THE ACCOUNTING REVIEW
Vol. 86, No. 3
2011
pp. 1087–1115

American Accounting Association
DOI: 10.2308/accr.00000043

The Effect of Annual Report Readability on Analyst Following and the Properties of Their Earnings Forecasts

Reuven Lehavy
Feng Li
Kenneth Merkley
University of Michigan

ABSTRACT: This study examines the effect of the readability of firms' written communication on the behavior of sell-side financial analysts. Using a measure of the readability of corporate 10-K filings, we document that analyst following, the amount of effort incurred to generate their reports, and the informativeness of their reports are greater for firms with less readable 10-Ks. Additionally, we find that less readable 10-Ks are associated with greater dispersion, lower accuracy, and greater overall uncertainty in analyst earnings forecasts. Overall, our results are consistent with the prediction of an increasing demand for analyst services for firms with less readable communication and a greater collective effort by analysts for firms with less readable disclosures. Our results contribute to the understanding of the role of analysts as information intermediaries for investors and the effect of the complexity of written financial communication on the usefulness of this information.

Keywords: *financial analysts; readability; Gunning Fog Index; analyst following*.

Data Availability: *The data used in this study are available from the sources indicated in the text.*

I. INTRODUCTION

Over the past two decades, changes in financial and reporting regulations (e.g., changes in segment disclosures, employee stock option reporting, and Sarbanes-Oxley disclosures) have significantly increased the amount of required disclosures by firms to external users. In addition, technological advancement and new developments in financial engineering have made it more challenging for firms to communicate information about the underlying fundamentals of their businesses in a clear and informative manner. The increase in the amount of required dis-

We thank Beverly Walther (editor), Steven Kachelmeier (senior editor), and two anonymous reviewers. We also thank workshop participants from The University of Chicago, University of Michigan, University of Minnesota, Erasmus University, University of Melbourne, Tel Aviv University, the 2010 AAA Annual Meeting, the 2010 EEA Annual Congress, and a joint presentation of HEC-Paris and INSEAD for their helpful comments and suggestions.

Editor's note: Accepted by Beverly Walther.

Submitted: December 2009
Accepted: September 2010
Published Online: April 2011

closure accompanied by the challenging task of communicating increasingly complex business transactions to investors has led to concerns about the effectiveness of management communication and the ability of interested users to make informed decisions based on this information. As a result, both the SEC and the popular press have routinely criticized firms for the complexity of the language in their 10-K filings (Schroeder 2002). These groups have expressed concerns about the ability of financial statement users, especially small investors, to understand the complicated writing found in firms' financial documents (SEC 1998; Schroeder 2002; Cox 2007).

Given the increasing complexity of firm disclosures and the related concerns about their usability, a natural question arises as to the role of financial analysts in intermediating such information to investors. Specifically, do financial analysts use their expertise to examine this complex communication and provide useful information to financial statement users, or do they prefer to focus their efforts on firms with less complex communication? We attempt to answer this question by examining how the behavior of financial analysts relates to the readability of firms' written communication.

Prior literature examines the relation between the properties of firms' financial disclosures and the behavior of security analysts. These properties include the informativeness of disclosures (Lang and Lundholm 1996; Healy et al. 1999), the use of segment disclosures (Botosan and Harris 2000), and the effect of intangible assets (Barth et al. 2001). These studies generally document that firms with better disclosure quality tend to attract greater analyst following. Other studies examine the effect of certain financial reporting items on the properties of analyst earnings forecasts. These studies find that analyst forecasts are influenced by firms' accounting choices (Hopkins et al. 2000; Bradshaw et al. 2008), changes in the tax law (Plumlee 2003), the clarity of the income effects of specific financial items (Hirst and Hopkins 1998; Hirst et al. 2003), and international diversification (Duru and Reeb 2002).

While each of these studies contributes to our understanding of the effect of firms' financial reporting and disclosure choices on analyst behavior, they generally focus on the effect of a single financial statement item (e.g., interest rates, taxes) or disclosure (e.g., segment reporting).[1] This approach, however, does not explicitly incorporate the potential effect of the properties of other financial statement items or disclosures on analyst following or the properties of their earnings forecasts. More important, this approach does not address the potential impact on analyst behavior of the *overall* complexity of firms' financial communication, that is, the cost of processing and interpreting the entirety of firms' disclosures. This potential limitation is particularly important, given the broad nature of corporate disclosures such as 10-K filings that describe a significant number of interrelated financial items.

We attempt to address this issue by examining the relation between a comprehensive measure of the overall readability of corporate 10-K filings and analyst behavior. This measure, known as the Gunning Fog Index (hereafter, Fog Index), incorporates the number of words per sentence and the number of complex words in a document to derive a measure of the readability or syntactic complexity of firms' 10-K filings (Jones and Shoemaker 1994; Courtis 1995; Li 2008).[2] The Fog Index has been widely used in social science research for several decades to examine the relation

[1] One notable exception is Lang and Lundholm (1996) who use AIMR scores as a measure of overall disclosure quality. The potential limitations of this measure, however, are that it is based on a subjective survey of analysts, limited to a subset of large firms, and no longer available after 1995.

[2] Specifically, we focus on the readability of firms' written communication as a measure of the syntactic complexity of such communication. Stimulated by Chomsky's (1965) seminal generative grammar, there has been a great deal of theoretical and empirical work on what constitutes syntactic complexity and how it functions as a determinant of both sentence comprehension (Gibson 1998) and sentence production (Thompson and Faroqi-Shah 2002). According to Stone et al. (2005, 341–343), sentence complexity is affected by the features of open-class words (nouns and verbs) and their relationships and the number and type of syntactic operations.

between the readability of written information and various decisions or outcomes.[3] Further, recent studies use the Fog Index to examine the readability of annual reports in connection with earnings persistence (Li 2008), timely price adjustment (Callen et al. 2009), and investment efficiency (Biddle et al. 2009).

The Fog Index offers several important advantages. First, it allows us to study a large and diverse group of firms. Second, it is an objective measure, not based on analyst surveys or opinion, and can be calculated for any narrative disclosure. Finally, it allows us not only to capture the effect of the readability of a variety of financial items, but, more important, to directly examine the overall syntactic complexity of firms' written public communication, over and above its specific content.[4]

We focus on the readability of the 10-K filing for several reasons. First, 10-K filings are required for all publicly traded companies and are frequently cited as an input in the decisions of investors and financial analysts (Previts et al. 1994; Rogers and Grant 1997). Second, both the SEC and popular press have criticized firms for the complexity of their language in these filings and the effect of this complexity on investors (Schroeder 2002). You and Zhang (2009) provide evidence that suggests that investors underreact to the information provided in 10-K filings and that this effect is stronger for firms with more complex 10-K reports. Finally, the 10-K filing contains a significant amount of written communication or narrative to use in interpreting its readability. These reasons make the 10-K filing an interesting setting to examine the effects of the readability of firm communication on the activities of financial analysts.

We test the relation between the readability of 10-K filings and several measures of analyst behavior, including analyst following, analyst forecast revision response time, the information content of analysts' reports, and the properties of analyst earnings forecasts. We begin our analysis by examining the relation between analyst following and 10-K filing readability. We argue that, on the one hand, lower readability of firm financial disclosures increases the cost of processing the information in these disclosures and therefore will increase the demand for analyst services. If the cost of obtaining analyst reports is less than the information-processing costs of firm disclosures, then analyst following should be greater for firms with less readable 10-K filings. On the other hand, less readable disclosure could also increase the costs of analyst coverage. That is, analysts could bear greater information-processing costs and higher private search costs, which could lead to less accurate forecasts. If these costs are significant, then analyst following should be lower for firms with less readable disclosures.

Consistent with the prediction of an increasing demand for analyst services for firms with less readable 10-K filings, we document a positive and significant association between a firm's 10-K Fog Index and the number of analysts who cover the firm, even after controlling for other factors related to analyst coverage. We interpret this result as being consistent with a greater collective effort by analysts for firms with less readable disclosures.

Next, we examine whether individual analyst effort is associated with the readability of the 10-K filing. We measure individual analyst effort as the length of time required for analysts to issue their first forecast revision following the 10-K filing. If analysts bear costs in following firms

[3] The Fog Index has been used in a variety of applications including medical error, consumer drug use, consumer warranties, mutual fund prospectuses, jury instructions, and academic research prestige. For specific examples, see Charrow and Charrow (1979), Armstrong (1980), Shuptrine and Moore (1980), Ott and Hardie (1997), Gazmararian et al. (1999), Koo et al. (2003), and Lee et al. (2006).

[4] In a recent speech, former SEC Chairman Christopher Cox specifically identifies readability measures as a tool to examine communication complexity. He states, "Just as the Black-Scholes model is commonplace when it comes to compliance with the stock option compensation rules, we may soon be looking to the Gunning-Fog and Flesch-Kincaid models to judge the level of compliance with the plain English rules" (Cox 2007). Similarly, Core (2001) proposes the use of computational linguistics methods in accounting to measure corporate disclosure quality.

with less readable disclosures, then it should take them more time, on average, to issue reports following less readable disclosures. Using a variety of empirical approaches, we find that analysts who cover firms with less readable 10-K filings take a longer time to issue their first report following the 10-K filing. This evidence is consistent with the argument that analysts exert more effort to follow firms with less readable disclosures.

We then turn to an examination of the relation between 10-K readability and the information content of analyst reports. Similar to Frankel et al. (2006), we measure information content as the proportion of a firm's stock returns related to analyst forecast revisions to the total firm's stock return during the time period between the 10-K filing and the subsequent fiscal year-end. We predict that, if investors place greater reliance on analyst reports issued for firms with less readable 10-K filings (potentially due to the higher costs associated with processing and interpreting a complex report), then the information content of analyst reports, or the proportion of firm returns associated with them, should be higher for firms with less readable disclosures. Consistent with this prediction, we find evidence that the proportion of firm returns associated with analysts' reports is higher for firms with less readable 10-K reports. This evidence suggests that investors find analysts' reports more informative for firms with less readable corporate disclosures.

Finally, we examine how 10-K filing readability relates to commonly studied properties of analyst earnings forecasts: dispersion, accuracy, and two measures of earnings consensus forecast uncertainty (overall and common uncertainty) based on Barron et al. (1998). Prior literature finds that the 10-K filing represents a major source of information used by analysts (Previts et al. 1994; Rogers and Grant 1997). However, studies also find that the complexity of items in the 10-K filing affects analysts' use of this information. For example, Plumlee (2003) finds that analyst forecasts are less accurate if they are associated with complex changes in the tax law. Bradshaw et al. (2008) find that differences in accounting choice negatively affect forecast accuracy and increase dispersion. Our tests relate to these findings, because we examine the effect of complexity through the ability of analysts to incorporate less readable financial information into their forecasts. If lower readability increases the costs of processing and interpreting firm disclosures, then it could limit the ability of analysts to correctly incorporate all of the pertinent information and could result in disagreement or ambiguity. We predict and find that analyst earnings forecasts for firms with less readable 10-K reports have greater dispersion, are less accurate, and are associated with greater overall analyst uncertainty. We also find that analyst common uncertainty is increasing in 10-K Fog. These results suggest that analyst forecasts are affected by 10-K readability and provide indirect evidence to support the notion that the information contained in the 10-K filing is used by analysts.

This study contributes to the literature in the following ways. First, our overall findings that lower readability is associated with greater levels of analyst coverage, effort, and information content, but with lower accuracy, higher dispersion, and greater uncertainty of their earnings forecasts complements and contributes to the literature on how analysts respond to firms' disclosure. While prior studies generally focus on the effect of the complexity of specific attributes of firms' disclosure on analyst behavior, we provide evidence that suggests that analyst behavior is associated with overall disclosure readability.[5] While the collection of the evidence in this study is consistent with this inference, the evidence obtained from each individual test is subject to limitations common to our research design.

[5] We note that covering firms with less readable communication is likely to result in a trade-off. Investors are more likely to find analyst coverage useful for firms with less readable communication; however, earnings forecasts for these firms will potentially be less accurate. Our findings suggest that the benefits of coverage outweigh the costs, because we document that both greater coverage and less accurate forecasts are associated with less readable communication (see also Li et al. 2009). This is consistent with the notion that investors value more than just forecast accuracy (Schipper 1991; Clement and Tse 2003).

Second, our findings relate to the debate about the intended audience of public corporate filings. Since its organization, the SEC has made consistent efforts to encourage firms to make their regulatory filings accessible to the average or "lay" investor (Firtel 1999).[6] The most recent of these efforts is the plain English disclosure rules adopted by the SEC on January 22, 1998. Our evidence, that less readable corporate disclosures are associated with greater analyst following and more informative analyst reports, suggests that analysts could serve as an available information intermediary for investors and relates to the SEC's concerns about the accessibility of these reports.

Section II discusses the role of the readability of firm communication and the hypotheses development. Section III describes the data and sample, and section IV presents our empirical evidence. Section V concludes.

II. HYPOTHESES DEVELOPMENT

Less readable communication is more difficult to interpret and process by investors because it requires that investors devote more time and effort to identify and extract relevant information (Bloomfield 2002). In this study we focus on the written complexity of firm disclosures, as measured by readability rather than content. We interpret disclosure readability to be a measure of the costs incurred by users to process and interpret a firm's written communication after controlling for the operational complexity of the business.[7] Firms with similar operations provide disclosures with varying levels of readability. For example, Berkshire Hathaway and AIG both provide a section about their reinsurance businesses in the management's discussion and analysis section of their 10-K filings for fiscal year 2003. While the subject matter and underlying business complexity of the reinsurance business sections are very similar across the two firms, Berkshire Hathaway's explanation is more readable as evidence by a lower Fog Index.[8] This is not surprising, insofar as Berkshire Hathaway CEO Warren Buffett is a strong proponent of more readable communication.

We begin our analysis by examining the effect of disclosure readability on analyst following. We assume that users of financial information have different abilities to process complex communication (Indjejikian 1991; Ball 1992). These differences provide opportunities for information intermediaries, such as financial analysts, to profit from their private analysis of firms by selling their opinions to users with greater information-processing costs (Schipper 1991). These profitable opportunities are arguably greater for firms with less readable disclosures because of the greater cost to users of processing the firms' information. As such, if analysts respond to this increased demand for their services, then we expect analyst following to be greater for firms with less readable disclosures.

---

6 Some academic scholars and practitioners have argued that the primary users of this information should be market professionals such as analysts and not the "lay" investor (Kripke 1970; Schipper 1991). This is because nonprofessionals might not possess the skill or expertise to read and understand the complex financial information contained in disclosure documents and, therefore, any effort to gear disclosure toward the layperson is a waste of time and money (Kripke 1970). Further, critics argue that the SEC's efforts to appeal to the average investor are not only inefficient, but also are hazardous to the disclosure regime because simplified and concise disclosure often leaves out many issues that are potentially valuable to the professional and leads to potential legal liability.

7 It is important to note that more readable communication can come at a cost. Legal scholars point out that "it is much harder to simplify than to complicate" (Kimble 1994, 53) and it requires significant skill, work, and time to compose documents in plain language. Additionally, some critics argue that plain English communication increases the risk of litigation (Kripke 1973).

8 The Fog Index for AIG's reinsurance business section of the MD&A is 18.51, while Berkshire Hathaway's is 17.23. The difference between these scores (1.28) can be interpreted as the number of additional years of formal education required to understand the text on a first reading. Further details regarding the Fog Index are provided in Section III.

We note, however, that analysts also face a variety of costs in covering firms with less readable communication. First, analysts bear the direct costs of processing the information provided by management. Second, less readable communication can result in confusion, requiring analysts to incur greater private search costs as they obtain additional information to evaluate and interpret management's communication. Third, greater syntactic complexity could lead to inaccurate forecasts and recommendations that adversely affect analysts' careers (e.g., Mikhail et al. 1999; Plumlee 2003; Hong and Kubik 2003).[9] Finally, Li (2008) finds that firms attempt to obfuscate bad news by increasing the complexity of their communication, and Lang et al. (2004) find that analysts are less likely to cover firms with incentives to withhold or manipulate information. These potential manipulations along with the costs to analysts of covering firms with less readable communication discourage analyst following. Because there are both potential positive and negative consequences of syntactic complexity on analyst following, we test the null hypothesis that syntactic complexity has no association with analyst following.

Given the difficulty of following firms with less readable disclosures, analysts who choose to follow these firms likely exert greater effort to do so. One way of measuring the effort that analysts exert in following firms with less readable disclosures is to examine the average time from the firm's 10-K filing to the analyst's first report subsequent to the filing. If lower readability results in higher processing costs to analysts, then firms with less readable communication will have greater average response time from the analysts that cover them as compared to the analysts covering firms with more readable communication. Similarly, analysts who cover a portfolio of firms that have less readable communication should take longer, on average, to issue their reports than other analysts. We test the hypothesis that analysts exert greater individual effort to cover firms with less readable disclosures.

Readability could also affect the properties of analyst earnings forecasts. Specifically, we examine its effect on the information content, accuracy, and dispersion of analyst forecasts and the uncertainty in analysts' information environment. Analyst reports for firms with less readable communication likely provide information that is more useful to investors due to greater costs in processing public information. Analyst reports for firms with less readable communication would also be more informative if analysts choose to acquire and incorporate more private information due to the difficulties involved in interpreting less readable disclosures. Accordingly, we predict that analyst reports for firms with less readable communication will have greater information content, on average, than firms with more readable communication.

Finally, because analysts who follow firms with less readable communication bear the costs of processing and interpreting such disclosures, syntactic complexity could also affect the dispersion and accuracy of analyst earnings forecasts as well as the degree of the overall and common uncertainty embedded in analyst forecasts. If less readable communication increases analysts' cost to process and interpret the information, then it is likely to lead to a more diverse set of interpretations about firm disclosures, resulting in higher analyst forecast dispersion. In addition, if less readable communication makes it more difficult to forecast earnings, then the accuracy of the analyst consensus forecast will be lower for firms with less readable communication. However, this difference in accuracy can be fully or partially offset by greater analyst effort in response to greater syntactic complexity. Accordingly, we predict that analyst forecasts will have greater dispersion and be less accurate for firms with less readable communication.

[9] Li et al. (2009) find that once analysts become all-stars they are more willing to cover firms with greater potential earnings management because investors could find their coverage more valuable for these firms. This suggests that coverage decisions can be related to favorable career outcomes; however, it is not clear how analysts assess the trade-offs between the demand for investment information and the cost of potentially less accurate forecasts.

Barron et al. (1998) point out that analyst earnings forecast dispersion and accuracy do not directly capture the theoretical properties of analysts' information environment such as uncertainty because these measures relate in different ways to the idiosyncratic and the common components of error in analysts' forecasts. They view idiosyncratic uncertainty as uncertainty in private information that affects each analyst differently, and common uncertainty as uncertainty in information that is common to all analysts. Accordingly, lower readability can be associated with higher dispersion and lower accuracy in earnings forecasts for reasons related to either or both common and idiosyncratic uncertainty in the analysts' information environment.

To further examine the association between readability and properties of analyst earnings forecasts, we follow the model derived in Barron et al. (1998) and construct empirical measures of the overall uncertainty (modeled as the sum of the idiosyncratic and the common uncertainty) and the common uncertainty in analyst forecasts. These measures combine the accuracy, the dispersion, and the number of analyst forecasts, allowing us to examine more directly how readability relates to analysts' information environment uncertainty. We predict that analyst forecasts for firms with less readable reports will be associated with greater overall uncertainty. The effect of readability on common uncertainty (i.e., how much the average belief reflects common versus private information) is less clear because analysts can make trade-offs between public and private information sources based on the relative precision and importance of these sources.

III. SAMPLE AND VARIABLE DEFINITIONS

Sample Selection

Our initial sample is based on the intersection of firm/years available on the Compustat Fundamental Annual table and the SEC's EDGAR filings database for fiscal years 1995–2006. These databases are joined based on Compustat GVKEY and the SEC's Central Index Key (CIK). Firms without matches are dropped from the sample. For each firm-year observation, we download the corresponding 10-K filing. Filings with less than 3,000 words or 100 lines are dropped to ensure that a complete filing is examined and that no errors were made in the filing transmission. This procedure results in 57,642 observations. We obtain stock return data from CRSP, analyst data from I/B/E/S, institutional holdings data from Thomson Reuters (CDA/Spectrum), and information on management earnings guidance from the First Call Company Issued Guidelines database. This procedure yields a sample of 33,704 observations.

Disclosure Readability

Similar to Li (2008), we measure the readability of 10-K filings using the Fog Index (*FOG*). This index, developed in the computational linguistics literature, captures the written complexity of a document as a function of the number of syllables per word and the number of words per sentence. Specifically, we calculate the readability of the 10-K report for firm *i* in year *t* as follows:

$$FOG_{i,t} = (\text{average words per sentence} + \text{percent of complex words}) \times 0.4 \tag{1}$$

where a complex word is defined as one with three or more syllables. The index is interpreted as the number of years of formal education required for a person of average intelligence to read the document once and understand it.[10] It is important to note that the Fog Index is a measuring tool,

[10] Before computing the Fog Index, we remove all tables, tabulated text, and financial statements from the 10-K. The constant of 0.4 found in Equation (1) was chosen by Robert Gunning based on the scores of a set of literary benchmarks in order that this specific interpretation could be made. Our inferences remain unchanged using alternative measures of readability such as the Kincaide grade level formula, the Flesch Reading Ease Index, and 10-K length. In addition, our main inferences are similar after including in the analysis the length of the 10-K filing as a control variable for the amount of disclosure provided by the firm (Leuz and Schrand 2009).

not a rule or formula for good writing; as stated by Gunning (1969, 12), "Nonsense written simply is still nonsense." It predicts the readability of a document, but does not provide information about whether the writing is interesting or informative. Despite these limitations, it is objective and simple to calculate. It allows us to study the disclosure characteristics of a large and diverse group of firms and does not depend on analyst surveys or opinions. It also provides us with a comprehensive measure of the overall syntactic complexity of 10-K filings as opposed to the complexity of individual financial items.[11]

Table 1 provides descriptive statistics on the Fog Index for firms in the intersection of Compustat and EDGAR. As can be seen in Panel A, the overall mean and median of the Fog Index are 19.52 and 19.38, respectively. As a comparison to help validate our measure, it is interesting to note that the average Fog Index for FASB Statement Nos. 1–122 is 22, the CPA Exams is 16, the CMA Exams is 17, *Wall Street Journal* is 12, and *Reader's Digest* is 8 (Phillips et al. 2007; Cox 2007). The relatively high mean and median Fog Index in our sample is consistent with concerns that financial reports are written in complex language.

While the variation in the mean Fog Index over the sample period is modest, there is large variation within each year. For example, over the total sample years, the interquartile range for the Fog Index is from 1.58 to 1.88 (about 9 percent of the mean). There is also significant variation in the Fog Index within industries, despite similarities in the underlying business complexity within each industry. Table 1, Panel B provides examples of specific industries that have high and low levels of the Fog Index (industries are classified using the Fama and French 48-industry classification). The healthcare, insurance, trading (e.g., security brokers, investment offices, etc.), utilities, and telecommunications industries comprise the group with the highest Fog Index. The fact that these industries have a low level of disclosure readability is not surprising because they are characterized by complex contracts and business models that are difficult to communicate. However, there is also significant variation within each industry. Industries with a low Fog Index comprise precious metals, shipping containers, food products, agriculture, and defense industries. While the precious metals and shipping container industries have small within-industry variation, most industries have an interquartile range of the Fog Index above 1.60. The relative rankings of these industries also help to validate the ability of the Fog Index to measure the readability of annual reports.

Variable Definitions

Our tests focus on examining the relation between the Fog Index and analyst following, their forecast revision response time, the information content of their reports, and the properties of their earnings forecasts. A description of the construction of these variables as well as our control variables follows.

Analyst Following

Similar to prior research (O'Brien and Bhushan 1990; Brennan and Subrahmanyam 1995), we define analyst following as the number of analysts (*#ANALYSTS*) that comprise the first I/B/E/S consensus annual earnings forecast after the filing date of the 10-K report. We follow Bhushan (1989) and interpret this measure as a proxy for the collective effort of the financial analyst community in the analysis of an individual firm.[12] Because some firms are not covered by I/B/E/S,

[11] Prior literature used the Fog Index and other similarly constructed measures to examine the readability of the overall annual report (Jones and Shoemaker 1994), management's discussion and analysis (Schroeder and Gibson 1990), and the notes to the financial statements (Smith and Smith 1971; Healy 1977).

[12] Similar to Bhushan (1989), we acknowledge that our proxy is not a perfect measure because it assumes homogeneity among analyst effort levels. For example, there are differences in individual analysts' effort levels based on differences in compensation, brokerage houses, etc.

TABLE 1

Descriptive Statistics for Disclosure Readability Using the Fog Index of the 10-K Filing

Panel A: Fog Index by Fiscal Year (All Compustat Firms)

Year	n	Mean	Std. Dev.	Q1	Median	Q3	Interquartile
1995	3,107	19.25	1.31	18.38	19.17	20.04	1.66
1996	5,221	19.45	1.31	18.53	19.38	20.24	1.71
1997	5,530	19.48	1.30	18.59	19.39	20.28	1.70
1998	5,329	19.44	1.25	18.59	19.35	20.20	1.62
1999	5,524	19.26	1.29	18.40	19.14	20.02	1.62
2000	5,494	19.19	1.29	18.34	19.05	19.92	1.58
2001	4,955	19.30	1.42	18.39	19.14	20.00	1.61
2002	4,723	19.57	1.52	18.60	19.38	20.33	1.73
2003	4,546	19.91	1.79	18.80	19.63	20.68	1.88
2004	4,425	19.78	1.86	18.81	19.64	20.67	1.87
2005	4,449	19.86	1.50	18.88	19.68	20.56	1.68
2006	4,339	19.90	1.49	18.94	19.72	20.61	1.67
All Years	57,642	19.52	1.47	18.58	19.38	20.29	1.70

Panel B: Fog Index by Industry (All Compustat Firms)

	n	Mean	Std. Dev.	Q1	Median	Q3	Interquartile
Industries with High Fog Index							
Healthcare	1,115	20.22	1.46	19.30	20.04	20.99	1.69
Insurance	2,143	20.16	1.29	19.32	20.10	20.89	1.57
Trading	3,279	19.88	1.67	18.78	19.75	20.75	1.97
Utilities	2,405	19.84	1.56	18.92	19.64	20.53	1.61
Telecommunications	2,021	19.80	1.45	18.89	19.59	20.45	1.56
Industries with Low Fog Index							
Precious Metals	135	18.43	0.94	17.75	18.38	19.12	1.37
Shipping Containers	159	18.64	1.25	17.87	18.43	19.20	1.33
Food Products	483	18.86	1.56	17.95	18.65	19.63	1.68
Agriculture	162	18.95	1.47	17.90	18.85	19.58	1.68
Defense	108	18.99	1.45	18.00	18.87	19.88	1.88

(continued on next page)

Panel A reports descriptive statistics on our measure of disclosure readability, the Fog Index, for observations available on the SEC's EDGAR and Compustat databases. The Fog Index is computed as: (average words per sentence + percent of complex words) × 0.4, using the text of the 10-K filings for fiscal years 1995–2006.

Panel B provides descriptive statistics for the five industries with the highest and lowest mean Fog Index. Industries are based on the Fama and French 48-industry classification.

we conduct tests for both the I/B/E/S sample and the full sample of firms, where missing coverage is coded as zero analyst coverage (Barth et al. 2001). Because these results are similar, we report our main results based on the I/B/E/S sample.

Analyst Forecast Revision Response Time

We define the analyst forecast revision response time as the time from the 10-K filing to the first annual or quarterly earnings forecast issued by each *individual* analyst following the firm. To ensure that we include only analysts who actively follow the firm, we require that each analyst issue a forecast in the 90 days prior to the 10-K filing and then another report within 90 days of the 10-K filing (this process eliminates 10,458 observations related to analysts who stopped coverage). Because earnings announcements can prompt analysts to issue reports, we exclude reports made after any earnings announcement that occurs after the 10-K filing, but before the end of the 90-day window. We then define analyst report duration as the length of time in days between the 10-K filing and the first report following the filing. In addition, we average individual analyst report duration at both the individual firm and analyst level for each year. We interpret this duration variable as a measure of the required amount of effort for an individual analyst to read, understand, and process the information contained in the 10-K filing and to issue an updated earnings forecast. As mentioned above, we expect analyst response time to be longer for less readable 10-K filings.

Information Content of Analyst Reports

Similar to Frankel et al. (2006), we measure the information content of analyst reports as the proportion of a firm's stock returns related to analyst forecast revisions to the total stock return during the time period between the 10-K filing and the subsequent fiscal year-end. This measure is constructed as the sum of the one-day, absolute size-adjusted returns on the analyst forecast revision day, divided by the sum of the one-day, absolute size-adjusted returns over the entire window.[13] We exclude analyst reports that coincide with earnings announcements. We treat multiple reports issued on the same day as a single report. Also, an observation must have a minimum of 90 trading days with available data to be considered in our tests. Similar to Frankel et al. (2006), we interpret this measure as the percentage of total firm information provided to investors that is related to analyst reports.

Properties of Analyst Earnings Forecasts

We define analyst forecast dispersion (*DISPERSION*) as the standard deviation of the individual analyst forecasts in the first analyst consensus annual earnings forecast issued after the 10-K filing for the fiscal period following the 10-K filing, scaled by share price 90 days before the consensus forecast date. Analyst forecast accuracy (*ACCURACY*) is computed as the squared difference between I/B/E/S reported earnings and the analyst consensus forecast, scaled by share price 90 days before the consensus forecast date.[14] We define analyst overall uncertainty

[13] Frankel et al. (2006) divide this measure by the number of forecast revision dates to obtain a measure of the average informativeness of an analyst report date. We omit this final step for two reasons. First, our variable of interest is the overall firm information that comes from analysts, not the average information content. Second, firms with less readable disclosures are likely to obtain greater amounts of information from analysts because analysts are more likely to revise their forecasts more often for such firms. We confirm this supposition in untabulated results. Thus, dividing by the number of forecast revisions would induce a negative correlation between the analyst informativeness measure and 10-K readability. Our results are robust to the inclusion of the number of analysts following the firm as a control for the expected number of analyst revisions.

[14] Our inferences are not sensitive to scaling the measures of analyst forecast accuracy and dispersion by the absolute value of earnings or using the unscaled variables.

($UNCERTAINTY_{OVERALL}$) and common uncertainty ($UNCERTAINTY_{COMMON}$) using the following equations derived by Barron et al. (1998):[15]

$$UNCERTAINTY_{OVERALL} = \left(1 - \frac{1}{\#ANALYSTS}\right) * DISPERSION + ACCURACY \tag{2}$$

$$UNCERTAINTY_{COMMON} = \frac{ACCURACY - \frac{DISPERSION}{\#ANALYSTS}}{UNCERTAINTY_{OVERALL}}. \tag{3}$$

Following Barron et al. (1998), we interpret overall uncertainty ($UNCERTAINTY_{OVERALL}$) as the sum of the idiosyncratic uncertainty (i.e., uncertainty associated with analysts' private information) and common uncertainty (i.e., uncertainty associated with information common to all analysts). $UNCERTAINTY_{COMMON}$ is calculated as the proportion of common uncertainty to overall uncertainty. Thus, we interpret $UNCERTAINTY_{COMMON}$ to be a measure of how much the average analyst's belief reflects common versus private information. As shown in Equations (2) and (3), as *DISPERSION* approaches 0, $UNCERTAINTY_{OVERALL}$ approaches *ACCURACY* and $UNCERTAINTY_{COMMON}$ approaches 1. This suggests that, when there is no disagreement among analysts (i.e., *DISPERISON* equals 0), the total uncertainty is only associated with analysts' common information and all information impounded in analyst forecasts is public. Empirically, because the uncertainty measures emphasize information across analysts, we require that each firm have at least four analysts following it.

Control Variables

Our analysis controls for a variety of variables that have been shown by prior literature to be associated with firms' information environment and business complexity and therefore relate to analyst behavior. Prior work finds that firm size is the most important determinant of analyst following (Bhushan 1989; O'Brien and Bhushan 1990; Brennan and Hughes 1991; Lang and Lundholm 1996; Barth et al. 2001). These studies find that larger firms have greater analyst following and suggest that large firms have better information environments, potentially more complex operations, and greater demand for investment advice. We use the natural logarithm of market value as of the year ending prior to the 10-K filing (*LOGSIZE*) as a proxy for size.

Following Barth et al. (2001), we include controls for growth. High-growth firms tend to attract greater analyst following due to investor interest and the potential for future investment banking deals. Further, analysts likely find it more difficult to accurately forecast earnings for firms with high growth, leading to greater disagreement among analysts and less accurate forecasts. We define the variable *GROWTH* as the compounded average growth rate in sales over the prior three to five fiscal years. Our inferences are robust to the inclusion of the natural logarithm of the book-to-market ratio as a proxy for growth.

Similar to Bradshaw et al. (2008), we include the natural logarithm of the number of business segments reported in the Compustat Segment File as a control for the underlying complexity of the firm. We also include a control for the level of institutional holdings, following the evidence in Bhushan (1989), Brennan and Subrahmanyam (1995), and Frankel et al. (2006). These studies find that institutional ownership is positively associated with analyst following and with the informa-

[15] We thank the reviewer for suggesting that we incorporate the Barron et al. (1998) measures into our analysis. Note that the variables we term overall and common uncertainty are termed analyst uncertainty and consensus, respectively, in Barron et al. (1998). We use the term common uncertainty as opposed to consensus to avoid confusion regarding our use of analyst consensus forecasts.

tion content of their reports.[16] Institutional ownership can also be associated with higher analyst forecast accuracy and lower dispersion because firms with high levels of institutional holdings tend to have better information environments. We define the variable *PINST* as the percentage of a firm's shares that are held by institutions from the 13F disclosures for the most recent quarter prior to the 10-K filing. Motivated by the evidence in Lang and Lundholm (1996), who document that analyst following increases with the quality of disclosures, we include the number of management earnings forecasts made during the prior year (*MFCOUNT*) as a proxy for firm discretionary disclosure (Nagar et al. 2003; Cotter et al. 2006), as well as the absolute value of the cumulative, market-adjusted return for the two-day window around the 10-K filing (*10-K NEWS*) as a proxy for the disclosure informativeness.[17]

Similar to Barth et al. (2001), we include variables to control for firms' information environment and business complexity, as these variables likely affect the properties of analyst earnings forecasts. Barth et al. (2001) examine the association between analysts' incentives to follow firms and the extent of their intangible assets. Because many intangible assets are generally not recognized and estimates of their fair value are not disclosed, Barth et al. (2001) argue that analysts have increased incentives to follow firms with greater intangibles due to increased demand from investors. They find that analyst following is greater for firms with larger research and development and advertising expenses. They interpret these results as evidence that analysts respond to the demands of investors for more information because of the difficulty in evaluating firms' intangible assets. Other studies examine the effect of intangibles on the properties of analyst forecasts. Barron et al. (2002) find that analyst uncertainty increases with the level of a firm's intangible assets. Gu and Wang (2005) find that analyst forecast errors are increasing in firm intangible intensity. Intangibles are important for our tests because they could be associated with 10-K readability. For example, it is likely more difficult to explain the operations of firms with high levels of research and development costs in a less complex manner. Similar to Barth et al. (2001), we define *R&D* as the ratio of research and development expense to operating expense and *ADV* as the ratio of advertising expense to operating expense.[18]

Finally, we include the standard deviation of firm monthly stock returns from the prior year (*STD_RET*) as a measure of information uncertainty.[19] Bhushan (1989) suggests that private information is more valuable for firms with higher return volatility and thus positively related to the demand for analyst services. However, it is likely that analysts bear increased costs for following firms with higher return volatility. Additionally, most of our tests include industry fixed affects (based on the Fama and French classification) and year fixed effects to account for variation in analyst following across specific industries and over time. This approach is used to help control for variation in business complexity or information uncertainty that is driven by industry or time.

---

[16] O'Brien and Bhushan (1990) suggest that analyst following and institutional ownership are potentially endogenous. Frankel et al. (2006) assume that the variables are exogenous. Similar to Barth et al. (2001), we examine our results with and without the inclusion of institutional ownership and also with the inclusion of lagged institutional ownership. The coefficient on *FOG* remains positive and significant under these specifications.

[17] We thank a reviewer for pointing out the importance of controlling for the informativeness of the 10-K disclosure.

[18] Our results are robust to the inclusion of the amount of recognized intangibles and depreciation expense, which are also examined by Barth et al. (2001).

[19] While some of our other control variables (e.g., sales growth and R&D) are likely correlated with firms' information uncertainty, we include this measure as a more explicit, market-based measure of this construct. In addition, the inclusion of an alternative measure, earnings volatility, does not significantly affect our results. However, the inclusion of this measure reduces our sample by about 20 percent due to the need for a sufficient time-series of earnings.

IV. RESULTS

Summary Statistics

Table 2 presents summary statistics for the sample of 33,704 firm-year observations that remain after imposing the availability of data on CRSP and I/B/E/S. Similar to Table 1 and consistent with concerns raised by the SEC about disclosure complexity, the mean (median) *FOG* score is 19.53 (19.38). The standard deviation and interquartile range are 1.42 and 1.71, respectively, with an interquartile range from 18.58 to 20.29. The mean (median) number of analysts per firm-year observation is 6.14 (4.00). The mean (median) average analyst forecast revision response time is 17.71 (16.00) days at the firm-year level and 18.77 (17.50) days at the analyst-year level. This suggests that, on average, analysts do not respond immediately to 10-K filings due to the significant amount of information that must be processed and interpreted. The mean (median) analyst report information content (*AI*) is 0.130 (0.092). This indicates that, on average, almost 13 percent of the information reflected in stock returns during the period between the 10-K filing and the end of the fiscal period is derived from analyst reports. The mean (median) of forecast dispersion is 0.008 (0.003) and that of squared forecast error (*ACCURACY*) is 0.078 (0.002). The mean (median) of $UNCERTAINTY_{OVERALL}$ and $UNCERTAINTY_{COMMON}$ are 0.051 (0.005) and 0.328 (0.267), respectively. This suggests that about 33 percent of analyst uncertainty about future earnings following the 10-K filing is based on public information.

Table 2 also provides statistics on the control variables. The mean (median) size of our sample firms is $2.1 billion ($381 million), and mean (median) compound averages growth rate of sales (*GROWTH*) is 0.20 (0.11). The mean (median) number of business segment is 1.83 (1), and mean (median) percent of institutional ownership is 46 (45). The mean (median) number of management earnings forecasts is 1.44 (0) and the mean (median) *10-K NEWS* is 0.03 (0.02). The mean (median) ratio of research and development expense (*R&D*) and advertising (*ADV*) to operating expense are 0.08 (0.00) and 0.01 (0.00), respectively.[20] The mean (median) standard deviation of returns (*STD_RET*) is 0.14 (0.12).

Similar to Li (2008), we find in untabulated results that, while *FOG* is significantly correlated with many of the control variables, the extent of these correlations is relatively small for most variables. This is consistent with the conclusion that it is difficult to explain a significant proportion of the variation in *FOG* using firm characteristics. This evidence is important because in this study we assume that the Fog Index measures the readability of 10-K disclosures rather than their content. Consistent with this assertion, the correlation between *FOG* and the absolute value of the 10-K event return is small and insignificant.

As a caveat, we note that disclosure readability can be decomposed into innate and discretionary components. While we include a variety of variables to control for innate readability, analysts' information costs are based on total readability, so the relative magnitude of these components is less relevant. Our inferences are limited by the extent to which the Fog Index measures readability and is not confounded with other (uncontrolled) firm characteristics. It is also important to note that, for a specific variable to influence the readability measure, it must either increase the length of the average sentence in the 10-K document or increase the percentage of complex words. Simply increasing the length of the 10-K will not directly affect the measure. However, we examine the effect of disclosure length on our main results in sensitivity tests because length also can be a form of complexity (You and Zhang 2009; Loughran and McDonald 2010; Miller 2010).

[20] Note that the medians for the number of management earnings forecasts, research and development expense, and advertising expense are zero. This is primarily because we follow prior literature and code missing items to be zero.

TABLE 2

Sample Descriptive Statistics

Variable	n	Mean	Std. Dev.	Q1	Median	Q3
FOG	33,704	19.53	1.42	18.58	19.38	20.29
#ANALYSTS	33,704	6.14	5.91	2.00	4.00	8.00
$RESPONSE_{Firm}$	17,868	17.71	10.76	9.50	16.00	24.25
$RESPONSE_{Analyst}$	5,737	18.77	8.49	12.50	17.50	23.50
AI	30,716	0.130	0.139	0.037	0.092	0.177
DISPERSION	26,078	0.008	0.018	0.001	0.003	0.008
ACCURACY	29,055	0.078	0.447	0.000	0.002	0.015
$UNCERTAINTY_{OVERALL}$	17,241	0.051	0.280	0.002	0.005	0.016
$UNCERTAINTY_{COMMON}$	17,222	0.328	0.401	−0.038	0.267	0.702
SIZE	33,704	2100	5600	122	381	1300
GROWTH	33,704	0.20	0.34	0.04	0.11	0.24
SEGMENTS	33,704	1.83	1.32	1.00	1.00	3.00
PINST	33,704	45.98	26.78	22.95	45.31	67.60
MFCOUNT	33,704	1.44	2.69	0.00	0.00	2.00
10-K NEWS	33,704	0.03	0.05	0.01	0.02	0.04
ADV	33,704	0.01	0.03	0.00	0.00	0.01
R&D	33,704	0.08	0.17	0.00	0.00	0.08
STD_RET	33,704	0.14	0.10	0.07	0.12	0.18

This table reports descriptive statistics for observations that are available on the SEC's EDGAR, Compustat, CRSP, and I/B/E/S databases. The variables are pooled across fiscal years 1995–2006.

Variable Definitions:

FOG = Fog Index of the 10-K filing calculated as (average words per sentence + percent of complex words) × 0.4;
#ANALYSTS = number of analysts in the first consensus annual earnings forecast following the 10-K filing;
$RESPONSE_{Firm}$ = average number of days that it takes a firm's analysts to issue their first report following the 10-K filing;
$RESPONSE_{Analyst}$ = average number of days that it takes an analyst to issue his/her first report following the 10-K filing;
AI = sum of the one-day, absolute size-adjusted returns from analyst reports between the 10-K filing and the next fiscal year-end divided by the sum of the one-day, absolute size-adjusted returns over the entire window;
DISPERSION = standard deviation of the individual analyst forecasts in the first analyst consensus annual earnings forecast issued after the 10-K filing for the fiscal period following the 10-K filing, scaled by share price 90 days before the consensus forecast date;
ACCURACY = squared difference between I/B/E/S reported earnings and the first analyst consensus annual earnings forecast issued after the 10-K filing for the fiscal period following the 10-K filing, scaled by share price 90 days before the consensus forecast date;
$UNCERTAINTY_{OVERALL}$ = sum of common and idiosyncratic uncertainty among analysts computed following Barron et al. (1998, equation 15), using the measures of accuracy, dispersion, and analyst following as previously defined in this table;
$UNCERTAINTY_{COMMON}$ = ratio of common uncertainty to total uncertainty among analysts computed following Barron et al. (1998, equation 16) using the measures of accuracy, dispersion, and analyst following as previously defined in this table;
SIZE = number of shares outstanding (Compustat item CSHO) times the share prices at the most recent fiscal year-end (Compustat item PRCC_F);
GROWTH = compound average growth rate of firm sales (Compustat item SALE) over the prior 3–5 years $(sales_{t-1} / sales_{t-(1+i)})^{1/i}$;
SEGMENTS = number of reported business segments in the Compustat segment file for prior fiscal year;
PINST = percentage of institutional ownership from the quarter prior to the 10-K filing;

(continued on next page)

TABLE 2 (continued)

MFCOUNT = number of management earnings forecasts issued in the prior year;
10-K NEWS = absolute value of the cumulative market-adjusted return for the 10-K filing event window [0,1];
ADV = advertising expense (Compustat item XAD) as a percentage of operating expense (Compustat item XOPR) from the prior fiscal year;
R&D = research and development expense (Compustat item XRD) as a percentage of operating expense (Compustat item XOPR) from the prior fiscal year; and
STD_RET = standard deviation of the firm's monthly stock returns from the previous fiscal year.

Analyst Following and the Readability of 10-K Filings

Our first prediction is that analyst following is affected by the level of readability of firms' 10-K filings, as measured by the Fog Index. To control for other factors that can affect analyst following, we estimate the following regression:

$$\#ANALYSTS_{i,t} = \beta_0 + \beta_1 FOG_{i,t} + \beta_2 LOGSIZE_{i,t} + \beta_3 GROWTH_{i,t} + \beta_4 LOGSEGMENTS_{i,t} + \beta_5 PINST_{i,t} + \beta_6 MFCOUNT_{i,t} + \beta_7 10\text{-}K\ NEWS_{i,t} + \beta_8 ADV_{i,t} + \beta_9 R\&D_{i,t} + \beta_{10} STD\_RET_{i,t} + \varepsilon_{i,t}. \quad (4)$$

The estimation is performed using ordinary least-squares regression with industry and time fixed effects.

Column 1 of Table 3 (model 1) reports the results of the linear model regression. t-statistics, presented in brackets, are based on standards errors that are robust to heteroscedasticity and are clustered at the firm level. The coefficient on *FOG* is positive and statistically significant, suggesting that analyst following is greater for firms with less readable disclosures. This is consistent with the notion that analysts respond to investors' demand for investment information for firms whose disclosures are more costly to process. The coefficients on the control variables are consistent with prior research. Larger firms are associated with greater analyst following as well as firms with higher growth rates and higher institutional ownership. Consistent with Lang and Lundholm (1996), we find that disclosure practice (*MFCOUNT*) and disclosure informativeness (*10-K NEWS*), are positively related to analyst following. We find that analyst following is negatively associated with the number of business segments.[21] Similar to Barth et al. (2001), we document that analyst following is greater for firms with higher amounts of advertising and research and development expenses. Consistent with Bhushan (1989), we also find that analyst following is positively associated with firm stock return volatility, suggesting that analysts provide greater support to investors when private information is valuable.

In addition to examining the linear association between *FOG* and analyst following, we also examine the association of analyst following and *FOG* using nonlinear and semi-parametric specifications (second and third columns of Table 3, respectively). In the nonlinear specification (model 2), we add a square term (FOG^2) to Equation (2) to capture the possibility that the association between readability and analyst following differs based on the level of *FOG*. In model 3, we examine a semi-parametric specification by including indicator variables for the various quartiles of *FOG*. The coefficient on FOG^2 (model 2) is negative and significant, indicating that the role of

[21] This effect becomes negative after controlling for firm size, but is positive if size is omitted. The negative coefficient in the multivariate results is consistent with the findings of Bhushan (1989) and Johnston et al. (2009).

TABLE 3
The Association between Annual Report Readability and Analyst Following

Panel A: OLS Regression

Variable	**Model 1**	**Model 2**	**Model 3**
Intercept	−13.067***	−20.190***	−11.625***
	[−16.79]	[−6.29]	[−18.23]
FOG	0.085***	0.798***	
	[3.74]	[2.58]	
FOG^2		−0.018**	
		[−2.33]	
FOG Q2			0.168**
			[2.41]
FOG Q3			0.331***
			[4.11]
FOG Q4			0.369***
			[4.18]
LOGSIZE	2.623***	2.624***	2.623***
	[62.20]	[62.23]	[62.22]
GROWTH	0.361***	0.356***	0.356***
	[3.90]	[3.85]	[3.85]
LOGSEGMENTS	−0.367***	−0.368***	−0.368***
	[−4.68]	[−4.70]	[−4.70]
PINST	0.011***	0.011***	0.011***
	[5.20]	[5.20]	[5.22]
MFCOUNT	0.086***	0.086***	0.086***
	[5.82]	[5.82]	[5.83]
10-K NEWS	2.391***	2.380***	2.371***
	[5.46]	[5.43]	[5.40]
ADV	7.555***	7.564***	7.553***
	[4.57]	[4.57]	[4.57]
R&D	1.696***	1.673***	1.662***
	[5.61]	[5.54]	[5.50]
STD_RET	0.932***	0.924***	0.931***
	[2.61]	[2.59]	[2.61]
Time Fixed Effects	Yes	Yes	Yes
Industry Fixed Effects	Yes	Yes	Yes
n	33704	33704	33704
Adjusted R^2	0.63	0.63	0.63

Panel B: Comparative Interquartile Effects

	Model 1	**Model 2**	**Model 3**
FOG	0.146	0.188	0.201
LOGSIZE	6.236	6.238	6.238
GROWTH	0.071	0.070	0.070
LOGSEGMENTS	−0.403	−0.404	−0.405
PINST	0.473	0.473	0.478
MFCOUNT	0.172	0.172	0.172
10-K NEWS	0.081	0.080	0.080
ADV	0.042	0.042	0.042

(continued on next page)

Panel B: Comparative Interquartile Effects

	Model 1	Model 2	Model 3
RD	0.130	0.129	0.128
STD_RET	0.098	0.097	0.174

**, *** $p < 0.05$ and $p < 0.01$, respectively, two-tailed t-tests.

Panel A reports coefficient estimates and t-statistics (in brackets) from the regression of analyst following on *FOG* and control variables. Analyst following is defined as the number of analysts contained in the first I/B/E/S consensus annual earnings forecast following the 10-K filing. *FOG QX* is an indicator variable set to 1 if the firm's 10-K Fog Index is in quartile *X* of the sample, and 0 otherwise. All other variables are as defined in Table 2. Industry fixed effects are based on the Fama and French 48-industry classification. t-statistics are robust to heteroscedasticity and clustered at the firm level. Panel B reports the estimated effect of an interquartile change in the variable of interest.

readability increases at a declining rate. Also, the coefficients on the second, third, and fourth *FOG* quartiles (model 3) are positive and significant and monotonically decreasing across the quartiles. These results are consistent with analysts making a trade-off between the benefits of covering firms with less readable disclosures and the costs of following them. While these costs include the processing costs of covering firms with less readable disclosures, analysts likely find it more difficult to accurately forecast the earnings of firms with more complex disclosures. These costs can be important, as prior research has found that forecast accuracy has important career effects for analysts (Mikhail et al. 1999; Hong and Kubik 2003). The effects of *FOG* on the properties of analyst earnings forecasts will be examined later.[22]

Table 3, Panel B compares the effects of an interquartile change in the independent variables on analyst following based on the marginal effect estimates in Panel A. Consistent with prior literature, size (*LOGSIZE*) has the largest association with analyst following. An interquartile change in the natural logarithm of size is associated with an increase of 6.2 analysts, while that of all other variables is less than 1. The interquartile effect of *FOG* is 0.15 in model 1, 0.19 in model 2, and 0.20 in model 3. While the incremental effect of the complexity of 10-K filings on analyst following is moderate, it is greater than that of most of the other variables included in the regressions. These are the same variables that have been shown in the literature to be related to analyst following. Based on these comparisons, we conclude that disclosure readability is an important determinant of analyst following.

Similar to Lang and Lundholm (1996), we also examine the association between *changes* in *FOG* and *lead changes* in analyst following. If analysts do, in fact, respond to changes in *FOG* by increasing their coverage, then we expect to observe a positive relation between current changes in *FOG* and future changes in analysts following. We investigate this relation by estimating multivariate models using the changes in the independent variables from Table 3. We define the current change in the Fog Index and other independent variables as the difference between their current values and those of the prior fiscal year. We define the lead change in analyst following as the difference between the number of analysts following the firm after the next 10-K filing and the number of analysts following the firm after the current report. Table 4 presents results from these tests. We find that contemporaneous changes in *FOG* are positively related to lead changes in analyst following. We also find that changes in firm size and institutional ownership are positively

[22] In a related study, You and Zhang (2009) examine stock returns following 10-K filings and find that investors underreact to firms with more complex annual filings. This apparent underreaction provides additional motivation for analyst coverage because prior work has found that analysts help improve market efficiency (Barron et al. 2002; Brown and Sivakumar 2003; Gu and Chen 2004).

TABLE 4

The Association between Changes in Annual Report Readability and Lead Changes in Analyst Following

Variable	Model 1	Model 2	Model 3
Intercept	−0.038***	−0.051***	−0.073***
	[−3.79]	[−4.68]	[−3.03]
ΔFOG	0.024***	0.023**	
	[2.66]	[2.55]	
$(\Delta FOG)^2$		0.007**	
		[2.58]	
ΔFOG $Q2$			−0.009
			[−0.28]
ΔFOG $Q3$			0.045
			[1.36]
ΔFOG $Q4$			0.110***
			[3.20]
$\Delta LOGSIZE$	0.480***	0.479***	0.480***
	[21.89]	[21.83]	[21.91]
$\Delta GROWTH$	0.048	0.045	0.050
	[0.66]	[0.62]	[0.68]
$\Delta LOGSEGMENTS$	−0.103**	−0.103**	−0.102**
	[−2.12]	[−2.11]	[−2.11]
$\Delta PINST$	0.008***	0.008***	0.008***
	[5.89]	[5.82]	[5.84]
$\Delta MFCOUNT$	−0.015	−0.015	−0.014
	[−1.61]	[−1.61]	[−1.59]
Δ*10-K NEWS*	−0.138	−0.138	−0.136
	[−0.91]	[−0.91]	[−0.90]
ΔADV	0.333	0.327	0.323
	[0.50]	[0.49]	[0.48]
Δ*R&D*	−0.075	−0.075	−0.074
	[−0.57]	[−0.57]	[−0.56]
ASTD_RET	−0.637***	−0.631***	−0.631***
	[−5.18]	[−5.14]	[−5.14]
n	27226	27226	27226
Adjusted R^2	0.03	0.03	0.03

**, *** $p < 0.05$ and $p < 0.01$, respectively, two-tailed t-tests.

This table reports coefficient estimates and t-statistics (in brackets) from the regression of lead changes in analyst following on changes in *FOG* and control variables. Lead change in analyst following is the difference between the number of analysts following the firm after the next 10-K filing and the number of analysts following the firm after the current report. Δ*FOG QX* is an indicator variable set to 1 if the current change in a firm's 10-K Fog Index is in quartile *X* of the sample, and 0 otherwise. All other variables are current period changes of the variables defined in Table 2. t-statistics are robust to heteroscedasticity and clustered at the firm level.

related to lead changes in analyst following and that changes in management earnings guidance, number of business segments, and return volatility are negatively related.[23] In untabulated results, we find that *FOG* has the third-highest effect of the variables employed in our analysis. The results from this test are important because they help to alleviate concerns about endogeneity or prior period information shocks relating to the levels results in Table 3.

Analyst Report Duration

One potential measure of the costs or effort that analysts bear in following firms with less readable disclosures is the amount of time it takes them to issue reports following the 10-K filing. We label this length of time "analyst report duration." In this section, we examine the analyst report duration at both the analyst and firm level. We expect that analyst report duration will be longer for firms with less readable 10-K filings. As noted previously, to ensure that we capture analyst reports issued in response to the filing of the 10-K (and not in response to other corporate events), we include in this analysis only firm-analyst observations for which the analyst has issued at least one report during the 90 days prior to the 10-K filing and issues a report in the 90 days subsequent to the 10-K filing.[24] Table 5, Panel A presents univariate results based on individual analyst reports and the Fog Index of the firms they cover. We classify each firm-specific *FOG* as high (low) if it is greater (less) than the median *FOG* value for all firms in the sample. We find that analysts covering firms with a high *FOG* require 1.45 days longer, on average, to issue their reports. Panels B and C of Table 5 present the average analyst report duration per firm-year and analyst-year, respectively. The value of *FOG* at the firm level is the individual firm 10-K Fog for each specific year. The *FOG* score at the analyst level is the average 10-K Fog score of the firms that the analyst covered in each year.[25] The univariate results suggest that the average (median) analyst report duration for firms with *FOG* scores higher than the median is 1.11 (1.50) days (Panel B). At the analyst level, we find that analysts who cover a portfolio of firms with an average *FOG* score higher than the median analyst portfolio average take 2.29 (2.17) days longer on average (median) to issue their first reports. These differences are statistically significant at a 1 percent level and are likely to be economically important given the rapid pace at which markets impound new information.

To further examine the analyst report duration, we estimate ordinary least-squares regressions on the average analyst report durations at both the firm and analyst level. Table 5, Panel D presents the results of these analyses. Similar to the univariate results, we find that firms with higher Fog scores and analysts who cover stocks with higher Fog scores are associated with longer analyst report duration. Specifically, controlling for a variety of other factors that can affect firms' information environment, the coefficient on *FOG* is 0.345 for the firm-level analysis and 1.138 for the analyst-level analysis, both of which are statistically significant. This suggests that firms and analyst portfolios with one unit higher of *FOG* have analyst response times that are 0.345 and 1.138 days longer. Although it is difficult to assess the economic significance of these results, we note that given the pace with which new analyst information is incorporated in stock prices, even such short delays in producing a report could reduce the information content of the report. The results in Table 5 further indicate that firms with greater institutional ownership, discretionary

[23] While less intuitive, one potential concern is that analyst following might lead *FOG*. For example, management might provide more complex information in response to a greater analyst following. In untabulated tests we find no relation between changes in *FOG* and lag analyst following.

[24] As noted previously, we exclude analyst observations made after earnings announcements that follow the 10-K filing and reports made after 90 days after the 10-K. Our results are similar if we use reports within a year after the 10-K filing; however, these observations are likely unrelated to the *FOG* score of the 10-K filing.

[25] To be included in our analysis each analyst must cover a minimum of four firms. Similar results are obtained by requiring only a single firm.

TABLE 5

The Association between Annual Report Readability and Analyst Report Duration

Panel A: Individual Analyst Reports

Group	n	Mean	Median
High Fog Index	30,892	19.19	16.00
Low Fog Index	30,884	17.74	15.00
Difference		1.45***	1.00***

Panel B: Average Analyst Report Duration Per Firm

Group	n	Mean	Median
High Fog Index	8,701	18.33	16.75
Low Fog Index	8,701	17.22	15.25
Difference		1.11***	1.50***

Panel C: Average Analyst Report Duration Per Analyst

Group	n	Mean	Median
High Fog Index	2,869	19.91	18.67
Low Fog Index	2,868	17.62	16.50
Difference		2.29***	2.17***

Panel D: Regression Analysis

	Mean Duration at Firm Level	Mean Duration at Analyst Level
Intercept	9.430†††	−8.750†††
	[7.62]	[−2.62]
FOG	0.345†††	1.138†††
	[5.79]	[6.96]
LOGSIZE	0.034	−0.013
	[0.51]	[−0.11]
GROWTH	−0.453	2.634†††
	[−1.62]	[3.93]
LOGSEGMENTS	0.405††	0.844†††
	[2.57]	[2.67]
PINST	0.039†††	0.105†††
	[9.22]	[11.50]
MFCOUNT	0.170†††	0.019
	[5.68]	[0.33]
10-K NEWS	−25.409†††	−43.743†††
	[−12.33]	[−8.01]
ADV	10.053††	30.936†††
	[2.52]	[4.06]
R&D	1.925†††	0.544
	[3.06]	[0.56]
STD_RET	−6.300†††	−13.345†††
	[−5.53]	[−5.86]
n	17868	5737
Adjusted R^2	0.04	0.11

(continued on next page)

*** $p < 0.01$, two-tailed t-tests for differences of means and Wilcoxon for medians.

††† $p < 0.01$, two-tailed t-tests.

Analyst report duration is defined as the length of time between the 10-K filing and the first report for each analyst. Panels A, B, and C report descriptive statistics of the differences in the average analyst report duration at the individual analyst report, firm-year, and analyst-year levels. The firm and analyst levels are computed by averaging variables over the unit of observation. The High and Low groups are based on the median of the sample in each group. Statistical differences of means (medians) are computed based on t-tests (Wilcoxon tests).

Panel D reports coefficient estimates based on ordinary least square regressions at both the firm and analyst levels. All variables are defined as in Table 2. On the analyst level, the variables are averaged over the portfolio of each analyst-year observation. t-statistics (in brackets) are robust to heteroscedasticity and clustered at the firm and analyst levels, respectively.

disclosure, and firm intangibles are associated with longer average analyst response times. In contrast, firms with greater reaction to the release of the 10-K report and greater prior return volatility are associated with shorter analyst report response time.[26] Overall, the evidence is consistent with the notion that analysts exert more effort to cover firms with less readable disclosures.

Information Content of Analyst Reports

Our third hypothesis predicts that the informativeness of analyst reports is positively related to the readability of firms' 10-K reports. To test this hypothesis, we estimate the following ordinary least-squares regression with industry and time fixed effects:

$$AI_{i,t} = \beta_0 + \beta_1 FOG_{i,t} + \beta_2 LOGSIZE_{i,t} + \beta_3 GROWTH_{i,t} + \beta_4 LOGSEGMENTS_{i,t} + \beta_5 PINST_{i,t} + \beta_6 MFCOUNT_{i,t} + \beta_7 10\text{-}K\ NEWS_{i,t} + \beta_8 ADV_{i,t} + \beta_9 R\&D_{i,t} + \beta_{10} STD\_RET_{i,t} + \varepsilon_{i,t}. \quad (5)$$

Similar to the regression reported in Table 3, we also examine the association of the information content of analyst reports and *FOG* using nonlinear and semi-parametric specifications.

Table 6 reports the regression findings. t-statistics, presented in brackets, are based on standard errors that are robust to heteroscedasticity and clustered at the firm level. Consistent with our hypothesis, the coefficient on *FOG* is positive and significant, suggesting that the informativeness of analyst reports is increasing in the complexity of the 10-K disclosure. This evidence is consistent with the notion that investors find analyst reports for firms with less readable disclosures more useful because of greater processing costs. Analysts' private information searches could also be more valuable in such cases. In model 2, the square term (FOG^2) is negative but insignificant, whereas the coefficients on the individual *FOG* quartiles in model 3 are significant and monotonically increasing. These findings are consistent with our prior results and support the notion that analysts trade off costs and benefits in their coverage of firms with less readable disclosure. The effects of the control variables on the information content of analyst reports are similar in direction to their effects on analyst following. Specifically, the informativeness of analyst reports is increasing in firm size, growth, institutional ownership, discretionary disclosure, firm intangibles, and return volatility. In untabulated results, we find that an interquartile change in *FOG* increases the

[26] In untabulated results, we include a control for the number of reports issued by an analyst for the analyst level results. The coefficient on this variable is not statistically significant and does not quantitatively affect the results as presented. We also estimate two proportional hazard models to examine the influence of readability on the time to an analyst's first report by conditioning at the firm and analyst level, similar to O'Brien et al. (2005). We find that, on average, the probability of an analyst making a report at any given time is between 1.7 percent and 3.2 percent lower for each unit change of *FOG*.

TABLE 6

The Association between Annual Report Readability and the Information Content of Analysts' Reports

Variable	Model 1	Model 2	Model 3
Intercept	−0.256***	−0.302***	−0.227***
	[−16.90]	[−4.09]	[−20.43]
FOG	0.002***	0.006	
	[3.25]	[0.86]	
FOG^2		−0.000	
		[−0.62]	
FOG Q2			0.003**
			[1.99]
FOG Q3			0.006***
			[3.33]
FOG Q4			0.007***
			[3.51]
LOGSIZE	0.049***	0.049***	0.049***
	[65.99]	[66.03]	[66.02]
GROWTH	0.012***	0.011***	0.011***
	[5.54]	[5.52]	[5.50]
LOGSEGMENTS	−0.011***	−0.011***	−0.011***
	[−6.47]	[−6.47]	[−6.48]
PINST	0.000***	0.000***	0.000***
	[10.09]	[10.09]	[10.12]
MFCOUNT	0.003***	0.003***	0.003***
	[8.70]	[8.70]	[8.72]
10-K NEWS	0.049***	0.049***	0.049***
	[4.34]	[4.33]	[4.30]
ADV	0.089***	0.090***	0.089***
	[2.65]	[2.65]	[2.64]
R&D	0.006	0.005	0.005
	[0.92]	[0.89]	[0.83]
STD_RET	0.036***	0.036***	0.036***
	[4.41]	[4.41]	[4.42]
Time Fixed Effects	Yes	Yes	Yes
Industry Fixed Effects	Yes	Yes	Yes
n	30716	30716	30716
Adjusted R^2	0.51	0.51	0.51

**, *** $p < 0.05$, and $p < 0.01$, respectively, two-tailed t-tests.

This table reports coefficient estimates and t-statistics (in brackets) from the regression of the information content of analyst reports on *FOG* and control variables. Analysts information content is defined as the proportion of firm stock returns related to analyst forecast revisions to the total firm stock return during the time period between the 10-K filing and the subsequent fiscal year-end. *FOG QX* is an indicator variable set to 1 if the firm's 10-K Fog Index is in quartile *X* of the sample, and 0 otherwise. All other variables are as defined in Table 2. Industry fixed effects are based on the Fama and French 48-industry classification. t-statistics are robust to heteroscedasticity and clustered at the firm level.

information content of analyst reports by about 0.29 percent to 0.46 percent, on average, across the three models. Consistent with our finding for analyst following, we find that the effect of an interquartile change in disclosure readability (*FOG*) is similar to many of the other variables in our analysis.[27]

Properties of Analyst Earnings Forecasts

Given our findings regarding the association of disclosure readability with analyst following, analyst report duration, and the information content of analyst reports, a reasonable next step is to examine its association with properties of analyst forecasts. In particular, we focus on the dispersion, accuracy, and the overall and common uncertainty associated with analyst forecasts of annual earnings. As previously explained, we hypothesize that less readable disclosures will be associated with greater forecast dispersion and lower analyst forecast accuracy. We also predict that overall analyst uncertainty will be higher for firms with less readable disclosures, but make no directional prediction regarding analyst common uncertainty.

For each of these hypotheses, we estimate a model of the following form:

$$\begin{aligned}\text{Forecast Property}_{i,t} = {} & \beta_0 + \beta_1 FOG_{i,t} + \beta_2 LOGSIZE_{i,t} + \beta_3 GROWTH_{i,t} + \beta_4 LOGSEGMENTS_{i,t} \\ & + \beta_5 PINST_{i,t} + \beta_6 MFCOUNT_{i,t} + \beta_7 10\text{-}K\ NEWS_{i,t} + \beta_8 ADV_{i,t} \\ & + \beta_9 R\&D_{i,t} + \beta_{10} STD\_RET_{i,t} + \varepsilon_{i,t},\end{aligned} \tag{6}$$

where "Forecast Property" represents the previously defined variable of interest (*DISPERSION*, *ACCURACY*, $UNCERTAINTY_{OVERALL}$, or $UNCERTAINTY_{COMMON}$). Each model is estimated using ordinary least-squares regression with industry and time fixed effects. Table 7 presents the results. t-statistics, presented in brackets, are based on standards errors that are robust to heteroscedasticity and clustered at the firm level.

As reported in the first two columns of Table 7, we find that the coefficient on *FOG* is positive and significant, indicating that less readable disclosures are associated with more dispersed and less accurate analyst earnings forecasts. Consistent with prior research, these results, in conjunction with our findings about the information content of analysts' reports, suggest that analysts provide information that investors find valuable at the cost of less accurate forecasts (Schipper 1991; Clement and Tse 2003). It is, however, difficult to assess the significance of these benefits because we cannot observe the appropriate benchmark—i.e., the effect on investors of analysts choosing not to follow firms with less readable disclosures.[28]

We further examine these results by testing the effect of 10-K readability on measures of overall and common analyst forecast uncertainty, as derived by Barron et al. (1998). We report our results in columns three and four of Table 7. Consistent with our prediction, we find that overall analyst uncertainty is increasing in *FOG*, suggesting that there is higher overall uncertainty in the analyst information environment for firms with less readable 10-K filings. Further, we find that the proportion of common analyst uncertainty to overall uncertainty is also increasing in *FOG*, suggesting that as readability decreases, publicly available information, such as the 10-K, becomes

[27] As an additional sensitivity test, we follow the suggestion of a reviewer and examine the effect of analysts' conflicts of interest associated with the Global Research Analyst Settlement to support our inferences. Because analysts with conflicts of interest are likely to place less importance on firm disclosures, we predict that our results for analyst following and report information content should be weaker when analysts' conflicts of interest are higher. We note, however, that other important events such as Regulation Fair Disclosure occurred around this time and could limit our inferences from this test. In untabulated results, we find that, consistent with our prediction, the association between analyst following and analyst report information content is stronger in the years following the Global Research Analyst Settlement, suggesting greater analyst reliance and attention to 10-K filings in the post period. This test further supports our overall conclusion regarding the association between analyst behavior and 10-K readability.

[28] Related to this point, Lee (2010) finds that less readable 10-Q filings are associated with lower information efficiency and greater information asymmetry. She also finds that these negative consequences of lower readability are mitigated by analyst coverage, which provides evidence of the benefits of analysts following firms with less readable reports.

TABLE 7

The Association between Annual Report Readability and Properties of Analysts' Information Environment

Variable	*DISPERSION*	*ACCURACY*	$UNCERTAINTY_{OVERALL}$	$UNCERTAINTY_{COMMON}$
Intercept	0.0192***	0.2887	−0.0142	−0.0440
	[4.04]	[1.35]	[−0.37]	[−0.56]
FOG	0.0003***	0.0041**	0.0044***	0.0053**
	[3.96]	[2.10]	[2.58]	[2.23]
LOGSIZE	−0.0027***	−0.0315***	−0.0133***	−0.0012
	[−22.00]	[−9.89]	[−7.19]	[−0.43]
GROWTH	0.0008	0.0077	−0.0005	0.0638***
	[1.08]	[0.45]	[−0.04]	[5.69]
LOGSEGMENTS	0.0008***	0.0134**	0.0025	0.0081
	[3.74]	[1.96]	[0.54]	[1.24]
PINST	−0.0001***	0.0001	−0.0001	0.0010***
	[−8.45]	[0.49]	[−0.77]	[5.48]
MFCOUNT	−0.0001***	0.0010	−0.0001	0.0034***
	[−2.92]	[1.10]	[−0.11]	[2.79]
10-K NEWS	0.0511***	0.7263***	0.5591***	0.3838***
	[7.48]	[5.80]	[4.75]	[4.37]
ADV	0.0266***	0.3901**	0.0933	0.0255
	[4.13]	[2.13]	[1.03]	[0.18]
R&D	0.0081***	−0.0266	−0.0036	−0.1394***
	[5.97]	[−0.87]	[−0.14]	[−4.39]
STD_RET	0.0242***	0.3793***	0.2209***	0.3683***
	[9.27]	[6.03]	[4.36]	[6.72]
Time Fixed Effects	Yes	Yes	Yes	Yes
Industry Fixed Effects	Yes	Yes	Yes	Yes
n	26078	29055	17241	17222
Adjusted R^2	0.20	0.05	0.04	0.06

**, *** $p < 0.05$, and $p < 0.01$, respectively, two-tailed t-tests.

This table reports coefficient estimates and t-statistics (in brackets) from regressions of analyst earnings forecast dispersion, accuracy, and uncertainty on *FOG* and control variables. Analyst earnings forecast dispersion is defined as the standard deviation of the individual analyst earnings forecasts in the most recent consensus earnings forecast following the annual report scaled by price. Analyst earnings forecast accuracy is defined as the squared value of the difference between the I/B/E/S actual reported earnings and the most recent analyst consensus earnings forecast following the annual report filing scaled by price. Analyst overall and common uncertainty are calculated following the equations derived by Barron et al. (1998). All other variables are as defined in Table 2. Industry fixed effects are based on the Fama and French 48-industry classification. t-statistics are robust to heteroscedasticity and clustered at the firm level.

more important to analysts relative to private information. This suggests that analysts have difficulty producing private information for firms with less readable reports. Overall, the evidence from the four regressions in Table 7 suggests that analyst earnings forecasts are affected by 10-K readability, supporting the notion that analysts use the information contained in the 10-K filings.

The results for the control variables are largely consistent with the notion that analysts perform better for firms with better information environments and less complex operations. We find that firms with better information environments, as measured by size and institutional ownership, are associated with greater accuracy, lower dispersion, and lower uncertainty. Institutional ownership is also positively associated with analyst common uncertainty, consistent with these investors improving the importance of public information. We find that firms with greater uncertainty and complexity, as measured by the volatility of returns, the number of business segments, and the extent of intangible investment, are generally associated with lower forecast accuracy, greater dispersion, and greater overall uncertainty. However, their relation to analyst common uncertainty is less clear. While *R&D* investment is negatively related to common uncertainty (because less public information on these investments is available), return volatility is positively related. These findings suggest that private information is less important relative to common information for firms with high return volatility. While the measures of firm disclosure, *MFCOUNT* and *10-K NEWS*, are both positively associated with analyst common uncertainty (i.e., public disclosures increase the importance of common information), they differ with respect to analyst forecast dispersion. *MFCOUNT* is associated with lower forecast dispersion, consistent with managers providing guidance to reduce analyst disagreement, while *10-K NEWS* is associated with greater dispersion and lower accuracy.

V. SUMMARY AND CONCLUSION

From the passage of the Securities Acts in 1933 and 1934 to the present, regulators, legal scholars, and various other parties have weighed in on the debate about the complexity of financial communication to external users. The SEC has gone so far as to require that the prospectuses of all registered public offerings meet the requirements of the Plain English Rules, and has suggested that similar procedures should be applied to other mandatory filings. Former SEC Chairman Christopher Cox has even suggested the use of readability models such as the Fog Index to measure the complexity of financial communication (Cox 2007). However, other users are concerned about the effects that these types of actions could have on the disclosure regime. These parties argue that, at a minimum, these actions are a waste of time and effort, and could lead to a reduction in disclosure as firms are forced to simplify their public communication (e.g., Kripke 1970, 1973; Firtel 1999). This study sheds insight relevant to this debate by examining the effect of readability on one important financial information intermediary—sell-side financial analysts.

We find evidence consistent with the notion that, because less readable firm communication is more costly to process and interpret, investors demand greater amounts of analyst services for firms with less readable communication. We find that analyst following is greater for firms with higher levels of syntactic complexity as measured by the Fog Index. We also find that analysts who cover firms with less readable communication take longer time on average to issue reports in response to 10-K filings. We interpret this evidence as analysts exerting greater effort to cover these firms. In addition, we find that analyst reports of firms with less readable 10-K reports are more informative to investors, but that the earnings forecasts of such firms have greater analyst dispersion or disagreement, are less accurate, and are associated with greater levels of uncertainty.

Our results suggest that analyst behavior is related to the readability of firms' communication. While prior studies have found that analysts are affected by the complexity of individual financial items, we provide evidence that the *overall* linguistic complexity of firms' communication incrementally influences analyst behavior over and above the effects of the content of the document (e.g., taxes, interest rates). Finally, our results that analysts provide greater amounts of information to investors for firms with less readable communication and that investors consider this information informative are relevant to the SEC's debate about the intended audience of financial information and the SEC's concerns on the accessibility of these reports. While the SEC moves to reduce the complexity of firms' communication, further research is needed to examine the explicit costs and benefits of such actions. As a final caveat, while the collection of the evidence from our various analyses of analysts' behavior conducted in this study is consistent with the prediction that 10-K readability influences analyst behavior, our individual findings still are subject to the limitation that the documented relations reflect associations and may not be fully causal.

REFERENCES

Armstrong, J. S. 1980. Unintelligible management research and academic prestige. *Interfaces* 10 (2): 80–86.

Ball, R. 1992. The earnings price anomaly. *Journal of Accounting and Economics* 15 (2–3): 319–345.

Barron, O. E., D. Byard, O. Kim, and E. J. Riedl. 2002. High-technology intangibles and analysts' forecasts. *Journal of Accounting Research* 40 (2): 289–312.

———, ———, O. Kim, and D. E. Stevens. 1998. Using analysts' forecasts to measure properties of analysts' information environment. *The Accounting Review* 73 (4): 421–433.

Barth, M. E., R. Kasznik, and M. F. McNichols. 2001. Analyst coverage and intangible assets. *Journal of Accounting Research* 39 (1): 1–34.

Bhushan, R. 1989. Firm characteristics and analyst following. *Journal of Accounting and Economics* 11 (2–3): 255–274.

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

Bloomfield, R. J. 2002. The "incomplete revelation hypothesis" and financial reporting. *Accounting Horizons* 16 (3): 233–243.

Botosan, C. A., and M. 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 (2): 329–353.

Bradshaw, M. T., G. S. Miller, and G. Serafeim. 2008. Accounting method heterogeneity and analysts' forecasts. Working paper, Harvard University and University of Michigan.

Brennan, M. J., and P. J. Hughes. 1991. Stock prices and the supply of information. *The Journal of Finance* 46 (5): 1665–1691.

———, and A. Subrahmanyam. 1995. Investment analysis and price formation in securities markets. *Journal of Financial Economics* 38 (3): 361–381.

Brown, L. D., and S. Sivakumar. 2003. Comparing the value relevance of two operating income measures. *Review of Accounting Studies* 8 (4): 561–572.

Callen, J. L., M. Khan, and H. Lu. 2009. Accounting quality, stock price delay and future stock returns. Working paper, University of Toronto and Massachusetts Institute of Technology.

Charrow, R. P., and V. R. Charrow. 1979. Making legal language understandable—Psycholinguistic study of jury instructions. *Columbia Law Review* 79 (7): 1306–1374.

Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Clement, M. B., and S. Y. Tse. 2003. Do investors respond to analysts' forecast revisions as if forecast accuracy is all that matters? *The Accounting Review* 78 (1): 227–249.

Core, J. E. 2001. A review of the empirical disclosure literature: Discussion. *Journal of Accounting and Economics* 31 (1–2): 441–456.

Cotter, J., I. Tuna, and P. D. Wysocki. 2006. Expectations management and beatable targets: How do analysts react to explicit earnings guidance? *Contemporary Accounting Research* 23 (3): 593–628.

Courtis, J. K. 1995. Readability of annual reports: Western versus Asian evidence. *Accounting, Auditing & Accountability Journal* 8 (2): 4–17.

Cox, C. 2007. Closing remarks to the Second Annual Corporate Governance Summit. Delivered at the USC Marshall School of Business, Los Angeles, CA, March 23. Available at http://www.sec.gov/news/speech/2007/spch032307cc.htm.

Duru, A., and D. M. Reeb. 2002. International diversification and analysts' forecast accuracy and bias. *The Accounting Review* 77 (2): 415–433.

Firtel, K. B. 1999. Plain English: A reappraisal of the intended audience of disclosure under the Securities Act of 1933. *Southern California Law Review* 72 (2–3): 851–897.

Frankel, R., S. P. Kothari, and J. Weber. 2006. Determinants of the informativeness of analyst research. *Journal of Accounting and Economics* 41 (1–2): 29–54.

Gazmararian, J. A., D. W. Baker, and M. V. Williams. 1999. Health literacy among Medicare enrollees in a managed care organization. *Journal of the American Medical Association* 281 (6): 545–551.

Gibson, E. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68 (1): 1–76.

Gu, Z. Y., and T. Chen. 2004. Analysts' treatment of nonrecurring items in street earnings. *Journal of Accounting and Economics* 38 (1–3): 129–170.

———, and W. M. Wang. 2005. Intangible assets, information complexity, and analysts' earnings forecasts. *Journal of Business Finance & Accounting* 32 (9–10): 1673–1702.

Gunning, R. 1969. The Fog Index after twenty years. *Journal of Business Communication* 6 (2): 3–13.

Healy, P., 1977. Can you understand the footnotes to financial statements? *Accountants Journal* July: 219–222.

Healy, P. M., A. P. Hutton, and K. G. Palepu. 1999. Stock performance and intermediation changes surrounding sustained increases in disclosure. *Contemporary Accounting Research* 16 (3): 485–520.

Hirst, D. E., and P. E. Hopkins. 1998. Comprehensive income reporting and analysts' valuation judgments. *Journal of Accounting Research* 36: 47–75.

———, ———, and J. M. Wahlen. 2004. Fair values, income measurement, and bank analysts' risk and valuation judgments. *The Accounting Review* 79 (2): 453–472.

Hong, H., and J. D. Kubik. 2003. Analyzing the analysts: Career concerns and biased earnings forecasts. *The Journal of Finance* 58 (1): 313–351.

Hopkins, P. E., R. W. Houston, and M. F. Peters. 2000. Purchase, pooling, and equity analysts' valuation judgments. *The Accounting Review* 75 (3): 257–281.

Indjejikian, R. J. 1991. The impact of costly information interpretation on firm disclosure decisions. *Journal of Accounting Research* 29 (2): 277–301.

Johnston, R., S. Markov, and S. Ramnath. 2009. Sell-side debt analysts. *Journal of Accounting and Economics* 47: 91–107.

Jones, M. J., and P. A. Shoemaker. 1994. Accounting narratives: A review of empirical studies of content and readability. *Journal of Accounting Literature* 13: 142–184.

Kimble, J. 1994. Answering the critics of plain language. *The Scribes Journal of Legal Writing* 5: 51–85.

Koo, M. M., I. Krass, and P. Aslani. 2003. Factors influencing consumer use of written drug information. *The Annals of Pharmacotherapy* 37 (2): 259–267.

Kripke, H. 1970. The SEC, the accountants, some myths and some realities. *New York University Law Review* 45 (6): 1151–1205.

———. 1973. The myth of the informed layman. *Business Law* 28: 631–638.

Lang, M. H., and R. J. Lundholm. 1996. Corporate disclosure policy and analyst behavior. *The Accounting Review* 71 (4): 467–492.

———, K. V. Lins, and D. P. Miller. 2004. Concentrated control, analyst following, and valuation: Do analysts matter most when investors are protected least? *Journal of Accounting Research* 42 (3): 589–623.

Lee, S. Y. D., J. A. Gazmararian, and A. M. Arozullah. 2006. Health literacy and social support among elderly Medicare enrollees in a managed care plan. *Journal of Applied Gerontology* 25 (4): 324–337.

Lee, Y. 2010. The effects of quarterly reported readability on information efficiency of stock prices. Working paper, National Taiwan University.

Leuz, C., and C. M. Schrand. 2009. Disclosure and the cost of capital: Evidence from firms' response to the Enron shock. Working paper, University of Chicago and University of Pennsylvania.

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

Li, Y., P. R. Rau, and J. Xu. 2009. All-star analyst coverage and corporate transparency. Working paper, Purdue University.

Loughran, T., and B. McDonald. 2010. Measuring readability in financial text. Working paper, University of Notre Dame.

Mikhail, M. B., B. R. Walther, and R. H. Willis. 1999. Does forecast accuracy matter to security analysts? *The Accounting Review* 74 (2): 185–200.

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

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

O'Brien, P. C., M. F. McNichols, and H. W. Lin. 2005. Analyst impartiality and investment banking relationships. *Journal of Accounting Research* 43 (4): 623–650.

———, and R. Bhushan. 1990. Analyst following and institutional ownership. *Journal of Accounting Research* 28: 55–76.

Ott, B. B., and T. L. Hardie. 1997. Readability of advance directive documents. *Image—the Journal of Nursing Scholarship* 29 (1): 53–57.

Phillips, T. J., C. M. Daily, and M. S. Luehlfing. 2007. A note on the readability of professional materials for management accountants. *Advances in Management Accounting* 16: 311–318.

Plumlee, M. A. 2003. The effect of information complexity on analysts' use of that information. *The Accounting Review* 78 (1): 275–296.

Previts, G. J., R. J. Bricker, T. R. Robinson, and S. J. Young. 1994. A content analysis of sell-side financial analysts company reports. *Accounting Horizons* 8 (2): 55–70.

Rogers, R. K., and J. Grant. 1997. Content analysis of information cited in reports of sell-side financial analysts. *Journal of Financial Statement Analysis* 3 (1): 17–30.

Schipper, K. 1991. Analysts' forecasts. *Accounting Horizons* 5 (4): 105–121.

Schroeder, M. 2002. SEC proposes rules to improve disclosure by public companies. *Wall Street Journal* (May 1): C5.

Schroeder, N. and C. Gibson. 1990. Readability of management's discussion and analysis. *Accounting Horizons* 4 (4): 78–87.

Securities and Exchange Commission (SEC). 1998. *A Plain English Handbook: How to Create Clear SEC Disclosure Documents*. Washington, D.C.: U.S. Securities and Exchange Commission.

Shuptrine, F. K., and E. M. Moore. 1980. Even after the Magnuson-Moss Act of 1975, warranties are not easy to understand. *The Journal of Consumer Affairs* 14 (2): 394–404.

Smith, J. E., and N. P. Smith. 1971. Readability: A measure of the performance of the communication function of financial reporting. *The Accounting Review* 46 (3): 552–561.

Stone, C. A., E. R. Silliman, B. J. Ehren, and K. Apel. 2005. *Handbook of Language and Literacy: Development and Disorders*. New York, NY: Guilford Press.

Thompson, C. K., and Y. Faroqi-Shah. 2002. Models of sentence production. In *Handbook of Adult Language Disorders: Integrating Cognitive Neuropsychology, Neurology, and Rehabilitation*, edited by A. E. Hillis, 311–330. Philadelphia, PA: Psychology Press, Taylor and Francis.

You, H., and X. Zhang. 2009. Financial reporting complexity and investor underreaction to 10-K information. *Review of Accounting Studies* 14 (4): 559–586.