Advances in Accounting 59 (2022) 100627

Contents lists available at ScienceDirect

Advances in Accounting

journal homepage: www.elsevier.com/locate/adiac

Accounting estimation intensity, analyst following, and earnings forecast properties☆

Jeff P. Boone [a,*], Inder K. Khurana [b], K.K. Raman [a]

[a] Alvarez College of Business, The University of Texas at San Antonio, San Antonio, TX 78249, United States of America
[b] Robert J. Trulaske, Sr. College of Business, University of Missouri, Columbia, MO 65211, United States of America

ARTICLE INFO

Keywords:
Accounting estimation intensity
Textual analysis
Analyst forecasts

ABSTRACT

We specify two measures of accounting estimation intensity (AEI) based on the textual analysis of the qualitative disclosures in the critical accounting policies (CAP) section of firms' MD&A. We then examine how these measures relate to financial analyst following and earnings forecast properties. Using a narrow dictionary definition of accounting estimates, we find AEI is positively associated with analyst following. It is also associated with increasing levels of private information in analysts' forecasts and the informativeness of analysts' reports when analysts engage in greater scrutiny of accounting estimates. Using a broader definition of accounting estimates yields a statistically significant relation with the informativeness of analysts' reports. Overall, our results are consistent with AEI stimulating investor demand for analysts' services and increasing the informativeness of these services.

1. Introduction

Estimation of the future is essential to and pervasive in the preparation of financial statements because accounting requires estimates and allocations across time periods (Dechow & Dichev, 2002). As a result, estimates embedded in accounting entail significant managerial judgment in deriving summary performance measures, such as earnings and the book value of shareholder equity. Prior research has documented the importance of accounting estimates embedded in financial reports and how these estimates are managed opportunistically to meet analyst forecasts. For example, Cohen, Darrough, Huang, and Zach (2011) finds that managers use warranty accruals to manage earnings opportunistically to meet earnings targets. Relatedly, Boone, Khurana, and Raman (2022) suggests that auditor estimation expertise does not limit the propensity of managers to use estimates to meet analyst forecasts.

We contribute to this literature by focusing on the effects of the extent of accounting estimation in the preparation of financial statements. Specifically, we examine the relation between accounting estimation intensity (AEI) and analyst following, the extent of private information in analysts' forecasts, and the informativeness of analyst reports. To measure AEI, we use textual analysis of the critical accounting policies (CAP) section of the management's discussion and analysis (MD&A) of firms' 10-Ks. The Securities and Exchange Commission (Securities and Exchange Commission (SEC), 2003) requires management to disclose in the CAP section of the MD&A any accounting estimates that are subjective, involve highly uncertain measurement issues, and are material to the financial statements. These qualitative CAP disclosures are distinct from the quantitative critical accounting estimate disclosures in the MD&A, where management estimates the sensitivity of the dollar effect on earnings from changes in assumptions (Glendening, 2017; Glendening, Mauldin, & Shaw, 2019).

We posit a positive relation between AEI and analyst following as well as the quality of analyst services. The higher the AEI, the higher the cost for investors to process the complex qualitative information in the CAP section of the company's MD&A, and the less informative the firm's prices in the absence of analyst coverage. To the extent that the cost of obtaining analyst reports is less than the cost of processing the qualitative information in the CAP section of the MD&A, a company's AEI can be expected to raise the need for analyst services. Additionally, the lower informativeness of stock prices provides analysts the incentive to

☆ Authors' names are listed alphabetically. We would like to thank the Editor, three reviewers, Matt Glendening, Shiva Sivaramakrishnan, Rustam Zufarov (Discussant) and participants of the AAA annual meeting for their helpfule comments. K. K. Raman acknowledges support from the Ramsdell Chair for Accounting at The University of Texas at San Antonio.

* Corresponding author.
E-mail addresses: Jeff.Boone@utsa.edu (J.P. Boone), Khurana@missouri.edu (I.K. Khurana), K.K.Raman@utsa.edu (K.K. Raman).

https://doi.org/10.1016/j.adiac.2022.100627
Received 21 April 2020; Received in revised form 9 August 2022; Accepted 16 August 2022
Available online 2 September 2022
0882-6110/

provide greater coverage to identify profitable investment recommendations and earn higher trading commissions (Barth, Kasznik, & McNichols, 2001). For these reasons, we expect higher AEI to be associated with increased analyst following. Alternatively, higher AEI could decrease the extent of analyst following if analysts' information acquisition and processing costs increase by more than the benefits of following a firm.

Further, the greater the AEI, the greater the amount of private information incorporated in analysts' earnings forecasts, as analysts draw divergent insights from the estimates. Also, to the extent that investors place greater reliance on analyst reports for companies with higher AEI, the greater the informativeness of analyst reports, as measured by the company stock returns associated with these reports.

Consistent with Chen, Chen, and Li (2021), we operationally define a disclosure of material accounting estimates as any instance within the CAP text in which a variation of the word "estimate" functions as a verb that directs action toward an object or as an adjective that modifies the meaning of an object. For example, in the phrase "estimated the useful life," the verb "estimated" directs action toward the object "life," and in the phrase "estimated loss," the adjective "estimated" modifies the meaning of the object "loss." We posit that the greater the number of material accounting estimates, the higher the AEI. Hence, our operational definition of AEI is the count of the number of instances in which variations of the word "estimate" functions as a verb or adjective within the CAP section of the MD&A. We also adopt an alternative proxy for AEI based on Chen and Li (2017), which is conceptually similar to but broader than the measure based on Chen et al. (2021).

We obtain our data from the intersection of 10-K filings on EDGAR, Compustat, CRSP, IBES, and CDS/Spectrum Institutional 13(f) filings. Our measure of the firm-year specific AEI is based on text-based analysis of the CAP section in the MD&A in 22,236 10-K filings for the period 2007–2017. Using a narrow dictionary definition of accounting estimates, we find that AEI is associated with an increase in analyst coverage, which is consistent with the idea that analysts respond to investors' need for assistance in processing and interpreting accounting estimates. We also find AEI to be positively related to the informativeness of analyst forecasts, as measured by the proportion of information flowing into a firm's stock price around the issuance of analyst reports, stated as a fraction of all information flowing into the stock price between the filing date of the prior year 10-K and the end of the fiscal year being forecast. In all these analyses, we control for previously identified determinants of the demand for analyst coverage and of the informativeness of analyst reports, including firm size, the readability of qualitative CAP disclosures, and the overall CAP length. Our results are robust to the exclusion of firms that are *not* followed by analysts.

In additional analyses, we exploit the cross-sectional variation in analysts' scrutiny of accounting estimates. An increasing number of analysts have been issuing cash flow forecasts in addition to earnings forecasts, although a majority of analysts continue to issue only earnings forecasts. The articulation of cash, earnings, and accruals suggests that an analyst who issues both cash flow and earnings forecasts implicitly is forecasting accruals, while an analyst who issues only an earnings forecast implicitly is *not* forecasting accruals. An accrual forecast necessarily requires the analyst to have a deeper understanding of the firm's accruals-generating process, including an understanding of the estimates embedded in those accruals as well as the accounting estimates that are excluded from the determination of accruals. We expect the impact of AEI on analyst forecasts to be amplified when a larger fraction of a firm's analysts issue a cash flow forecast in addition to an earnings forecast. We use this fraction as a proxy for analysts' scrutiny of a firm's accounting estimation intensity.

Our findings indicate that the relation between AEI and the private information in analyst forecasts becomes stronger as analyst scrutiny of accounting estimates increases, which suggests that analyst scrutiny of accounting estimates results in analysts' earnings forecasts incorporating more private information. We also find a positive relation between AEI and the informativeness of analyst forecasts as analyst scrutiny of accounting estimates increases. These cross-sectional results generally hold for both our narrow and broader dictionary definitions of accounting estimates.

Our study contributes to the sparse prior literature on the effects of the highly subjective and uncertain accounting estimates disclosed in the CAP section of the MD&A (Gordon, Ma, & Runesson, 2018; Levine & Smith, 2011). In particular, our study is closely related to Gordon et al. (2018), which finds an adverse effect of AEI on the properties of analyst forecasts (i.e., forecast error and dispersion), with such effects mitigated when CAP disclosures contain more specific language or when an accounting topic is specified in conjunction with estimation-related words. Our finding that AEI is associated with reduced precision of private information embedded in analysts' forecasts is fully consistent with the key findings in Gordon et al. (2018) in the sense that both studies suggest that accounting estimates challenge analysts. Our study contributes incrementally to Gordon et al. (2018) by finding that greater analyst scrutiny of accounting estimates yields earnings forecasts that incorporate more private information and are more informative to investors. Hence, we both confirm and extend Gordon et al. (2018).

Our study is different from Gordon et al. (2018) in three ways. First, we focus on the link between our accounting estimation intensity metric and a different set of analyst properties (i.e., analyst following, the extent of private information in analysts' forecasts, and the information content of those forecasts). Second, whereas Gordon et al. (2018) utilizes a simpler measure of AEI based on word counts, we derive our AEI metric by applying a more sophisticated machine learning technique (the Stanford machine-learning algorithm) to the qualitative disclosures in the CAP section of firms' MD&A. Finally, we document that AEI stimulates investor need for analysts' services, and that greater scrutiny of accounting estimates by analysts increases the informativeness of those services and results in more private information being embedded in analysts' forecasts. In our study, we show that this increased informativeness and private information stems from analysts engaging in additional information acquisition and processing to better understand the valuation implications of the subjective/uncertain accounting estimates.

A key insight from our study is that the association between CAP discussion of AEI and the properties of analysts' forecasts depends on the actions of analysts. Specifically, we find that the association between AEI and the properties of analysts' forecasts depends on the percentage of analysts issuing cash flow forecasts (as a proxy for analysts' attention to estimates). More broadly, our findings suggest that although accounting estimates can be problematic for investors (Chen et al., 2021; Roh, 2018), analysts can help ameliorate these problems by providing increased coverage, scrutinizing accounting estimates, and incorporating their own private information in more informative earnings forecasts.

Our findings inform policy making. The Financial Accounting Standards Board (FASB) views the understandability of accounting information as an important qualitative characteristic. To help investors better understand managerial judgments and accounting projections, the Securities and Exchange Commission (SEC) (2003) requires companies to discuss their critical accounting policies in the MD&A. Our findings suggest that those qualitative disclosures reveal the firm's underlying accounting estimation intensity, that this intensity is associated with increased analyst coverage, and that analyst scrutiny of accounting estimates enables analysts to incorporate their own private information in generating more informative earnings forecasts.

The rest of the paper is organized as follows. In Section 2, we discuss the related literature and develop testable hypotheses. Section 3 describes our sample, construction of variables, and the empirical model. Section 4 reports and discusses the empirical results, and Section 5 concludes the paper.

2. Background and hypothesis development

2.1. Background

Estimates are embedded in earnings. Financial reports, including summary performance measures such as earnings and the book value of shareholder equity, are heavily subject to managers' subjective judgments and projections of an uncertain and often turbulent future. Thus, inventories are reported at the lower of cost or market (usually an estimate), while accounts receivables are net of an estimated loss for delinquencies (bad debts). Restructuring charges incorporate forecasts of future employee severance payments and plant closure costs. Other estimates include those for warranty provisions, pensions and other post-retirement expenses, write-offs of assets and goodwill (impairment charges), stock option expenses, and more recently, the marking of assets and liabilities to fair values. While estimates of bad debts are based on prior factual experience, other more recent estimates (required by accounting standards) involve multiple layers of forecasts, including the range of interest rates, expected volatilities, and expected life (etc.), which are blended into and reported as one item, such as stock option expense. Furthermore, given that nontraded financial assets and liabilities have no market prices, they need to be adjusted to fair values and subject to write-downs based on subjective predictions of cash flows that are input into proprietary valuation models. Even on the revenue side (such as those for revenues from long-term contracts), revenue recognition is based on estimated profitability derived from cost projections.

A large number of estimates incorporated in the financial statements entail measurement subjectivity. To the extent that managers use judgment to measure the amount, estimates are derived from managers' assumptions about the occurrence or non-occurrence of future events. Often, a manager's assumptions are based on private beliefs, which could be biased if managers are overoptimistic, or misrepresented if managers behave opportunistically. In summary, managers' assumptions underpinning accounting estimates may be biased or misrepresented, yet they cannot be verified in the traditional sense of verifying a past event such as the historical cost of an asset.

To help investors better understand these managerial judgments and projections (estimates), the SEC requires companies to discuss their critical accounting policies in the MD&A. As part of that discussion, the SEC requires that companies disclose and discuss critical accounting estimates and assumptions when the nature of the estimate or assumption is subjective, creates highly uncertain measurement issues, and is material to the financial statements (Securities and Exchange Commission (SEC), 2003). The purpose of CAP disclosures is to make the impact of uncertain estimates on earnings more transparent (Securities and Exchange Commission (SEC), 2002, Securities and Exchange Commission (SEC), 2003, Securities and Exchange Commission (SEC), 2011). While CAP disclosures are about accounting estimates, they are intended to relate to underlying business uncertainty that is reflected in the estimates. In particular, companies should address such issues as "how the registrant arrived at the estimate/assumption, how accurate the estimate/assumption has been in the past; how much the estimate/assumption has changed in the past, and whether the estimate/assumption is reasonably likely to change in the future" (Ernst & Young, 2017). Levine and Smith (2011) finds that 80% of their sample firms identify, on average, six to seven qualitative critical accounting policies, suggesting that firms generally comply with SEC guidance.

The MD&A also includes *quantitative* disclosures of critical accounting estimates (CAE). The quantitative CAE disclosures estimate the dollar effect on earnings from likely percent changes in assumptions and have been previously studied by Glendening (2017) and Glendening et al. (2019). These studies suggest that quantitative CAE disclosures convey accounting measurement uncertainty and lower the predictive value of earnings for future cash flows, but also constrain earnings management and lower the likelihood of financial misstatements. Citing prior research (Levine & Smith, 2011), Glendening et al. (2019) notes relatively low rates of quantitative CAE disclosures by SEC registrants. In contrast to the quantitative CAE disclosures, the CAP disclosures we examine in our study are qualitative, indicating uncertainty and judgment in the application of GAAP without providing dollar effects on earnings. The distinction between the quantitative CAE disclosures and qualitative CAP disclosures is important because requiring a detailed discussion of accounting estimation intensity (AEI) under CAP can reduce uneven implementation of accounting standards. CAP can serve as a mechanism to better align managers' incentives with enforcers, encouraging full compliance with a unified accounting framework.

Prior research has investigated the consequences of CAP disclosures. Levine and Smith (2011) finds that CAP disclosures are associated with lower earnings valuation multiples and a reversal of returns around the 10-K filing. Gordon et al. (2018) finds that CAP disclosure of accounting estimates is positively associated with dispersion in analysts' forecasts and negatively associated with forecast accuracy. However, prior research does not address whether the firm's AEI as reflected in CAP disclosures is associated with analyst following, the extent of private information incorporated in analysts' forecasts, and the information content of those forecasts. We address these questions in our study.

Potentially, the qualitative CAP disclosures could improve the usefulness of accounting information by providing managers a venue for conveying forward-looking private information. Alternatively, the quality of information in these disclosures may be compromised by the complexity faced by managers in making reliable projections into the future in a rapidly changing business environment. Further, there is the potential for managers to misuse estimates to achieve self-serving goals, including higher compensation.

The available empirical evidence indicates that accounting measurement subjectivity often leads to both intentional (Fields, Lys, & Vincent, 2001; Healy & Wahlen, 1999) and unintentional (Gong, Li, & Xie, 2009; Peterson, 2012; Plumlee & Yohn, 2010) reporting errors. These errors that stem from subjective and highly uncertain accounting estimates potentially create uncertainty about the precision and reliability of earnings, increase the cost to investors and analysts of processing accounting estimation information, and stimulate a search for other information sources that can assist in evaluating the reliability of the accounting estimates. Below, we describe how these characteristics of subjective and highly uncertain accounting estimates could impact analyst following and various properties of analysts' earnings forecasts.

2.2. Hypothesis development

2.2.1. Analyst following

As information intermediaries and information providers, sell-side financial analysts are of interest to academic researchers because of their important role in acquiring, interpreting, and disseminating information to capital market participants. While prior research has focused by and large on the relation between the *quantitative* numeric data in financial reports and analyst behavior, we analyze the qualitative accounting estimation textual disclosures in the CAP section of the MD&A to test their impact on analyst following.

As noted previously, the qualitative information about accounting estimates in the CAP section of the MD&A is based on managers' subjective assumptions and projections of the future in an uncertain business environment. Needless to say, these qualitative disclosures impose additional time and demands on investors to identify, extract, and comprehend the pertinent information. To the extent that the cost of obtaining analyst reports is less than the cost of processing the qualitative information in the CAP section of the MD&A, the extent of estimation in the company's accruals can be expected to raise the need for analyst services. Additionally, the lower informativeness of stock prices associated with higher AEI provides an incentive for analysts to provide greater coverage to identify profitable investment recommendations and earn higher trading commissions (Barth et al., 2001). For these reasons, we expect higher AEI to be associated with increased analyst following.

Our first hypothesis, stated in the alternative form, is as follows:

H1. There is a positive relation between accounting estimation intensity (based on the firm's qualitative disclosures in the CAP section of the MD&A) and analyst following, *ceteris paribus*.

However, there are also reasons to believe that AEI may not be related to analyst following. CAP estimation disclosures could be uninformative, either because managers themselves may be unable to make reliable projections of a fast-changing business future, or managerial incentives to manage earnings for self-serving reasons (such as executive compensation and maintaining the stock price) could make the projections misleading. In either scenario, analysts will need to incur greater information acquisition and processing costs to extract the additional useful (if any) information from the qualitative disclosures, which could discourage analyst following. The implication is that noisy, potentially biased, and potentially value-irrelevant estimates increase analysts' information-processing costs, making the benefit of following the firm less than the cost of following the firm.

2.2.2. *Private information and informativeness of analyst forecasts*

In performing their function as information intermediaries and information providers, analysts interpret publicly available information and also generate private information by discovering and interpreting information not otherwise publicly available (Barron, Byard, Kim, & Riedl, 2002; Barth et al., 2001; Livnat & Zhang, 2012).

As noted earlier, prior research indicates that accounting measurement subjectivity often leads to both intentional and unintentional reporting errors. Other research suggests that analysts rely less on information that they perceive as being more likely to have been misreported (Elliott & Hanna, 1996; Hirst, 1994a; Hirst, 1994b; Williams, 1996). Barron et al. (2002) reports that when deriving earnings forecasts for firms with significant intangible assets, analysts seek to overcome the earnings distortion that arises from the expensing of investment in intangible assets by placing greater relative emphasis on their own idiosyncratic information (obtained through additional information acquisition and processing) as a supplement to firms' financial reports.

Collectively, these studies suggest that in developing their earnings forecasts, analysts can be expected to assess the precision and reliability of earnings when accruals contain material estimates that are subjective and highly uncertain. To the extent that analysts view accounting estimates as likely to be misreported or create distortion in earnings, analysts will supplement firms' financial disclosures by engaging in additional information acquisition and processing to better understand the valuation implications of the subjective and highly uncertain accounting estimates. Since these information acquisition and processing efforts likely will vary across analysts (depending upon an analyst's own information endowment and ability to critically evaluate accounting estimates), these efforts will cause analysts' forecasts to be informed by more analyst-specific (idiosyncratic) beliefs and to contain more idiosyncratic forecast error (since the idiosyncratic beliefs are not perfectly accurate). This diverse interpretation of a firm's performance from a common source of public information embeds analysts' forecasts with private information that stems from "informed judgments or opinions" (Kim & Verrecchia, 1994, p. 42).

In summary, if analysts supplement the CAP discussion of accounting estimates by acquiring and processing additional information in generating their earnings forecasts, there should be a positive relation between accounting estimation intensity and private information embedded in analysts' earnings forecasts. Further, to the extent that analysts' supplemental information acquisition and processing enable analysts to better understand the valuation implications of the subjective and highly uncertain accounting estimates, accounting estimates should also be associated with increased usefulness of analysts' forecasts. These predictions are formalized in the following two hypotheses stated in the alternative form.

H2. There is a positive relation between accounting estimation intensity (based on the firm's qualitative disclosures in the CAP section of the MD&A) and private information embedded in analysts' earnings forecasts.

H3. There is a positive relation between accounting estimation intensity (based on the firm's qualitative disclosures in the CAP section of the MD&A) and the informativeness of analyst reports.

In analyses related to our hypotheses H2 and H3, we conduct additional cross-sectional tests that exploit the variation in the level of scrutiny of accruals by analysts. Lev and Gu (2016) notes that an increasing number of analysts have started to issue cash flow forecasts in addition to earnings forecasts.[1] The articulation of cash flows, earnings, and accruals suggests that an analyst who forecasts cash flows in addition to earnings is implicitly forecasting accruals. An accrual forecast necessarily requires the analyst to have a deeper understanding of the firm's accruals-generating process, including the accounting estimates embedded as well as those excluded from the determination of the accruals. Consistent with this argument, prior research (e.g., Gordon, Petruska, & Yu, 2014; Mohanram, 2014; Radhakrishnan & Wu, 2014) finds accruals mispricing to be less severe for firms where analysts issue both cash flow forecasts and earnings forecasts, compared with firms where analysts only issue earnings forecasts. In a similar vein, Call, Chen, and Tong (2013, p. 440) finds that "the majority of analysts' cash flow forecasts reflect meaningful adjustments for working capital and other accruals, and that these adjustments are superior to those implied by time-series cash flow forecasts." Further, in a survey of financial analysts, Brown, Call, Clement, and Sharp (2015, p. 22) reports that 25% of analysts indicate the extent of estimates embedded in earnings is an important determinant of earnings quality, which suggests that a subset of analysts likely review estimates as part of their assessment of earnings quality. Consequently, the hypothesized impact of accounting estimation intensity on analyst forecasts can be expected to be amplified as a larger fraction of a firm's analysts forecast cash flows in addition to earnings, which reflects greater scrutiny of accounting estimates by analysts.[2]

3. Sample and research design

3.1. *Sample*

Table 1 summarizes the sample formation process. We begin by identifying 52,076 firm-years in both Compustat and IBES with 10-K filings available for the years 2007–2017. The ending period of 2017 represents the most recent year of data availability at the time of data collection. Our choice of 2007 as the starting period for our study represents a trade-off between external validity and information-processing costs. In particular, the derivation of our estimate test variable, described below, requires hundreds of hours of computer processing time, which necessitated that we select a subset of available years for analysis. Starting the analysis in 2007 kept the computer processing task feasible while still providing a reasonable number of years, in the interest of external validity. For these firm-years, we requested a third-party data vendor to extract the MD&A portion of the 10-K and

[1] Mohanram (2014) notes that cash flow forecasts were rare until 2001 and that by 2010, analysts in almost half of all firms forecast cash flows, as reported on IBES.

[2] Our approach of using cash flow forecasts as a sign of greater analyst scrutiny of accounting estimation is also consistent with that of Penman (2001) and Wild et al. (2001), which suggest that market participants find cash flows helpful in assessing whether earnings are likely to contain material misstatements. A similar conclusion was reached by financial analysts who were quoted by DeFond and Hung (2003, p. 76) as saying that "large accruals impact their decisions to forecast cash flows."

Table 1
Sample formation.

	Analysis of analyst following	Analysis of private information	Analysis of analyst informativeness
Firms in Computstat and IBES with 10-K filings 2007–2017.	52,076	52,076	52,076
MD&A algorithmically extractable from 10-K by data provider SeekEdgar	46,940	46,940	46,940
Critical Accounting Policies section of MD&A algorithmically extractable by authors	41,819	41,819	41,819
with IBES earnings forecasts	26,425	26,425	26,425
with FOG index from WRDS	26,229	26,229	26,229
with complete data on control variables	22,793	22,793	22,793
with at least 500 words in the Critical Accounting policies section of MD&A[1]	22,236	22,236	22,236
with at least 4 analysts issuing earnings forecasts[2]		17,774	
excluding fiscal year 2017[3]		16,209	
with analyst informativeness metric [4]			20,984

[1] Consistent with Chen and Li (2017) and Roh (2018), done to eliminate potential errors in the algorithmic extraction of Critical Accounting Policies section of the MD&A.
[2] Metrics that are based on standard deviation in analysts' forecasts require at least 4 analysts.
[3] Fiscal year 2017 requires 2018 actual earnings to measure private information metric, which was unavailable at the time the data were collected.
[4] Requires at least 90 days of return data to calculate analyst informativeness. Most of the lost observations relate to 2018 returns data (unavailable at time of data extraction).

eliminate the tabular information and html tags therein. The vendor provided us with extracted data for 46,940 firm-years. We were able to algorithmically identify and extract the critical accounting policies section of the MD&A[3] for 41,819 of the 46,940 firm-years. Eliminating firm-years (1) with fewer than 500 words in the CAP section, (2) without analysts' forecast data in IBES, and (3) with missing values of FOG and other control variables winnowed the 41,819 observations to a core sample of 22,236 firm-year observations.[4] This core sample forms the basis of our analyses, subject to additional attrition from data screens that vary from test to test.

3.2. Empirical model

We use the following model to test our hypotheses H1–H3

$$y_{i,t+1} = \alpha + \beta_1 ESTIMATES_{i,t} + \sum_{1}^{k} \delta_k X_{k,i,t} + \varepsilon_{i,t+1} \tag{1}$$

where $y_{i,t+1}$ is one of the analyst-related variables, $ESTIMATES_{i,t}$ is our test variable capturing accounting estimation intensity, and $X_{k,i,t}$ is one of k explanatory variables, controlling for a factor shown in prior research to be associated with the dependent variable. Detailed operational definitions for all variables and the related construct are provided in Appendix A. Model (1) is estimated by ordinary least squares, with standard errors clustered by firm, and includes year and industry (two-digit SIC) fixed effects.[5] The explanatory variables in model (1) are taken from the year t 10-K (filed in the early part of year $t+1$), while the dependent variables are measured in year $t+1$ immediately after the year t 10-K is filed.

As noted above, we also conduct additional tests that allow the effect of ESTIMATES to vary cross-sectionally as a function of the level of scrutiny that analysts give to accounting estimates. We proxy for that level of scrutiny using CF_FORECASTS, which is the fraction of a firm's analysts that issue cash flow forecasts. Model (2) below is used for these additional tests.

$$\begin{aligned} y_{i,t+1} = {} & \alpha + \beta_1 ESTIMATES_{i,t} + \beta_2 ESTIMATES \times CF\_FORECAST_{i,t} \\ & + \beta_3 CF\_FORECAST_{i,t} + \sum_{k} \delta_k X_{k,i,t} + \varepsilon_{i,t+1} \end{aligned} \tag{2}$$

3.3. Dependent variables

Hypotheses H1–H3 focus on analyst following, the properties of analysts' earnings forecasts, and the information content of their reports. We follow Lehavy, Li, and Merkley (2011) and Barron, Kim, Lim, and Stevens (1998) in deriving the dependent variables to test Hypotheses H1–H3.

Analyst following (NANALYSTS) is the number of distinct analysts issuing at least one forecast of year $t+1$ annual earnings following the filing of the year t Form 10-K. We also operationalize a dichotomous measure of analyst following (FOLLOW), where FOLLOW $=1$ if NANALYSTS >0, otherwise 0. Private information embedded in analysts' earnings forecasts (PRIVATE) is derived from Barron et al. (1998). Barron et al. (1998) analyze properties of analyst forecasts in a setting where analysts are assumed to rely upon both private and common imprecise information in developing their earnings forecasts. Common information is available to all analysts and moves their beliefs toward a common belief. Private information, arising either from private information search or from different interpretations of public information, potentially moves analysts' beliefs in different directions and embeds analyst-specific (i.e., idiosyncratic) beliefs into an analyst's earnings forecast.

Within the Barron et al. (1998) model, uncertainty (UNCERTAINTY) is the average of each analyst's expected squared forecast error, CONSENSUS is the across-analyst correlation in forecast errors, SE is the expected squared error in the mean forecast, and PRIVATE is the complement of CONSENSUS (1 − CONSENSUS), which measures the fraction of the total precision of information available to analysts that is private information. That is, PRIVATE measures the precision of private information as a fraction of the precision of all information (both common and private) available to analysts.

We measure CONSENSUS using Eq. (16) in Barron et al. (1998) as $CONSENSUS = \frac{SE - \frac{DISPERSION}{NANALYSTS}}{UNCERTAINTY}$, where SE is the squared difference between IBES actual fiscal year $t+1$ earnings and the mean analysts' forecast of fiscal year $t+1$ earnings, deflated by share price taken from the last IBES monthly update immediately preceding year t fiscal year-end; DISPERSION is the variance of analysts' forecasts of fiscal year $t+1$ earnings, deflated by share price taken from the last IBES monthly update immediately preceding fiscal year t fiscal year-end; and UNCERTAINTY

[3] Our algorithm is to search the CAP for the first instance of the phrase “critical accounting” and to extract the textual portion of the CAP from that point to the end of the CAP. We assessed the accuracy of our algorithm by selecting a sample of 1568 10-Ks, manually extracting the CAP, and comparing the estimate metric from our manually extracted CAP to the estimate metric derived from our algorithmically extracted CAP. The correlation between the two metrics was approximately 0.80.

[4] Eliminating observations with fewer than 500 words in the CAP is intended to exclude potential errors in the algorithmic extraction of the CAP section of the MD&A and is consistent with Chen et al. (2021) and Roh (2018).

[5] Except in Panels A and B of Table 4, where model (1) is estimated by probit and tobit, respectively.

is measured using Eq. (15) in Barron et al. (1998), which is $UNCERTAINTY = \left(1 - \frac{1}{NANALYSTS}\right) \times DISPERSION + SE$.

We measure the information content of an analyst report (INFORMATIVENESS) as the proportion of information flowing into a firm's stock price around the issuance of analysts' earnings forecast reports, stated as a fraction of all information flowing into stock price during the period between the fiscal year t 10-K filing and the fiscal year $t+1$ year–end. The numerator of the fraction is the summed one-day absolute market-adjusted returns (value-weighted index) on all analysts' forecast report dates using only forecast reports issued the day after the fiscal year t 10-K filing date through the $t+1$ fiscal year-end. The denominator of the fraction is the summed absolute daily market-adjusted returns beginning the day after the year t 10-K filing date and ending on the year $t+1$ fiscal year-end. Multiple analyst reports issued on a single day are treated as a single report, and we require at least 90 trading days in calculating the denominator.

3.4. Test variables

We use two operational definitions of accounting estimation intensity (ESTIMATES), which are ESTIMATES_NDICT and ESTIMATES_WDICT. Measurement of each begins by first tokenizing the text in the CAP section of each fiscal-year MD&A into sentences, using public-use software from the Stanford Natural Language Processing Group, then identifying the grammatical relationship between words in each CAP sentence. As noted earlier, we focus on the CAP disclosures in the MD&A section of the 10-K because these disclosures discuss critical accounting estimates and assumptions when the nature of the estimate or assumption is subjective, creates highly uncertain measurement issues, and is material to the financial statements.

Using each fiscal-year CAP, we measure ESTIMATES_NDICT as the total number of instances in which a CAP sentence (within that fiscal-year CAP) contains a grammatical relation between two words in which a variation of the word "estimate" serves as a verb that directs action toward an object, or as an adjective modifying the meaning of an object. For example, in the phrase "estimated the useful life," the verb "estimated" directs action toward the object "life," and in the phrase "estimated loss," the adjective "estimated" modifies the meaning of the object "loss." Our measurement of ESTIMATES_NDICT is consistent with Chen et al. (2021) and yields a firm-year measure of accounting estimation intensity.

Again using each fiscal-year CAP, we measure ESTIMATES_WDICT as the total number of instances in which a CAP sentence contains a grammatical relation between two words, in which a variation of the word "estimate" (or its synonyms) functions as a verb, adjective, or direct object of a verb. As compared to ESTIMATES_NDICT, ESTIMATES_WDICT is a broader definition in that it considers synonyms of "estimate" and also allows "estimate" (or its synonyms) to function as the object of a verb. Our measurement of ESTIMATES_WDICT is consistent with Chen and Li (2017).

ESTIMATES_NDICT and ESTIMATES_WDICT enter our empirical models in both percentile-rank form and in natural-log form. We use percentile-ranks to facilitate assessment of economic significance. We use natural-log form to account for skewness in the raw count total while also avoiding the loss of information content that comes from the rank transformation.

Panels A and B of Table 2 summarize the mean value of ESTIMATES_NDICT and ESTIMATES_WDICT by industry (Fama-French 30-industry classification) and year, respectively. Inspection of Panel A shows that tobacco products (textiles) ranks highest (lowest) in accounting estimation intensity. Inspection of Panel B shows an initial across-time upward trend in accounting estimation intensity, a flattening, and then a slight across-time decrease.

Appendix B provides more detail on the validation of ESTIMATES_NDICT and ESTIMATES_WDICT (henceforth referenced generically as ESTIMATES_xx). We validate the construct validity of ESTIMATES_xx by correlating it with a separate proxy for the estimation based on the extent of estimates required to comply with the Accounting Standards Codification guidance that governs the firm's accounting (ESTIMATES_ASC). The logic of the validity test is that firms whose governing GAAP requires more accounting estimates should have more discussion of accounting estimates in their CAP. Inspection of Table B3 in Appendix B reveals that ESTIMATES_xx is positively associated with ESTIMATES_ASC. In other words, firms whose governing GAAP requires the firm to make more accounting estimates tend to have more discussion of accounting estimates in their CAP section of the MD&A. This provides some evidence to confirm the construct validity of ESTIMATES.[6]

3.5. Control variables

The purpose of the control variables is to capture important elements of a firm's information environment and business complexity that affect analyst behavior. A firm's information environment and business complexity can enhance the payoff of analysts and therefore affect their incentives to follow a firm and the properties of their earnings forecasts. Following Lehavy et al. (2011), firm-level control variables include those typically found in prior empirical research to be associated with analyst behavior, such as the natural log of market value of equity (*SIZE*), mean annual sales growth over a five-year period (GROWTH), natural log of operating segments (SEGMENTS), percentage of institutional ownership (INST_OWNERSHIP), number of management earnings forecasts issued during a year (MGMT_GUIDANCE), absolute value of cumulative market-adjusted returns summed over a two-day window around the 10-K filing date (10-K_NEWS), advertising expense as a fraction of operating expenses (ADV), research and development expense as a fraction of operating expenses (R&D), and standard deviation of monthly stock returns during the year (STD_RET).

Lehavy et al. (2011) find that 10-K readability is associated with analyst following, private information embedded in analysts' forecasts, and the informativeness of their reports. Therefore, we control for the readability of 10-K filings using the Fog Index (FOG).

We control for the length of a firm's CAP section of the MD&A (LENGTH_CAP) to ensure that our test variable does not simply proxy for the length of these disclosures. In addition, we control for the extent of estimate-related disclosures in the footnotes to the financial statements (FN_NDICT and FN_WDICT) using the same approach used in the measurement of ESTIMATES_NDICT and ESTIMATES_WDICT. We also control for the length of the footnote disclosures (LENGTH_FN). Control for footnote-related disclosure of estimates is necessary to ensure that our results are attributable to CAP disclosures and not to similar disclosures made in the footnotes to the financial statements.

Moreover, we control for the overall tone of uncertainty in the 10-K using the Loughran and McDonald (2011) measure of financial uncertainty derived from content analysis of the 10-K textual discussion (LM_UNCERTAINTY). Control for overall uncertainty disclosed in the 10-K is necessary because some of our estimate-related cue words could conceivably be used in narrative form within the 10-K to describe generic business uncertainty, raising the risk that our estimate metric might simply proxy for generic business uncertainty described in the 10-K. Including LM_UNCERTAINTY as a control helps reduce that risk.

We include industry- and year-fixed effects in all our models to account for variation in analyst behavior across industries and over time. Given the extensive array of control variables in our estimations, we diagnose multicollinearity in the regressions using variance inflation factors (VIFs). The highest VIF in our regressions is around 5.5, indicating that collinearity is not likely to be a problem in interpreting the

[6] We do not use ESTIMATES_ASC in our tests because the metric is available for only approximately one-half of our sample. See discussion in Appendix B for details.

Table 2
Mean Values of ESTIMATES_NDICT and ESTIMATES_WDICT by Industry and Year.

Panel A. Accounting estimates by Fama-French 30 industry groups

Industry	N	ESTIMATES_NDICT	ESTIMATES_WDICT	N Words in CAP (LENGTH_CAP)
Aircraft, ships, and railroad equipment	219	11.85	57.48	2758
Apparel	269	9.30	43.70	2519
Automobiles and Trucks	373	7.20	45.92	2494
Banking, Insurance, Real Estate, Trading	4883	10.95	62.52	5900
Beer & Liquor	46	12.20	57.98	3541
Business Equipment	2419	10.21	52.84	3510
Business Supplies and Shipping Containers	289	9.38	60.87	3399
Chemicals	424	8.64	51.79	3358
Coal	66	9.17	57.29	2512
Communication	558	10.26	51.53	3837
Construction and Construction Materials	655	11.29	52.73	3267
Consumer Goods	259	9.14	55.72	3764
Electrical Equipment	285	8.74	44.12	2456
Fabricated Products and Machinery	723	8.59	46.90	2729
Food Products	441	8.75	46.51	3094
Healthcare, Medical Equipment, Pharmaceutical Products	1901	10.55	47.68	3151
Personal and Business Services	2689	8.97	48.18	3561
Petroleum and Natural Gas	909	10.46	51.70	2996
Precious Metals, Non-Metallic, and Industrial Metal Mining	131	12.10	66.21	3851
Printing and Publishing	150	9.05	51.40	3964
Recreation	358	9.51	47.33	3113
Restaraunts, Hotels, Motels	370	7.45	38.47	2254
Retail	1162	8.50	41.27	2682
Steel Works Etc	243	7.92	40.85	2450
Textiles	49	5.57	35.24	2241
Tobacco Products	34	15.47	73.00	9255
Transportation	600	8.77	44.03	2902
Utilities	598	9.09	60.16	4957
Wholesale	648	7.91	39.68	2889
Everything Else	485	9.19	48.03	3290
	22,236			

Panel B. ESTIMATES by fiscal year

Year	N	ESTIMATES_NDICT	ESTIMATES_WDICT	N Words in CAP (LENGTH)
2007	2202	8.58	50.18	3814
2008	2228	9.64	54.81	4096
2009	2206	9.77	53.77	3995
2010	2034	10.12	53.47	3924
2011	2004	10.12	53.54	3933
2012	2147	10.03	51.99	3797
2013	2096	10.15	52.17	3834
2014	2131	10.07	51.69	3758
2015	2184	10.03	50.92	3694
2016	1649	9.63	49.05	3647
2017	1355	9.08	46.53	3471
	22,236			

regression results.

4. Empirical results

4.1. Descriptive statistics and correlations

Table 3 – Panel A presents central tendency descriptive statistics for the dependent and independent variables. Central tendencies reported in this panel are remarkably similar to those reported by Lehavy et al. (2011) and hence are not discussed in detail for the sake of brevity. The mean value of ESTIMATES_NDICT is about 9.77, while the mean value of ESTIMATES_WDICT is about 51.88. The larger value of ESTIMATES_WDICT reflects its broader definition, as discussed earlier.

Table 3 – Panel B presents pairwise correlations among the variables. Each ESTIMATES_xx exhibits positive and significant pairwise correlation with each of the dependent variables except PRIVATE, for which the correlation is insignificant. Each ESTIMATES_xx is correlated with LENGTH_CAP (the total number of words in the CAP section of MD&A). Correlation between these variables is unsurprising (i.e., more discussion of estimates also increases the length of the document). None of the other pairwise correlations are large enough to raise concerns about multicollinearity. As noted above, we conduct routine multicollinearity diagnostic tests (i.e., data matrix condition indices and VIF scores) as part of the empirical analysis and verified the absence of problematic multicollinearity.

Tables 4, 5, and 6 present empirical results for the tests of H1, H2, and H3, respectively. Table 4 presents the regression results of estimating model (1) using analyst following as the dependent variable. Tables 5 and 6 present regression results of estimating models (1) and (2) using PRIVATE and INFORMATIVENESS, respectively, as dependent variables. Each of the three tables includes one or more panels presenting the basic regression estimates in several columns. Results reported under the “Narrow Dict” column heading are regressions in which the test variable is ESTIMATES_NDICT. Results reported under the “Wide Dict” column heading represent regressions in which the test variable is ESTIMATES_WDICT. Regressions reported under the “Ranked Estimates” column specify the test variable in percentile-rank form, while regressions reported under the “Logged Estimates” column specify

Table 3
Descriptive statistics and correlations.

Panel A. Descriptive statistics

	N	Mean	Median	Std	Q1	Q3
NANALYTSTS	22,236	10.9312	8.0000	9.1223	4.0000	15.0000
PRIVATE	16,209	0.4625	0.3655	0.3550	0.1544	0.7584
INFORMATIVENESS	20,984	0.1399	0.1096	0.1135	0.0552	0.1950
ESTIMATES_NDICT	22,236	9.7677	8.0000	7.2595	5.0000	13.0000
ESTIMATES_WDICT	22,236	51.8775	44.0000	35.6180	30.0000	64.0000
CF_FORECASTS	22,236	0.1737	0.1111	0.2229	0.0000	0.2500
LENGTH_CAP	22,236	3830.9930	2845.0000	3059.5010	1593.0000	5106.5000
FN_NDICT	22,236	17.2000	15.0000	9.3855	11.0000	22.0000
FN_WDICT	22,236	76.7749	71.0000	35.2830	52.0000	94.0000
LENGTH_FN	22,236	6694.5510	6206.0000	2819.2320	4775.0000	8058.0000
LM_UNCERTAINTY	22,236	0.0160	0.0160	0.0027	0.0142	0.0178
FOG	22,236	20.2544	20.1796	0.9497	19.6251	20.7907
SIZE	22,236	6.9989	6.9621	1.7238	5.7937	8.1264
GROWTH	22,236	0.1423	0.0887	0.2091	0.0246	0.1905
SEGMENTS	22,236	0.8788	0.6931	0.5925	0.6931	1.3863
INST_OWNERSHIP	22,236	0.4317	0.4547	0.4020	0.0000	0.8399
MGMT_GUIDANCE	22,236	1.1923	0.0000	2.1526	0.0000	2.0000
10-K_NEWS	22,236	0.0323	0.0178	0.0417	0.0076	0.0387
ADV	22,236	0.0137	0.0000	0.0318	0.0000	0.0121
R&D	22,236	0.0527	0.0000	0.1151	0.0000	0.0415
STD_RET	22,236	0.1114	0.0956	0.0640	0.0658	0.1391

Panel B. Correlation matrix

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	NANALYTSTS	1.00	0.03	0.81	0.17	0.18	0.27	0.07	0.15	0.19	0.21	0.02	0.07	0.69	0.09	0.07	0.19	0.21	−0.11	0.08	0.06	−0.14
2	PRIVATE	0.03	1.00	0.00	0.00	−0.02	0.07	0.00	0.03	0.02	0.02	0.00	0.03	0.02	0.03	−0.02	0.01	−0.04	0.00	−0.03	0.05	0.02
3	INFORMATIVENESS	0.81	0.00	1.00	0.18	0.20	0.36	0.10	0.15	0.22	0.27	−0.03	0.04	0.64	0.08	0.06	0.11	0.11	−0.08	0.01	−0.05	−0.12
4	ESTIMATES_NDICT	0.17	0.00	0.18	1.00	0.74	0.01	0.52	0.50	0.44	0.38	0.09	0.15	0.20	−0.01	0.06	0.06	0.06	−0.02	−0.01	0.03	−0.02
5	ESTIMATES_WDICT	0.18	−0.02	0.20	0.74	1.00	−0.01	0.73	0.37	0.52	0.46	0.13	0.14	0.24	−0.04	0.07	0.02	0.04	−0.04	−0.02	−0.02	−0.05
6	CF_FORECASTS	0.27	0.07	0.36	0.01	−0.01	1.00	−0.08	0.01	0.04	0.06	−0.04	0.00	0.30	0.12	0.08	0.07	0.03	−0.01	−0.06	−0.09	−0.04
7	LENGTH_CAP	0.07	0.00	0.10	0.52	0.73	−0.08	1.00	0.30	0.39	0.45	−0.01	0.16	0.15	−0.04	−0.07	−0.03	−0.05	−0.06	−0.01	−0.05	−0.08
8	FN_NDICT	0.15	0.03	0.15	0.50	0.37	0.01	0.30	1.00	0.76	0.63	0.01	0.14	0.20	−0.03	−0.03	0.10	0.08	−0.03	0.01	0.03	−0.05
9	FN_WDICT	0.19	0.02	0.22	0.44	0.52	0.04	0.39	0.76	1.00	0.81	0.07	0.17	0.29	−0.04	0.03	0.08	0.08	−0.04	−0.02	0.00	−0.06
10	LENGTH_FN	0.21	0.02	0.27	0.38	0.46	0.06	0.45	0.63	0.81	1.00	−0.10	0.19	0.36	−0.03	0.03	0.05	0.04	−0.05	−0.02	−0.06	−0.08
11	LM_UNCERTAINTY	0.02	0.00	−0.03	0.09	0.13	−0.04	−0.01	0.01	0.07	−0.10	1.00	−0.05	−0.07	0.14	−0.02	0.05	0.04	0.01	0.00	0.20	0.06
12	FOG	0.07	0.03	0.04	0.15	0.14	0.00	0.16	0.14	0.17	0.19	−0.05	1.00	0.06	0.10	0.00	0.06	0.03	0.01	−0.01	0.13	0.00
13	SIZE	0.69	0.02	0.64	0.20	0.24	0.30	0.15	0.20	0.29	0.36	−0.07	0.06	1.00	−0.01	0.14	0.26	0.28	−0.25	0.05	−0.04	−0.42
14	GROWTH	0.09	0.03	0.08	−0.01	−0.04	0.12	−0.04	−0.03	−0.04	−0.03	0.14	0.10	−0.01	1.00	−0.03	−0.11	−0.09	0.08	−0.01	0.16	0.13
15	SEGMENTS	0.07	−0.02	0.06	0.06	0.07	0.08	−0.07	−0.03	0.03	0.03	−0.02	0.00	0.14	−0.03	1.00	0.09	0.12	0.00	−0.08	−0.02	−0.01
16	INST_OWNERSHIP	0.19	0.01	0.11	0.06	0.02	0.07	−0.03	0.10	0.08	0.05	0.05	0.06	0.26	−0.11	0.09	1.00	0.39	−0.14	0.00	0.00	−0.27
17	MGMT_GUIDANCE	0.21	−0.04	0.11	0.06	0.04	0.03	−0.05	0.08	0.08	0.04	0.04	0.03	0.28	−0.09	0.12	0.39	1.00	−0.12	0.04	0.05	−0.20
18	10-K_NEWS	−0.11	0.00	−0.08	−0.02	−0.04	−0.01	−0.06	−0.03	−0.04	−0.05	0.01	0.01	−0.25	0.08	0.00	−0.14	−0.12	1.00	0.00	0.04	0.32
19	ADV	0.08	−0.03	0.01	−0.01	−0.02	−0.06	−0.01	0.01	−0.02	−0.02	0.00	−0.01	0.05	−0.01	−0.08	0.00	0.04	0.00	1.00	−0.03	0.00
20	R&D	0.06	0.05	−0.05	0.03	−0.02	−0.09	−0.05	0.03	0.00	−0.06	0.20	0.13	−0.04	0.16	−0.02	0.00	0.05	0.04	−0.03	1.00	0.14
21	STD_RET	−0.14	0.02	−0.12	−0.02	−0.05	−0.04	−0.08	−0.05	−0.06	−0.08	0.06	0.00	−0.42	0.13	−0.01	−0.27	−0.20	0.32	0.00	0.14	1.00

See Appendix A for variable definitions. ESTIMATES and LENGTH reported in Panels A and B are in raw (total count) format.
Correlations of 0.011, 0.013 and 0.018 (in absolute value) are significant at the 0.10, 0.05 and 0.01 levels (two-tailed test).

Table 4
Analysis of analyst following.

Panel A. Probit Analysis of FOLLOW

	Ranked estimates		Logged estimates	
	Narrow dict	Wide dict	Narrow dict	Wide dict
ESTIMATES	0.272**	0.454***	0.131**	0.240***
	(2.24)	(2.69)	(2.46)	(3.20)
LENGTH_CAP	−0.043	−0.253	−0.019	−0.109*
	(−0.40)	(−1.62)	(−0.50)	(−1.94)
FN	−0.041	−0.212	−0.027	−0.150
	(−0.31)	(−1.17)	(−0.37)	(−1.30)
LENGTH_FN	0.417***	0.546***	0.284***	0.382***
	(3.14)	(3.06)	(3.41)	(3.25)
LM_UNCERTAINTY	80.445***	79.925***	79.288***	78.618***
	(8.75)	(8.52)	(8.65)	(8.40)
FOG	0.068***	0.068***	0.063**	0.063***
	(2.75)	(2.78)	(2.56)	(2.58)
SIZE	0.514***	0.516***	0.513***	0.513***
	(22.22)	(22.34)	(22.49)	(22.59)
GROWTH	0.532***	0.526***	0.538***	0.533***
	(4.03)	(4.01)	(4.08)	(4.05)
SEGMENTS	−0.166***	−0.167***	−0.174***	−0.175***
	(−2.86)	(−2.87)	(−2.99)	(−3.02)
INST_OWNERSHIP	0.739***	0.744***	0.739***	0.741***
	(9.17)	(9.22)	(9.18)	(9.20)
MGMT_GUIDANCE	0.150***	0.150***	0.149***	0.149***
	(7.20)	(7.19)	(7.15)	(7.16)
10-K_NEWS	0.164	0.175	0.150	0.161
	(0.55)	(0.59)	(0.51)	(0.54)
ADV	1.193	1.264	1.218	1.253
	(1.40)	(1.48)	(1.42)	(1.46)
R&D	1.732***	1.759***	1.722***	1.755***
	(4.45)	(4.54)	(4.42)	(4.53)
STD_RET	0.807***	0.792***	0.787***	0.765***
	(3.03)	(2.97)	(2.95)	(2.86)
intercept	−5.724***	−5.721***	−8.109***	−8.256***
	(−8.49)	(−8.52)	(−9.19)	(−8.79)
N	24,675	24,675	24,675	24,675
pseudo R-sq	0.449	0.449	0.450	0.451

Panel B. Tobit Analysis of NANALYSTS

	Ranked estimates		Logged estimates	
	Narrow dict	Wide dict	Narrow dict	Wide dict
ESTIMATES	0.989**	0.990*	0.368**	0.432
	(2.46)	(1.80)	(2.04)	(1.64)
LENGTH_CAP	−0.283	−0.539	−0.085	−0.186
	(−0.76)	(−1.01)	(−0.63)	(−0.93)
FN	−0.095	−0.979	−0.095	−0.810**
	(−0.21)	(−1.59)	(−0.35)	(−1.99)
LENGTH_FN	−0.133	0.693	0.128	0.788*
	(−0.28)	(1.11)	(0.43)	(1.86)
LM_UNCERTAINTY	275.376***	288.809***	279.104***	294.361***
	(8.29)	(8.32)	(8.41)	(8.52)
FOG	0.268***	0.283***	0.264***	0.275***
	(3.16)	(3.32)	(3.10)	(3.25)
SIZE	4.183***	4.187***	4.173***	4.170***
	(48.67)	(48.70)	(49.31)	(49.49)
GROWTH	3.332***	3.300***	3.338***	3.296***
	(8.02)	(7.97)	(8.03)	(7.97)
SEGMENTS	−0.645***	−0.644***	−0.654***	−0.653***
	(−3.64)	(−3.63)	(−3.69)	(−3.69)
INST_OWNERSHIP	0.644***	0.663***	0.650***	0.669***
	(2.77)	(2.85)	(2.80)	(2.88)
MGMT_GUIDANCE	0.140***	0.141***	0.141***	0.142***
	(3.42)	(3.45)	(3.43)	(3.47)
10-K_NEWS	5.407***	5.447***	5.391***	5.416***
	(4.54)	(4.58)	(4.52)	(4.55)
ADV	11.501***	11.580***	11.624***	11.668***
	(3.93)	(3.95)	(3.96)	(3.98)
R&D	9.092***	9.248***	9.159***	9.299***
	(8.50)	(8.66)	(8.56)	(8.71)
STD_RET	9.148***	9.178***	9.025***	9.005***
	(9.09)	(9.10)	(9.01)	(8.95)

(*continued on next page*)

Table 4 (*continued*)

Panel B. Tobit Analysis of NANALYSTS

	Ranked estimates		Logged estimates	
	Narrow dict	Wide dict	Narrow dict	Wide dict
intercept	−34.615***	−35.032***	−35.595***	−38.570***
	(−15.36)	(−15.58)	(−11.46)	(−11.36)
N	24,675	24,675	24,675	24,675
Pseudo R-sq	0.176	0.176	0.176	0.176

Panel C. OLS analysis of NANALYSTS

	Ranked estimates		Logged estimates	
	Narrow dict	Wide dict	Narrow dict	Wide dict
ESTIMATES	0.742**	0.746	0.309**	0.244
	(2.25)	(1.63)	(2.11)	(1.05)
LENGTH_CAP	−0.613*	−0.788*	−0.219*	−0.231
	(−1.83)	(−1.81)	(−1.68)	(−1.35)
FN	−0.191	−0.886*	−0.159	−0.533*
	(−0.52)	(−1.89)	(−0.82)	(−1.73)
LENGTH_FN	−0.298	0.289	−0.121	0.271
	(−0.81)	(0.61)	(−0.46)	(0.77)
LM_UNCERTAINTY	170.752***	181.938***	172.291***	184.103***
	(5.57)	(5.61)	(5.62)	(5.64)
FOG	0.310***	0.317***	0.308***	0.314***
	(3.88)	(3.94)	(3.84)	(3.91)
SIZE	4.155***	4.150***	4.153***	4.150***
	(57.76)	(57.38)	(58.22)	(57.83)
GROWTH	2.722***	2.721***	2.731***	2.714***
	(7.16)	(7.12)	(7.18)	(7.10)
SEGMENTS	−0.711***	−0.710***	−0.716***	−0.714***
	(−4.57)	(−4.56)	(−4.60)	(−4.58)
INST_OWNERSHIP	0.199	0.213	0.208	0.219
	(0.71)	(0.76)	(0.75)	(0.79)
MGMT_GUIDANCE	0.111***	0.111***	0.111***	0.111***
	(2.80)	(2.81)	(2.80)	(2.82)
10-K_NEWS	5.968***	5.988***	5.964***	5.971***
	(5.27)	(5.29)	(5.26)	(5.27)
ADV	10.759***	10.730***	10.788***	10.733***
	(3.85)	(3.86)	(3.86)	(3.85)
R&D	6.832***	6.915***	6.851***	6.903***
	(7.88)	(8.01)	(7.91)	(8.00)
STD_RET	10.114***	10.105***	10.111***	10.103***
	(9.17)	(9.16)	(9.20)	(9.18)
intercept	−30.716***	−30.798***	−28.285***	−30.346***
	(−14.83)	(−14.95)	(−10.25)	(−10.15)
N	22,236	22,236	22,236	22,236
R-sq	0.647	0.647	0.647	0.647

Panel A reports Probit analysis of FOLLOW from an expanded sample in which firm-years not followed by analysts (NANALYSTS = 0) are added to the core sample of 22,236 shown in Table 1. Panel B reports Tobit analysis of NANALYSTS using the same expanded sample as in Panel A. Panel C reports OLS analysis of NANALYSTS using the core sample of 22,236 shown in Table 1.

The model is $y_{i,t+1} = \alpha + \beta_1 ESTIMATES_{i,t} + \sum_1^k \delta_k X_{k,i,t} + \varepsilon_{i,t+1}$ (1)

Standard errors are clustered by firm and the models include year and industry fixed effects.

NANALYSTS is the number of analysts issuing at least one forecast of year $t + 1$ annual earnings following the filing of the year t Form 10-K.

FOLLOW is a dichotomous variable =1 if NANALYSTS>0, =0 otherwise.

The test variable accrual estimation intensity (ESTIMATES) is either ESTIMATES_NDICT (reported in the "Narrow Dict" column) or ESTIMATES_WDICT (reported in the "Wide Dict" column). The test variable ESTIMATES is percentile-ranked in the "Ranked Estimates" column and is natural log-transformed in the "Logged Estimates" column.

$X_{k,\ i,\ t}$ is one of k control variables.

The control variable FN is either FN_NDICT (reported in the "Narrow Dict" column) or FN_WDICT (reported in the "Wide Dict" column). The control variable FN is percentile-ranked in the "Ranked Estimates" column and is natural log-transformed in the "Logged Estimates" column.

The control variables LENGTH_CAP and LENGTH_FN are percentile-ranked in the "Ranked Estimates" column and are natural log-transformed in the "Logged Estimates" column.

Appendix A provides detailed definitions for all variables.

*, **, and *** denote two-tailed significance levels of 0.10, 0.05 and 0.01, respectively.

Table 5
Analysis of analysts' private information.

Panel A. Model estimates

	Ranked estimates				Logged estimates			
	Narrow dict	Narrow dict	Wide dict	Wide dict	Narrow dict	Narrow dict	Wide dict	Wide dict
ESTIMATES	−0.005	−0.040**	−0.035*	−0.053**	−0.002	−0.016**	−0.017*	−0.025**
	(−0.36)	(−2.27)	(−1.86)	(−2.46)	(−0.24)	(−1.97)	(−1.79)	(−2.32)
ESTIMATES×CF_FORECASTS		0.188***		0.090*		0.078***		0.042
		(3.33)		(1.69)		(2.98)		(1.55)
CF_FORECASTS		−0.043		0.012		−0.115*		−0.103
		(−1.18)		(0.35)		(−1.90)		(−1.00)
LENGTH_CAP	0.005	0.006	0.025	0.026	0.002	0.002	0.009	0.010
	(0.39)	(0.45)	(1.41)	(1.48)	(0.29)	(0.35)	(1.37)	(1.44)
FN	0.012	0.013	0.049**	0.051***	0.008	0.008	0.030**	0.030**
	(0.79)	(0.88)	(2.57)	(2.65)	(1.00)	(1.04)	(2.40)	(2.42)
LENGTH_FN	−0.012	−0.012	−0.039*	−0.040**	−0.011	−0.010	−0.029**	−0.029*
	(−0.81)	(−0.79)	(−1.95)	(−2.00)	(−0.98)	(−0.91)	(−1.97)	(−1.96)
LM_UNCERTAINTY	−0.809	−0.775	−1.182	−1.183	−0.824	−0.775	−1.078	−1.060
	(−0.61)	(−0.59)	(−0.88)	(−0.88)	(−0.63)	(−0.59)	(−0.80)	(−0.78)
FOG	0.002	0.002	0.002	0.001	0.002	0.002	0.002	0.002
	(0.50)	(0.46)	(0.48)	(0.40)	(0.52)	(0.47)	(0.52)	(0.44)
SIZE	0.011***	0.009***	0.011***	0.010***	0.011***	0.009***	0.012***	0.010***
	(4.12)	(3.45)	(4.31)	(3.62)	(4.17)	(3.48)	(4.40)	(3.71)
GROWTH	0.001	0.001	0.001	−0.000	0.001	0.001	0.001	−0.001
	(0.07)	(0.04)	(0.09)	(−0.01)	(0.07)	(0.04)	(0.06)	(−0.04)
SEGMENTS	−0.005	−0.005	−0.005	−0.005	−0.005	−0.005	−0.005	−0.005
	(−0.81)	(−0.87)	(−0.77)	(−0.79)	(−0.81)	(−0.87)	(−0.74)	(−0.76)
INST_OWNERSHIP	−0.019	−0.018	−0.019	−0.019	−0.019	−0.018	−0.019	−0.019
	(−1.49)	(−1.39)	(−1.51)	(−1.47)	(−1.50)	(−1.39)	(−1.51)	(−1.48)
MGMT_GUIDANCE	−0.006***	−0.006***	−0.006***	−0.006***	−0.006***	−0.006***	−0.006***	−0.006***
	(−3.46)	(−3.49)	(−3.45)	(−3.47)	(−3.45)	(−3.48)	(−3.45)	(−3.46)
10-K_NEWS	0.005	0.007	0.007	0.008	0.005	0.006	0.009	0.008
	(0.07)	(0.08)	(0.09)	(0.09)	(0.07)	(0.07)	(0.11)	(0.10)
ADV	−0.240**	−0.239**	−0.232**	−0.227**	−0.242**	−0.241**	−0.232**	−0.228**
	(−2.26)	(−2.26)	(−2.17)	(−2.13)	(−2.28)	(−2.27)	(−2.18)	(−2.14)
R&D	0.158***	0.165***	0.157***	0.163***	0.158***	0.164***	0.157***	0.163***
	(4.13)	(4.33)	(4.14)	(4.31)	(4.12)	(4.31)	(4.13)	(4.30)
STD_RET	0.166***	0.155**	0.168***	0.157**	0.166***	0.155**	0.170***	0.159**
	(2.66)	(2.49)	(2.70)	(2.52)	(2.67)	(2.49)	(2.73)	(2.56)
intercept	0.218**	0.245***	0.218**	0.240**	0.281**	0.312***	0.327***	0.364***
	(2.29)	(2.60)	(2.26)	(2.51)	(2.38)	(2.65)	(2.60)	(2.89)
N	16,209	16,209	16,209	16,209	16,209	16,209	16,209	16,209
R-sq	0.034	0.035	0.034	0.035	0.034	0.035	0.034	0.035

Panel B. Marginal effects of ESTIMATES for lowest and highest scrutiny of accrual estimates

	Ranked estimates	
	Narrow dict	Wide dict
Difference in conditional mean of DV between highest percentile of ESTIMATES and lowest percentile of ESTIMATES in firm-years with least scrutiny of estimates (i.e, β_1)	−0.040**	−0.053**
Difference in conditional mean of DV between highest percentile of ESTIMATES and lowest percentile of ESTIMATES in firm-years with greatest scrutiny of estimates (i.e, $\beta_1 + \beta_2$)	0.147***	0.038

The models are $y_{i,t+1} = \alpha + \beta_1 ESTIMATES_{i,t} + \sum_1^k \delta_k X_{k,i,t} + \varepsilon_{i,t+1}$ (1)
and
$y_{i,t+1} = \alpha + \beta_1 ESTIMATES_{i,t} + \beta_2 ESTIMATES \times CF\_FORECAST_{i,t} + \beta_3 CF\_FORECAST_{i,t} + \sum_k \delta_k X_{k,i,t} + \varepsilon_{i,t+1}$ (2)

Both are estimated by ordinary least squares with standard errors clustered by firm and include year and industry fixed effects. The dependent variable y is PRIVATE.
PRIVATE is

$$1 - \frac{SE - \frac{DISPERSION}{NANALYSTS}}{UNCERTAINTY}.$$

where
UNCERTAINTY is

$$\left(1 - \frac{1}{NANALYSTS}\right) \times DISPERSION + SE$$

DISPERSION is the variance of analysts' estimate of fiscal year $t+1$ earnings using each analyst's first revised forecast of year $t+1$ earnings made in the 90-day period following the filing of the year t 10-K, deflated by share price taken from the last IBES monthly update immediately preceding fiscal year t fiscal year end.
SE is the squared difference between IBES actual fiscal year $t+1$ earnings and the mean analysts' forecast of fiscal year $t+1$ earnings, deflated by share price taken from the last IBES monthly update immediately preceding year t fiscal year end. The mean forecast is based on each analyst's first revised forecast of year $t+1$ earnings following the filing of the year t Form 10-K.
The test variable accrual estimation intensity (ESTIMATES) is either ESTIMATES_NDICT (reported in the "Narrow Dict" column) or ESTIMATES_WDICT (reported in the "Wide Dict" column). The test variable ESTIMATES is percentile-ranked in the "Ranked Estimates" column and is natural log-transformed in the "Logged Estimates"

column.
$X_{k,\ i,\ t}$ is one of k control variables.
The control variable FN is either FN_NDICT (reported in the "Narrow Dict" column) or FN_WDICT (reported in the "Wide Dict" column). The control variable FN is percentile-ranked in the "Ranked Estimates" column and is natural log-transformed in the "Logged Estimates" column.
The control variables LENGTH_CAP and LENGTH_FN are percentile-ranked in the "Ranked Estimates" column and are natural log-transformed in the "Logged Estimates" column.
Appendix A provides detailed definitions for all variables.
*, **, and *** denote two-tailed significance levels of 0.10, 0.05 and 0.01, respectively.

the test variable in natural log form.

In addition, Tables 5 and 6 include a Panel B, which reports (for the percentile-rank specification only) tests of the sum of coefficients, which measure the marginal effect of ESTIMATES_xx on the dependent variable under conditions of least analyst scrutiny of estimates (CF_FORECAST = 0) and highest scrutiny of accounting estimates by analysts (CF_FORECAST = 1). Panel B confines the marginal effect analysis to the percentile-rank specification of ESTIMATES_xx, since that specification lends itself to straightforward assessment of economic significance. More specifically, because percentile-rank ESTIMATES_xx ranges from 0 to 1, the marginal effect of ESTIMATES_xx reported in Panel B represents the difference in the conditional mean of the dependent variable between firm-years with the lowest and highest levels of ESTIMATES_xx.

4.2. Regression results for test of H1: Analyst following

Table 4 reports regression results of estimating model (1) using both a dichotomous measure of analyst following (FOLLOW) and a count measure of analyst following (NANALYSTS). FOLLOW = 1 if NANALYSTS >0, otherwise 0.

Table 4 includes three panels of model estimates. Panel A reports the probit analysis of FOLLOW from an expanded sample in which firm-years not followed by analysts (NANALYSTS = 0) are added to the core sample of 22,236 shown in Table 1. Panel B reports the tobit analysis of NANALYSTS using the same expanded sample as in Panel A. Panel C reports the OLS analysis of NANALYSTS using the core sample of 22,236 shown in Table 1.[7]

Statistical inferences are similar in all three panels. Firms with more 10-K discussion of uncertainty, larger firms, high-growth firms, firms with more voluntary disclosure (as proxied by management guidance), firms with higher amounts of advertising and R&D expenditures, firms with more informative 10-K reports, and firms with more stock return volatility are associated with greater analyst following. Analyst following is also greater for firms with less readable 10-Ks. Firms with more segments are associated with lower analyst following.

The coefficient β_1 on ESTIMATES forms a test of H1. In each of the four model specifications reported in the Panel A probit analysis, ESTIMATES is positively associated with the probability that a firm is followed by analysts. In three of the four specifications reported in the tobit analysis (Table 4, Panel B), ESTIMATES is at least marginally positively associated with analyst following (logged ESTIMATES based on the wide dictionary specification just misses statistical significance at p = .101). In the OLS analysis (Panel C), ESTIMATES is positively associated with analyst following when estimates is based on the narrow dictionary specification; i.e., the first and third columns. Percentile-ranked ESTIMATES based on the wide dictionary specification reported in the second column of Panel C, just misses statistical significance at the 10% level (p = .102). The natural log form of ESTIMATES based on the wide dictionary specification reported in the fourth column of Panel C does not approach statistical significance.

Given the percentile-rank specification and the narrow dictionary definition, the coefficient value of 0.989 (0.742) on ESTIMATES in Panel B (Panel C) implies that moving from the smallest to the largest sample value of ESTIMATES moves the conditional mean of NANALYSTS by 0.742 (0.989), that is, attracting approximately one additional analyst. In Panel A, the marginal effect of ESTIMATES on the probability that a firm is followed is around 3% (untabulated), implying that moving from the smallest to the largest sample value of ESTIMATES moves the probability by around 3%. Thus, although statistically significant, the economic effect of accounting estimation intensity on analyst following is relatively modest. Keeping in mind that estimates could increase investors' demand for analyst services (leading to *greater* analyst following) or increase analysts' information acquisition and processing costs (leading to *lower* analyst following), there are two possible interpretations of this finding. First, it is possible that accounting estimates have little impact on investors' demand for analyst services and also little impact on analysts' information acquisition and processing costs. Alternatively, it is possible that accounting estimates meaningfully affect both investors' demand for analyst services and analysts' information acquisition and processing costs, and the effects offset each other.

Ideally, we would pinpoint which of these two competing explanations explain our results through tests that would quantify the effects of accounting estimates on demand for analysts' services, and quantify the effects of accounting estimates on the supply of analysts' services (through effects on analysts' information acquisition and processing costs). To do so credibly, however, would require estimating structural-form models that tease out the demand and supply effects separately. Unfortunately, such structural-form models are missing from the literature, which rely upon reduced-form models of the type we use. Our approach and interpretation of results is consistent with the recent extant literature (e.g., Lehavy et al., 2011; Li & Yang, 2016), which similarly reports statistically significant yet economically modest marginal effects.

4.3. Regression results for test of H2: Analysts' private information

Table 5 – Panel A presents regression results of estimating models (1) and (2) using analysts' private information as the dependent variable. As a reminder, model (1) is the specification without the ESTIMATES×CF_FORECASTS interaction term, while model (2) includes the interaction term. Also, results reported under the "Narrow Dict" ("Wide Dict") columns use ESTIMATES_NDICT (ESTIMATES_WDICT) as the test variable, while results reported under the "Ranked Estimates" ("Logged Estimates") column specify the test variable in percentile rank (natural log) form.

Inspection of Table 5 – Panel A reveals that firms with more voluntary disclosure and advertising expenditures are associated with less private information embedded in analysts' forecasts. Larger firms, firms with more R&D investment, and return volatility are associated with increased private information embedded in analysts' forecasts.

In Table 5 – Panel A, the test variable ESTIMATES in the model (1)

[7] Due to constraints on computational resources, ESTIMATES derived from the machine learning technique (described above and in Appendix B) is available only for the core sample of 22,236 observations that are followed by analysts. It is not available for firms not followed by analysts. To overcome this limitation, in analyses that include firms not followed by analysts (i.e., Table 4 Panels A and B), ESTIMATES is based on the count of the number of estimate-related words in the CAP, which is available for both groups of firms. ESTIMATES based on word counts is consistent with some prior research (e.g., Gordon et al., 2018).

Table 6
Analysis of informativeness of analysts' reports.

Panel A. Model estimates

	Ranked estimates				Logged estimates			
	Narrow dict	Narrow dict	Wide dict	Wide dict	Narrow dict	Narrow dict	Wide dict	Wide dict
ESTIMATES	0.014***	0.006	0.020***	0.011**	0.006***	0.003	0.010***	0.006**
	(3.49)	(1.42)	(3.74)	(2.01)	(3.68)	(1.61)	(3.62)	(2.13)
ESTIMATES×CF_FORECASTS		0.045**		0.054**		0.020**		0.024**
		(2.05)		(2.46)		(2.03)		(2.13)
CF_FORECASTS		−0.006		−0.008		−0.025		−0.070*
		(−0.48)		(−0.79)		(−1.22)		(−1.76)
LENGTH_CAP	−0.008**	−0.008**	−0.016***	−0.015***	−0.002	−0.002	−0.005***	−0.005***
	(−2.07)	(−1.97)	(−3.18)	(−3.12)	(−1.63)	(−1.53)	(−2.68)	(−2.62)
FN	−0.004	−0.003	−0.008	−0.007	−0.004*	−0.004*	−0.007*	−0.007*
	(−0.83)	(−0.76)	(−1.40)	(−1.34)	(−1.73)	(−1.70)	(−1.83)	(−1.84)
LENGTH_FN	0.012***	0.012***	0.015***	0.014**	0.010***	0.010***	0.012***	0.012***
	(2.68)	(2.67)	(2.58)	(2.56)	(3.09)	(3.13)	(2.80)	(2.85)
LM_UNCERTAINTY	0.831**	0.848**	0.790**	0.789**	0.829**	0.850**	0.796**	0.800**
	(2.28)	(2.33)	(2.06)	(2.06)	(2.28)	(2.34)	(2.07)	(2.08)
FOG	0.002**	0.002**	0.002**	0.002**	0.002**	0.002*	0.002**	0.002*
	(2.07)	(2.01)	(2.09)	(1.99)	(2.00)	(1.93)	(2.04)	(1.94)
SIZE	0.045***	0.045***	0.045***	0.045***	0.045***	0.045***	0.045***	0.045***
	(56.08)	(54.63)	(55.83)	(54.21)	(56.92)	(55.50)	(56.66)	(55.09)
GROWTH	0.022***	0.021***	0.023***	0.022***	0.022***	0.022***	0.023***	0.022***
	(4.79)	(4.71)	(4.95)	(4.87)	(4.81)	(4.73)	(4.95)	(4.86)
SEGMENTS	−0.008***	−0.008***	−0.008***	−0.008***	−0.008***	−0.008***	−0.008***	−0.008***
	(−4.36)	(−4.35)	(−4.39)	(−4.37)	(−4.40)	(−4.40)	(−4.45)	(−4.42)
INST_OWNERSHIP	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001
	(−0.43)	(−0.33)	(−0.42)	(−0.34)	(−0.41)	(−0.31)	(−0.43)	(−0.37)
MGMT_GUIDANCE	−0.001**	−0.001**	−0.001**	−0.001**	−0.001**	−0.001**	−0.001**	−0.001**
	(−2.23)	(−2.26)	(−2.30)	(−2.41)	(−2.25)	(−2.29)	(−2.29)	(−2.40)
10-K_NEWS	0.104***	0.104***	0.104***	0.104***	0.104***	0.104***	0.104***	0.103***
	(7.37)	(7.44)	(7.41)	(7.46)	(7.36)	(7.42)	(7.36)	(7.39)
ADV	0.017	0.017	0.016	0.017	0.018	0.019	0.016	0.016
	(0.59)	(0.61)	(0.57)	(0.60)	(0.62)	(0.64)	(0.55)	(0.58)
R&D	−0.002	−0.000	−0.000	0.002	−0.002	−0.000	−0.000	0.001
	(−0.21)	(−0.01)	(−0.01)	(0.18)	(−0.20)	(−0.00)	(−0.04)	(0.15)
STD_RET	0.127***	0.125***	0.126***	0.124***	0.126***	0.124***	0.125***	0.123***
	(9.11)	(8.99)	(9.07)	(8.97)	(9.11)	(8.98)	(9.04)	(8.94)
intercept	−0.269***	−0.262***	−0.265***	−0.256***	−0.329***	−0.320***	−0.329***	−0.312***
	(−11.05)	(−10.88)	(−11.02)	(−10.79)	(−9.94)	(−9.68)	(−9.24)	(−8.64)
N	20,984	20,984	20,984	20,984	20,984	20,984	20,984	20,984
R-sq	0.617	0.618	0.617	0.618	0.617	0.618	0.617	0.618

Panel B. Marginal effects of ESTIMATES for lowest and highest scrutiny of accrual estimates

	Ranked estimates	
	Narrow dict	Wide dict
Difference in conditional mean of DV between highest percentile of ESTIMATES and lowest percentile of ESTIMATES in firm-years with least scrutiny of estimates (i.e, β_1)	0.006	0.011**
Difference in conditional mean of DV between highest percentile of ESTIMATES and lowest percentile of ESTIMATES in firm-years with greatest scrutiny of estimates (i.e, $\beta_1 + \beta_2$)	0.051***	0.065***

The models are $y_{i,t+1} = \alpha + \beta_1 ESTIMATES_{i,t} + \sum_1^k \delta_k X_{k,i,t} + \varepsilon_{i,t+1}$ (1)
and
$$y_{i,t+1} = \alpha + \beta_1 ESTIMATES_{i,t} + \beta_2 ESTIMATES \times CF\_FORECAST_{i,t} + \beta_3 CF\_FORECAST_{i,t} + \sum_k \delta_k X_{k,i,t} + \varepsilon_{i,t+1} \quad (2)$$

Both are estimated by ordinary least squares with standard errors clustered by firm and include year and industry fixed effects. The dependent variable y is INFORMATIVENESS.

INFORMATIVENESS is the proportion of information flowing into a firm's stock price around analysts' earnings forecast reports as a fraction of all information flowing into stock price during the period between the fiscal year *t* 10-K filing and the fiscal year $t+1$ year-end. The numerator of the proportion is the summed one-day absolute market-adjusted returns (value weighted index) on all analysts' forecast report dates using only forecast reports issued the day after the fiscal year *t* 10-K filing date through the $t+1$ fiscal year end. The denominator of the proportion is the summed absolute daily market-adjusted returns beginning the day after the year *t* 10-K filing date and ending on the year $t+1$ fiscal year end. Multiple analysts' reports issued on a single day are treated as a single report and we require at least 90 trading days in calculating the denominator.

The test variable accrual estimation intensity (ESTIMATES) is either ESTIMATES_NDICT (reported in the "Narrow Dict" column) or ESTIMATES_WDICT (reported in the "Wide Dict" column). The test variable ESTIMATES is percentile-ranked in the "Ranked Estimates" column and is natural log-transformed in the "Logged Estimates" column.

$X_{k,\ i,\ t}$ is one of k control variables.

The control variable FN is either FN_NDICT (reported in the "Narrow Dict" column) or FN_WDICT (reported in the "Wide Dict" column). The control variable FN is percentile-ranked in the "Ranked Estimates" column and is natural log-transformed in the "Logged Estimates" column.

The control variables LENGTH_CAP and LENGTH_FN are percentile-ranked in the "Ranked Estimates" column and are natural log-transformed in the "Logged Estimates" column.

Appendix A provides detailed definitions for all variables.
*, **, and *** denote two-tailed significance levels of 0.10, 0.05 and 0.01, respectively.

specification is insignificant when based on the narrow dictionary definition and, contrary to the positive association predicted in H2, is negative and marginally significant when ESTIMATES is based on the wide dictionary definition. One possible explanation for this unexpected negative association is that accounting estimates increase the precision of both analysts' private information and common information. PRIVATE is the ratio of the precision of analysts' private information to the precision of analysts' total information (i.e., both private and common information). Hence, if the effect of accounting estimates on the precision of common information dominates the effect on the precision of analysts' private information, a negative association could be observed between ESTIMATES and PRIVATE.

In the model (2) specification, the test variable ESTIMATES is negative and significant when ESTIMATES is based on either the narrow or wide dictionary definitions, while the interaction term ESTIMATES×CF_FORECASTS is positive and at least marginally significant in three of the four model (2) specifications. Hence, model (2) estimates suggest that ESTIMATES is associated with lower private information in analysts' forecasts when accounting estimates are given little scrutiny by analysts, but with increasing levels of private information in analysts' forecasts when analysts increase their scrutiny of accountings estimates. Table 5 – Panel B explores this in greater detail.

Table 5 – Panel B quantifies the marginal effect of ESTIMATES on PRIVATE using the percentile-rank specification of ESTIMATES in model (2). When accounting estimates are given little scrutiny by analysts, the conditional mean value of PRIVATE in the highest percentile rank of ESTIMATES_NDICT (ESTIMATES_WDICT) is −0.040 less (−0.053 less) than in the lowest percentile rank of ESTIMATES_NDICT (ESTIMATES_WDICT). That is, when analysts give little scrutiny to accounting estimates, these estimates are associated with less precise private information in analysts' forecasts. However, when analysts give the greatest scrutiny to accounting estimates (i.e., when CF_FORECASTS = 1), the conditional mean value of PRIVATE in the highest percentile rank of ESTIMATES_NDICT (ESTIMATES_WDICT) is 0.147 more (0.038 more) than in the lowest percentile rank of ESTIMATES_NDICT (ESTIMATES_WDICT). That is, accounting estimates are associated with more precise private information in analysts' forecasts when analysts give the greatest scrutiny to accounting estimates. The 0.147 increase in the conditional mean for the narrow dictionary definition is 31.8% of the unconditional mean value of PRIVATE, as reported in Table 3 – Panel A (i.e., 0.147/0.4625 = 0.3178), which we view as economically significant. The 0.038 increase in the conditional mean for the wide dictionary definition is 8.2% of the unconditional mean value of PRIVATE, as reported in Table 3 – Panel A (i.e., 0.038/0.4625 = 0.082), which is much less economically significant and also is statistically insignificant.

The Table 5 findings are consistent with the idea that the act of forecasting accruals causes analysts to scrutinize the estimates embedded in and excluded from those accruals, which prompts analysts to engage in additional information acquisition and processing to better understand the valuation implications of the accounting estimates. In turn, this causes analyst forecasts to be informed by more analyst-specific (idiosyncratic) beliefs, which is manifested in greater private information embedded in analysts' earnings forecasts.

In summary, the alternative form of H2 is not supported when the effects of ESTIMATES are held constant across firms, but it is supported when the effect of ESTIMATES is allowed to vary cross-sectionally (i.e., model 2) and ESTIMATES is based on the narrow dictionary definition. The evidence is consistent with the idea that greater scrutiny of estimates causes more idiosyncratic analyst beliefs to be embedded in analyst forecasts (per unit of estimates), which increases the level of private information embedded in analysts' forecasts.[8] However, our inferences are somewhat sensitive to the measure of ESTIMATES that is used.

4.4. Regression results for test of H3: Informativeness of analyst reports

Table 6 – Panel A presents regression estimates of models (1) and (2) using the informativeness of analysts' reports as the dependent variable. Firms with lengthier footnotes, firms with more discussion of uncertainty in their 10-K, larger firms, firms with less readable financial reports, growth firms, firms with greater reaction to the 10-K release, and firms with more return volatility are associated with more informative analyst reports. Firms with more operating segments and more voluntary disclosures are associated with less informative analyst reports.

In each of the model (1) specifications, ESTIMATES loads positive and significant, indicating that the informativeness of analysts' reports increases as accounting estimates increase. In the model (2) specifications, ESTIMATES loads positive and is either significant or approaches significance in each of the specifications. Further, the coefficient on the interaction term ESTIMATES×CF_FORECASTS is positive and significant in each of the model (2) specifications, indicating that analysts' scrutiny of accounting estimates increases the effect of estimates on the informativeness of analysts' reports. Table 6 – Panel B explores this in greater detail.

Table 6 – Panel B quantifies the marginal effect of ESTIMATES on INFORMATIVENESS using the percentile-rank specification of ESTIMATES in model (2). When accounting estimates are given little scrutiny by analysts, the conditional mean value of INFORMATIVENESS in the highest percentile rank of ESTIMATES_NDICT (ESTIMATES_WDICT) is 0.006 more (0.011 more) than in the lowest percentile rank of ESTIMATES_NDICT (ESTIMATES_WDICT). That is, accounting estimates are associated with more informative analyst reports when analysts give little scrutiny to accounting estimates, but this result is statistically significant only for ESTIMATES_WDICT. However, when analysts give the greatest scrutiny to accounting estimates (i.e., when CF_FORECASTS = 1), the conditional mean value of INFORMATIVENESS in the highest percentile rank of ESTIMATES_NDICT (ESTIMATES_WDICT) is 0.051 more (0.065 more) than in the lowest percentile rank of ESTIMATES_NDICT (ESTIMATES_WDICT). That is, accounting estimates are associated with even more informative analyst reports when analysts give the greatest scrutiny to accounting estimates, and this result holds for both ESTIMATES_NDICT and ESTIMATES_WDICT. The 0.051 increase in the conditional mean, noted above, is 36.4% of the unconditional mean value of INFORMATIVENESS as reported in Table 3 – Panel A (i.e., 0.051/0.1399 = 0.364). The 0.065 increase in the conditional mean, noted above, is 46.5% of the unconditional mean value of INFORMATIVENESS as reported in Table 3 – Panel A (i.e., 0.065/0.1399 = 0.465). We regard both increases as economically significant.

[8] Gordon et al. (2018) find that estimates are associated with more dispersed and less accurate analyst forecasts and that greater specificity in the CAP disclosure dampens this effect. In untabulated analysis, we similarly analyze the association between estimates and forecast dispersion and forecast accuracy. Our data confirm the findings in Gordon et al. that accounting estimates are associated with more forecast dispersion and greater forecast error. However, we find that greater scrutiny of estimates by analysts dampens the deleterious effects of estimates on dispersion and forecast accuracy. In other words, the more analysts scrutinize estimates, the smaller the increase in dispersion and forecast error. This finding is likely attributable to the increased precision of private information embedded in analysts' forecasts resulting from greater analyst scrutiny as documented in Table 5.

The evidence is consistent with the idea that accounting estimates are associated with incrementally more informative analysts' earnings forecasts, and this association increases as analysts give greater scrutiny to accounting estimates. Hence, null form H3 is rejected.

In summary, the data reject null-form H1 using the narrow dictionary definition of accounting estimates, reject null-form H2 using the narrow dictionary definition of accounting estimates under conditions of high analyst scrutiny of accounting estimates, and generally reject null-form H3 using all definitions of accounting estimates under both conditions of high and low scrutiny of accounting estimates, with the effect of accounting estimates amplified under conditions of high analyst scrutiny of accounting estimates. The results are consistent with accounting estimates stimulating investor demand for analysts' services and increasing the informativeness of those services. The enhanced informativeness of those services stem from analysts' assessment and interpretation of those estimates, which introduces idiosyncratic beliefs into earnings forecasts, leading to greater precision of private information in analysts' earnings forecasts.

4.5. Additional analysis: Event type specification

Our primary analysis of PRIVATE relates the *level* of accounting estimates reported in the year t 10-K (filed early in year $t+1$) to the *level* of PRIVATE as measured from analysts' first forecast of year $t+1$ earnings issued after the filing of the year t 10-K. A more demanding test would relate the *change* in PRIVATE (i.e., the level of PRIVATE immediately after the 10-K filing as compared to the level immediately before the 10-K filing) to the *change* in the level of accounting estimates (the level of accounting estimates measured from the period t 10-K as filed in year $t+1$ less the level of accounting estimates measured from the year $t-1$ 10-K as filed in year t). Results obtained in the event-type change-form analysis would provide compelling confirmatory evidence suggestive of causation. Failure to obtain confirmatory evidence would suggest that our main results should be regarded more cautiously.

Table 7 reports such event type tests in the analysis of PRIVATE.[9] Each of the explanatory variables in Table 7, models 1 and 2, are in change-form specification (i.e., the value from year t less the value from year $t-1$). The dependent variable ΔPRIVATE is measured as PRIVATE_AFTER less PRIVATE_BEFORE, where PRIVATE_AFTER is the value of PRIVATE calculated using analysts' first forecast of year $t+1$ earnings issued after the year t 10-K filing in early year $t+1$, while PRIVATE_BEFORE is the value of PRIVATE calculated using analysts' last forecast of year $t+1$ earnings issued immediately before the year t 10-K filing in early year $t+1$.

Inspection of Table 7 Panel A reveals that ΔESTIMATES is positive and marginally significant in model (1) based on the narrow dictionary definition of accounting estimates, and positive and significant in model (2) based only on the narrow dictionary definition of accounting estimates in natural log form. Hence, there is only modest evidence to reject null form H2 on average (i.e., model 1) and under conditions of low analyst scrutiny of accounting estimates (i.e., model 2). However, the interaction term ΔESTIMATES×ΔCF_FORECASTS is positive and at least marginally significant, suggesting that analysts' heightened scrutiny of accounting estimates impacts the marginal effect of ΔESTIMATES. Table 7 – Panel B explores this in greater detail.

Table 7 – Panel B reveals that under low analyst scrutiny of accounting estimates, the difference in the conditional mean value of ΔPRIVATE between the highest and lowest percentile of ΔESTIMATES is 0.056 (0.000) when ESTIMATES is based on the narrow dictionary (wide dictionary) definition. Neither of these values are significantly different from zero. Hence, null form H2 is not rejected in this analysis under low analyst scrutiny of accounting estimates. Under high analyst scrutiny of estimates, the difference in the conditional mean value of ΔPRIVATE between the highest and lowest percentile of ΔESTIMATES is 0.648 (1.141) when ESTIMATES is based on the narrow dictionary (wide dictionary) definition. Both are significantly different from zero. Hence, null form H2 is rejected under the condition of high analyst scrutiny of accounting estimates.

In brief, the event-type tests generally confirm the inferences drawn from the main analysis of PRIVATE as reported in Table 5.

4.6. Sensitivity analyses

Table 8 reports results of additional analysis intended to probe the sensitivity of our results to (1) selection bias (Panel A) and (2) control for the amount and mix of accruals (Panel B).

The first issue—the possibility of selection bias—arises because of the concern that unobserved factors that cause analysts to follow a firm might also be associated with unobserved factors that influence the values of PRIVATE and INFORMATIVENESS. If such unobserved factors also are correlated with estimate intensity, then the results we report in our analyses of PRIVATE and INFORMATIVENESS might be driven by selection bias. We address the issue of selection bias using a Heckman (1979) type analysis. Specifically, a first-stage probit model is estimated from the expanded sample (described earlier) using FOLLOW as a dependent variable, where FOLLOW = 1 if a firm is followed by an analyst, 0 otherwise. The estimate of this first-stage probit model is reported in Table 4 – Panel A. Parameters obtained in the probit model are used to construct the inverse Mills' ratio, which is added as an explanatory variable to the second-stage regression specification. We then repeat the main analyses of PRIVATE and INFORMATIVENESS with models (1) and (2) augmented with the inverse Mills' ratio as an additional explanatory variable and report these results in Table 8 – Panel A. For brevity, we report only the coefficients on the test variables from models (1,2) and the inverse Mills' ratio. Results are broadly consistent with the main results reported in Tables 5 and 6, albeit slightly weaker. Thus, the sensitivity test does not suggest that our main results are driven by selection bias.

The second issue is the concern that our ESTIMATE test variable may simply be proxying for the amount and mix of accruals and not the CAP-related disclosures related to accounting estimates. That is, the amount and mix of accruals loom as a possible correlated omitted variable. We address the issue by adding these variables to models (1) and (2) and repeating the main analyses. We measure the amount of accruals as the absolute value of asset-deflated total accruals, and measure the mix of accruals as the absolute value of the receivables and inventory component of accruals deflated by total assets. We focus on these two components of accruals because prior research shows these components are embedded with significant subjectivity and are associated with SEC enforcement actions (Dechow, Ge, Larson, & Sloan, 2011). Regression estimates of models (1, 2), as augmented with total accruals and the mix of accruals as additional controls, are reported in Table 8 – Panel B. For brevity, we report only the coefficients on the test variables. Our main results are robust to the inclusion of total accruals and the mix of accruals as additional control variables and thus do not appear meaningfully affected by the presence or the absence of these controls.

[9] We do not regard event-type tests of NANALYSTS or INFORMATIVENESS as useful and do not pursue them. Event-type analysis of NANALYSTS is not useful because a decision to initiate or discontinue coverage is unlikely to be spawned by a single 10-K filing. Rather, the decision to initiate or discontinue coverage likely occurs when the analyst's net benefit of coverage reaches a tipping point due to changes in the cost versus benefit of coverage that slowly accrue over time. Event-type analysis of INFORMATIVENESS is not useful because INFORMATIVENESS measures the fraction of information flowing into stock prices over a relatively wide time period (i.e., the fiscal year) that stems from analysts' reports. By its nature, this metric is less amenable to measurement in an event-test setting.

Table 7
Analysis of change in analysts' private information in the 90-day period following 10-K release.

Panel A. Model estimates

	Ranked estimates				Logged estimates			
	Narrow dict	Narrow dict	Wide dict	Wide dict	Narrow dict	Narrow dict	Wide dict	Wide dict
ΔESTIMATES	0.058*	0.056	0.001	0.000	0.032**	0.030**	0.016	0.014
	(1.68)	(1.61)	(0.02)	(0.00)	(2.08)	(1.96)	(0.74)	(0.67)
ΔESTIMATES×ΔCF_FORECASTS		0.593*		1.141***		0.254*		0.513***
		(1.87)		(3.48)		(1.77)		(2.93)
ΔCF_FORECASTS		−0.033		−0.039		−0.032		−0.033
		(−0.79)		(−0.96)		(−0.78)		(−0.81)
ΔLENGTH_CAP	−0.012	−0.012	0.008	0.004	−0.000	−0.000	−0.000	−0.000
	(−0.34)	(−0.32)	(0.19)	(0.10)	(−0.98)	(−0.93)	(−0.59)	(−0.59)
ΔFN	−0.019	−0.021	0.011	0.008	−0.012	−0.012	−0.006	−0.007
	(−0.71)	(−0.78)	(0.40)	(0.29)	(−1.03)	(−1.07)	(−0.47)	(−0.55)
ΔLENGTH_FN	−0.016	−0.016	−0.019	−0.017	−0.000	−0.000	−0.000	−0.000
	(−1.08)	(−1.06)	(−1.21)	(−1.13)	(−0.82)	(−0.81)	(−0.89)	(−0.82)
ΔLM_UNCERTAINTY	0.344	0.380	0.385	0.376	0.288	0.326	0.304	0.319
	(0.13)	(0.14)	(0.14)	(0.14)	(0.11)	(0.12)	(0.11)	(0.12)
ΔFOG	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007
	(1.13)	(1.16)	(1.15)	(1.18)	(1.12)	(1.16)	(1.14)	(1.17)
ΔSIZE	0.001	0.002	0.001	0.001	0.002	0.002	0.001	0.001
	(0.14)	(0.19)	(0.10)	(0.12)	(0.16)	(0.19)	(0.10)	(0.13)
ΔGROWTH	0.087*	0.089**	0.088*	0.092**	0.088*	0.090**	0.088*	0.090**
	(1.92)	(1.97)	(1.93)	(2.03)	(1.94)	(1.98)	(1.94)	(1.99)
ΔSEGMENTS	−0.007	−0.006	−0.007	−0.006	−0.007	−0.006	−0.006	−0.006
	(−0.46)	(−0.43)	(−0.44)	(−0.40)	(−0.46)	(−0.43)	(−0.43)	(−0.38)
ΔINST_OWNERSHIP	0.023	0.023	0.023	0.023	0.023	0.022	0.023	0.022
	(1.14)	(1.12)	(1.14)	(1.13)	(1.13)	(1.10)	(1.13)	(1.11)
ΔMGMT_GUIDANCE	0.000	0.000	0.000	0.000	0.000	0.000	−0.000	0.000
	(0.03)	(0.06)	(0.00)	(0.02)	(0.04)	(0.07)	(−0.00)	(0.02)
Δ10-K_NEWS	0.081	0.083	0.082	0.083	0.081	0.083	0.081	0.083
	(0.92)	(0.95)	(0.93)	(0.94)	(0.92)	(0.94)	(0.92)	(0.94)
ΔADV	−0.245	−0.250	−0.232	−0.241	−0.252	−0.254	−0.234	−0.241
	(−0.45)	(−0.46)	(−0.43)	(−0.44)	(−0.46)	(−0.47)	(−0.43)	(−0.44)
ΔR&D	0.145	0.144	0.148	0.149	0.146	0.145	0.147	0.148
	(0.93)	(0.93)	(0.95)	(0.95)	(0.94)	(0.93)	(0.94)	(0.95)
ΔSTD_RET	0.070	0.072	0.071	0.073	0.071	0.072	0.072	0.074
	(0.88)	(0.90)	(0.89)	(0.91)	(0.88)	(0.91)	(0.91)	(0.93)
intercept	−0.139	−0.139	−0.138	−0.137	−0.144	−0.143	−0.141	−0.141
	(−1.39)	(−1.38)	(−1.37)	(−1.36)	(−1.43)	(−1.43)	(−1.40)	(−1.40)
N	11,986	11,986	11,986	11,986	11,986	11,986	11,986	11,986
R-sq	0.013	0.014	0.013	0.014	0.014	0.014	0.013	0.014

Panel B. Marginal effects of change in ESTIMATES for lowest and highest scrutiny of accrual estimates

	Ranked estimates	
	Narrow dict	Wide dict
Difference in conditional mean of DV between highest percentile of ESTIMATES and lowest percentile of ESTIMATES in firm-years with least scrutiny of estimates (i.e, β_1)	0.056	0.000
Difference in conditional mean of DV between highest percentile of ESTIMATES and lowest percentile of ESTIMATES in firm-years with greatest scrutiny of estimates (i.e, $\beta_1+\beta_2$)	0.648**	1.141***

The models are $\Delta y_{i,t+1}=\alpha+\beta_1\Delta ESTIMATES_{i,t}+\sum_1^k\delta_k\Delta X_{k,i,t}+\varepsilon_{i,t+1}$ Changes specification of Model (1)
and $\Delta y_{i,t+1}=\alpha+\beta_1\Delta ESTIMATES_{i,t}+\beta_2\Delta ESTIMATES\times\Delta CF_{FORECAST\,i,t}+\beta_3\Delta CF_{FORECAST\,i,t}+\sum_k\delta_k\Delta X_{k,i,t}+\varepsilon_{i,t+1}$ Changes specification of Model (2)

Both are estimated by ordinary least squares with standard errors clustered by firm and include year and industry fixed effects. The dependent variable y is ΔPRIVATE, where ΔPRIVATE is the value of PRIVATE in the 90 days following the filing of the year t (PRIVATE_AFTER) 10-K less the value of PRIVATE in the 90 days preceding the 10-K filing (PRIVATE_BEFORE). Appendix A defines PRIVATE_AFTER and PRIVATE_BEFORE.
For test and control variables, Δ denotes the change in value from fiscal year *t*-1 to fiscal year *t*.
The test variable accrual estimation intensity (ΔESTIMATES) is either ΔESTIMATES_NDICT (reported in the "Narrow Dict" column) or ΔESTIMATES_WDICT (reported in the "Wide Dict" column). The test variable ΔESTIMATES is percentile-ranked in the "Ranked Estimates" column and is natural log-transformed in the "Logged Estimates" column.
$\Delta X_{k,\ i,\ t}$ is one of k control variables.
The control variable ΔFN is either ΔFN_NDICT (reported in the "Narrow Dict" column) or ΔFN_WDICT (reported in the "Wide Dict" column). The control variable ΔFN is percentile-ranked in the "Ranked Estimates" column and is natural log-transformed in the "Logged Estimates" column.
The control variables ΔLENGTH_CAP and ΔLENGTH_FN are percentile-ranked in the "Ranked Estimates" column and are natural log-transformed in the "Logged Estimates" column.
Appendix A provides detailed definitions for all variables.
*, **, and *** denote two-tailed significance levels of 0.10, 0.05 and 0.01, respectively.

Table 8
Sensitivity analyses.

Panel A. Sensitivity to selection bias: heckman analysis of PRIVATE and INFORMATIVENESS

	DV=PRIVATE							
	Ranked estimates				Logged estimates			
	Narrow dict	Narrow dict	Wide dict	Wide dict	Narrow dict	Narrow dict	Wide dict	Wide dict
ESTIMATES	−0.017	−0.043**	−0.037*	−0.048**	−0.007	−0.017**	−0.020**	−0.024**
	(−1.08)	(−2.20)	(−1.86)	(−2.05)	(−0.95)	(−1.97)	(−2.02)	(−2.05)
ESTIMATES×CF_FORECASTS		0.134**		0.055		0.056**		0.020
		(2.25)		(0.94)		(2.04)		(0.68)
IMR	0.037	0.035	0.042*	0.042*	0.037	0.036	0.040	0.041
	(1.48)	(1.41)	(1.66)	(1.67)	(1.50)	(1.46)	(1.58)	(1.61)

	DV=INFORMATIVENESS							
	Ranked estimates				Logged estimates			
	Narrow dict	Narrow dict	Wide dict	Wide dict	Narrow dict	Narrow dict	Wide dict	Wide dict
ESTIMATES	0.020***	0.018***	0.033***	0.024***	0.009***	0.009***	0.016***	0.013***
	(6.36)	(4.68)	(8.10)	(5.13)	(6.65)	(5.05)	(7.65)	(5.45)
ESTIMATES×CF_FORECASTS		0.011		0.043***		0.004		0.015***
		(0.91)		(3.80)		(0.70)		(2.65)
IMR	0.053***	0.054***	0.057***	0.056***	0.051***	0.052***	0.054***	0.054***
	(10.09)	(10.27)	(10.51)	(10.37)	(9.80)	(10.01)	(10.22)	(10.19)

Panel B. Sensitivity to control for the amount and mix of accruals

	DV=NANALYSTS			
	Ranked estimates		Logged estimates	
	Narrow dict	Wide dict	Narrow dict	Wide dict
ESTIMATES	0.682**	0.777*	0.289**	0.261
	(2.09)	(1.73)	(1.98)	(1.13)

	DV=PRIVATE							
	Ranked estimates				Logged estimates			
	Narrow dict	Narrow dict	Wide dict	Wide dict	Narrow dict	Narrow dict	Wide dict	Wide dict
ESTIMATES	−0.003	−0.036**	−0.036*	−0.054**	−0.001	−0.014*	−0.018*	−0.025**
	(−0.24)	(−2.03)	(−1.91)	(−2.45)	(−0.14)	(−1.76)	(−1.85)	(−2.31)
ESTIMATES×CF_FORECASTS		0.176***		0.087		0.073***		0.039
		(3.09)		(1.59)		(2.77)		(1.42)

	DV=INFORMATIVENESS							
	Ranked estimates				Logged estimates			
	Narrow dict	Narrow dict	Wide dict	Wide dict	Narrow dict	Narrow dict	Wide dict	Wide dict
ESTIMATES	0.014***	0.006	0.021***	0.012**	0.006***	0.003*	0.010***	0.007**
	(3.47)	(1.46)	(3.92)	(2.23)	(3.65)	(1.68)	(3.80)	(2.36)
ESTIMATES×CF_FORECASTS		0.043**		0.051**		0.019*		0.022**
		(1.97)		(2.35)		(1.92)		(2.01)

Panel A reports analyses of PRIVATE, and INFORMATIVENESS that use Heckman (1979) control for selection bias. For brevity, results are reported for the test variables and the inverse Mills' ratio only.

In Panel A, a first-stage probit model is estimated using FOLLOW as a dependent variable, where FOLLOW = 1 if a firm is followed by an analyst, =0 otherwise (see Table 4 Panel A). The parameters obtained in the probit model are used to construct the inverse Mills' ratio which is added as an explanatory variable to the second-stage regression specification. For brevity, Panel A reports coefficient estimates on the test variables of interest and the inverse Mills' ratio for estimates of models (1, 2) as augmented with the inverse Mills' ratio as an explanatory variable.

Panel B reports analyses of NANALYSTS, PRIVATE, and INFORMATIVENESS in which the regression model is augmented by the addition of the absolute value of total accruals and the absolute value of the receivables and inventory components of total accruals, where both accrual metrics are deflated by total assets. For brevity, results are tabulated for the test variables only.

In each panel, the test variable accrual estimation intensity (ESTIMATES) is either ESTIMATES_NDICT (reported in the "Narrow Dict" column) or ESTIMATES_WDICT (reported in the "Wide Dict" column). The test variable ESTIMATES is percentile-ranked in the "Ranked Estimates" column and is natural log-transformed in the "Logged Estimates" column.

5. Concluding remarks

This paper contributes to the literature on qualitative disclosures in the critical accounting policies section of the MD&A. These disclosures pertain to managers' subjective accounting estimates of future uncertain events and are pervasive in financial reporting. Although intended to communicate managers' private informed beliefs, they can be uninformative (because of the difficulties in making reliable projections of the future) or even misleading (due to possible managerial opportunism).

We utilize textual analysis of the qualitative CAP disclosures to specify two new measures of accounting estimation intensity (AEI) and examine the relation between AEI of firms and analyst following, the private information incorporated in analysts' forecasts, and the informativeness of analyst reports. Prior research (Gordon et al., 2018; Levine & Smith, 2011) does not address whether the firm's AEI as reflected in CAP disclosures is associated with analyst following, the extent of private information incorporated in analysts' forecasts, and the information content of those forecasts. Using a narrow dictionary definition of accounting estimates, our findings indicate that AEI is associated with increased analyst following, consistent with the notion that estimation intensity raises the need for analyst coverage. Also, we find that estimation intensity increases the private information in analyst forecasts as analyst scrutiny of accounting estimates increases. Lastly, consistent with increased collective analyst activity (as proxied by analyst coverage) and the incorporation of private information in earnings forecasts, we find estimation intensity to be associated with increased informativeness of analyst reports, particularly as analyst scrutiny of estimates increases. In sum, although Chen et al. (2021) and Roh (2018) suggest that accounting estimates can be problematic for investors, our study finds that analysts can help ameliorate these problems by providing increased coverage, by scrutinizing accounting estimates, and by incorporating their own private information in more informative earnings forecasts.

Declaration of Competing Interest

The authors have no conflict of interest to disclose.

Appendix A. Model variables

Variable	Definition	Source/Prior research
Dependent variables		
NANALYSTS	Number of analysts issuing at least one forecast of year $t+1$ annual earnings following the filing of the year t Form 10-K. Proxy for the collective effort of analyst community in following a firm	IBES Detail file Bhushan (1989) O'Brien and Bhushan (1990) Brennan and Subrahmanyam (1995) Barth et al. (2001)
FOLLOW	=1 if NANALYSTS>0, =0 otherwise	IBES Detail file Bhushan (1989) O'Brien and Bhushan (1990) Brennan and Subrahmanyam (1995) Barth et al. (2001)
PRIVATE	The fraction of total precision of information available to analysts that is private information (Barron et al., 1998), defined as: $1-\frac{SE-\frac{DISPERSION}{NANALYSTS}}{UNCERTAINTY}$ The greater the precision of analysts' private information, the higher the metric (PRIVATE). SE is the squared difference between IBES actual fiscal year $t+1$ earnings and the mean analysts' forecast of fiscal year $t+1$ earnings, deflated by share price taken from the last IBES monthly update immediately preceding year t fiscal year end. The mean forecast is based on each analyst's first revised forecast of year t + 1 earnings following the filing of the year t Form 10-K. DISPERSION is the variance of analysts' estimates of fiscal year $t+1$ earnings using each analyst's first revised forecast of year $t+1$ earnings following the filing of the year t 10-K, deflated by share price taken from the last IBES monthly update immediately preceding fiscal year t fiscal year end. UNCERTAINTY is the average of analysts' individual uncertainty (i.e., each analyst's expected squared forecast error) measured as $\left(1-\frac{1}{NANALYSTS}\right)\times DISPERSION+SE$	IBES Detail file Barron et al. (1998)
INFORMATIVENESS	A proxy for the fraction of total information flowing into a firm's stock price that stems from analysts' reports, measured as the proportion of information flowing into a firm's stock price around analysts' earnings forecast reports as a fraction of all information flowing into stock price during the period between the fiscal year t 10-K filing and the fiscal year $t+1$ year-end. The numerator of the proportion is the summed one-day absolute market-adjusted returns (value weighted index) on all analysts' forecast report dates using only forecast reports issued the day after the fiscal year t 10-K filing date through the $t+1$ fiscal year end. The denominator of the proportion is the summed absolute market-adjusted daily market-adjusted returns beginning the day after the year t 10-K filing date and ending on the year $t+1$ fiscal year end. Multiple analyst reports issued on a single day are treated as a single report and we require at least 90 trading days in calculating the denominator.	IBES Detail file and CRSP Frankel, Kothari, and Weber (2006)
Test variable		
ESTIMATES_NDICT	Accrual estimation intensity (the extent of estimations used during the accruals generating process) measured as the count of the number of instances in which a variation of the word "estimate" functions as a verb or adjective within the Critical Accounting Policies (CAP) section of the MD&A. This count is derived from an analysis of each sentence in the CAP section of the MD&A. Consistent with Chen et al. (2021), we use public-use software from the Stanford Natural Language Processing Group to identify instances in which a variation of the word "estimate" functions as a verb or adjective within the CAP sentences.	Authors' analysis of CAP. Chen et al. (2021)

(continued on next page)

(*continued*)

Variable	Definition	Source/Prior research
ESTIMATES_WDICT	Accrual estimation intensity (the extent of estimations used during the accruals generating process) measured as the count of the number of instances in which a variation of the word "estimate" (or related similar words) functions as a verb, adjective, or direct object of a verb within the Critical Accounting Policies (CAP) section of the MD&A. This count is derived from an analysis of each sentence in the CAP section of the MD&A. Consistent with Chen and Li (2017), we use public-use software from the Stanford Natural Language Processing Group to identify instances in which a variation of the word "estimate" (or related similar words) functions as a verb or adjective within the CAP sentences.	Authors' analysis of CAP. Chen and Li (2017)
Control variables		
CF_FORECAST	A proxy for the level of scrutiny of accruals (and by extension a scrutiny of accrual estimation intensity) by analysts, measured as the fraction of a firm's analysts issuing a cash flow forecast in addition to an earnings forecast.	IBES detail file
LENGTH_CAP	Length of the Critical Accounting Policies (CAP) section of the MD&A, measured as the number of non-stop words in the CAP.	Authors' analysis of CAP. Chen et al. (2021)
FN_NDICT	Count of the number of instances in which a variation of the word "estimate" functions as a verb or adjective within the financial statement footnotes. This count is derived from an analysis of each sentence in the financial statement footnotes. Consistent with Chen et al. (2021), we use public-use software from the Stanford Natural Language Processing Group to identify instances in which a variation of the word "estimate" functions as a verb or adjective within the sentences of the financial statement footnotes.	Authors' analysis of financial statement footnotes. Chen et al. (2021)
FN_WDICT	Count of the number of instances in which a variation of the word "estimate" (or related similar words) functions as a verb, adjective, or direct object of a verb within the financial statement footnotes. This count is derived from an analysis of each sentence in the financial statement footnotes. Consistent with Chen and Li (2017), we use public-use software from the Stanford Natural Language Processing Group to identify instances in which a variation of the word "estimate" (or related similar words) functions as a verb or adjective within the sentences of the financial statement footnotes.	Authors' analysis of financial statement footnotes. Chen and Li (2017)
LENGTH_FN	Length of financial statement footnotes, measured as the number of non-stop words within the footnotes.	Authors' analysis of financial statement footnotes. Chen et al. (2021)
LM_UNCERTAINTY	Loughran and McDonald (2011) measure of financial uncertainty obtained from WRDS SEC Analytics Suite. A proxy for the extent of financial uncertainty evidenced in the language of the 10-K filing.	
FOG	Gunning Fog index of 10-K readability. A proxy for the readability of the 10-K.	WRDS (WRDS_NLPSA file) Lehavy et al. (2011)
SIZE	Natural log of market value of equity at the balance sheet date (Compustat PRCC_F x CSHO). A proxy for the quality of the firm's information environment, the firm's complexity of operations, and investors' demand for investment advice.	Compustat Bhushan (1989), O'Brien and Bhushan (1990), Brennan and Hughes (1991), Lang and Lundholm (1996), Barth et al. (2001).
GROWTH	Mean annual sales growth across the current and preceding 4 years. A proxy for investor interest, potential for future investment banking services, analysts' forecasting difficulty.	Compustat Barth et al. (2001).
SEGMENTS	A proxy for complexity of operations, measured as the natural log of the number of operating segments.	Compustat Bradshaw, Miller, and Serafeim (2008)
INST_OWNERSHIP	A proxy for demand for investment advice and the quality of the firm's information environment, measured as the shares owned by institutional investors divided by total number of shares outstanding.	Thomson Reuters Institutional Managers (13F) Holdings Bhushan (1989) Brennan and Subrahmanyam (1995) Frankel et al. (2006).
MGMT_GUIDANCE	A proxy for quality of voluntary disclosure, measured as the number of management earnings forecasts issued during the fiscal year.	IBES Guidance file Lang and Lundholm (1996) Nagar, Nanda, and Wysocki (2003) Cotter, Tuna, and Wysocki (2006)
10-K_NEWS	A proxy for disclosure informativeness, measured as the absolute value of cumulative market-adjusted returns summed across event days 0 and 1 relative to the 10-K filing date.	CRSP Lehavy et al. (2011)
ADV	Advertising expenses (Compustat XAD) as a fraction of operating expenses (Compustat XOPR). Missing values of XAD are set to zero. A proxy for investor demand for analysts' services and analysts' forecasting difficulty arising from the opacity in financial reports created by investment in intangible assets.	Compustat Barth et al. (2001) Barron et al. (2002) Gu and Wang (2005)
R&D	R&D expense (Compustat XRD) as a fraction of operating expenses (Compustat XOPR). Missing values of XRD are set to zero. A proxy for investor demand for analysts' services and analysts' forecasting difficulty arising from the opacity in financial reports created by investment in intangible assets.	Compustat Barth et al. (2001) Barron et al. (2002) Gu and Wang (2005)
STD_RET	A proxy for investor demand for analysts' services, measured as the standard deviation of monthly stock returns during the fiscal year.	CRSP Bhushan (1989)
Other		
Δ	Change operator. For all variables except PRIVATE, the change in the variable is calculated by subtracting the year *t*-1 value from the year *t* value. For the variable PRIVATE, the change in the variable is calculated by subtracting PRIVATE_BEFORE from PRIVATE_AFTER. PRIVATE_AFTER is $1-\frac{SE\_AFTER-\frac{DISPERSION\_AFTER}{NANALYSTS\_AFTER}}{UNCERTAINTY\_AFTER}$ where UNCERTAINTY_AFTER is $\left(1-\frac{1}{NANALYSTS\_AFTER}\right)\times DISPERSION\_AFTER+SE\_AFTER$ DISPERSION_AFTER is the variance of analysts' estimate of fiscal year $t+1$ earnings using	IBES Detail file Barron et al. (1998)

(*continued on next page*)

(*continued*)

Variable	Definition	Source/Prior research
	each analyst's first revised forecast of year $t+1$ earnings made in the 90 day period following the filing of the year t 10-K, deflated by share price taken from the last IBES monthly update immediately preceding fiscal year t fiscal year end. SE_AFTER is the squared difference between IBES actual fiscal year $t+1$ earnings and the mean analysts' forecast of fiscal year $t+1$ earnings, deflated by share price taken from the last IBES monthly update immediately preceding year t fiscal year end. The mean forecast is based on each analyst's first revised forecast of year $t+1$ earnings issued in the 90-day period following the filing of the year t Form 10-K. PRIVATE_BEFORE is $1-\frac{SE\_BEFORE-\frac{DISPERSION\_BEFORE}{NANALYSTS\_BEFORE}}{UNCERTAINTY\_BEFORE}$ DISPERSION_BEFORE is the variance of analysts' estimate of fiscal year $t+1$ earnings using each analyst's last revised forecast of year $t+1$ earnings made in the 90 day period preceding the filing of the year t 10-K, deflated by share price taken from the last IBES monthly update immediately preceding fiscal year t fiscal year end. SE_BEFORE is the squared difference between IBES actual fiscal year $t+1$ earnings and the mean analysts' forecast of fiscal year $t+1$ earnings, deflated by share price taken from the last IBES monthly update immediately preceding year t fiscal year end. The mean forecast is based on each analyst's last revised forecast of year $t+1$ earnings issued in the 90-day period preceding the filing of the year t Form 10-K.	

Appendix B. Measurement and validation of accounting estimates

Extraction of critical accounting policies

A third-party data vendor extracted the MD&A portions of the 10-K, eliminated tabular information and html tags, and delivered the excerpted text to us. From these files, we algorithmically extracted the Critical Accounting Policies (CAP) section of the MD&A and parsed the text into sentences using the open-source Stanford Parser software. These sentences were further analyzed using Stanford Parser software to measure accounting estimation intensity.

Measurement of ESTIMATES_WDICT

We follow the approach of Chen and Li (2017) to measure ESTIMATES_WDICT by tabulating (for each firm-year) the total "number of objects that are said to be estimated and the number of times that estimates are said to be used" within the textual discussion of the Critical Accounting Policies (CAP). This approach entails first assigning part-of-speech function to each word in a CAP, then identifying instances in which an estimate-related word functions as an adjective, verb, or noun within the sentence. We use Stanford Parser software to assign part-of-speech function.

Estimation-related words, derived from Chen and Li (2017), are organized into dictionaries, which are detailed in Table B1. The Estimation Verbs Dictionary consists of words that can function as a verb to denote an estimation action is directed toward an object. The Estimation Adjectives Dictionary consists of words that can function as an adjective to denote a noun that is being estimated. The Estimation Nouns Dictionary consists of words that can function as a noun to denote an estimate. The Use Verbs Dictionary consists of words that can function as a verb to denote that action is directed toward an estimate. We infer the presence of an accounting estimate when these dictionary words assume any of five sentence roles (R1-R5). These roles are summarized in Table B2 and are explained below.

The first role (R1) is when an Estimation Verb directs action toward some object. For example, in the sentence

> We **assess** the potential **impairment** of intangible and fixed assets whenever events or changes in circumstances indicate that the carrying values may not be recoverable.

The Estimation Verb "assess" directs action toward the noun "impairment" (i.e., impairment is the direct object of assess). This signals that impairment is the object to be estimated.

The second role (R2) is when a Use Verb directs action toward an Estimation Noun. For example, in the sentence

> When products are sold directly to retailers or end-users, we **make** an **estimate** of sales returns based on historical experience.

the Estimation Noun "estimate" is the direct object of the Use Verb "make." This signals that an estimate is to be used.

The third role (R3) is when an Estimation Adjective modifies a noun. For example, in the sentence

> We recognize stock-based compensation expense over the requisite service period, net of **estimated forfeitures**.

the Estimation Adjective ("estimated") modifies a noun ("forfeitures"). This signals that forfeitures is the object to be estimated.

The fourth role (R4) is when a word from the Estimation Adjective Dictionary modifies a dollar value. For example, in the sentence

> For the year ended December 31, 2008, depreciation expense aggregated **approximately $146,000**.

the Estimation Adjective ("approximately") modifies a dollar value ("$146,000"). This signals that a dollar amount is the object to be estimated.

The fifth role (R5) is when an Estimation Noun is modified by another noun. For example, in the sentence

> Each reporting period, changes in the actuarial assumptions resulting from changes in actual claims experience and other trends are incorporated into the Company's workers compensation claims **cost estimates**.

an Estimation Noun ("estimates") is modified by another noun ("cost"). This signals that cost is the object to be estimated.

As noted above, we use Stanford Parser software to identify instances of R1 to R5 within the CAP. This software recognizes approximately fifty grammatical relationships between words in sentences, called "Stanford Typed Dependencies," of which five are pertinent in our analysis because they collectively encompass roles R1-R5 above. Those grammatical relationships are DOBJ (direct object), NSUBJ (nominal subject), AMOD (adjectival modifier), QUANTMOD (quantifier modifier), and NN (noun compound modifier).

For each firm-year, we count the number of R1-R5 found within the CAP and assign that count to the variable ESTIMATES_WDICT. As an example, assume a firm's CAP consisted of the following (i.e., the five example sentences from above):

> We **assess** the potential **impairment** of intangible and fixed assets whenever events or changes in circumstances indicate that the carrying values may not be recoverable. When products are sold directly to retailers or end-users, we **make** an **estimate** of sales returns based on historical experience. We recognize stock-based compensation expense over the requisite service period, net of **estimated forfeitures**. At December 31, 2008, the Company's future minimum lease payments due under non-cancelable leases aggregated **approximately $ 146,000**. Each reporting period, changes in the actuarial assumptions resulting from changes in actual claims experience and other trends are incorporated into the Company's workers compensation claims **cost estimates**.

There are five instances of R1 to R5 within this hypothetical CAP, so ESTIMATES_WDICT = 5.

Measurement of ESTIMATES_NDICT

ESTIMATES_NDICT, based on Chen et al. (2021), is derived the same as ESTIMATES_WDICT but with a narrower set of sentence roles and dictionary elements. In terms of roles, ESTIMATE_NDICT counts only roles R1, R3 and R4. In terms of dictionary elements, ESTIMATES_NDICT includes only words which are variations on the root "estimat" (i.e., estimate, estimating, estimated, estimates). For the hypothetical CAP above, ESTIMATES_NDICT = 1 (the sentence that illustrates R3 is the only sentence in which a variation of "estimat" functions as either a verb or adjective).

Validation of ESTIMATES_WDICT and ESTIMATES_NDICT

As detailed above, ESTIMATES_WDICT and ESTMATES_NDICT are based on the part-of-speech function of estimate-related words found in the CAP section of the 10-K Management's Discussion and Analysis. We validate the construct validity of ESTIMATES by correlating it with a separate, independently-derived proxy for the estimation intensity, ESTIMATES_ASC, described below.

Our approach in deriving ESTIMATES_ASC is conceptually similar to that used by Chychyla, Leone, and Minutti-Meza (2019) to measure financial statement complexity from XBRL tags. For each dollar value amount reported in an XBRL filing that is coded as appearing in a firm's financial statements and footnotes, we determine the Accounting Standards Codification (ASC) Topic and Subtopic that is mapped to the item as disclosed in the XBRL taxonomy guide issued by the FASB. This represents the ASC Topic and Subtopic that governs the accounting and disclosure for the item. Next, we count the number of estimation-related cue words (variations of the word "estimate") that appear within the relevant ASC Topic/Subtopic and assign this word count to the particular dollar value amount from the XBRL filing that is governed by the ASC Topic/Subtopic.[10] The process is repeated for each dollar value amount reported in the XBRL filing. Then, we calculate the dollar-weighted mean number of estimation-related cue words (for each firm-year) by averaging across all financial statement and footnote dollar value amounts reported in the XBRL filing for that year. We interpret the dollar-weighted mean number of estimation-related cues as representing the extent of accounting estimates permitted or required by the collective set of GAAP that applies to the firm. We percentile rank the dollar-weighted mean number of estimation-related cue words and assign this percentile rank to the variable ESTIMATES_ASC.

Finally, we regress ESTIMATES_xx (where _xx denotes _WDICT or _NDICT) on ESTIMATES_ASC, LENGTH_CAP (as defined in Appendix A), and SIZE (as defined in Appendix A, used as general control for size effects). All else equal, firms whose governing GAAP requires more accounting estimates should have more discussion of accounting estimation in their CAP. We include industry and year fixed effects and base statistical inferences on standard errors that control for clustering at the firm level. We note that the sample in the validation test is approximately one-half the sample used in the main tests. This is due to two factors. First, detailed XBRL filings, which are needed to measure ESTIMATES_ASC, are available only for fiscal years 2012 and forward, causing the loss of observations from 2007–2011. Second, because we are unable to determine the governing ASC subtopic for non-standard (i.e., "custom") XBRL tags, we exclude firm-years in which more than 10% of dollar value amounts are given non-standard XBRL tags.

Table B3 below reports results obtained from the validation test. Inspection of Table B3 reveals that ESTIMATES_ASC is positively and significantly associated with ESTIMATES_xx. In other words, firms for which governing GAAP requires the firm to make more accounting estimates tend to have more discussion of accounting estimates in their CAP. This provides evidence supporting the construct validity of ESTIMATES_WDICT and ESTIMATES_NDICT.

Table B1
Dictionaries.

Dictionary	Description of dictionary
Estimation Verbs Dictionary:	
Estimate, Estimating, Estimated, Anticipate, Anticipates, Anticipating, Anticipated, Approximate, Approximates, Approximated, Approximating, Assess, Assesses, Assessed, Assessing, Believe, Believed, Believes, Believing, Determine, Determined, Determining, Determines, Evaluate, Evaluated, Evaluating, Evaluates, Expect, Expects, Expected, Expecting, Forecast, Forecasts, Forecasted, Forecasting	Words that can function as a verb to denote an estimation action is directed toward an object.

(continued on next page)

[10] We measure ESTIMATES_ASC using a "bag of words" approach rather than the grammatical relation approach used to measure ESTIMATES because the code-like structure of the codification (lengthy enumerated subparagraph phrases embedded within a single sentence) does not lend itself to use of the Stanford Parser software, which requires reasonably normal sentence structure.

Table B1 (*continued*)

Dictionary	Description of dictionary
Estimation Adjectives Dictionary: Estimated, Anticipated, Approximate, Approximately, Expected, Forecasted, Likely, Probable	Words that can function as an adjective to denote that a noun is being estimated.
Estimation Nouns Dictionary: Estimate, Estimates, Estimation, Estimations, Approximation, Approximations, Assumption, Assumptions, Belief, Beliefs, Forecast, Forecasts	Words that can function as a noun to denote an estimate is being made.
Use Verbs Dictionary: Make, Makes, Made, Making, Use, Uses, Used, Using, Include, Includes, Included, Including	Words that can function as a verb to denote that action is directed toward an estimate.

Table B2
Grammatical roles.

Grammatical role	Description
R1: Estimation Verb targets a noun	A noun is the object of a word in the Estimation Verb dictionary.
R2: Use Verb targets an Estimation Noun	A word in the Estimation Noun dictionary is the object of a word in the Use Verb dictionary
R3: Estimation Adjective modifies a noun.	A word from the Estimation Adjectives dictionary modifies a noun.
R4: Estimation Adjective modifies a value.	A word from the Estimation Adjectives dictionary modifies a dollar value.
R5: Estimation Noun modified by another noun.	A noun modifies a word from the Estimation Nouns dictionary.

Table B3
Validation of ESTIMATES_NDICT and ESTIMATES_WDICT.

	Ranked estimates		Logged estimates	
	Narrow dict	Wide dict	Narrow dict	Wide dict
ESTIMATES_ASC	0.042**	0.077***	18.093***	22.236***
	(2.50)	(5.89)	(2.89)	(5.20)
LENGTH_CAP	0.453***	0.711***	0.400***	0.556***
	(30.39)	(64.39)	(31.19)	(56.27)
SIZE	0.019***	0.026***	0.040***	0.051***
	(7.46)	(13.08)	(7.16)	(12.46)
intercept	0.385***	−0.025	−0.862***	−1.029***
	(4.90)	(−0.37)	(−4.62)	(−5.66)
N	12,525	12,525	12,525	12,525
R-sq	0.279	0.580	0.308	0.616

The dependent variable in the "Narrow Dict" ("Wide Dict") column is ESTIMATES_NDICT (ESTIMATES_WDICT). The dependent variable and LENGTH_CAP are percentile-ranked in the "Ranked Estimates" columns. The dependent variable and LENGTH_CAP are natural log-transformed in the "Logged Estimates" columns. The value 1 is added to all log-transformed metrics prior to logging. The variable SIZE is defined in Appendix A.
*, **, *** denote two-tailed significance levels of 0.10, 0.05, and 0.01, respectively.

References

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

Barron, O., Kim, O., Lim, S., & Stevens, D. (1998). Using analysts' forecasts to measure properties of analysts' information environment. *The Accounting Review, 73*(4), 421–433.

Barth, M., Kasznik, R., & McNichols, M. (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.

Boone, J., Khurana, I., & Raman, K. K. (2022). Accounting estimation intensity, auditor estimation expertise, and managerial bias. *Accounting Horizons*. https://doi.org/10.2308/HORIZONS-2020-193. Forthcoming.

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

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

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

Brown, L., Call, A., Clement, M., & Sharp, N. (2015). Inside the "black box" of sell-side financial analysts. *Journal of Accounting Research, 53*(1), 1–47.

Call, A., Chen, S., & Tong, Y. (2013). Are analysts' cash flow forecasts naive extensions of their own earnings forecasts? *Contemporary Accounting Research, 30*(2), 438–485.

Chen, H., Chen, J., & Li, F. (2021). The number of estimates in footnotes and accruals. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2738842.

Chen, J. V., & Li, F. (2017). Estimating the amount of estimation in accruals. In *Working paper. June 27, 2017*.

Chychyla, R., Leone, A. J., & Minutti-Meza, M. (2019). Complexity of financial reporting standards and accounting expertise. *Journal of Accounting and Economics, 67*(1), 226–253.

Cohen, D., Darrough, M., Huang, R., & Zach, T. (2011). Warranty reserve: Contingent liability, information signal, or earnings management tool? *The Accounting Review, 86*(2), 569–604.

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

Dechow, P., & Dichev, I. (2002). The quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review, 77*(s-1), 35–59.

Dechow, P., Ge, W., Larson, C., & Sloan, R. (2011). Predicting material accounting misstatements. *Contemporary Accounting Research, 28*(1), 17–82.

DeFond, M., & Hung, M. (2003). An empirical analysis of analysts' cash flow forecasts. *Journal of Accounting and Economics, 35*(1), 73–100.

Elliott, J. A., & Hanna, J. D. (1996). Repeated accounting write-offs and the information content of earnings. *Journal of Accounting Research, 34*(1), 135–155.

Ernst and Young. (2017). SEC comments and trends. https://www.ey.com/publication/vwluassetsdld/seccommentstrends_05443-171us_25september2017/$file/seccommentstrends_05443-171us_25september2017.pdf.
Fields, T. D., Lys, T. Z., & Vincent, L. (2001). Empirical research on accounting choice. *Journal of Accounting and Economics, 31*(1–3), 255–307.
Frankel, R., Kothari, S., & Weber, J. (2006). Determinants of the informativeness of analyst research. *Journal of Accounting and Economics, 41*(1–2), 29–54.
Glendening, M. (2017). Critical accounting estimate disclosures and the predictive value of earnings. *Accounting Horizons, 31*(4), 1–12.
Glendening, M., Mauldin, E., & Shaw, K. W. (2019). Determinants and consequences of quantitative critical accounting estimate disclosures. *The Accounting Review, 94*(5), 189–218.
Gong, G., Li, L. Y., & Xie, H. (2009). The association between management earnings forecast errors and accruals. *The Accounting Review, 84*(2), 497–530.
Gordon, E., Ma, X., & Runesson, E. (2018). Information uncertainty and critical accounting policies and estimates. In *Presented at the 2018 AAA Annual Meeting*.
Gordon, E., Petruska, K., & Yu, M. (2014). Do analysts' cash flow forecasts mitigate the accrual anomaly? International evidence. *Journal of International Accounting Research, 13*(1), 61–90.
Gu, F., & Wang, W. M. (2005). Intangible assets, information complexity, and analysts' earnings forecasts. *Journal of Business Finance and Accounting, 32*(9–10), 1673–1702.
Healy, P. M., & Wahlen, J. M. (1999). A review of the earnings management literature and its implications for standard setting. *Accounting Horizons, 13*(4), 365–383.
Heckman, J. (1979). Sample selection bias as a specification error. *Econometrica., 47*(1), 153–161.
Hirst, D. E. (1994a). Auditors' sensitivity to source reliability. *Journal of Accounting Research, 32*(1), 113–126.
Hirst, D. E. (1994b). Auditor sensitivity to earnings management. *Contemporary Accounting Research, 11*(1), 405–422.
Kim, O., & Verrecchia, R. (1994). Market liquidity and volume around earnings announcements. *Journal of Accounting and Economics, 17*(1), 41–67.
Lang, M., & Lundholm, R. (1996). Corporate disclosure policy and analyst behavior. *The Accounting Review, 71*(4), 467–492.
Lehavy, R., Li, F., & Merkley, K. (2011). The effect of annual report readability on analyst following and the properties of their earnings forecasts. *The Accounting Review, 86*(3), 1087–1115.
Lev, B., & Gu, F. (2016). *The end of accounting and the path forward for investors and managers*. Hoboken, New Jersey: John Wiley & Sons.
Levine, B., & Smith, M. (2011). Critical accounting policy disclosures. *Journal of Accounting, Auditing and Finance, 26*(1), 39–75.
Li, X., & Yang, H. (2016). Mandatory financial reporting and voluntary disclosure: The effect of mandatory IFRS adoption on management forecasts. *The Accounting Review, 91*(3), 933–953.
Livnat, J., & Zhang, Y. (2012). Information interpretation or information discovery: Which role of analysts do investors value more? *Rev. Account. Study., 17*(3), 612–641.
Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance, 66*(1), 35–65.
Mohanram, P. (2014). Analysts' cash flow forecasts and the decline of the accruals anomaly. *Contemporary Accounting Research, 31*(4), 1143–1170.
Nagar, V., Nanda, D., & Wysocki, P. (2003). Discretionary disclosure and stock-based incentives. *Journal of Accounting and Economics, 34*(1–3), 283–309.
O'Brien, P., & Bhushan, R. (1990). Analyst following and institutional ownership. *Journal of Accounting Research, 28*(1), 55–76.
Penman, S. (2001). *Financial statement analysis and security valuation*. New York, NY: McGraw-Hill Irwin.
Peterson, K. (2012). Accounting complexity, misreporting, and the consequences of misreporting. *Rev. Account. Study., 17*(1), 72–95.
Plumlee, M., & Yohn, T. L. (2010). An analysis of the underlying causes attributed to restatements. *Accounting Horizons, 24*(1), 41–64.
Radhakrishnan, S., & Wu, S. L. (2014). Analysts' cash flow forecasts and accrual mispricing. *Contemporary Accounting Research, 31*(4), 1191–1219.
Roh, Y. (2018). Does accounting estimation improve stock price informativeness?. In *Presented at the 2018 AAA Annual Meeting*.
Securities and Exchange Commission (SEC). (2002). *Release Nos. 33–8098; 34–45907. Disclosure in management's discussion and analysis about the application of critical accounting policies*. Washington, DC: SEC.
Securities and Exchange Commission (SEC). (2003). *Release Nos. 33–8350; 34–48960; FR-72. Commission guidance regarding management's discussion and analysis of financial condition and results of operations*. Washington, DC: SEC.
Securities and Exchange Commission (SEC). (2011). Roundtable subject: Measurement uncertainty in financial reporting. Transcript of the inaugural public roundtable for financial reporting series (November 8). Available at https://www.sec.gov/about/offices/oca/frs/ocafrseries110811-transcript.txt.
Wild, J., Bernstein, L., & Subramanyam, K. (2001). *Financial statement analysis*. New York, NY: McGraw-Hill Irwin.
Williams, P. A. (1996). The relation between a priori earnings forecast by management and analyst response to a current management forecast. *The Accounting Review, 71*(1), 103–115.