanes-Oxley Act does not seem to have an immediate impact on management's forward-looking tone.[21]

Several possible reasons may explain the puzzling finding that the tone in MD&As dropped significantly in the 1990s when the stock market was going up. First, because of anxiety over Y2000 issues, managers may have devoted more discussion to these issues in an uncertain tone, leading to a more negative aggregate tone. Second, managers might have been anticipating the equity market downturn and thus have become more cautious in communicating to stock investors. One possible explanation is the "market expectation hypothesis"—when investors have high expectations for a firm's future performance, managers tend to become more cautious in their MD&A discussions. The cross-sectional test results in Table 4 show that, consistent with this hypothesis, firms with a high market-to-book ratio tend to have more negative MD&As.

This hypothesis can also be tested in a time-series setting at the macro level. The hypothesis is that, as the equity market index increases (which is the case between 1994 and 1999), managers face more pressure from the market and begin to show more caution in their MD&As and hence a significant drop in the aggregate MD&A tone. To test this

[21]Unreported results show the aggregate MD&A tone using the GI and LIWC dictionaries. The overall trend remains similar to that based on the machine learning algorithm.

hypothesis, I divide my sample into five quintiles based on their market-to-book ratios and then plot the time-series trend of the MD&A tone for each portfolio. Consistent with the hypothesis, Figure 3 shows that firms in the top MTB quintile experienced the biggest drop in MD&A tone over this period. During the mid-1990s, the aggregate tone of this portfolio is around 0, but by the end of the 1990s, it drops to about -0.4. In contrast, as shown in Figure 4, the bottom MTB quintile portfolio firms experienced a drop of about 0.2 (from 0 in the mid-1990s to slightly less than -0.2 at the end of the 1990s). Unreported results show that the differences in the drop in MD&A tone between portfolios with different market-to-book ratios are statistically significant. Overall, the aggregate evidence suggests that market expectation is negatively correlated with manager's tone in the MD&A section. A more careful examination of these explanations would be an interesting topic for future research.

5.7 MD&A tone and the accrual anomaly

This section investigates the implications of MD&A tone for the accrual anomaly. To examine whether MD&A disclosures mitigate the mis-pricing of accruals, I examine the Fama-French three-factor excess monthly returns (in percentage) of the decile portfolios formed on accruals, where the excess returns are the intercepts from the time-series regressions of excess portfolio returns (raw returns minus the risk-free returns) on the market excess returns and the Fama-French size and book-to-market mimicking portfolio returns.

In Table 9, I first replicate the accrual anomaly using the full sample. The results show that the monthly hedge portfolio return (buying firms with the lowest decile accruals and shorting the highest decile firms) is 0.70% (t=1.77). The sample is then divided into two sub-samples based on the value of $WARN$, which is a dummy variable that equals 1 if, based on the most recent data, a firm (1) has accruals below the median of all firms and an MD&A tone above the median, or (2) has above-median accruals and below-median MD&A tone, and 0 otherwise.[22] The negative correlation between accruals and MD&A tone (Table

[22] This variable is constructed using the most recent data for a given firm and is recalculated every month.

4) suggests that, on average, managers "warn" about the future performance implications of accruals. Hence, the sub-sample with $WARN = 1$ has more observations (1910 firms) than the sub-sample with $WARN = 0$ (1670 firms).

There are two ways to sort firms into accruals decile portfolios for the two sub-samples. First, firms can be sorted into the portfolios *within* each sub-sample. In this case, the average number of observations is the same across portfolios (i.e., 167 firms per portfolio for the $WARN = 0$ sub-sample and 191 for the $WARN = 1$ sub-sample). The results show that under this sorting approach, accruals are negatively associated with future returns only when managers do not provide discussions in MD&A that are consistent with the implications of accruals (i.e., the accrual anomaly holds only for the sub-sample of firms with $WARN = 0$). When managers send signals about future performance in MD&As (i.e., $WARN = 1$), there is no abnormal returns from the accrual anomaly with the hedge portfolio return has a mean of -0.12% per month (t-statistic=-0.19). In contrast, when $WARN = 0$, there are strong abnormal returns from the accrual strategy: Buying firms in the bottom accruals decile and shorting those in the top decile generate a monthly return of 1.38% (t=3.58).

The second approach to sort firms based on accruals is to do full-sample sorting, i.e., firms are sorted into accruals decile portfolios first and then are divided into the two sub-samples. This sorting procedure leads to different numbers of observations across portfolios. The results are qualitatively similar: When $WARN = 0$ ($WARN = 1$), the hedge portfolio return is 0.98% (0.03%) with a t-value of 1.71 (0.05). Interestingly, as accruals become more positive, the number of firms with managers that "warn" about future performance increases. For instance, the bottom (top) accruals decile portfolio where managers issue such a warning in the MD&A disclosures has 160 (254) firms. This suggests that when firms report extreme positive accruals, managers are more likely to discuss future event in negative tone.[23] Overall, the evidence suggests that information from MD&As helps mitigate the mis-pricing of accruals.

[23]One possible explanation for this finding is that, when firms report positive accruals, managers are more likely to discuss future outlook because of litigation concerns.

5.8 Additional analysis

To examine whether conditioning the tone of MD&A on the content (i.e., profitability versus liquidity) makes a difference in forecasting future performance, I include profitability-related tone ($PROFIT\_TONE$), liquidity-related tone ($LIQUIDITY\_TONE$), and other sentence tone ($OTHER\_TONE$) separately in the statistical analysis. Table 10 reports the regression of $EARN(t+1)$ and $CFRATIO(t+1)$ on the current quarter's tone and control variables (coefficients unreported). In column (1), the coefficient on $PROFIT\_TONE$ is 0.003 (t=4.93), $LIQUIDITY\_TONE$ 0.013 (t=8.49), and $OTHER\_TONE$ -0.001 (t=-1.67). This suggests that both profitability-related FLS and liquidity-related FLS have implications for future earnings, while other types of FLS do not contain much information. Surprisingly, in predicting future earnings, the effect of liquidity-related FLS is much bigger than that of profitability-related FLS. In column (2), where the dependent variable is $CFRATIO(t+1)$, similar effects are observed: Liquidity-related FLS have more predictive power in forecasting future liquidity situations than profitability-related FLS.

Next, I compare the information content of 10-K MD&As with that of 10-Qs by separating the sample into 10-Qs and 10-Ks. Untabulated results show that, while both have substantial information content about future earnings and liquidity, 10-Q MD&As tend to have more information content than those from 10-K filings. One explanation for this finding is that in annual reports, management may be more likely to discuss longer-term events. Therefore, the implications for short-term earnings and liquidity are not as strong.

One potential concern with this study is that the research assistants may code a sentence as "uncertain" if they are uncertain about whether it is positive or negative, thus introducing noise into the training data. However, an unreported empirical analysis excluding the "uncertain" tone in the training data yields similar results.[24] This is consistent with the argument that the sentences in the "uncertain" category coded by the RAs are neither random noise nor a reflection of RA uncertainty about the categorization of the sentences,

[24] Dropping the "uncertain" category reduces the training data from 30,000 sentences to about 23,000 sentences.

since they appear to improve the information content of the tone measure generated by the Bayesian learning algorithm.

Another potential concern of the empirical results is that different companies have a different number of forward-looking statements in their MD&As. The different number of forward-looking sentences can lead to different levels of precision in the MD&A tone estimates. For instance, a firm that has five forward-looking statements may have a high variance in the MD&A tone measure. To mitigate any potential concerns about this issue, I re-run the main tests by weighting the observations in the regressions using the number of forward-looking sentences in the MD&A. The intuition is that lower weighting for observations with fewer sentences in the MD&A section may mitigate any concerns about their influences on the empirical results. Unreported results indicate that greater weighting for MD&As with more sentences leads to qualitatively similar results and does not change the inferences of the paper.

I also examine the implications of within-firm variations of tone for future profitability by including firm fixed effects (instead of industry fixed effects) in the regressions. Unreported results indicate that the coefficients on $TONE$ are slightly smaller compared with those in the main analysis in Table 5A. Nonetheless, MD&A tone is still positively and significantly associated with future earnings, even after controlling for firm and year fixed effects and other control variables, and the economic magnitudes are still substantial.

Finally, unreported results indicate that sell-side financial analysts' consensus earnings forecasts are higher when MD&A FLS are more positive. However, they do not appear to fully utilize the information in the MD&A tone. Even after controlling for the latest analyst forecasts before the announcement of future earnings, the tone of the MD&A forward-looking statements still has significant predictive power for future earnings and liquidity.

6 Conclusions

This paper examines the implications of the forward-looking statements from the MD&A section of corporate 10-Q and 10-K filings for future performance. I use a Naïve Bayesian machine learning algorithm to categorize the tone and content of forward-looking statements from more than 140,000 corporate 10-Q and 10-K filings between 1994 and 2007. I find that the tone of the forward-looking statements is a function of current performance, accruals, firm size, market-to-book ratio, and return volatility. The tone of the forward-looking statements is positively correlated with future performance and has explanatory power incremental to other variables, but the informativeness of the MD&As has not changed systematically over time despite continuous efforts from the SEC to strengthen MD&A disclosures. Furthermore, when managers warn about the future performance implications of accruals in MD&As, accruals are less likely to be mis-priced by investors.

Appendix A1: Sample papers on the implications of corporate disclosures

	Independent variable (Disclosure measures)		
Dependent variable	(1) "How much you say" (e.g., Level / Amount)	(2) "What you mean" (e.g., Tone)	(3) "How you say it" (e.g., Transparency / Truthfulness)
Cost of capital	Botosan (1997) Botosan and Plumlee (2002)*	Kothari, Li, and Short (2008)	
Future earnings		Bryan (1997) Miller and Piotroski (2000) Callahan and Smith (2004) Davis, Piger, and Sedor (2005)	Li (2008) Mayew and Venkatachalam (2008)
Analyst behavior	Lang and Lundholm (1996)* Barron, Kile, and O'Keefe (1999)		
Other		Levine and Smith (2006) Feldman, Govindaraj, Livnat, and Segal (2009)	

*Note: The AIMR score used by these papers to measure disclosure quality, which mainly covers detail disclosure (i.e., "level") as ranked by analysts, also partly covers the "candor" of the disclosures and thus can also be regarded as a measure of disclosure transparency.

Appendix A2: Sample textual analysis papers

Paper	Text analyzed	Name of method	Method	Firm-level measure
This paper	Forward-looking MD&A	Naïve Bayes	Classify sentences as positive/negative/neutral/uncertain	Average of sentence tones
Davis, Piger, and Sedor (2005)	Press release	DICTION	Classify words as optimistic/pessimistic	% of words
Kothari, Li, and Short (2008)	MD&A, analyst reports etc.	General Inquirer	Classify words as optimistic/pessimistic	% of words
Tetlock, Saar-Tsechansky, and Macskassy (2007)	News articles	General Inquirer	Classify words as optimistic/pessimistic	% of words
Li (2008)	10-Ks and different sections	Fog / word count and LIWC	Classify at word level and averaged across sentences	% of words and average length of sentence
Henry (2008)	Earnings release	Customized dictionary	Classify words as positive/negative	% of words
Matsumoto, Pronk, and Roelofsen (2008)	Conference calls	Customized dictionary and LIWC	Classify words into forward-looking	% of words
Feldman, Govindaraj, Livnat, and Segal (2009)	MD&A	Customized dictionary	Classify words as positive/negative	% of words
Rogers, Buskirk, and Zechman (2009)	Earnings announcements	DICTION	Classify words as optimistic/pessimistic	% of words

Appendix B: Data preparation

I define forward-looking statements as all those sentences that contain: "will," "should," "can," "could," "may," "might," "expect," "anticipate," "believe," "plan," "hope," "intend," "seek," "project," "forecast," "objective," or "goal." I do not include the word "shall" in the searching process because it is usually associated with legal language and boilerplate disclosures. For instance, I randomly select 5% (or 601 filings) of all the 10-Ks filed in 1998 by all public issuers in the U.S. with a file size greater than 10K bytes. These 601 filings contain 68,878 sentences with the word "shall" in it. On the other hand, these filings have 62,783 sentences that contain one of the "forward-looking" words listed above, which suggests that including "shall" in the search process will more than double the number of forward-looking statement search results. I then randomly checked 100 of the 68,878 sentences with the word "shall" and conclude that almost all of them can be classified as boilerplate disclosures and are not of interest to this paper. Hence, I "shall" not include "shall" in the search for forward-looking statements.

I also exclude any sentence that contains the word "undersigned," sentences that consist of all capital letters, and sentences containing words such as "herein," "hereinafter," "hereof," "hereon," "hereto," "theretofore," "therein," "thereof," and "thereon," because they are almost certain to be legal boilerplate. Furthermore, I exclude all sentences that contain "expected," "anticipated," "forecasted," "projected," or "believed" when such words follow "was," "were," "had," and "had been." Situations like these typically indicate a sentence that is not *forward-looking* in nature. The search for forward-looking statements has both type I and type II errors. I expect the type II errors to be small in this case; given the long list of words that are used in the search process, sentences that are not flagged as forward-looking are unlikely to be forward-looking. On the other hand, it is possible that some sentences are flagged as forward-looking that may not be.

I carry out the content and tone analysis at the sentence level. The disadvantage of sentence-level (rather than paragraph-level or document-level) analysis is that, in some instances, the tone and content of a sentence depend on its context. That is, not knowing the context makes it impossible to determine the tone of the sentence "The increase is primarily the result of lower-than-anticipated sales volume." This is because we do not know what (e.g., cost or other items) increases because of the low sales volume. However, there are several advantages to a sentence-level analysis. First, it can significantly reduce the amount of labor in coding the text. Second, different sentences in a paragraph or article can have different tones and contents and bundling them together introduces noise. Sentence-level analysis can therefore increase the power of the classification.

Finally, before doing the Naïve Bayesian classification, I implement "stemming" and "stopwording" processes to further clean up the text. Stemming is a process for reducing inflected (or sometimes derived) words to their base or root form (e.g., "dependent" to "depend") to increase the power of textual analysis. I use the Lingua::Stem::En module from Perl, which implements Porter's stemming algorithm, invented by Martin Porter at Cambridge University and first described in Porter (1980). "Stopwords" are a class of words that are typically the short, frequently occurring words in a language. Typical stopwords usually have only a grammatical function within a sentence, and do not add to the meaning. Stopwords include articles, case particles, conjunctions, pronouns, auxiliary verbs, and common prepositions. Some examples of stopwords for English are: "the," "and," "it," "is," and "of." To conduct a stopwords cleanup, I use the Lingua::EN::Stopwords modules from Perl. The stop-words list used in the Lingua::EN::StopWords has a list of 213 words (refer to "http://wiki.christophchamp.com/index.php/Perl/Modules/Lingua" for a complete list). I modify this list by deleting the following words from the stop words list: "cannot," "no," "none," "nor," and "not" (i.e., include them in the statistical analysis). The reason for this modification is that one of the goals of this paper is to categorize the tones of statements, and words like "cannot" can impact the tone. Empirically, however, including or excluding these words in the analysis does not affect any of the results in this paper.[25]

[25]This is likely due to the learning nature of the analysis. For example, when managers talk about "material effect," most of the time they will include "no" or "not" in the same sentence. Consequently, even if "no" and "not" are included in the stop-words list, the algorithm can still capture the positive tone by conditioning on the words "material effect."

Appendix C: Classification categories

- Category 1: Sales / revenues / market condition / market position / consumer demand / competition / pricing / new contract
- Category 2: Cost / expense / reserves for contingent liability / asset impairment / goodwill impairment
- Category 3: Profit / income / performance results / margin
- Category 4: Operations / productions / general business
- Category 5: Liquidity: interest coverage / cash balance / working capital conditions
- Category 6: Investment - general capital expenditure; M&A / divestiture / discontinued operation
- Category 7: Financing - debt / equity / dividend / repurchase
- Category 8: Litigation / lawsuit (material impact or not)
- Category 9: Employee relations / retention / hiring / union relations
- Category 10: Regulations (e.g., environment laws) / income tax / government relation
- Category 11: Accounting method / accounting estimation assumptions / auditing / internal control
- Category 12: Other: Boilerplate / legal statement / standard statement

Appendix D: Training data preparation

The training data construction is done with the help of 15 research assistants; most of the research assistants are students of the Master of Accounting program, BBA program or MBA program at the University of Michigan Ross School of Business. To make sure that the training dataset is of high quality, I impose the following search criteria for research assistants: (1) Native English speakers are given higher priority; (2) If a person is not a native English speaker, then she has to rank within the top 10% on the TOEFL test or GMAT/GRE verbal test; and (3) the student should have received an A- or above (or equivalent) in an intermediate financial accounting course or holds an American, British, or Canadian CPA certificate.

The manual classification process is challenging because many forward-looking statements cannot be easily categorized into a content or tone category. A particular challenge lies in the unobserved expectation of the financial statement readers. Suppose a manager discloses that the adverse effect of an environment liability will not be material for the firm. This can be positive news if the reader has been expecting a material effect or neutral if the immateriality is fully anticipated. As another example, an increase in capital expenditure may signal a negative tone for shareholders if the company is already over-investing, but it could also be seen as positive if the investment is part of a turnaround plan. Without a more specific context, it is almost impossible to figure out these subtle differences in the training data construction process. I ask the research assistants to keep a neutral prior—i.e., assuming that the reader has no information about the topic—in making classification judgments. In the environment liability example above, the sentence would be classified as positive because I assume that the reader has no prior information about the potential environmental liability. In later empirical analysis, the content and tone of the FLS are examined conditional on other observable variables such as contemporaneous stock returns. To the extent that any expectations of the readers/investors about the subject are reflected in these observable variables and, therefore, are controlled, my "uninformative prior" research design in the training process should be effective.

References

Antweiler, Werner, and Murray Z. Frank, 2004, Is all that talk just noise? The information content of internet stock message boards, *Journal of Finance* pp. 1259–1294.

Bainbridge, Stephen M., 2007, *A Complete Guide to Sarbanes-Oxley: Understanding How Sarbanes-Oxley Affects Your Business* (Adams Media: Cincinnati, OH).

Barron, Orie E., Charles O. Kile, and Terrence B. O'Keefe, 1999, MD&A quality as measured by the SEC and analysts' earning forecasts, *Contemporary Accounting Research* 16, 75–109.

Bellman, Richard E., 1961, *Adaptive Control Processes* (Princeton University Press: Princeton, NJ).

Botosan, Christine A., 1997, Disclosure level and the cost of equity capital, *The Accounting Review* 72, 323–349.

———, and Mary S. Harris, 2000, Motivations for a change in disclosure frequency and its consequences: An examination of voluntary quarterly segment disclosures, *Journal of Accounting Research* 38, 329–353.

Botosan, Christine A., and Marlene A. Plumlee, 2002, A re-examination of disclosure level and the expected cost of equity capital, *Journal of Accounting Research* 40, 21–40.

Bryan, Stephen H., 1997, Incremental information content of required disclosures contained in management discussion and analysis, *The Accounting Review* 72, 285–301.

Butler, Marty, Andrew J. Leone, and Michael Willenborg, 2004, An empirical analysis of auditor reporting and its association with abnormal accruals, *Journal of Accounting and Economics* 37, 139–165.

Callahan, Carolyn M., and Rodney E. Smith, 2004, Firm performance and management's discussion and analysis disclosures: An industry approach, Working paper, University of Arkansas-Fayetteville.

Core, John E., 2001, A review of the empirical disclosure literature: Discussion, *Journal of Accounting and Economics* 31, 441–456.

———, Wayne Guay, and David Larcker, 2008, The power of the pen and executive compensation, *Journal of Financial Economics* 88, 1–25.

Das, Somnath, and Pervin K. Shroff, 2002, Fourth quarter reversals in earnings changes and earnings management, Working paper, University of Minnesota.

Davis, Angela K., Jeremy M. Piger, and Lisa M. Sedor, 2005, Beyond the numbers: An analysis of optimistic and pessimistic language in earnings press releases, Working paper, Washington University.

Dichev, Ilia, and Wei Tang, 2007, The link between earnings volatility and earnings predictability, Working paper, University of Michigan.

Dietrich, J. Richard, Steven J. Kachelmeier, Don N. Kleinmuntz, and Thomas J. Linsmeier, 1997, An experimental investigation of forward-looking non-financial performance disclosures, Working paper, University of Illinois at Urbana-Champaign and University of Texas at Austin.

Feldman, Ronen, Suresh Govindaraj, Joshua Livnat, and Benjamin Segal, 2009, Management's tone change, post earnings announcement drift and accruals, Working paper, New York University.

Geisser, Seymour, 1975, The predictive sample reuse method with applications, *Journal of The American Statistical Association* 70, 320–328.

Grundfest, Joseph A., and Michael A. Perino, 1997, Ten things we know and ten things we don't know about the Private Securities Litigation Reform Act of 1995: Testimony Before the Subcommittee on Securities of the Committee on Banking, Housing, and Urban Affairs of the United States Senate, http://securities.stanford.edu/research/articles/19970723sen1.html.

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

Henry, Elaine, 2008, Are investors influenced by how earnings press releases are written?, Working paper, University of Miami.

Hirshleifer, D., and S. Teoh, 2003, Limited attention, financial reporting and disclosure, *Journal of Accounting and Economics* pp. 337–386.

Huddart, Steve, and Henock Louis, 2009, Insider selling, earnings management, and the 1990s stock market bubble, Working paper, Pennsylvania State University.

Hüfner, Bernd, 2007, The SEC's MD&A: Does it meet the informational demands of investors? A conceptual evaluation, *Schmalenbach Business Review* 59, 58–84.

Hutton, Amy, Gregory Miller, and Douglas Skinner, 2003, The role of supplementary statements with management earnings forecasts, *Journal of Accounting Research* 41, 867–890.

Knutson, Peter H., 1993, *Financial Reporting in the 1990's and Beyond* (NY: Association for Investment Management and Research: Charlottesville, VA).

Kothari, S.P., Xu Li, and James E. Short, 2008, The effect of disclosures by management, analysts, and financial press on the equity cost of capital: A study using content analysis, *Working paper* MIT.

Lang, Mark, and Russell Lundholm, 1993, Cross-sectional determinants of analyst ratings of corporate disclosures, *Journal of Accounting Research* Autumn, 246–271.

———, 1996, Corporate disclosure policy and analyst behavior, *Accounting Review* 71, 467–492.

———, 2003, Voluntary disclosure and equity offerings: Reducing information asymmetry or hyping the stock?, *Contemporary Accounting Research* 17, 623–662.

Leuz, Christian, and Robert E. Verrecchia, 2000, The economic consequences of increased disclosure, *Journal of Accounting Research* 38, 91–124.

Levine, Carolyn B., and Michel J. Smith, 2006, Critical accounting policy disclosures, Working paper, Carnegie Mellon University and Boston University.

Li, Feng, 2007, Do stock market investors understand the risk sentiment of corporate annual reports?, University of Michigan Working Paper.

———, 2008, Annual report readability, current earnings, and earnings persistence, *Journal of Accounting and Economics.*

Manning, Chris, and Hinrich Schütze, 1999, *Foundations of Statistical Natural Language Processing* (The MIT Press: Cambridge, MA).

Martindale, Colin, 1990, *The Clockwork Muse: The Predictability of Artistic Change* (Basic Books: New York).

Matsumoto, Dawn, Maarten Pronk, and Erik Roelofsen, 2008, What makes conference calls useful? The role of analyst participation in conference calls, Working paper, University of Washington.

Mayew, William J., and Mohan Venkatachalam, 2008, The power of voice: Managerial affective states and future firm performance, Working paper, Duke University.

Miller, Gregory S., 2002, Earnings performance and discretionary disclosure, *Journal of Accounting Research* 40, 173–204.

———, and Joseph D. Piotroski, 2000, Forward-looking earnings statements: Determinants and market response, Working paper, Harvard University and University of Chicago.

Mitchell, Tom M., 2006, The discipline of machine learning, Working paper, Carnegie Mellon University CMU-ML-06-108.

Muslu, Volkan, Suresh Radhakrishnan, K. R. Subramanyam, and Dongkuk Lim, 2008, Causes and consequences of forward looking disclosures in the management discussion and analysis (MD&A) sections: An exploratory study, Working paper, University of Texas at Dallas and University of Southern California.

Patell, J., 1976, Corporate forecasts of earnings per share and stock price behavior: Empirical tests, *Journal of Accounting Research* pp. 246–276.

Pava, Moses L., and Marc J. Epstein, 1993, How good is MD&A as an investment tool?, *Journal of Accountancy* 175, 51–53.

Penman, Stephen, 1980, An empirical investigation of the voluntary disclosure of corporate earnings forecasts, *Journal of Accounting Research* pp. 132–160.

Pennebaker, James W., and Lori D. Stone, 2001, Contribution to a CATA workshop presented at the 51st annual meeting of the International Communication Association, .

Porter, Martin F., 1980, An algorithm for suffix stripping, program, *Program* 14, 130–137.

Pownell, Grace, Charles Wasley, and Gregory Waymire, 1993, The stock price effect of alternative types of management earnings forecasts, *Accounting Review* 68, 869–912.

Richardson, Scott, Item Tuna, and Peter Wysocki, 2009, Accounting anomalies and fundamental analysis: A review of recent research advances, Working paper, London Business School.

Rogers, Jonathan L., Andrew Van Buskirk, and Sarah L.C. Zechman, 2009, Disclosure tone and shareholder litigation, Working paper, University of Chicago.

Rogers, Rodney K., and Julia Grant, 1997, Content analysis of information cited in reports of sell-side financial analysts, *Journal of Financial Statement Analysis* 3, 17–30.

SEC, 1968, Securities Act Release No. 4936 (December 9), .

———, 1974, Securities Act Release No. 5520 (August 14), .

———, 1980, Securities Act Release No. 6231 (September 2), .

———, 2003, Commission Guidance Regarding Management's Discussion and Analysis of Financial Condition and Results of Operations; Interpretation; Rule, .

Skinner, Douglas J., 1994, Why firms voluntarily disclose bad news, *Journal of Accounting Research* 32, 38–60.

Sloan, Richard, 1996, Do stock prices fully reflect information in accruals and cash flows about future earnings?, *Accounting Review* 71, 289–315.

Tavcar, Lawrence R., 1998, Make the MD&A more readable, *CPA Journal Online* http://nysscpa.org/cpajournal/1998/0198/newsViews/0198NV7.htm (1.12.2005).

Tetlock, Paul, M. Saar-Tsechansky, and S. Macskassy, 2007, More than words: Quantifying language to measure firm's fundamentals, *Journal of Finance*.

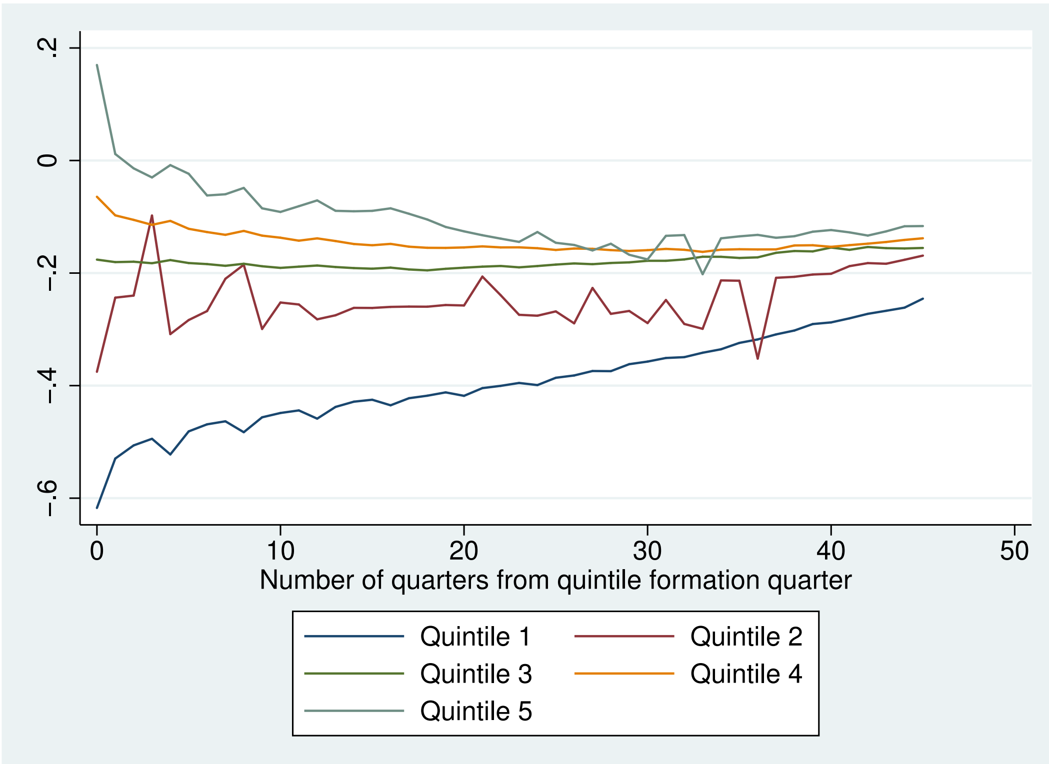
Verrecchia, Robert E., 1983, Discretionary disclosure, *Journal of Accounting and Economics* 5, 179–194.

Watts, Ross L., and Jerry L. Zimmerman, 1986, *Positive Accounting Theory* (Prentice-Hall: Englewood Cliffs, NJ).

Waymire, Gregory, 1984, Additional evidence on the information content of management earnings forecasts, *Journal of Accounting Research* 22, 703–719.

FIGURE 1

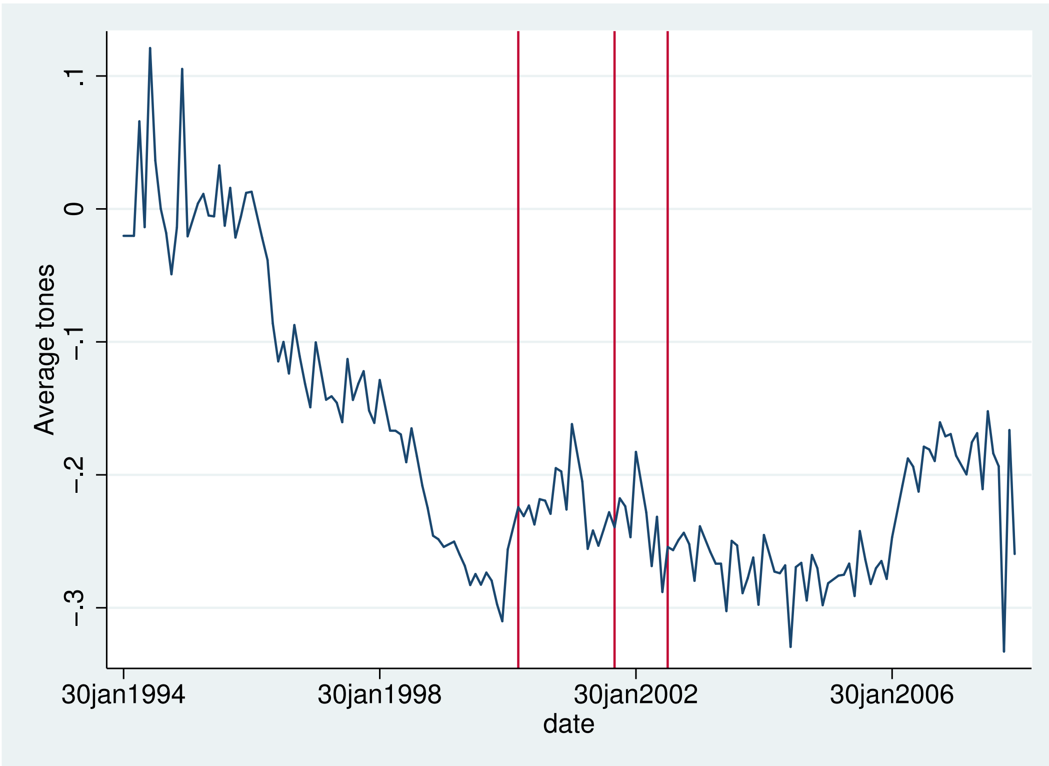
MD&A Tone Over Time by Quintile



This figure shows the average tone of the 10-K and 10-Q MD&As for the five quintile portfolios sorted on $TONE$ in the 48 quarters after the portfolios are formed. Every quarter, five portfolios are formed by sorting $TONE$, the average tone of all the forward-looking statements in a firm's MD&A (a forward-looking sentence's tone has a value of 1 if the Bayesian learning algorithm classifies the sentence as positive, 0 if neutral, and -1 if negative or uncertain.) In quarter 0, firms are sorted into 5 quintiles (quintile 1 to quintile 5, with firms in quintile 1 having the most negative MD&A tone and quintile 5 firms having the most positive tone) and the average tone of the five quintile firms is plotted over time for the next 48 quarters.

FIGURE 2

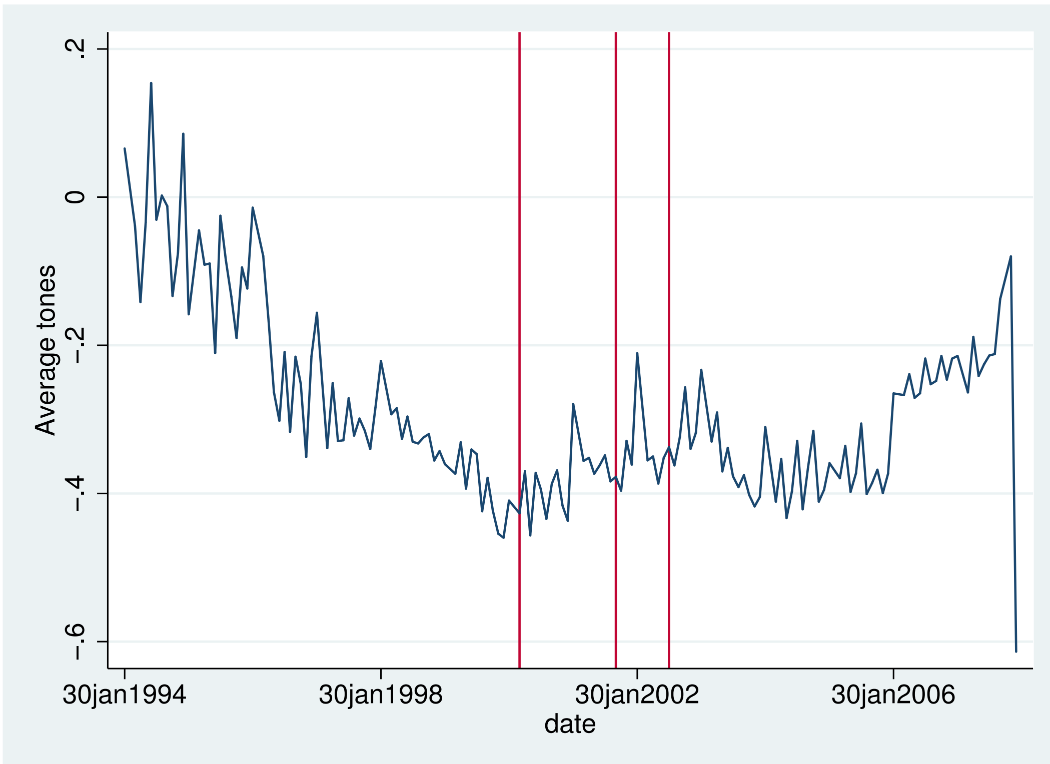
Aggregate Tone of MD&As Over Time (All Firms)



This figure shows the average $TONE$ of all MD&As from the 10-Ks and 10-Qs filed in a given month between January 1994 and December 2007. $TONE$ is the average tone of all the forward-looking statements in a firm's MD&A (a forward-looking sentence's tone has a value of 1 if the Bayesian learning algorithm classifies the sentence as positive, 0 if neutral, and -1 if negative or uncertain. The first vertical line indicates the month of March 2000 (when the NASDAQ index peaked); the second indicates the month of September 2001 (the terrorist attack event); the third indicates the month of July 2002 (passage of the Sarbanes-Oxley Act).

FIGURE 3

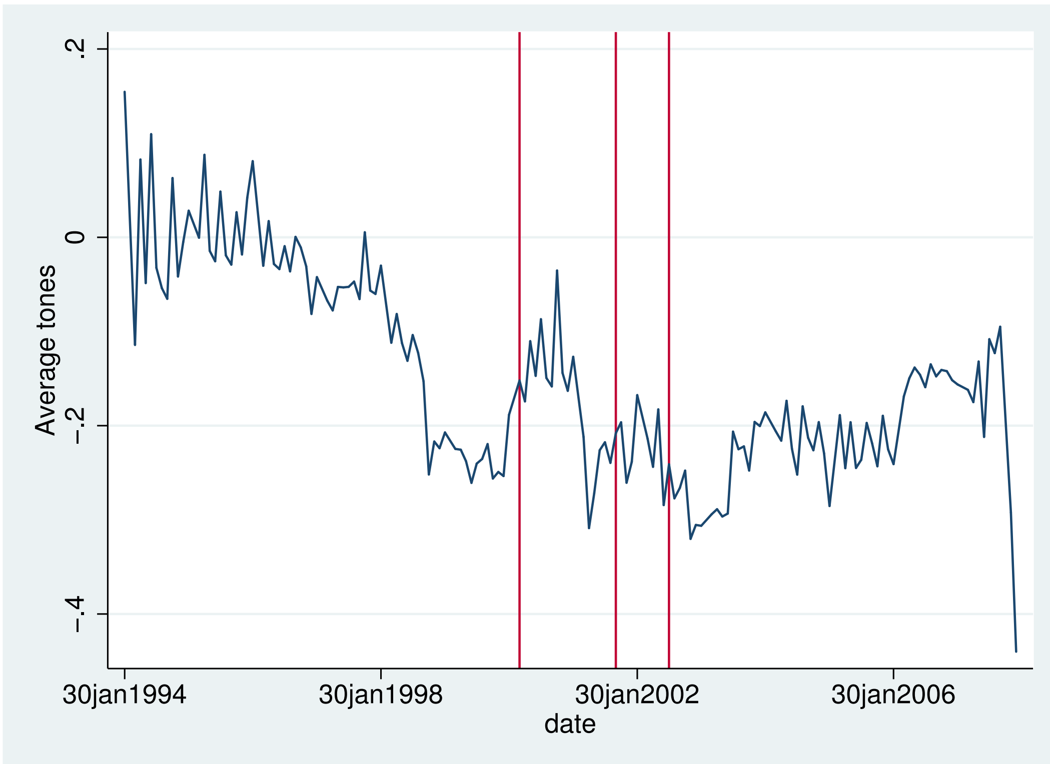
Aggregate Tone of MD&As Over Time (High MTB Firms)



This figure shows the average $TONE$ of the MD&As from the 10-Ks and 10-Qs filed by high market-to-book ratio firms in a given month between January 1994 and December 2007. High market-to-book firms are those firms in the top quintile market-to-book (MTB) portfolios for a month. $TONE$ is the average tone of all the forward-looking statements in a firm's MD&A (a forward-looking sentence's tone has a value of 1 if the Bayesian learning algorithm classifies the sentence as positive, 0 if neutral, and -1 if negative or uncertain. The first vertical line indicates the month of March 2000 (when the NASDAQ index peaked); the second indicates the month of September 2001 (the terrorist attack event); the third indicates the month of July 2002 (passage of the Sarbanes-Oxley Act).

FIGURE 4

Aggregate Tone of MD&As Over Time (Low MTB Firms)



This figure shows the average $TONE$ of the MD&As from the 10-Ks and 10-Qs filed by low market-to-book ratio firms in a given month between January 1994 and December 2007. Low market-to-book firms are those firms in the bottom quintile market-to-book (MTB) portfolios for a month. $TONE$ is the average tone of all the forward-looking statements in a firm's MD&A (a forward-looking sentence's tone has a value of 1 if the Bayesian learning algorithm classifies the sentence as positive, 0 if neutral, and -1 if negative or uncertain. The first vertical line indicates the month of March 2000 (when the NASDAQ index peaked); the second indicates the month of September 2001 (the terrorist attack event); the third indicates the month of July 2002 (passage of the Sarbanes-Oxley Act).

TABLE 1

Percentage Distributions of MD&A FLS Tone and Content

Positive tone	19.59	Cetegory 1: Revenues	15.06
Neutral tone	39.97	Category 2: Cost	10.45
Negative tone	17.82	Category 3: Profits	8.72
Uncertain tone	22.55	Category 4: Operations	28.58
		Sum of 1-4	*62.81*
		Category 5: Liquidity	11.57
		Category 6: Investing	10.79
		Category 7: Financing	16.45
		Sum of 5-7	*38.81*
		Category 8: Litigation	2.14
		Category 9: Employees	1.41
		Category 10: Regulation	4.05
		Category 11: Accounting	2.78
		Category 12: Other	3.32
		Sum of 8-12	*13.70*

This table shows the percentage distributions of the 30,000 sentences (i.e., the training data) that are manually coded into different tone and content categories. The 30,000 sentences are extracted randomly from the MD&As that are classified as forward-looking statements. Fifteen research assistants manually categorize them into two dimensions: tone and content. The details of the procedures are presented in Appendix C and Appendix D.

TABLE 2

N-fold Cross-validation Tests

Bayesian learning				
	Tone		Content	
N	4 categories	3 categories	12 categories	3 categories
3	59.15	66.95	62.52	82.31
5	59.30	67.00	62.76	82.37
10	59.31	67.02	62.91	82.42
25	59.27	66.99	62.88	82.40
50	59.37	67.11	63.02	82.46

Informed guessing				
	Tone		Content	
N	4 categories	3 categories	12 categories	3 categories
3	32.44	40.47	15.21	44.64
5	32.05	40.17	15.54	44.50
10	32.25	40.19	15.47	44.51
25	32.22	40.26	15.92	44.37
50	31.77	39.80	15.32	44.25

This table reports the N-fold cross-validation test results for the machine learning algorithm and for an "informed guessing" strategy.

In the "Bayesian learning" section of the table, for each N, the 30,000 training sentences that are manually coded (details described in Table 1 note and Appendices C and D) are randomly divided into N equal parts. N experiments are then carried out, with $N-1$ parts used as learning data to classify the remaining data using the Bayesian learning algorithm. The average percentage of correct classifications is reported for each N. For the tone, the algorithm classifies each sentence into four possible categories (positive, neutral, negative, and uncertain) and the N-fold tests for the four categories are reported under "Tone—4 categories" and "Tone—3 categories," with the negative and uncertain categories combined. For the content, the algorithm classifies each sentence into 12 possible categories (details in Section 4.4) and the N-fold tests for the 12 categories are reported under "Content—12 categories" and "Content—3 categories," with Categories 1 to 4 combined as a "profitability" category; Categories 5 to 7 combined as a "liquidity" category, and Categories 8 to 12 combined as an "other" category.

In the "Informed guessing" section of the table, for each N-fold cross-validation tests, classification success rates are calculated using an "informed guessing" strategy. In this strategy, the probability for each tone/content category is calculated using the learning data (i.e., $(N-1)/N$ of the 30,000 sentences) and the sentences in the predicting set (i.e., $1/N$ of the 30,000 sentences) are classified following these probabilities.

TABLE 3A

Descriptive Statistics

N	Mean	Pr(=0)	P5	P25	Median	P75	P95	STDEV
TONE	-0.23	0.000	-0.75	-0.41	-0.21	-0.03	0.23	0.29
PROFIT_TONE	-0.42	0.000	-0.94	-0.67	-0.44	-0.21	0.14	0.35
LIQUIDITY_TONE	0.16	0.000	-0.38	-0.03	0.13	0.33	1.00	0.35
OTHER_TONE	-0.26	0.000	-1.00	-0.50	-0.13	0.00	0.20	0.40
PROFIT_PCT (%)	53.65	-	18.18	40.00	55.32	68.09	84.00	19.93
LIQUIDITY_PCT (%)	32.99	-	8.33	18.37	29.63	44.44	69.23	19.03
OTHER_PCT (%)	13.35	-	0.00	4.17	12.00	20.00	33.33	11.47
EARN	-0.02	-	-0.17	-0.02	0.01	0.02	0.05	0.13
RET	.04	-	-0.44	-0.14	0.01	0.16	0.59	0.38
CFRATIO	0.00	-	-1.76	-0.12	0.11	0.39	1.15	1.16
ACC	-0.02	-	-0.17	-0.07	-0.02	0.01	0.16	0.16
SIZE	5.54	-	2.41	4.10	5.46	6.86	9.02	2.01
MTB	2.17	-	0.80	1.08	1.46	2.30	5.79	2.47
RETVOL	0.16	-	0.05	0.09	0.13	0.20	0.37	0.12

	P-value
Test of PROFIT_TONE=LIQUIDITY_TONE	0.000
Test of PROFIT_TONE=OTHER_TONE	0.000
Test of LIQUIDITY_TONE=OTHER_TONE	0.000

This table shows the descriptive statistics for 145,479 sample firm-quarters. The column labeled "Pr(=0)" is the P-value of the t-test indicating whether a variable equals 0. In the second part of Panel A, the column "P-value" shows the P-value of the t-test testing whether the mean of two variables are the same. The variables are defined as follows. *TONE* is the average tone of the forward-looking statements of a firm-quarter. A forward-looking sentence's tone has a value of 1 if the learning algorithm classifies the sentence as positive, 0 if neutral, and -1 if negative or uncertain. *PROFIT_TONE* is the average tone of the forward-looking statements of a firm-quarter that are about profits or operations (i.e., the statements that are classified as categories 1 to 4 as defined in Section 4.4). *LIQUIDITY_TONE* is the average tone of the forward-looking statements of a firm-quarter that are about liquidity or capital resources (i.e., the statements that are classified as categories 5 to 7 as defined in Section 4.4). *OTHER_TONE* is the average tone of the forward-looking statements of a firm-quarter that are about other topics (i.e., the statements that are classified as categories 8 to 12 as defined in Section 4.4). *PROFIT_PCT* is the percentage of the forward-looking statements of a firm-quarter that are about profits or operations (i.e., the statements that are classified as categories 1 to 4 as defined in Section 4.4). *LIQUIDITY_PCT* is the percentage of the forward-looking statements of a firm-quarter that are about liquidity or capital resources (i.e., the statements that are classified as categories 5 to 7 as defined in Section 4.4). *OTHER_PCT* is the percentage of the forward-looking statements of a firm-quarter that are about other topics (i.e., the statements that are classified as categories 8 to 12 as defined in Section 4.4). *EARN* is the quarterly earnings (Compustat Quarterly file data item 69) scaled by the book value of assets (Compustat Quarterly file data item 44), winsorized at -3 and 3. *RET* is the contemporaneous stock returns in the fiscal quarter calculated using CRSP monthly return data. *CFRATIO* is the quarterly cash flows from operations (Compustat data item 108) scaled by book value of current liability

(Compustat Quarterly file data item 49). ACC is the accruals (earnings subtract cash flow from operations) scaled by the book value of assets(Compustat Quarterly file data item 44). $SIZE$ is the logarithm of the market value of equity at the end of the quarter (Compustat Quarterly file data item 14 times item 61). MTB is the market value of equity (Compustat Quarterly file data item 14 times item 61) plus the book value of total liabilities (Compustat Quarterly file data item 54) scaled by the book value of total assets(Compustat Quarterly file data item 44). $RETVOL$ is the stock return volatility calculated using twelve months of monthly return data before the fiscal quarter ending date.

TABLE 3B

Correlation Matrix (P-values in Parentheses)

Variables	TONE	PROFIT_TONE	LIQUIDITY_TONE	OTHER_TONE	PROFIT_PCT	LIQUIDITY_PCT	EARN	RET	ACC	SIZE	MTB	RETVOL
TONE	1.000											
PROFIT_TONE	0.779	1.000										
	(0.000)											
LIQUIDITY_TONE	0.541	0.279	1.000									
	(0.000)	(0.000)										
OTHER_TONE	0.486	0.262	0.231	1.000								
	(0.000)	(0.000)	(0.000)									
PROFIT_PCT	-0.467	-0.206	0.028	-0.140	1.000							
	(0.000)	(0.000)	(0.000)	(0.000)								
LIQUIDITY_PCT	0.520	0.230	0.008	0.215	-0.828	1.000						
	(0.000)	(0.000)	(0.003)	(0.000)	(0.000)							
EARN	0.142	0.103	0.159	0.065	-0.028	0.017	1.000					
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)						
RET	-0.005	-0.008	-0.001	-0.000	-0.013	-0.002	0.077	1.000				
	(0.043)	(0.003)	(0.841)	(0.873)	(0.000)	(0.495)	(0.000)					
ACC	-0.082	-0.066	-0.078	-0.045	0.020	0.001	0.118	0.004	1.000			
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.604)	(0.000)	(0.126)				
SIZE	0.040	0.062	-0.015	-0.013	-0.042	-0.026	0.213	0.103	-0.107	1.000		
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)			
MTB	-0.169	-0.122	-0.114	-0.078	0.097	-0.087	-0.200	0.208	0.150	0.152	1.000	
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		
RETVOL	-0.263	-0.204	-0.176	-0.091	0.154	-0.110	-0.252	0.164	0.110	-0.323	0.198	1.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	

This table shows the pair-wise Pearson correlation coefficients of the variables with the P-values testing whether the correlation coefficients are significantly different from 0 in the parentheses. The variables are defined as follows. $TONE$ is the average tone of the forward-looking statements of a firm-quarter. A forward-looking sentence's tone has a value of 1 if the learning algorithm classifies the sentence as positive, 0 if neutral, and -1 if negative or uncertain. $PROFIT\_TONE$ is the average tone of the forward-looking statements of a firm-quarter that are about profits or operations (i.e., the statements that are classified as categories 1 to 4 as defined in Section 4.4). $LIQUIDITY\_TONE$ is the average tone of the forward-looking statements of a firm-quarter that are about liquidity or capital resources (i.e., the statements that are classified as categories 5 to 7 as defined in Section 4.4). $OTHER\_TONE$ is the average tone of the forward-looking statements of a firm-quarter that are about other topics (i.e., the statements that are classified as categories 8 to 12 as defined in Section 4.4). $PROFIT\_PCT$ is the percentage of the forward-looking statements of a firm-quarter that are about profits or operations (i.e., the statements that are classified as categories 1 to 4 as defined in Section 4.4). $LIQUIDITY\_PCT$ is the percentage of the forward-looking statements of a firm-quarter that are about liquidity or capital resources (i.e., the statements that are classified as categories 5 to 7 as defined in Section 4.4). $OTHER\_PCT$ is the percentage of the forward-looking statements of a firm-quarter that are about other topics (i.e., the statements that are classified as categories 8 to 12 as defined in Section 4.4). $EARN$ is the quarterly earnings (Compustat Quarterly file data item 69) scaled by the book value of assets (Compustat Quarterly file data item 44), winsorized at -3 and 3. RET is the contemporaneous stock returns in the fiscal quarter calculated using CRSP monthly return data. $CFRATIO$ is the quarterly cash flows from operations (Compustat data item 108) scaled by book value of current liability (Compustat Quarterly file data item 49). ACC is the accruals (earnings subtract cash flow from operations) scaled by the book value of assets(Compustat Quarterly file data item 44). $SIZE$ is the logarithm of the market value of equity at the end of the quarter (Compustat Quarterly file data item 14 times item 61). MTB is the market value of equity (Compustat Quarterly file data item 14 times item 61) plus the book value of total liabilities (Compustat Quarterly file data item 54) scaled by the book value of total assets(Compustat Quarterly file data item 44). $RETVOL$ is the stock return volatility calculated using twelve months of monthly return data before the fiscal quarter ending date.

TABLE 4

Determinants of MD&A tone

COEFFICIENT	TONE
EARN	0.119***
	(9.59)
RET	0.034***
	(4.74)
ACC	-0.062***
	(-5.69)
SIZE	-0.005***
	(-3.48)
MTB	-0.007***
	(-10.61)
RETVOL	-0.393***
	(-9.04)
Q2	0.001
	(0.44)
Q3	-0.005
	(-1.02)
Q4	0.001
	(0.15)
Observations	145479
R^2	0.22

This table shows the regression results of MD&A tone on its potential determinants. The dependent variable is $TONE$, which is the average tone of the forward-looking statements of a firm-quarter. A forward-looking sentence's tone has a value of 1 if the Bayesian learning algorithm classifies the sentence as positive, 0 if neutral, and -1 if negative or uncertain. The independent variables include $EARN$, RET, ACC, $SIZE$, MTB, and $RETVOL$, all of which are defined in the note to Table 3. $Q2$ ($Q3$ or $Q4$) is a dummy variable which is set to 1 if the current reporting quarter is the second (third or fourth) fiscal quarter. Year and 2-digit SIC industry fixed effects are included in the regressions, but are not reported. T-statistics clustered by year are reported in parentheses. *** indicates $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

TABLE 5A

Future Earnings and MD&A Tone

COEFFICIENT	(1) EARN(t+1)	(2) EARN(t+2)	(3) EARN(t+3)	(4) EARN(t+4)
TONE	0.010***	0.009***	0.009***	0.005***
	(5.99)	(4.93)	(6.79)	(3.54)
EARN	0.388***	0.348***	0.356***	0.362***
	(16.32)	(15.82)	(14.73)	(12.66)
RET	0.011***	0.011***	0.006***	0.005***
	(4.64)	(7.09)	(4.40)	(3.11)
ACC	-0.185***	-0.179***	-0.165***	-0.188***
	(-14.04)	(-12.94)	(-13.24)	(-16.68)
SIZE	0.003***	0.003***	0.003***	0.003***
	(11.92)	(11.14)	(7.40)	(7.67)
MTB	-0.003***	-0.004***	-0.005***	-0.006***
	(-5.09)	(-4.86)	(-5.33)	(-5.86)
RETVOL	-0.062***	-0.067***	-0.060***	-0.053***
	(-8.53)	(-10.08)	(-6.62)	(-6.69)
Q2	-0.000	-0.009***	0.010***	0.001
	(-0.66)	(-6.16)	(4.35)	(1.02)
Q3	-0.009***	0.000	0.010***	-0.001
	(-5.44)	(0.12)	(4.54)	(-1.01)
Q4	0.001	0.001	0.009***	-0.009***
	(0.52)	(0.60)	(4.37)	(-5.65)
Observations	122919	118100	112116	109597
R^2	0.35	0.30	0.27	0.27

This table shows the regression results of future earnings on MD&A tone and other control variables. The dependent variables are the earnings in the next four quarters (Compustat Quarterly file data item 69) scaled by the book value of assets at the end of the current quarter (Compustat Quarterly file data item 44). Independent variable $TONE$ is the average tone of the forward-looking statements of a firm-quarter. A forward-looking sentence's tone has a value of 1 if the Bayesian learning algorithm classifies the sentence as positive, 0 if neutral, and -1 if negative or uncertain. Other independent variables include $EARN$, RET, ACC, $SIZE$, MTB, $RETVOL$, $Q2$, $Q3$, and $Q4$. $EARN$ is the quarterly earnings (Compustat Quarterly file data item 69) scaled by the book value of assets (Compustat Quarterly file data item 44), winsorized at -3 and 3. RET is the contemporaneous stock returns in the fiscal quarter calculated using CRSP monthly return data. ACC is the accruals (earnings subtract cash flow from operations) scaled by the book value of assets(Compustat Quarterly file data item 44). $SIZE$ is the logarithm of the market value of equity at the end of the quarter (Compustat Quarterly file data item 14 times item 61). MTB is the market value of equity (Compustat Quarterly file data item 14 times item 61) plus the book value of total liabilities (Compustat Quarterly file data item 54) scaled by the book value of total assets(Compustat Quarterly file data item

44). $RETVOL$ is the stock return volatility calculated using twelve months of monthly return data before the fiscal quarter ending date. $Q2$ ($Q3$ or $Q4$) is a dummy variable which is set to 1 if the current reporting quarter is the second (third or fourth) fiscal quarter. Year and 2-digit SIC industry fixed effects are included in the regressions, but are not reported. T-statistics clustered by year are reported in parentheses. *** indicates $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

TABLE 5B

Future Earnings Changes and MD&A tone

COEFFICIENT	(1) DEARN(t+1)	(2) DEARN(t+2)	(3) DEARN(t+3)	(4) DEARN(t+4)
TONE	0.010***	0.009***	0.009***	0.005***
	(6.23)	(4.80)	(6.20)	(3.18)
EARN	-0.617***	-0.651***	-0.640***	-0.633***
	(-24.84)	(-26.10)	(-22.91)	(-17.16)
RET	0.011***	0.011***	0.006***	0.005***
	(4.67)	(7.41)	(4.24)	(2.68)
ACC	-0.187***	-0.186***	-0.176***	-0.200***
	(-11.66)	(-14.46)	(-18.05)	(-18.50)
SIZE	0.004***	0.003***	0.003***	0.003***
	(11.68)	(11.44)	(7.01)	(7.75)
MTB	-0.003***	-0.004***	-0.005***	-0.006***
	(-5.51)	(-5.23)	(-5.31)	(-5.76)
RETVOL	-0.061***	-0.066***	-0.058***	-0.051***
	(-8.42)	(-9.53)	(-5.88)	(-5.55)
Q2	-0.000	-0.009***	0.010***	0.001
	(-0.37)	(-6.14)	(4.41)	(0.73)
Q3	-0.009***	0.000	0.010***	-0.002*
	(-5.32)	(0.04)	(4.62)	(-1.31)
Q4	0.001	0.001	0.009***	-0.009***
	(0.45)	(0.52)	(4.59)	(-6.33)
Observations	122919	118100	112116	109597
R^2	0.46	0.44	0.36	0.37

This table shows the regression results of future earnings changes on MD&A tone and other control variables. The dependent variables are the earnings in the next four quarters (Compustat Quarterly file data item 69) minus the earnings in the current quarter (Compustat Quarterly file data item 69) scaled by the book value of assets at the end of the current quarter (Compustat Quarterly file data item 44). Independent variable $TONE$ is the average tone of the forward-looking statements of a firm-quarter. A forward-looking sentence's tone has a value of 1 if the Bayesian learning algorithm classifies the sentence as positive, 0 if neutral, and -1 if negative or uncertain. Other independent variables include $EARN$, RET, ACC, $SIZE$, MTB, $RETVOL$, $Q2$, $Q3$, and $Q4$. $EARN$ is the quarterly earnings (Compustat Quarterly file data item 69) scaled by the book value of assets (Compustat Quarterly file data item 44), winsorized at -3 and 3. RET is the contemporaneous stock returns in the fiscal quarter calculated using CRSP monthly return data. ACC is the accruals (earnings subtract cash flow from operations) scaled by the book value of assets(Compustat Quarterly file data item 44). $SIZE$ is the logarithm of the market value of equity at the end of the quarter (Compustat Quarterly file data item 14 times item 61). MTB is the market value of equity (Compustat Quarterly file data item 14 times item 61) plus the book value of total liabilities (Compustat Quarterly file data item 54) scaled by the book value of total assets(Compustat Quarterly file data item 44). $RETVOL$ is the stock return volatility calculated using twelve months of monthly return data before the fiscal quarter ending date. $Q2$ ($Q3$ or $Q4$) is a dummy variable which is set to 1 if the current reporting quarter is the second (third or fourth) fiscal quarter. Year and 2-digit SIC industry fixed effects are included in the regressions, but are not reported. T-statistics clustered by year are reported in parentheses. *** indicates $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

TABLE 6

Future Liquidity and MD&A tone

COEFFICIENT	(1) CFRATIO(t+1)	(2) CFRATIO(t+2)	(3) CFRATIO(t+3)	(4) CFRATIO(t+4)
TONE	0.280***	0.278***	0.241***	0.194***
	(10.37)	(11.62)	(11.02)	(9.05)
EARN	2.643***	2.469***	2.423***	2.654***
	(18.62)	(13.96)	(14.03)	(14.91)
RET	0.081***	0.071***	0.045***	0.035
	(3.34)	(3.26)	(1.76)	(0.93)
ACC	-2.094***	-1.453***	-1.635***	-2.433***
	(-19.34)	(-13.52)	(-18.57)	(-25.00)
SIZE	0.044***	0.050***	0.046***	0.032***
	(12.02)	(12.88)	(10.26)	(6.04)
MTB	-0.015*	-0.019**	-0.014**	-0.004
	(-2.14)	(-2.85)	(-2.25)	(-0.61)
RETVOL	-0.875***	-0.972***	-0.964***	-0.882***
	(-12.65)	(-13.75)	(-11.24)	(-14.51)
Q2	0.019***	0.062***	-0.074***	0.005
	(2.30)	(4.67)	(-2.75)	(0.71)
Q3	0.102***	-0.009	-0.070***	0.028*
	(8.98)	(-0.33)	(-3.06)	(2.12)
Q4	0.015	0.000	-0.061***	0.088***
	(0.81)	(0.02)	(-3.58)	(5.88)
Observations	105626	101410	96272	94029
R^2	0.33	0.28	0.28	0.33

This table shows the regression results of liquidity measures on MD&A tone and other control variables. The dependent variables are the cash flows from operations in the next four quarters (Compustat Quarterly file data item 108) scaled by the book value of current liabilities at the end of the current quarter (Compustat Quarterly file data item 49). Independent variable $TONE$ is the average tone of the forward-looking statements of a firm-quarter. A forward-looking sentence's tone has a value of 1 if the Bayesian learning algorithm classifies the sentence as positive, 0 if neutral, and -1 if negative or uncertain. Other independent variables include $EARN$, RET, ACC, $SIZE$, MTB, and $RETVOL$, all of which are defined in the note to Table 3. $Q2$ ($Q3$ or $Q4$) is a dummy variable which is set to 1 if the current reporting quarter is the second (third or fourth) fiscal quarter. Year and 2-digit SIC industry fixed effects are included in the regressions, but are not reported. T-statistics clustered by year are reported in parentheses. *** indicates $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

TABLE 7A

Descriptive Statistics of Tone Measures Based on Dictionaries

Variable	Mean	P5	P25	Median	P75	P95	Std.
DICTION_POS	5.64	1.30	3.25	5.03	7.24	9.84	3.80
DICTION_NEG	10.05	2.50	6.14	9.36	13.07	19.70	5.68
DICTION_TONE	-4.41	-15.59	-8.16	-4.04	-0.36	5.72	7.08
GI_POSITIV	5.81	3.64	4.97	5.80	6.61	8.04	1.36
GI_NEGATIV	2.87	1.05	2.06	2.79	3.63	4.90	1.18
GI_TONE	2.94	0.00	1.74	2.89	4.07	6.04	1.85
LIWC_POSEMO	2.00	0.74	1.48	1.95	2.44	3.42	0.83
LIWC_NEGEMO	0.79	0.00	0.37	0.70	1.14	1.82	0.58
LIWC_TONE	1.21	-0.36	0.56	1.16	1.80	2.93	1.03
TONE	-0.23	-0.75	-0.41	-0.21	-0.03	0.23	0.29

This table shows the descriptive statistics for the MD&A tone based on Diction, General Inquirer (GI), and Linguistic Inquiry and Word Count (LIWC). The variables are defined as follows. $DICTION\_POS$ is the percentage of positive words in the MD&A forward-looking statements classified by Diction. $DICTION\_POS = praise + satisfaction + inspiration$, where $praise$ is the percentage of words in the praise word list of Diction, $satisfaction$ is the percentage of words in the satisfaction word list of Diction, and $inspiration$ is the percentage of words in the inspiration word list of Diction. $DICTION\_NEG$ is the percentage of negative words in the MD&A forward-looking statements classified by Diction. $DICTION\_NEG = blame + hardship + denial$, where $blame$ is the percentage of words in the blame word list of Diction, $hardship$ is the percentage of words in the hardship word list of Diction, and $denial$ is the percentage of words in the denial word list of Diction. $DICTION\_TONE$ is calculated as $(DICTION\_POS - DICTION\_NEG)$. $GI\_POSITIV$ is the percentage of positive words of the MD&A forward-looking statements classified by the General Inquirer. $GI\_NEGATIV$ is the percentage of negative words of the MD&A forward-looking statements classified by the General Inquirer. GI_TONE is calculated as $(GI\_POSITIV - GI\_NEGATIV)$. $LIWC\_POSEMO$ is the percentage of positive emotion words in the MD&A forward-looking statements classified by the LIWC software. $LIWC\_NEGEMO$ is the percentage of negative emotion words in the MD&A forward-looking statements classified by the LIWC software. $LIWC\_TONE$ is calculated as $(LIWC\_POSEMO - LIWC\_NEGEMO)$. $TONE$ is the tone of the MD&A forward-looking statements classified by the Bayesian learning algorithm.

TABLE 7B

Correlation Matrix of the Different Tone Measures (P-values in Parentheses)

Variables	TONE	DICTION_POS	DICTION_NEG	DICTION_TONE	GI_POSITIV	GI_NEGATIV	GI_TONE	LIWC_POSEMO	LIWC_NEGEMO	LIWC_TONE
TONE	1.000									
DICTION_POS	0.025	1.000								
	(0.000)									
DICTION_NEG	-0.292	-0.080	1.000							
	(0.000)	(0.000)								
DICTION_TONE	0.247	0.600	-0.845	1.000						
	(0.000)	(0.000)	(0.000)							
GI_POSITIV	-0.025	0.246	-0.011	0.141	1.000					
	(0.000)	(0.000)	(0.000)	(0.000)						
GI_NEGATIV	-0.500	-0.042	0.339	-0.295	-0.064	1.000				
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)					
GI_TONE	0.299	0.207	-0.223	0.290	0.774	-0.681	1.000			
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)				
LIWC_POSEMO	-0.016	0.240	-0.017	0.142	0.483	-0.053	0.388	1.000		
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)			
LIWC_NEGEMO	-0.476	-0.038	0.440	-0.373	-0.073	0.627	-0.451	-0.042	1.000	
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		
LIWC_TONE	0.254	0.214	-0.261	0.324	0.429	-0.395	0.565	0.828	-0.596	1.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	

This table shows the Pearson correlation coefficients between $TONE$ and MD&A tone based on Diction, GI, and LIWC. The variables are defined as follows. $DICTION\_POS$ is the percentage of positive words in the MD&A forward-looking statements classified by Diction. $DICTION\_POS = praise + satisfaction + inspiration$, where *praise* is the percentage of words in the praise word list of Diction, $satisfaction$ is the percentage of words in the satisfaction word list of Diction, and *inspiration* is the percentage of words in the inspiration word list of Diction. $DICTION\_NEG$ is the percentage of negative words in the MD&A forward-looking statements classified by Diction. $DICTION\_NEG = blame + hardship + denial$, where *blame* is the percentage of words in the blame word list of Diction, *hardship* is the percentage of words in the hardship word list of Diction, and *denial* is the percentage of words in the denial word list of Diction. $DICTION\_TONE$ is calculated as $(DICTION\_POS - DICTION\_NEG)$. $GI\_POSITIV$ is the percentage of positive words of the MD&A forward-looking statements classified by the General Inquirer. $GI\_NEGATIV$ is the percentage of negative words of the MD&A forward-looking statements classified by the General Inquirer. GI_TONE is calculated as $(GI\_POSITIV - GI\_NEGATIV)$. $LIWC\_POSEMO$ is the percentage of positive emotion words in the MD&A forward-looking statements classified by the LIWC software. $LIWC\_NEGEMO$ is the percentage of negative emotion words in the MD&A forward-looking statements classified by the LIWC software. $LIWC\_TONE$ is calculated as $(LIWC\_POSEMO - LIWC\_NEGEMO)$. $TONE$ is the tone of the MD&A forward-looking statements classified by the Bayesian learning algorithm.

TABLE 7C

Information Content of Tone Measures Based on the Dictionary Approach

COEFFICIENT	(1) EARN(t+1)	(2) EARN(t+1)	(3) EARN(t+1)	(4) EARN(t+1)	(5) EARN(t+1)	(6) EARN(t+1)
DICTION_POS	-0.000					
	(-1.13)					
DICTION_NEG	-0.000**					
	(-2.03)					
DICTION_TONE		0.000				
		(1.06)				
GI_POSITIV			-0.002***			
			(-9.05)			
GI_NEGATIV			-0.001**			
			(-2.57)			
GI_TONE				-0.001***		
				(-5.72)		
LIWC_POSEMO					-0.002***	
					(-5.94)	
LIWC_NEGEMO					-0.001	
					(-1.15)	
LIWC_TONE						-0.001***
						(-4.49)
Observations	109749	109749	113802	113802	113207	113207
R^2	0.36	0.36	0.36	0.36	0.36	0.36

This table shows the regression results of future earnings on the tone measures based on Diction, GI, and LIWC and other control variables. The dependent variables are the earnings in the next quarter scaled by the book value of assets at the end of this quarter. The independent variables are defined as follows. $DICTION\_POS$ is the percentage of positive words in the MD&A forward-looking statements classified by Diction. $DICTION\_POS = praise + satisfaction + inspiration$, where $praise$ is the percentage of words in the praise word list of Diction, $satisfaction$ is the percentage of words in the satisfaction word list of Diction, and $inspiration$ is the percentage of words in the inspiration word list of Diction. $DICTION\_NEG$ is the percentage of negative words in the MD&A forward-looking statements classified by Diction. $DICTION\_NEG = blame + hardship + denial$, where $blame$ is the percentage of words in the blame word list of Diction, $hardship$ is the percentage of words in the hardship word list of Diction, and $denial$ is the percentage of words in the denial word list of Diction. $DICTION\_TONE$ is calculated as $(DICTION\_POS - DICTION\_NEG)$. $GI\_POSITIV$ is the percentage of positive words of the MD&A forward-looking statements classified by the General Inquirer. $GI\_NEGATIV$ is the percentage of negative words of the MD&A forward-looking statements classified by the General Inquirer. GI_TONE is calculated as $(GI\_POSITIV - GI\_NEGATIV)$. $LIWC\_POSEMO$ is the percentage of positive emotion words in the MD&A forward-looking statements classified by the LIWC software. $LIWC\_NEGEMO$ is the percentage of negative emotion words in the MD&A forward-looking statements classified by the LIWC software. $LIWC\_TONE$ is calculated as $(LIWC\_POSEMO - LIWC\_NEGEMO)$. $TONE$ is the tone of the MD&A forward-looking statements classified by the Bayesian learning algorithm. Year and 2-digit SIC industry fixed effects are included in the regressions, but are not reported. The following control variables are also included in the regressions but the coefficients are not reported: $EARN$, RET, ACC, $SIZE$, MTB, $RETVOL$, $Q2$, $Q3$, and $Q4$, all of which are as defined in the notes to Table 3 and Table 5. T-statistics clustered by year are reported in parentheses. *** indicates $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

TABLE 8

Information content of MD&A before and after 2003

COEFFICIENT	(1) EARN(t+1)	(2) EARN(t+2)	(3) EARN(t+3)	(4) EARN(t+4)	(5) CFRATIO(t+1)	(6) CFRATIO(t+2)	(7) CFRATIO(t+3)	(8) CFRATIO(t+4)
TONE	0.010***	0.009***	0.008***	0.005***	0.229***	0.233***	0.207***	0.182***
	(5.37)	(4.61)	(7.16)	(3.29)	(5.76)	(6.90)	(8.36)	(10.73)
POST2003	-0.000	0.001	0.003**	0.004***	0.025	0.023	0.018	0.008
	(-0.05)	(1.30)	(2.46)	(3.89)	(1.46)	(1.43)	(1.13)	(0.61)
TONE×POST2003	-0.004	-0.003	-0.002	0.002	0.003	-0.011	-0.031	-0.076*
	(-1.43)	(-1.38)	(-0.62)	(0.39)	(0.04)	(-0.23)	(-0.86)	(-2.15)
Observations	119859	115159	109290	106865	102840	98703	93670	91529
R^2	0.33	0.28	0.24	0.24	0.33	0.28	0.27	0.33

This table shows the regression results of future earnings and liquidity measures on MD&A tone, a dummy for post-2003, the interaction of MD&A tone and the post-2003 dummy and other control variables. The dependent variables are the earnings in the next four quarters (Compustat Quarterly file data item 69) scaled by the book value of assets at the end of the current quarter (Compustat Quarterly file data item 44) in columns (1) to (4) and the cash flows from operations in the next four quarters (Compustat Quarterly file data item 108) scaled by the book value of current liabilities at the end of that quarter (Compustat Quarterly file data item 49) in columns (5) to (8). $POST2003$ is a dummy variable that equals 1 if the 10-K or 10-Q report is filed in or after 2003 and 0 otherwise. Year and 2-digit SIC industry fixed effects are included in the regressions, but are not reported. Other control variables (coefficients unreported) are the same as those in Table 5A and Table 6. T-statistics clustered by year are reported in parentheses. *** indicates $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

TABLE 9

Future Excess Returns of Portfolios Sorted on Accruals

	Full-sample tests	Sub-sample tests (Within-sub-sample sorting)		Sub-sample tests (Within-full-sample sorting)	
Portfolios		$WARN = 0$	$WARN = 1$	$WARN = 0$	$WARN = 1$
1	0.59	0.92	0.48	0.83	0.51
	[358]	[167]	[191]	[197]	[160]
2	0.27	0.40	0.47	0.45	0.33
	[358]	[167]	[191]	[185]	[173]
3	0.28	0.50	0.21	0.42	0.49
	[358]	[167]	[191]	[180]	[179]
4	-0.05	0.37	-0.11	0.25	-0.05
	[358]	[167]	[191]	[170]	[189]
5	0.06	0.21	0.17	0.33	-0.11
	[358]	[167]	[191]	[162]	[198]
6	0.12	0.01	0.03	-0.10	0.26
	[358]	[167]	[191]	[158]	[166]
7	-0.16	-0.23	-0.29	-0.21	-0.17
	[358]	[167]	[191]	[179]	[180]
8	-0.22	-0.19	-0.04	-0.25	-0.07
	[358]	[167]	[191]	[160]	[199]
9	-0.39	-0.46	0.03	-0.71	-0.06
	[358]	[167]	[191]	[145]	[216]
10	-0.12	-0.48	0.60	-0.52	0.42
	[358]	[167]	[191]	[143]	[254]
Hedge portfolio (1-10)	0.70*	1.40***	-0.12	0.98*	0.03
	(1.77)	(3.61)	(-0.19)	(1.71)	(0.05)
Average number of observations	358	167	191	167	191

This table shows the monthly excess returns (in percentage) of the decile portfolios sorted on accruals as a function of the MD&A tone. Every month, a firm is assigned into one of ten portfolios based on its most recent accruals, ACC, which is calculated as earnings (Compustat Quarterly file data item 69) subtract cash flow from operations (Compustat Quarterly file data item 108) scaled by the book value of assets(Compustat Quarterly file data item 44). Portfolio 1 has the lowest ACC and portfolio 10 has the highest ACC. The excess returns of the portfolios are returns controlling for the market returns and size and book-to-market factors, which are calculated as the intercepts from the time-series regressions of portfolio returns on $R_m - R_f$ (market returns subtract risk-free returns), SMB (the Fama-French small-minus-big mimicking portfolio returns), and HML (the Fama-French low-minus-high mimicking portfolio returns).

The table reports results for three samples. The "full-sample tests" include all firms; the full sample is also divided into two sub-samples: firms with $WARN = 1$ and those with $WARN = 0$. $WARN$ is a dummy variable that equals 0 if ACC and $TONE$ (both variables measured using the most recent financial statement data) are both either above the median or below the median based on data from the most recent quarter, and 1 otherwise (i.e., when one is above median and the other is below median). For the columns titled "within-sub-sample sorting", firms are sorted into decile portfolios based on ACC within the sub-sample. For the columns titled "within-full-sample sorting", firms are sorted into decile portfolios based on ACC within the full-sample. The average numbers of observations for each portfolio are reported in the bracket. The hedge portfolio return is calculated as the return from the portfolio that longs portfolio 1 (i.e., the portfolio of firms with the lowest ACC) and shorts portfolio 10 (i.e., the portfolio of firms with the highest ACC). T-statistics indicating whether the hedge portfolio returns are significantly different from zero are reported in parentheses. *** indicates $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

TABLE 10

Decomposing MD&A Tone by Content Category

	(1)	(2)
COEFFICIENT	EARN(t+1)	CFRATIO(t+1)
PROFIT_TONE	0.003***	0.096***
	(4.93)	(11.74)
LIQUIDITY_TONE	0.013***	0.250***
	(8.49)	(11.86)
OTHER_TONE	-0.001	0.006
	(-1.67)	(0.48)
Observations	122922	105637
R^2	0.36	0.33

This table shows the regression results of future earnings and liquidity on MD&A tone of the forward-looking statements in different categories. The dependent variables are the earnings in the next quarter (Compustat Quarterly file data item 69) scaled by the book value of assets at the end of the current quarter (Compustat Quarterly file data item 44) in column (1) and the cash flows from operations in the next quarter (Compustat Quarterly file data item 108) scaled by the book value of current liabilities at the end of the current quarter (Compustat Quarterly file data item 49) in column (2). $PROFIT\_TONE$ is the average tone of the forward-looking statements of a firm-quarter that are about profits or operations (i.e., the statements that are classified as categories 1 to 4 as defined in Section 4.4). $LIQUIDITY\_TONE$ is the average tone of the forward-looking statements of a firm-quarter that are about liquidity or capital resources (i.e., the statements that are classified as categories 5 to 7 as defined in Section 4.4). $OTHER\_TONE$ is the average tone of the forward-looking statements of a firm-quarter that are about other topics (i.e., the statements that are classified as categories 8 to 12 as defined in Section 4.4). Control variables (results unreported) include $EARN$, RET, ACC, $SIZE$, MTB, $RETVOL$, $Q2$, $Q3$, and $Q4$. $Q2$ ($Q3$ or $Q4$) is a dummy variable which is set to 1 if the current reporting quarter is the second (third or fourth) fiscal quarter. Other independent variables are as defined in the note to Table 3. Year and 2-digit SIC industry fixed effects are included in the regressions, but are not reported. T-statistics clustered by year are reported in parentheses. *** indicates $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.