

10-K Disclosure Repetition and Managerial Reporting Incentives

Richard A. Cazier

The University of Texas at El Paso

Ray J. Pfeiffer

Texas Christian University

ABSTRACT: Standard setters and investors have expressed concern about the substantial amount of repetition of information found within firms' 10-K disclosures. Ostensibly, such repetition is the result of overlapping disclosure requirements that solicit similar information in separate sections of the 10-K. Our evidence, however, suggests that repetition of information within the 10-K is also a strategic response to managers' reporting incentives. We find evidence that, on average, discretionary repetition of information within the 10-K is driven by both litigation concerns and by managers' incentives to report information more opaquely when earnings performance is poor. We find no evidence that repetition of information within the 10-K is associated with improved price discovery following the 10-K filing date, but that repetition is in some cases associated with slower price discovery. Overall, our results suggest that current overlap in disclosure requirements allows managers to repeat information in ways that serve their strategic reporting purposes but that do not result in increased clarity for investors.

Keywords: disclosure regulation; disclosure effectiveness; narrative disclosure; repetition; textual analysis.

JEL Classifications: D8; G38; M4.

INTRODUCTION

In recent years, the SEC and the FASB have begun to examine the extent to which corporate disclosure volume has become unduly excessive (e.g., FASB 2012; Higgins 2014). Practitioners, regulators, investors, and academics have noted that one significant source of disclosure volume is the amount of information that gets repeated within the 10-K (e.g., FASB 2001; Higgins 2014; KPMG 2011; PwC 2014; Cazier and Pfeiffer 2016). Current reporting standards embed overlapping or similar disclosure requirements across different sections of the 10-K, in part because existing standards reflect an accumulation of requirements set by different bodies at different times and with different objectives. Critics of current reporting practices argue that the SEC and the FASB should work together to eliminate overlapping disclosure requirements to reduce the considerable amount of repetition contained within firms' 10-K filings (e.g., KPMG 2011; Standard & Poor's 2012).

Managers have significant discretion, however, in the extent to which they actually repeat information because standard setters permit and even encourage cross-referencing rather than duplicating disclosures within the 10-K in many instances. Cazier and Pfeiffer (2016) document that repetitious disclosures contribute substantially not only to overall 10-K length but also to variation in 10-K length across firms. Thus, managers' reporting incentives and strategies are a key co-determinant of the level and nature of the repetitious disclosure that appears in financial reports. An important but unanswered question relates to the factors that influence managers' discretion in repeating disclosures within their financial reports.

We examine three alternative, non-mutually exclusive hypotheses that potentially explain managers' choice to repeat information within the 10-K. The litigation-risk hypothesis predicts that firms repeat more information when their securities litigation risk is higher. KPMG's (2011) survey of financial executives indicates that the vast majority of public firms view

We gratefully acknowledge helpful comments and suggestions from John Campbell, Josh Filzen, In-Mu Haw, Todd Kravet, Ken Merkley, Kyle Peterson, Elizabeth Plummer, Chad Proell, Shana Proell, Anup Srivastava, Wendy Wilson, and participants at the 2014 Brigham Young University Accounting Research Symposium, attendees at the 2015 AAA Annual Meeting, and anonymous reviewers.

Editor's note: Accepted by Anne Beatty.

Submitted: April 2016
Accepted: August 2017
Published Online: September 2017

litigation risk as having a significant impact on their disclosure practices. Firms that face greater litigation risk have incentives to provide more disclosure to preempt allegations that their disclosures were inadequate, and this increase in disclosure may lead to overlapping discussions in different parts of the 10-K. Because the costs of litigation are high and the costs of repeating information across sections of the 10-K are low, repetition of disclosures that could potentially be deemed relevant may be a rational managerial response to litigation risk.

The *obfuscation hypothesis* predicts that managers' choice to infuse the 10-K with more repetitious disclosure is driven by managers' efforts to mask poor performance. The notion that managers obfuscate poor performance by providing lower-quality disclosures has found broad support in prior literature (e.g., Bloomfield 2008; Kothari, Shu, and Wysocki 2009; Li 2008; Miller 2002; Schrand and Walther 2000). Managers who seek to avoid a straightforward discussion of poor current performance may prefer to repeat mundane information reported elsewhere in the 10-K, under the pretext of complying with current reporting requirements. Disclosure repetition can mask relevant information by increasing investors' processing costs and redirecting investors' attention toward less relevant information.

The *salience hypothesis* predicts that repetition in the 10-K is primarily a manifestation of managers' efforts to emphasize favorable information. The role of repetition in increasing message salience has been examined in various other domains such as marketing and pedagogy (e.g., Belch 1981; Darley and Glass 1975; Bergerbest, Ghahremani, and Gabrieli 2004; Rock 1957; Luotoniemi, Service, and Maury 2007). Consistent with prior literature, we expect managers' incentives to make information more salient is increasing in the favorability of that information. The salience hypothesis of disclosure repetition conjectures that disclosure repetition reflects strongly performing firms repeating information to increase message salience, and that disclosure repetition is increasing in firm performance.

We examine over 60,000 10-Ks filed from 1994 through 2013 to test the descriptive validity of each of these hypotheses regarding managers' repetition of disclosures within the 10-K. We measure disclosure repetition as the percentage of ten-word sequences within the 10-K that are repeated more than once within a single 10-K, although we find that our results are robust to measuring disclosure repetition as the total number of ten-word sequences that are repeated or as the longest string of repeated text within the 10-K. We estimate our models with firm-fixed effects to capture the within-firm, time-varying determinants of firms' provision of repetitive disclosures.

We find strong evidence that 10-K disclosure repetition is correlated with firms' litigation risk, providing support for the litigation-risk hypothesis. However, we also find persistent evidence across a broad set of proxies for firms' earnings performance that disclosure repetition is higher when earnings performance is worse, consistent with the obfuscation hypothesis. To ensure that the relationship between our firm performance measures and disclosure repetition is not driven by firms' litigation concerns, we re-examine this relationship after splitting our sample between firms with higher versus lower litigation risk. We find that the relationship between firms' earnings performance and 10-K disclosure repetition is actually strongest among firms that face the lowest litigation risk, providing further support that the relationship between earnings and repetition is driven by obfuscation incentives rather than by firms' responses to litigation risk.

We also examine separately the repetition of disclosures that are more likely to be informative to investors versus the repetition of disclosures that are more likely to be mundane or uninformative. We use earnings-related disclosures to proxy for disclosures that are more likely to be informative to investors, and we use the repeated description of recent FASB pronouncements within the 10-K as our proxy for relatively more mundane disclosures. We find a positive relationship between the repetition of earnings-related disclosures and firm performance, suggesting that at least some redundancy is driven by firms' efforts to highlight favorable news (the salience hypothesis). In contrast, when earnings performance is worse, firms are more likely to repeat mundane disclosures, consistent with the obfuscation hypothesis.

We find some evidence that 10-K disclosure repetition is somewhat negatively correlated with the speed with which 10-K information is incorporated into investors' assessment of equity value. On average, 10-K disclosure repetition is associated with slower price discovery around 10-K filing dates, particularly among poorly performing firms. Specifically, when 10-K disclosure repetition is higher, a greater percentage of the total abnormal returns during the 20 trading days after the 10-K filing date accrues later in that event window.

Overall, our findings suggest that 10-K disclosure repetition increases in response to managers' litigation concerns and obfuscation incentives, and to a lesser extent by managers' incentives to highlight favorable news. Our findings regarding how managers exercise discretion to selectively repeat 10-K disclosures have implications for standard setters. The FASB has indicated that reducing redundancy in the 10-K is one objective of its ongoing Disclosure Framework project, which is consistent with recommendations made by a host of observers (e.g., Center for Audit Quality 2012; FASB 2001; Investors Technical Advisory Committee 2007; KPMG 2011). Recently, however, both the SEC and the FASB have questioned whether repetition in SEC filings is problematic (FASB 2013; White 2013), and regulators have yet to make significant revisions to 10-K requirements that would eliminate redundant disclosure. Our findings suggest that managers exploit the overlap in reporting requirements by repeating information when doing so will reduce the transparency of the most relevant content in the 10-K. Our findings also illustrate negative consequences for market participants of repetitious 10-K disclosure, which may help to

provide a further impetus for regulators to find ways to reduce opportunities for repetitive disclosure in the 10-K. At a minimum, our research helps regulators and standard setters to better understand some of the outcomes associated with overlapping 10-K reporting requirements that currently facilitate repetitious disclosures.

This study contributes to recent academic literature that examines how firm performance shapes qualitative properties of financial disclosures. [Dyer, Lang, and Stice-Lawrence \(2016\)](#) examine factors that contribute to 10-K length and conclude that a primary driver of both 10-K length and repetition within the 10-K is firms' compliance with disclosure regulations. Our study suggests that an additional important determinant of repetition within the 10-K is managers' discretion and that managers' choice to repeat various disclosures is driven, at least in part, by their incentives to enhance or impair the transparency of those disclosures. Although [Li \(2008\)](#) and [Loughran and McDonald \(2014\)](#) find that worse-performing firms are associated with longer 10-K reports, little evidence exists on the mechanisms that managers use to make 10-Ks long in response to poor earnings. Our study suggests that discretionary repetition of disclosures is one such mechanism.

HYPOTHESIS DEVELOPMENT

Overlap in Disclosure Requirements

Overlap in disclosure requirements in the 10-K is at least partly attributable to the independently developed reporting rules mandated by the SEC and the FASB. The FASB prescribes much of the format and requirements for reporting information in the financial statement notes for all entities that report under U.S. GAAP. The SEC prescribes reporting requirements for certain aspects of the financial statement notes for public registrants as well as for the other sections of the 10-K. In their comment letter on the FASB's Conceptual Framework project, [Standard & Poor's \(2012\)](#) argues that the placement and organization of financial reporting has been driven by whichever standard setter has mandated the relevant requirements rather than on how the information can best be presented to an external user. They, along with several other observers, call for the SEC and the FASB to work jointly to reduce overlapping disclosure requirements (e.g., [KPMG 2011](#); [Standard & Poor's 2012](#)).

[Dyer et al. \(2016\)](#) argue that firms' regulatory compliance explains a significant percentage of the repetition, or redundancy, within firms' 10-K filings. Current disclosure regulation and SEC guidance, however, permit firms to cross-reference between different sections of the 10-K in lieu of providing duplicate information in many instances (e.g., [SEC 2014](#)). Thus, firms have significant discretion whether to avoid or include repetitious information in their financial reports. Whereas some firms respond to a 10-K disclosure requirement by referencing the related financial statement note, other firms respond to the same disclosure requirement by simply repeating, often word for word, the same information in both the notes and in other sections of the 10-K. This evidence suggests that firms use repetition selectively in preparing their financial reports.

In recent years, standard setters at both the SEC and the FASB have repeatedly expressed concern that the amount of disclosure duplicated within firms' 10-K reports makes those reports less readable (e.g., [FASB 2014](#); [Gallagher 2013](#)). Recent academic literature documents that these redundancies are pervasive and contribute significantly to firms' total 10-K disclosure volume ([Cazier and Pfeiffer 2016](#)). The current study contributes to the discussion regarding the costs and benefits of 10-K repetition by shedding light on the factors that drive managers' decisions to repeat or not repeat information found within the 10-K.

We note that disclosure repetition is conceptually and empirically distinct from disclosure fog and length, which have been examined in prior research. One important distinction is that, unlike disclosure fog, the repetition of disclosures can be used to enhance or detract from the transparency of relevant information. Thus, [Li \(2008\)](#) finds that managers with incentives to obfuscate poor performance produce longer and foggier 10-Ks, but his results shed little light on whether or how obfuscation incentives affect managers' responses to overlapping disclosure requirements. Accordingly, [Li's \(2008\)](#) results are silent regarding the costs and benefits of retaining overlapping disclosure requirements, as they provide no insight into how reporting incentives interact with managers' discretion in reporting duplicate information.¹

Transparent 10-K disclosure is important because, even after earnings announcements have been made, 10-Ks contain price-relevant information that investors trade on (e.g., [Griffin 2003](#); [Huddart, Ke, and Shi 2007](#)), and less readable 10-Ks are associated with a slower market response to 10-K information (e.g., [You and Zhang 2009](#)). More generally, research indicates that the market may underreact to public information such as that in earnings announcements but that additional subsequent information, such as that contained in 10-K reports, can accelerate price discovery (e.g., [Hong and Stein 1999](#)).

¹ The correlation between disclosure repetition and 10-K length is 0.43 and statistically significant. Although repeating 10-K disclosure mechanically increases 10-K length, 10-K length may also be increased by disclosing additional non-repetitive information. Thus, the direction of correlation between 10-K length and the percentage of 10-K length consisting of repetitious disclosure is not obvious *ex ante*, as it depends on whether longer 10-Ks are characterized primarily by excess repetitious disclosure or excess non-repetitious disclosure. The correlation between fog and repetition in our sample is negative and small in magnitude (−0.065).

Potential Determinants of 10-K Disclosure Repetition

Disclosure repetition may reflect firms' rational response to the risk of litigation that alleges inadequate disclosure. KPMG's (2011) survey indicates that most public financial executives consider litigation risk as having a significant impact on their disclosure choices. Because managers do not know which disclosures will be deemed material *ex post*, they have incentives to "disgorge information whether it is useful or not" (Loughran and McDonald 2010), and these incentives may extend to the decision of whether to repeat potentially relevant information across various sections of the 10-K. The costs to managers of repeating a disclosure are very low, but the costs of a lawsuit that alleges insufficient disclosure can be considerable. The litigation-risk hypothesis of 10-K disclosure repetition posits a positive relationship between firms' securities litigation risk and the amount of disclosure that is repeated in firms' 10-Ks.

H1: 10-K disclosure repetition is positively associated with firms' litigation risk.

In many contexts, verbal repetition is used to increase the salience of a narrative or a key point. Prior research examines the benefits of repetition for the classroom (e.g., Bergerbest et al. 2004; Darley and Glass 1975; Rock 1957; Luotoniemi et al. 2007), for marketing campaigns (e.g., Dardis 2009), and even for increasing subjects' affinity for musical pieces (Hargreaves 1984). Disclosure repetition in the 10-K may be symptomatic of managers' efforts to accentuate favorable information in the minds of investors. Because strong-performing firms have the greatest incentives to increase the salience of the information in the 10-K, the salience hypothesis of 10-K disclosure repetition predicts 10-K repetition will be greatest among better-performing firms.

H2: 10-K disclosure repetition is positively associated with firm performance.

Li (2008) finds evidence consistent with managers making 10-Ks longer and more difficult to read when performance is worse. Similarly, it is possible that 10-K disclosure repetition is reflective of managers choosing to duplicate mundane or less relevant information to detract attention from firm performance when performance is poor. Repetitive disclosures may impair 10-K readability in several ways. Investors or analysts may skip over disclosures that appear familiar and may miss any relevant, non-redundant information that is buried within predominantly repetitious text (KPMG 2011). Research indicates that repeating information can adversely affect judgment and decision making, both by biasing decision makers' assessments of the likelihood of events (Bonner 2007; Joe 2003) and by increasing the total quantity of information (Davis, Lohse, and Kottemann 1994). Perhaps for these reasons, the SEC (2003) has specifically discouraged firms from disclosing information in the MD&A that is redundant with the financial statement notes.

The obfuscation hypothesis predicts that 10-K repetition is highest when managers have incentives to obfuscate relevant 10-K information. We follow prior literature on the obfuscation hypothesis that argues that obfuscation incentives are greatest when firm performance is poor. We state the obfuscation hypothesis of 10-K disclosure repetition as follows:

H3: 10-K disclosure repetition is negatively associated with firm performance.

Consequences of Repetitious 10-K Disclosures

We also consider whether disclosure repetition in the 10-K affects the speed with which 10-K information becomes embedded into stock price. Prior research documents that 10-K financial information is not fully impounded into stock prices immediately, but that price discovery occurs gradually as investors process the information content in the 10-K (e.g., You and Zhang 2009). Recent research confirms that less readable 10-Ks exacerbate investors' information-processing problem (e.g., Lawrence 2013; Lehavy, Li, and Merkley 2011; Miller 2010; Rennekamp 2012). Cazier and Pfeiffer (2016) examine 10-K disclosure repetition for a relatively small sample of 620 firm-years and find no effect of long-term price drift associated with repetitious disclosures following 10-K filings. Their failure to find an association between repetitious disclosures and subsequent price drift, however, may be due to low power, given their small sample size and the length of the window over which they test for market mispricing. The market price tests in the current study are likely to be significantly more powerful, as we examine a sample of over 60,000 observations and test for short-term price inefficiencies immediately after the 10-K filing.

Although repetition likely contributes to overall 10-K length, repetition may expedite price discovery if providing the same information in different sections of the 10-K helps to get relevant information into the market more quickly. Greater disclosure repetition, however, may impair investors' information-processing ability if the redundant content obscures the information that is most relevant to market participants, regardless of whether the repetition is driven by litigation concerns or obfuscation incentives. Thus, we make no directional prediction regarding whether, on average, disclosure repetition is likely to increase or decrease the speed with which relevant 10-K information becomes impounded into stock price.

H4: There is no association between 10-K disclosure repetition and the speed of price formation following 10-K filings.

TABLE 1
Sample Selection

Selection Criteria	Firm-Years Remaining
All firm-years between 1994 and 2013 listed on Compustat	232,406
Remove 73,738 firm-years without sufficient Compustat data to compute control variables	158,668
Remove 45,704 firm-years missing CRSP data to compute control variables	112,964
Remove 9,879 firm-years missing Compustat or CRSP data to measure litigation risk variable	103,085
Remove 39,364 firm-years with no 10-K filed within 120 days after fiscal year-end ^a	63,721
Remove 26 firm-years with 10-Ks containing less than 1,000 words	63,695
Final Sample	63,695

^a This exclusion captures firms filing late 10-Ks as well as companies filing non-standard annual reports such as the 10KSBs issued by small companies or 20-Fs issued by foreign listings.

EMPIRICAL METHOD

Sample Selection

Table 1 depicts the evolution of our sample from the population of available CRSP and Compustat firm-years. Our sample selection process begins with 232,406 firm-years that are listed on Compustat from 1994 through 2013. We lose 73,738 firm-year observations lacking sufficient Compustat data to compute our control variables. We lose an additional 45,704 firm-years that lack the necessary CRSP data required to compute our control variables. We lose an additional 9,879 observations that lack sufficient data from either CRSP or Compustat to construct our proxy for litigation risk.

We use Perl to extract cleaned text, after removing tables and other HTML encoding, for all 10-Ks filed on EDGAR up through 2013. We eliminate 39,364 firm-years that do not file a standard 10-K within 120 days of their fiscal year-end. This exclusion eliminates small businesses that file Form 10-KSB and foreign private issuers that do not file a 10-K at all, as well as domestic firms that fail to file a timely 10-K. We eliminate 26 additional firms whose 10-Ks contain fewer than 1,000 words to mitigate the effect of bad text parses or problematic 10-K filings. Our final sample consists of 63,695 firm-years.

Measuring Disclosure Repetition within the 10-K

We build on methods developed in research on plagiarism detection to identify instances of repetitious disclosure within firms' 10-K reports. These methods infer plagiarism between two documents based on the number of short word sequences ("*n*-grams"), typically two to four words in length, that are found in both documents. Prior research in linguistics notes that independently written text typically shares few matching three-word sequences (e.g., Gibbon, Moore, and Winski 1997; Lyon, Malcolm, and Dickerson 2001). Related work in accounting has used three- or four-word sequences to identify text that is essentially copied from one document to another (e.g., Cazier and Pfeiffer 2016; Lang and Stice-Lawrence 2015; McMullin 2014; Nelson and Pritchard 2014).

The choice of *n* in determining an optimal *n*-gram length to use to identify repeated information constitutes a trade-off between Type I and Type II errors. Shorter *n*-grams, such as two- to four-word sequences, may lead to a higher rate of Type I errors, resulting in incorrect identification of redundant disclosures. Higher *n*-grams risk a higher rate of Type II errors, whereby the method fails to detect repeated disclosure because a single word in the phrase or sentence has been modified. Our study focuses on ten-word sequences (10-grams), which helps minimize the incidence of false positives in identifying repeated text. We believe that imposing a high "*n*" in our *n*-gram sequence is appropriate in our setting for two reasons. First, unlike plagiarism, in which text is typically modified to avoid detection, anecdotal evidence and casual observation suggests that the repetition found in 10-K reports is typically "cut-and-paste" repetition, in which longer sequences are more likely to be retained. Second, we examine repetitious content for a significant subsample of our firms and find that the vast majority of duplicated content identified that uses shorter *n*-grams consists of sequences of ten words or longer. In fact, it is common to observe multiple sentences or even paragraphs that are copied and pasted between sections of the 10-K.

Our primary measure of disclosure repetition is the total number of repeated 10-grams (i.e., ten-word sequences that appear more than once in the same 10-K filing), scaled by the total number of 10-grams within the 10-K. This variable, *REPETITION*, is an estimate of the percentage of 10-K disclosure that is repeated in the 10-K. We supplement this measure with two

alternatives. The first is the unscaled number of repeated 10-grams that occurs within each 10-K (*REPEATED_N_GRAMS*). The second is the length of the longest repeated sequence of words in the 10-K (*LONGEST_REPETITION*).²

Appendix A presents the differences in disclosure repetition practices between two sample companies, LSB Industries and Ford Motor Company, for fiscal year 2012. Appendix A also presents disclosure statements in regard to accounting policies related to PP&E taken from each company's financial statement notes. Underlined text denotes sequences of ten or more words repeated verbatim in prior sections of the 10-K. Appendix A illustrates that LSB Industries's disclosure is significantly more redundant than that of Ford Motor Company. The higher incidence of disclosure repetition is pervasive throughout LSB Industries's 10-K, reflected by a total 10-K *REPETITION* score of 0.34 for LSB Industries and 0.072 for Ford Motor Company. We find that at least some level of cut-and-paste repetition is pervasive in our sample of 10-Ks.³

The last part of Appendix A is an excerpt from American States Water Company's 10-K for fiscal year 2009. This excerpt, taken from the MD&A, is repeated nearly completely verbatim from disclosures made in the financial statement notes. The sentence in bold italics, which indicates that the fair value of the pension plan assets decreased to \$48.6 million due to changes in market conditions, however, is not repeated elsewhere. Rather, that potentially relevant disclosure is sandwiched between passages of redundant text, possibly increasing the likelihood of its being skipped by readers of the 10-K.

Measuring Litigation Risk and Reporting Incentives

Our first three hypotheses posit how repetition within the 10-K relates to firms' securities litigation risk and reporting incentives. We measure *LITIGATION_RISK* for each firm-year as the fitted values from Kim and Skinner's (2012, Model (3)) coefficient estimates:

$$SUED_{i,t} = -7.883 + (0.566)*FPS_{i,t} + (0.518)*LNASSETS_{i,t-1} + (0.982)*SALESGROWTH_{i,t-1} + (0.379)*RETURN_{i,t-1} + (-0.108)*RETURN_SKEWNESS_{i,t-1} + (25.635)*RETURN_STD_DEV_{i,t-1} + (0.00007)*TURNOVER_{i,t-1}$$

This measure of firm-year-specific litigation risk is used commonly in accounting research (e.g., Ahmed and Duellman 2012; Billings and Cedergren 2015; Iliev, Miller, and Roth 2014).

Following prior research (e.g., Li 2008; Miller 2002; Schrand and Walther 2000), we proxy for firms' 10-K financial reporting incentives using firms' earnings performance. Our primary performance variable is *EARNINGS*, measured as operating income scaled by total assets. For robustness, we also examine two other measures of firm performance designed to capture reporting incentives. The first is *LOSS*, which is an indicator variable equal to 1 for firm-years in which the company reports negative income before extraordinary items, and 0 otherwise. The second is *MISS_AF*, which is an indicator variable equal to 1 for firm-years in which the company missed the consensus analyst earnings forecast, and 0 otherwise. Because *MISS_AF* requires I/B/E/S data, analyses that incorporate this variable are run on a smaller sample of observations than is used in the main analyses.

Control Variables

We are interested in examining the relationship between managers' discretionary provision of repetitious disclosure and earnings performance. We control for a host of firm characteristics that may affect firms' exposure to overlapping disclosure requirements and that may thus affect the level of repetition due to factors besides managers' reporting preferences. We also control for several determinants of 10-K length, as suggested by Li (2008), who finds that 10-Ks that convey worse earnings news tend to be longer. We also include firm-fixed effects to ensure that our results are not driven by unmodeled differences across firms.⁴ Thus, our tests measure the association between within-firm variation in disclosure repetition and within-firm variation in earnings performance and litigation risk.

Our control variables are as follows: *SIZE* is the natural logarithm of the market value of equity; *MTB* is the market-to-book ratio; *FIRM_AGE* is number of years that the firm has been covered on CRSP; *SPECIAL_ITEMS* are special items as reported by Compustat, scaled by assets; *FY_VOLAT* is the standard deviation of monthly returns during the fiscal year, multiplied by the square root of 12; *EARN_VOL* is the standard deviation of operating earnings over the last five fiscal years;

² Anecdotal evidence suggests that much of the repetition within the 10-K occurs between the financial statement notes and earlier sections of the 10-K. To validate our measure of redundancy, we test for a correlation between *REPETITION* and the percentage of repetitious trigrams that occur specifically between the financial statement notes and earlier sections of the 10-K for a sample of 580 firms randomly selected from 2003 to 2012. We find that *REPETITION* is approximately 74 percent correlated with this measure.

³ Various algorithms exist for measuring document redundancy and similarity. Brown and Tucker (2011) compare MD&A disclosures by the same firm in successive years using a vector space model, which measures document similarity, but that requires a comparison of two distinct documents and does not necessarily highlight cut-and-paste repetition.

⁴ All inferences from our analyses are unchanged if we use industry-fixed effects in place of firm-fixed effects.

NBSEG is the natural logarithm of 1 plus the number of operating segments reported by the firm; *NGSEG* is the natural logarithm of 1 plus the number of geographic segments reported by the firm; *INTANG* is equal to intangible assets scaled by total assets; *DERIVATIVES* is an indicator variable equal to 1 if the firm reports gains or losses from derivatives in other comprehensive income or accumulated other comprehensive income, and 0 otherwise; *ACQUISITION* is an indicator variable equal to 1 if the firm acquired another entity during the fiscal year, and 0 otherwise; *DLW* is an indicator variable equal to 1 for firms incorporated in Delaware, and 0 otherwise; *RD* is research and development scaled by assets; *FY_RET* is the fiscal-year raw return; *CAP_LEASE* is equal to 1 if the firm reports capital lease obligations, and 0 otherwise; *DEBT_RATIO* is total liabilities scaled by total assets; and *CURRENT_RATIO* is the firm's current assets scaled by current liabilities. Finally, we include year-fixed effects to capture the influence of time-varying disclosure requirements that affect all firms.

EMPIRICAL RESULTS

Before examining the results of our multivariate hypothesis tests, we first present descriptive evidence regarding time-series patterns of disclosure repetition in the 10-K over the past two decades. We also provide some descriptive evidence regarding the topics that are most commonly repeated within the 10-K.

Time-Series Evolution of Repeated Disclosures within the 10-K

To provide insight into how the repetition of 10-K disclosures has evolved over time, we plot the annual means of each of our three repetition variables (*REPETITION*, *REPEATED_N_GRAMS*, and *LONGEST_REPETITION*) for each year in our sample period and display these as Figure 1, Panel A, Panel B, and Panel C. Panel A indicates that disclosure repetition as a percentage of total 10-K length has been gradually increasing, on average, over our sample period. There is a dramatic spike in repeated disclosures in 2002, which partially reverses in subsequent years. This spike suggests that firms responded to regulations associated with or contemporaneous to the Sarbanes-Oxley Act (SOX) by providing increased repetition of disclosures in their 10-Ks, consistent with findings in Dyer et al. (2016).

Panel B of Figure 1 shows a more marked increase in disclosure repetition when repetition is measured in unscaled terms. Panel B indicates that the average number of repeated 10-grams has increased nearly fourfold over the sample period. Because total 10-K length also has increased substantially over the sample period, the increase in repeated disclosure, as a percentage of total disclosure (Figure 1, Panel A), is more modest. In Panel C, we see that the mean length of the longest repeated sequence of words in each 10-K has increased somewhat in parallel with *REPETITION*. The mean has increased from approximately 200 words to approximately 350 words. To provide some context, a page of 12-point double-spaced text in Microsoft Word tends to be roughly 300 words. In all three figures, the uptick in redundancy around 2002 is pronounced.⁵

Descriptive Analysis of Duplicated Content

We provide descriptive statistics for each of the measures of repetition examined in this paper in Panel A of Table 2. To provide some insight into the tone and content of disclosures that are repeated in the 10-K, we provide a descriptive analysis of linguistic attributes for the entire 10-K and repeated content separately and display the results in Panel B of Table 2. We partition the descriptive analysis for Panels A and B between the pre-SOX and post-SOX periods because Figure 1 indicated a change in disclosure repetition levels around the passage of SOX.⁶ Panel C of Table 2 presents a correlation matrix for each measure or repetition examined in this paper, along with our two independent variables of interest.

Table 2 confirms, as documented in prior literature, that 10-K length has increased substantially over recent years. Mean 10-K word length is double in the post-SOX period relative to that in the pre-SOX period. In addition, mean *REPETITION* increased from 0.148 to 0.199 over the same time period. In untabulated analyses, we also find that 10-Ks with the most disclosure volume tend to have the highest levels of *REPETITION*. Mean *REPETITION* is approximately 0.23 in our sample among firms in the top quintile of 10-K length.

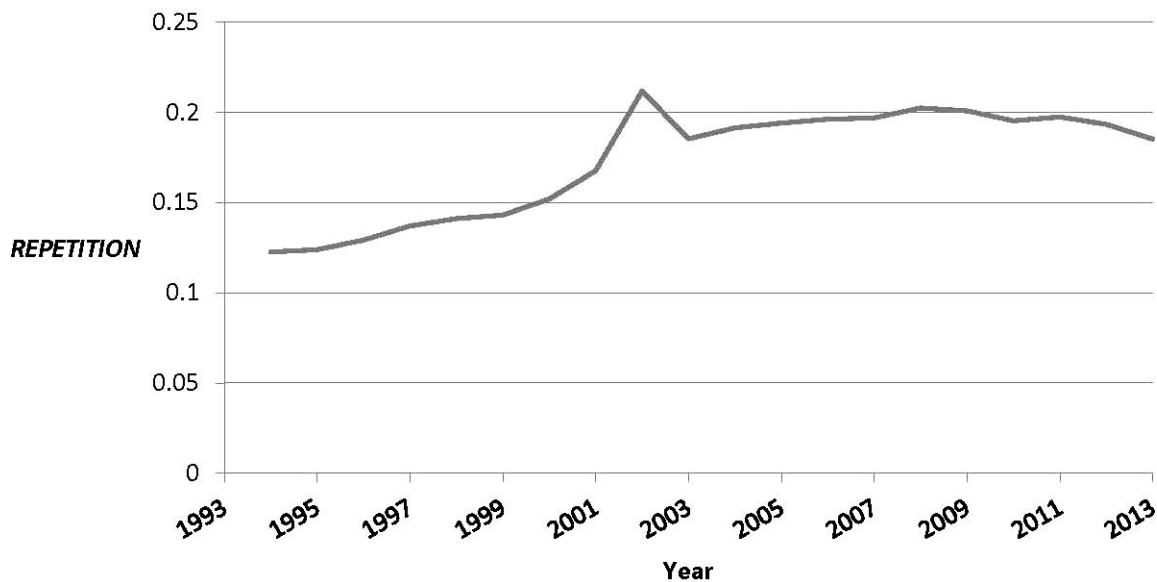
POSITIVE, *NEGATIVE*, *UNCERTAIN*, *FORWARD*, and *LITIGIOUS* are all measured as the word count for words that appear in Loughran and McDonald's (2011) word dictionaries for positive words, negative words, uncertain words, forward-looking words, and litigious words, scaled by total words. *ACCOUNTING_POLICY* is the equivalent measure from a list of

⁵ In each figure, a decrease in mean *REPETITION* is observed in fiscal year 2013. We note, however, that we have only partial coverage of firms for fiscal year 2013, so the distribution of firms in 2013 may not be representative of that of earlier years.

⁶ To facilitate the measure of certain variables, such as *FOG*, this analysis of redundant content is based on entire repeated sentences rather than only on repeated 10-grams. *FOG* requires inputs, such as sentence length, that cannot be performed on word sequences that do not comprise an entire sentence. We note that repeated content measured on the basis of entire verbatim sentences is significantly less than that based on 10-grams due to small changes in wording that result in Type II errors when measuring repetition in terms of identical sentences.

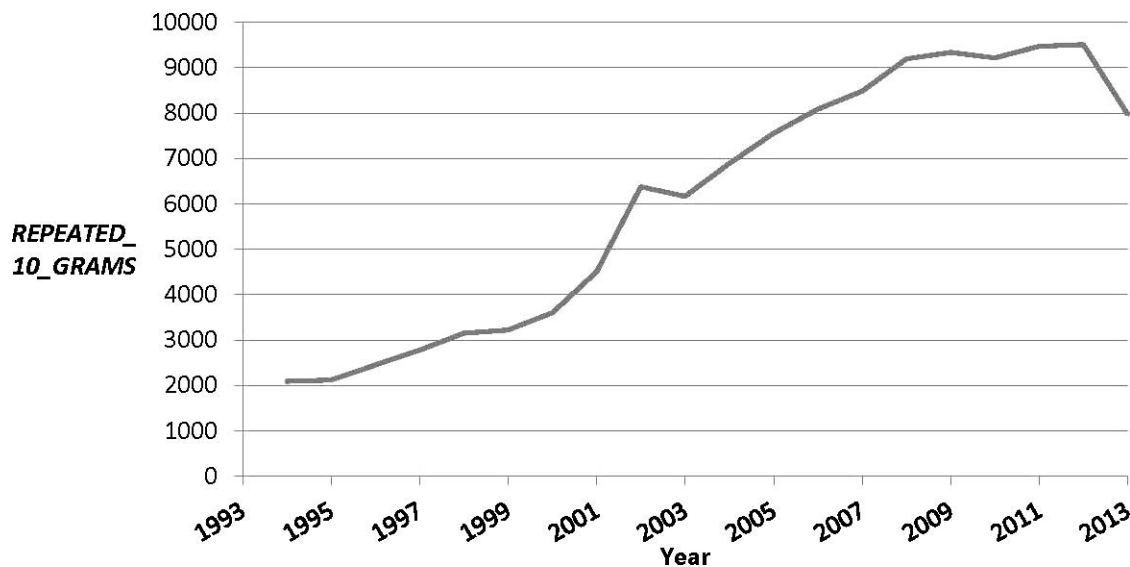
FIGURE 1
Time Trends in Measures of 10-K Repetition

Panel A: Mean Number of Repeated Ten-Word Sequences as a Percentage of Total Ten-Word Sequences in the 10-K



Panel A shows the annual mean of *REPETITION*, which is *REPEATED_N_GRAMS* as a percentage of the total number of ten-word sequences in the 10-K, for each year in the sample period (1994–2013). A more detailed explanation of the construction of *REPETITION* can be found in the “Sample Selection” section.

Panel B: Mean Number of Repeated Ten-Word Sequences in 10-K Filings from 1994–2013

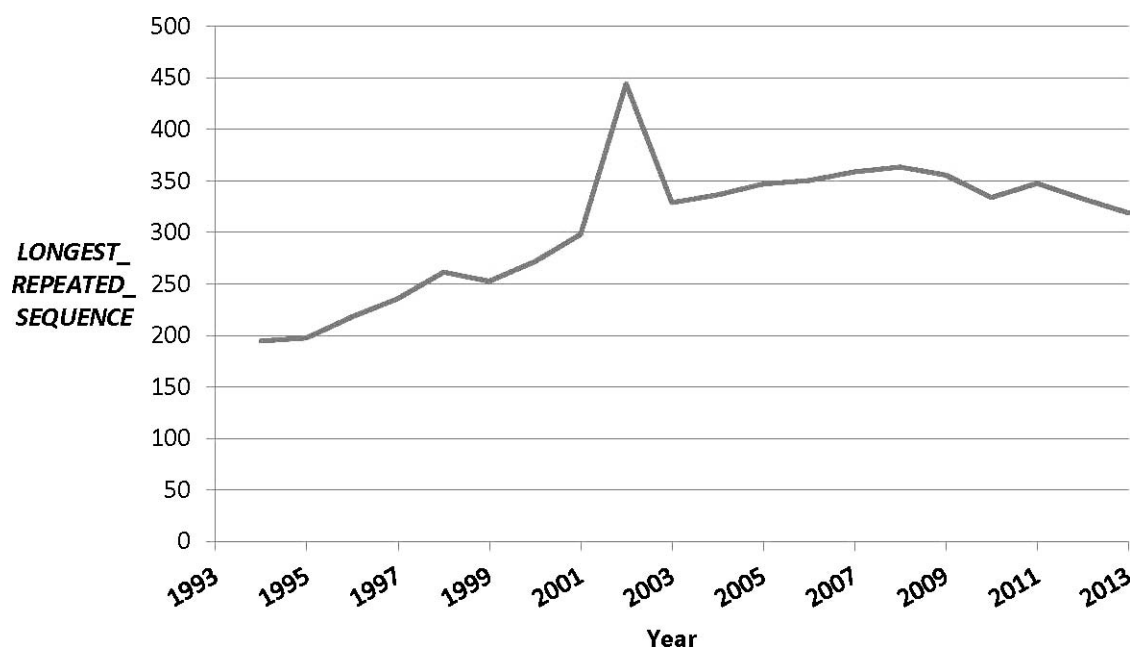


Panel B shows the mean of *REPEATED_10_GRAMS* for each year during the sample period (1994–2013). *REPEATED_N_GRAMS* is the number of ten-word sequences in the 10-K that are repeated at least once in the 10-K. A more detailed explanation of the construction of *REPEATED_N_GRAMS* can be found in the “Sample Selection” section.

(continued on next page)

FIGURE 1 (continued)

Panel C: Mean Length in Words of the Longest Continuous Sequence of Repeated Text in the 10-K



Panel C shows the mean length of the longest sequence of repeated text in 10-K filings (*LONGEST_REPETITION*) for each year from 1994 through 2013. A more detailed explanation of the construction of *LONGEST_REPETITION* can be found in the “Sample Selection” section.

words that tend to be used in firms’ summary of significant accounting policies in the notes to the financial statements.⁷ Untabulated t-tests indicate that the means of the linguistic features in Table 2, Panel B are all statistically significantly different from their counterparts in Panel A (repetitious content versus the rest of the 10-K), although visual examination suggests that not all of the differences are large in magnitude. The two largest differences between repeated content and the rest of the 10-K are in the higher frequency of litigious words and accounting policy words in repetitious content. The frequency of accounting policy words in redundant content increased by 70 percent in the post-SOX period, from a mean of 0.017 to 0.029.

Table 3 provides further insight into the content of repeated disclosures. For both the pre- and post-SOX eras, we display in the left-hand column the most frequently used words in repeated disclosures, in descending order of frequency. Because many of these words may be common to multiple sections of the 10-K and, thus, not reflective of the topic being discussed, we display the words that are most distinctive to repeated disclosures in the right-hand column for both pre- and post-SOX disclosures. The lists in the right-hand column are formed by scaling the word frequency for each word in repeated content by its frequency in the 10-K, which provides an indication of how much more likely a word is to appear in repeated disclosures relative to its frequency elsewhere in the 10-K.⁸

Table 3 indicates that, in the pre-SOX era, words most distinctive to repeated disclosures were often related to litigation (in italic in the pre-SOX column in Table 3). In the post-SOX era, words distinctive to accounting policies became much more common (in italic in the post-SOX column in Table 3). In the last column of Table 3, we list the words whose frequency in repeated content increased the most from the pre-SOX to post-SOX eras. This column underscores the notion that words associated with accounting policies (in italic in the last column of Table 3) have become increasingly prevalent in disclosures that are repeated. This evidence supports anecdotal claims that, despite the SEC’s admonition not to repeat information in the

⁷ We constructed the accounting policy words, displayed in Appendix B, because anecdotal evidence and some research evidence assert that the accounting policies tend to comprise a significant portion of repeated content (Li 2014).

⁸ We have excluded common “stop words” from this list, such as “the,” “an,” and “a,” as well as other words, such as “calendar months,” that convey no topical meaning.

TABLE 2
Descriptive Analyses

Panel A: Descriptive Statistics for Dependent and Key Independent Variables

Variable	Pre-SOX (1994–2001)		Post-SOX (2002–2013)	
	Median	Mean	Median	Mean
<i>10K_WORD_COUNT</i>	17,872.000	19,486.330	37,490.000	39,241.470
<i>REPETITION</i>	0.134	0.148	0.193	0.199
<i>REPEATED_N_GRAMS</i>	2,375.000	3,274.870	7,177.000	8,251.880
<i>LONGEST_REPETITION</i>	165.000	159.034	292.000	348.128
<i>REPETITION_FASB_PRONOUNCEMENTS</i>	0.000	0.163	1.000	0.633
<i>REPETITION_EARNINGS_RELATED</i>	0.459	0.529	0.716	0.743
<i>EARNINGS</i>	0.067	0.008	0.050	0.010
<i>LITIGATION_RISK</i>	0.020	0.035	0.028	0.045

Panel B: Textual Features of Repeated Content versus the Entire 10-K

Variable	Pre-SOX (1994–2001)		Post-SOX (2002–2013)	
	Median	Mean	Median	Mean
Entire 10-K				
<i>FOG</i>	18.705	18.702	18.598	18.568
<i>POSITIVE</i>	0.005	0.006	0.006	0.006
<i>NEGATIVE</i>	0.009	0.010	0.013	0.014
<i>UNCERTAIN</i>	0.009	0.010	0.012	0.012
<i>FORWARD</i>	0.007	0.007	0.008	0.009
<i>LITIGIOUS</i>	0.007	0.008	0.007	0.008
<i>ACCOUNTING_POLICY</i>	0.014	0.014	0.021	0.020
Repeated Content Only				
<i>FOG</i>	17.130	17.952	18.020	17.726
<i>POSITIVE</i>	0.001	0.004	0.005	0.005
<i>NEGATIVE</i>	0.007	0.013	0.013	0.015
<i>UNCERTAIN</i>	0.006	0.009	0.011	0.011
<i>FORWARD</i>	0.006	0.008	0.008	0.009
<i>LITIGIOUS</i>	0.006	0.013	0.008	0.011
<i>ACCOUNTING_POLICY</i>	0.009	0.017	0.025	0.029

(continued on next page)

MD&A and the financial statement notes, firms persist in repeating their significant accounting policies in both the financial statement notes and in the critical accounting estimates section of the MD&A.

Determinants of 10-K Disclosure Repetition

We next examine the results of our primary tests of our first three hypotheses, which posit a relationship between 10-K disclosure repetition and firms' securities litigation risk and performance. Table 4 presents results of regressing alternative measures of disclosure repetition on *EARNINGS* and *LITIGATION_RISK*, along with control variables. The first column shows the results when using our primary measure of the dependent variable *REPETITION*. The results indicate a strongly negative coefficient estimate on *EARNINGS*, consistent with repetitions being greater when firms' performance-related incentives to report transparently are low. This result is consistent with the obfuscation hypothesis of 10-K repetition, and inconsistent with the salience hypothesis. In addition, the coefficient estimate on *LITIGATION_RISK* is positive and strongly significant, suggesting that incentives to preempt allegations of insufficient disclosure also may help to drive some 10-K repetition.⁹

⁹ To ensure that the positive coefficient estimate on *LITIGATION_RISK* is not reflecting repetitive disclosures of actual lawsuits, we re-estimate this regression after excluding firms that had experienced an actual securities lawsuit within the prior two years. The revised coefficient estimate remains positive and strongly significant ($p < 0.01$), and the magnitude of the revised coefficient estimate is similar (0.0209 versus 0.0240).

TABLE 2 (continued)

Panel C: Pearson Correlation Coefficients (n = 63,695)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) <i>REPETITION</i>	1	0.76	0.63	0.40	0.44	-0.06	0.07
		< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
(2) <i>REPEATED_N_GRAMS</i>		1	0.52	0.43	0.71	-0.06	0.25
			< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
(3) <i>LONGEST_REPETITION</i>			1	0.24	0.24	-0.06	0.06
				< 0.001	< 0.001	< 0.001	< 0.001
(4) <i>REPETITION_FASB_PRONOUNCEMENTS</i>				1	0.36	-0.05	0.04
					< 0.001	< 0.001	< 0.001
(5) <i>REPETITION_EARNINGS_RELATED</i>					1	0.18	0.27
						< 0.001	< 0.001
(6) <i>EARNINGS</i>						1	0.08
							< 0.001
(7) <i>LITIGATION_RISK</i>							1

Panel A presents the means and medians of our measures of 10-K repetition and key independent variables, broken out between pre- and post-SOX periods. Panel B presents means and medians of certain lexical features of disclosure for the entire 10-K and, separately, for repeated content. Panel C presents the Pearson correlation coefficients for all measures of 10-K disclosure repetition used in the paper, as well as our two primary explanatory variables *LITIGATION_RISK* and *EARNINGS*. Correlation coefficients are in bold. p-values are listed below each correlation coefficient. *POSITIVE*, *NEGATIVE*, and *LITIGIOUS* are all based on the word lists used in Loughran and McDonald (2014). *ACCOUNTING_POLICY* is from a word list we construct for the purposes of this study, and is available upon request.

Variable Definitions:

10K_WORD_COUNT = the total number of words in the 10-K;

REPETITION = the total number of repeated ten-word sequences (10-grams) in the 10-K, scaled by the total number of ten-word sequences in the 10-K;

LONGEST_REPETITION = the longest sequence of words in the 10-K that is duplicated at least once elsewhere in the 10-K, measured in words;

REPEATED_N_GRAMS = the unscaled equivalent of *REPETITION*, or just the total number of repeated ten-word sequences in the 10-K;

REPETITION_FASB_PRONOUNCEMENTS = an indicator variable equal to 1 if the firm repeats disclosures relating to the FASB, and 0 otherwise;

REPETITION_EARNINGS_RELATED = the number of repeated passages in the 10-K that contain the words “income” or “earnings,” scaled by the total number of words in the 10-K;

FOG = the Gunning Fog Index of readability;

EARNINGS = operating income after depreciation, scaled by assets;

LITIGATION_RISK = the fitted value of litigation risk using the parameter estimates from Kim and Skinner’s (2012, Model (3));

POSITIVE = the total number of positive words, scaled by the total number of words;

NEGATIVE = the total number of negative words, scaled by the total number of words;

FORWARD = the total number of forward-looking words, where forward-looking words are as identified by Li (2010), scaled by total words;

LITIGIOUS = the total number of litigious words scaled by total words; and

ACCOUNTING_POLICY = the total number of words typically associated with firms’ accounting policy disclosures, scaled by the total number of words.

The estimated associations between 10-K repetition and certain of our control variables in Table 4 are also informative. First, the coefficient estimate on *SPECIAL_ITEMS* is negative and significant. Since negative values of *SPECIAL_ITEMS* are income increasing, this negative coefficient is consistent with firms repeating more information when net income is reduced either through operating income (*EARNINGS*) or through one-time charges. The significantly positive coefficient estimate on *FY_VOLAT* suggests higher return volatility corresponds to greater repetition, even though this association could be consistent with the obfuscation or the salience hypothesis. Finally, we note that the coefficient estimate on *FY_RET* is insignificant, suggesting that repetition correlates more with measures of earnings performance than with return performance. This may not be surprising given that the 10-K is largely devoted to discussing factors relating to annual earnings performance.

The second and third columns of Table 4 show that our main results are robust alternative measures of 10-K disclosure repetition. The second column displays results based on using *REPEATED_N_GRAMS* as the dependent variable. The third column displays results using *LONGEST_REPETITION* as the dependent variable. In all specifications, the coefficient estimate is significantly negative on *EARNINGS* and significantly positive on *LITIGATION_RISK*, providing support for the obfuscation hypothesis and the litigation-risk hypothesis of 10-K disclosure repetition.

In Table 5, we present results based on our primary measure of 10-K disclosure repetition, *REPETITION*, as the dependent variable, and alternative measures of earnings performance as explanatory variables. The first column of Table 5 presents the results when we replace *EARNINGS* with the *LOSS* indicator. The second column of Table 5 provides the results when using *LOSS* and *MISS_AF* together as earnings-performance indicators. The third column of Table 5 presents the results from including all three proxies for earnings performance as explanatory variables. In all specifications, the coefficient estimates on

TABLE 3**Which Topics Are Disclosed Most Repetitiously?**

Pre-SOX		Post-SOX		Pre- versus Post-SOX
Most Common Words in Repetitious Disclosures	Words Distinctive to Repetitious Disclosures	Most Common Words in Repetitious Disclosures	Words Distinctive to Repetitious Disclosures	Words Whose Frequency Distribution in Repetitious Disclosures Increased Most after SOX
agreement	<i>damages</i>	agreement	servicing	<i>guidance</i>
incorporated	<i>court</i>	value	court	<i>controls</i>
interest	<i>district</i>	interest	<i>guidance</i>	servicing
plan	audit	assets	certificate	WMMSC
management	<i>action</i>	fair	<i>FASB</i>	<i>criteria</i>
costs	registration	accounting	district	<i>evaluation</i>
income	derivative	incorporated	<i>adoption</i>	<i>interim</i>
common	exhibit	based	shall	pooling
shares	indenture	income	<i>SFAS</i>	<i>pass-through</i>
operations	reasonable	costs	entity	<i>FIN</i>
assets	incorporated	credit	action	<i>measurement</i>
net	<i>settlement</i>	operations	series	certification
accounting	free	issued	registrants	<i>ASC</i>
credit	<i>shall</i>	related	registration	<i>variable</i>
registration	actions	tax	measurement	<i>ASU</i>
purchase	<i>claims</i>	SFAS	arrangement	residential
corporation	agent	U.S.	judgment	<i>valuation</i>
partnership	perform	material	<i>ASC</i>	<i>interpretation</i>
rate	amended	effective	mortgage	<i>assessment</i>
shall	<i>file</i>	rate	entities	<i>historical</i>
services	letter	net	beginning	<i>assumptions</i>
value	<i>case</i>	shares	settlement	<i>FASB</i>
amended	covenants	services	case	<i>EITF</i>
material	against	common	<i>standard</i>	entities
operating	consent	plan	<i>criteria</i>	measurements
related	amendment	bank	applied	recent
court	<i>litigation</i>	management	amendment	<i>FSP</i>
based	presentation	future	association	procedures
products	trustee	market	final	entity
bank	agreement	results	agreed	intangible
general	conducted	revenue	agreement	expect
file	audits	price	incorporated	disclosure
results	SFAS	loan	certification	chief
issued	warrant	current	claim	<i>allowance</i>
capital	final	reporting	requires	record
additional	agreed	purchase	bankruptcy	circumstances
time	pending	mortgage	covenants	internal
audit	notice	recognized	disclosures	<i>impairment</i>
commission	restated	additional	early	likely
facility	standards	time	<i>effective</i>	covered
through	test	facility	recognize	<i>goodwill</i>
option	disclosures	required	notice	experience
amendment	taken	upon	recognition	considered
basis	commission	basis	<i>amended</i>	arrangements
price	LLC	through	evaluation	items

(continued on next page)

TABLE 3 (continued)

Table 3 shows the words with the highest frequency distribution in 10-K disclosures repeated in the pre-SOX versus post-SOX era, listed in descending order of frequency. To identify the words most distinctive to redundant disclosures in each time period (as opposed to words most frequent throughout the entire 10-K), we divided the number of times each word appeared in repeated sentences by the number of times the word appeared throughout the entire 10-K. These lists reflect the removal of certain stop words, or commonly appearing words that do not convey information about the topic being discussed, such as articles, months of the year, and dates. Litigation-related disclosures are particularly likely to be repeated in the pre-SOX era, as reflected by the relative frequency of litigation-related words (shown in *italic*). In the post-SOX era, words often found in the accounting policies disclosures become significantly more prevalent (shown in *italic*). This topical shift is seen most strongly in the last column, which identifies the words that exhibited the greatest increase in frequency distribution in repetitious content between the pre- and post-SOX eras.

LOSS and *MISS_AF* are significantly positive, and the coefficient estimate on *EARNINGS* remains significantly negative, consistent with greater disclosure repetition when performance-related incentives to report transparently are low. The coefficient estimate on *LITIGATION_RISK* remains significantly positive across all specifications. Together, these results provide further support for the obfuscation hypothesis and the litigation-risk hypothesis of 10-K disclosure repetition and no support for the salience hypothesis.

The 10-K Repetition-Earnings Relationship between High and Low Litigation Risk Firms

A potential concern with our analyses in Tables 4 and 5 is that our earnings-performance measures could just be additional proxies for firms' litigation concerns rather than for their market-based reporting incentives. To determine whether the negative relationship between our proxies for earnings performance and 10-K disclosure repetition is driven by litigation risk concerns not captured by *LITIGATION_RISK*, we split our sample at the median level of *LITIGATION_RISK* and repeat our analysis. The results are presented in Table 6. The first column of Table 6 displays results for firms with relatively high litigation risk, and the second column of Table 6 displays results for firms with relatively low litigation risk. The results indicate that the coefficient estimate on *EARNINGS* is statistically insignificant in the first column but is significantly negative in the second column. Thus, even firms with a relatively low likelihood of being sued are more likely to repeat 10-K disclosures when earnings performance is poor, suggesting this behavior is not driven by litigation risk. However, the coefficient estimates on *LOSS* is significantly positive in both specifications, suggesting that the tendency for poorly performing firms to provide more repetitious disclosure persists for firms, regardless of their litigation concerns. Although the coefficient estimate on *EARNINGS* is not statistically significant among high litigation risk firms, we find that this coefficient estimate is not statistically different from that for low litigation risk firms.

Separating Repetition of Relevant from Mundane Disclosure

It is possible that earnings performance relates not only to the quantity but also to the type of disclosure that is duplicated in the 10-K. The obfuscation hypothesis predicts that poorly performing firms are more likely to repeat mundane or less useful disclosures. The salience hypothesis predicts that strong performers are more likely to repeat relevant or favorable disclosures. We next examine whether the type of disclosure repeated in the 10-K varies with firms' earnings performance. We first measure the amount of repetition in the 10-K that relates to earnings, which we expect reflects disclosure that is of particular interest to investors. *REPETITION_EARNINGS_RELATED* is equal to the number of repeated passages in the 10-K that mentions either "earnings" or "income," scaled by the number of words in the 10-K. We next proxy for firms' repetition of mundane disclosure with *REPETITION_FASB_PRONOUNCEMENTS*, which is an indicator variable equal to 1 if firms repeat disclosures regarding recent FASB pronouncements, and 0 otherwise.¹⁰ Firms sometimes describe recent FASB pronouncements in their discussion of critical accounting estimates within the MD&A or in the significant accounting policies note. However, these accounting pronouncements apply to all firms and are not specific to a company's own accounting choices. Regulators, audit firms, and financial executives have expressed the view that repetition of generic accounting policies in multiple sections of the 10-K is unlikely to be helpful to investors (e.g., Higgins 2014; EY 2015; KPMG 2011).

Panel A of Table 7 shows results from regressing *REPETITION_EARNINGS_RELATED* on our earnings measures and *LITIGATION_RISK*. All control variables from prior tables are included but are suppressed from Table 7 for expositional

¹⁰ Specifically, we code *REPETITION_FASB_PRONOUNCEMENTS* equal to 1 if at least one repeated passage of the 10-K mentions the FASB. An examination of instances of repeated disclosures within the 10-K that mention the FASB indicates that the vast majority of these disclosures relate either to recent FASB pronouncements or to a description of the firm's significant accounting policies. We model the repetition of FASB pronouncements as a binary choice because firms typically repeat these disclosures no more than once within the 10-K.

TABLE 4
The Association between Earnings Performance, Litigation Risk, and Alternative Measures of 10-K Disclosure Repetition

	<i>REPETITION</i> Parameter Estimate Standard Error	<i>REPEATED_N_GRAMS</i> Parameter Estimate Standard Error	<i>LONGEST_REPETITION</i> Parameter Estimate Standard Error
Independent Variables			
<i>EARNINGS</i>	−0.0133*** 0.0022	−0.2351*** 0.0212	−0.1418*** 0.0274
<i>LITIGATION_RISK</i>	0.0240*** 0.0074	0.6783*** 0.0716	0.2799*** 0.0924
Control Variables			
<i>SIZE</i>	−0.0040***	−0.0127***	−0.0326***
<i>MTB</i>	0.0000	−0.0003	0.0001
<i>FIRM_AGE</i>	0.0005***	0.0011	0.0041**
<i>SPECIAL_ITEMS</i>	−0.0341***	−0.3961***	−0.2133***
<i>FY_VOLAT</i>	0.0100***	0.1795***	0.1200***
<i>EARN_VOL</i>	0.0000*	0.0001***	0.0000
<i>NBSEG</i>	0.0025***	0.0556***	0.0134
<i>NGSEG</i>	0.0036***	0.1101***	0.0595***
<i>INTANG</i>	0.0104***	0.2888***	0.1179***
<i>DERIVATIVES</i>	−0.0004	0.0921***	0.0302***
<i>ACQUISITION</i>	−0.0012**	0.0150**	−0.0070
<i>RD</i>	−0.0256***	−0.3124***	−0.1663**
<i>FY_RET</i>	0.0001	−0.0046	0.0052
<i>CAP_LEASE</i>	0.0011	0.0495***	0.0201**
<i>OP_LEASE</i>	−0.0015*	−0.0082	0.0081
<i>DEBT_RATIO</i>	0.0161***	0.1841***	0.0787***
<i>CURRENT_RATIO</i>	−0.0003***	−0.0021**	−0.0002
Year-fixed effects included	Yes	Yes	Yes
Firm-fixed effects included	Yes	Yes	Yes
n	63,695	63,695	63,695
R ²	0.680	0.934	0.586

*, **, *** Denote statistical significance at the 0.10, 0.05, and 0.01 levels of significance, respectively.

Table 4 shows the relationship between annual earnings and three alternative measures of 10-K disclosure repetition. All continuous variables are winsorized at the 1st and 99th percentiles.

Variable Definitions:

REPETITION = the total number of repeated ten-word sequences (10-grams) in the 10-K, scaled by the total number of ten-word sequences in the 10-K;

REPEATED_N_GRAMS = the total number of ten-word sequences that are duplicated at least once elsewhere in the 10-K;

LONGEST_REPETITION = the length (in words) of the longest sequence of words that is duplicated at least once elsewhere in the 10-K;

EARNINGS = operating income after depreciation, scaled by assets;

LITIGATION_RISK = the fitted value of litigation risk using the parameter estimates from Kim and Skinner's (2012, Model (3));

SIZE = the natural logarithm of the market value of equity;

MTB = the market value of equity scaled by the book value of equity;

FIRM_AGE = the number of years since the firm was first listed on CRSP;

SPECIAL_ITEMS = equal to special items as reported by Compustat, scaled by assets;

FY_VOLAT = the standard deviation of 12-month returns ending on the last day of the fiscal year;

EARN_VOL = the standard deviation of annual earnings over the five-year period ending in the current fiscal year;

NBSEG = the natural logarithm of 1 plus the number of operating segments reported by Compustat;

NGSEG = the natural logarithm of 1 plus the number of reported geographic segments;

INTANG = intangible assets reported on the balance sheet scaled by assets;

DERIVATIVES = an indicator variable equal to 1 if the firm reports current or accumulated gains or losses from derivatives in other comprehensive income, and 0 otherwise;

ACQUISITION = an indicator variable equal to 1 if the firm reports cash or shares issued for an acquisition during the year, and 0 otherwise;

DLW = an indicator equal to 1 if the firm is incorporated in Delaware, and 0 otherwise;

RD = research and development expenditures scaled by assets;

(continued on next page)

TABLE 4 (continued)

FY_RET = the fiscal year buy-and-hold return;

CAP_LEASE = an indicator variable equal to 1 if the firm reports a capital lease on its balance sheet, and 0 otherwise;

OP_LEASE = an indicator variable equal to 1 if the firm reports a minimum operating lease obligation due in the next year of at least 1 percent of assets, and 0 otherwise;

DEBT_RATIO = total liabilities scaled by assets; and

CURRENT_RATIO = current assets scaled by current liabilities.

convenience. Panel A of Table 7 shows that the coefficient estimates on our earnings variables switch signs relative to those when total repetition is the dependent variable of interest. Specifically, the coefficient estimate on *EARNINGS* is significantly positive, and the coefficient estimates on *LOSS* and *MISS_AF* are significantly negative. This evidence suggests that the salience hypothesis explains at least some disclosure repetition, in that firms with stronger earnings performance are more likely

TABLE 5

The Association between Disclosure Repetition and Alternative Measures of Earnings Performance

Independent Variables	<i>REPETITION</i>		
	Parameter Estimate Standard Error		
<i>EARNINGS</i>			−0.0049* 0.0027
<i>LOSS</i>	0.0062*** 0.0007	0.0062*** 0.0007	0.0053*** 0.0008
<i>MISS_AF</i>		0.0016*** 0.0005	0.0011** 0.0005
<i>LITIGATION_RISK</i>	0.0238*** 0.0074	0.0248*** 0.0082	0.0250*** 0.0082
Control Variables			
<i>SIZE</i>	−0.0038***	−0.0037***	−0.0030***
<i>MTB</i>	0.0000	0.0000	0.0000
<i>FIRM_AGE</i>	0.0005***	0.0006***	0.0005***
<i>SPECIAL_ITEMS</i>	−0.0273***	−0.0375***	−0.0290***
<i>FY_VOLAT</i>	0.0093***	0.0113***	0.0100***
<i>EARN_VOL</i>	0.0000*	0.0000***	0.0000***
<i>NBSEG</i>	0.0024***	0.0020**	0.0020**
<i>NGSEG</i>	0.0036***	0.0039***	0.0038***
<i>INTANG</i>	0.0104***	0.0111***	0.0103***
<i>DERIVATIVES</i>	−0.0003	−0.0010	−0.0010
<i>ACQUISITION</i>	−0.0012**	−0.0010	−0.0009
<i>RD</i>	−0.0119**	−0.0148***	−0.0207***
<i>FY_RET</i>	0.0001	−0.0004	−0.0002
<i>CAP_LEASE</i>	0.0010	0.0014*	0.0013
<i>OP_LEASE</i>	−0.0015*	−0.0013	−0.0014***
<i>DEBT_RATIO</i>	0.0161***	0.0182***	0.0174***
<i>CURRENT_RATIO</i>	−0.0003***	−0.0004***	−0.0004
Year-fixed effects included	Yes	Yes	Yes
Firm-fixed effects included	Yes	Yes	Yes
n	63,695	51,092	51,092
R ²	0.680	0.679	0.679

*, **, *** Denote statistical significance at the 0.10, 0.05, and 0.01 levels of significance, respectively.

Table 5 reports the association between *REPETITION* and alternative measures of earnings performance. *LOSS* is an indicator variable equal to 1 for firm-years in which income before extraordinary items is less than 0, and 0 otherwise. *MISS_AF* is an indicator variable equal to 1 for firm-years in which the firm's reported earnings are less than the consensus analyst forecast, and 0 otherwise. All other variables are defined in the notes to Table 4. All continuous variables are winsorized at the 1st and 99th percentiles.

TABLE 6
The Association between Firm Performance and Disclosure Repetition in High versus Low Litigation Risk Firms

	<i>REPETITION</i>	
	High Litigation Risk Parameter Estimate <i>Standard Error</i>	Low Litigation Risk Parameter Estimate <i>Standard Error</i>
Independent Variables		
<i>EARNINGS</i>	0.0036 <i>0.0050</i>	−0.0112*** <i>0.0036</i>
<i>LOSS</i>	0.0050*** <i>0.0013</i>	0.0036*** <i>0.0012</i>
<i>MISS_AF</i>	0.0017** <i>0.0008</i>	0.0007 <i>0.0007</i>
<i>LITIGATION_RISK</i>	0.0294*** <i>0.0100</i>	0.0448 <i>0.0757</i>
Control Variables		
<i>SIZE</i>	−0.0040***	−0.0037***
<i>MTB</i>	−0.0001	0.0001
<i>FIRM_AGE</i>	0.0005***	0.0001
<i>SPECIAL_ITEMS</i>	−0.0380***	−0.0214***
<i>FY_VOLAT</i>	0.0107***	0.0041**
<i>EARN_VOL</i>	0.0000***	0.0000
<i>NBSEG</i>	0.0008	0.0084***
<i>NGSEG</i>	0.0004	0.0092***
<i>INTANG</i>	0.0137***	0.0037
<i>DERIVATIVES</i>	−0.0019	0.0033**
<i>ACQUISITION</i>	−0.0001	−0.0007
<i>RD</i>	−0.0107	−0.0231***
<i>FY_RET</i>	0.0005	0.0003
<i>CAP_LEASE</i>	0.0003	0.0031**
<i>OP_LEASE</i>	−0.0026*	0.0003
<i>DEBT_RATIO</i>	0.0122***	0.0217***
<i>CURRENT_RATIO</i>	−0.0003**	−0.0001
Year-fixed effects included	Yes	Yes
Firm-fixed effects included	Yes	Yes
n	25,546	25,546
R ²	0.690	0.735

*, **, *** Denote statistical significance at the 0.10, 0.05, and 0.01 levels of significance, respectively.

Table 6 displays results from regressing *REPETITION* on measures of earnings performance, *LITIGATION_RISK*, and control variables for two separate subsamples. The first subsample, reported on the left-hand side of Table 6, consists of firms with litigation risk above the sample median, where litigation risk for each firm is measured using the coefficient estimates from Kim and Skinner (2012, Table 7, Model (3)). The second subsample, reported on the right-hand side of Table 6, consists of firms with litigation risk lower than the sample median. All variables are defined in Tables 4 and 5. Standard errors are clustered by firm. All continuous variables are winsorized at the 1st and 99th percentiles.

to repeat disclosures that discuss earnings. The coefficient estimate on *LITIGATION_RISK* remains positive and significant across all specifications, suggesting that litigation risk also explains earnings-related disclosures.

Panel B of Table 7 shows the results from a logistic regression that estimates the probability of firms repeating disclosures regarding recent FASB pronouncements. Panel B of Table 7 indicates that this proxy for mundane disclosure is negatively associated with *EARNINGS*, consistent with firms being more likely to repeat mundane disclosures when earnings performance is poor. This evidence provides further corroboration for the obfuscation hypothesis. The coefficient estimate on *LOSS*, however, is significantly negative, and the coefficient estimate on *MISS_AF* is statistically insignificant. Together, these results suggest that firms are more likely to repeat mundane disclosures when earnings performance is poor but not necessarily in response to a loss or to a missed analyst forecast. In addition, the coefficient estimate on *LITIGATION_RISK* is statistically

TABLE 7

The Association between Earnings Performance and the Type of Disclosure Repeated

Panel A: Firm Performance and the Repetition of Earnings-Related Disclosures

Independent Variables	<i>REPETITION_EARNINGS-RELATED</i>			
	Parameter Estimate		Standard Error	
<i>EARNINGS</i>	0.1891***		0.1274***	
	0.0125		0.0158	
<i>LOSS</i>		−0.0818***		−0.0646***
		0.0040		0.0048
<i>MISS_AF</i>			−0.0077**	−0.0012
			0.0031	0.0031
<i>LITIGATION_RISK</i>	0.0898**	0.0931**	0.1293***	0.1262**
	0.0422	0.0421	0.0479	0.0477
Table 4 control variables included	Yes	Yes	Yes	Yes
Year-fixed effects included	Yes	Yes	Yes	Yes
Firm-fixed effects included	Yes	Yes	Yes	Yes
n	63,695	63,695	51,092	51,092
R ²	0.668	0.670	0.676	0.678

Panel B: Firm Performance and the Repetition of Mundane Disclosures

Independent Variables	<i>Pr(REPETITION_FASB-PRONOUNCEMENTS)</i>			
	Parameter Estimate		Standard Error	
<i>EARNINGS</i>	−0.2612***		−0.2991***	
	0.0934		0.1115	
<i>LOSS</i>		−0.1194***		−0.1539***
		0.0321		0.0378
<i>MISS_AF</i>			−0.0265	−0.0198
			0.0240	0.0241
<i>LITIGATION_RISK</i>	0.1572	0.2161	0.1833	0.1618
	0.3438	0.3434	0.3782	0.3784
Table 4 control variables included	Yes	Yes	Yes	Yes
Year-fixed effects included	Yes	Yes	Yes	Yes
Firm-fixed effects included	Yes	Yes	Yes	Yes
n	63,695	63,695	51,092	51,092
Area under ROC	0.796	0.796	0.797	0.797

*, **, *** Denote statistical significance at the 0.10, 0.05, and 0.01 levels of significance, respectively.

Panel A of Table 7 displays results from regressing the number of repeated passages in the 10-K that mention “earnings” or “income” (*REPETITION_EARNINGS-RELATED*) on measures of earnings performance and litigation risk. *REPETITION_EARNINGS-RELATED* is scaled by the total number of words in the 10-K. Panel B of Table 7 displays results from a logistic regression of the probability that *REPETITION_FASB-PRONOUNCEMENTS* is equal to 1. *REPETITION_FASB-PRONOUNCEMENTS* is an indicator variable equal to 1 for firm-years in which a company repeats its disclosure of recent FASB pronouncements, and 0 otherwise. All other variables are as defined in Table 4. Coefficient estimates and standard errors reported in Panel A have been multiplied by 1,000 for expositional clarity. All continuous variables are winsorized at the 1st and 99th percentiles.

insignificant, suggesting that firms’ decision to repeat their discussion of recent FASB pronouncements is not tied to litigation concerns.

Disclosure Repetition and the Efficiency of Price Formation

Finally, we examine tests of our fourth hypothesis, which concerns the association between disclosure repetition and the efficiency of price formation subsequent to 10-K filings. We measure inefficiencies in stock price discovery using *DRIFT_REV*,

which we compute as the difference between the short-window (two-day) price reaction to the 10-K filing and the long-window (20-day) reaction to the 10-K filing. More precisely, *DRIFT_REV* is computed as the absolute value of the difference between the cumulative abnormal return over the two-day window and the 20-day window immediately subsequent to the 10-K filing date, scaled by the cumulative abnormal return over the 20-day window, starting on the 10-K filing date. The assumption behind this measure is that the 20-day window cumulative abnormal return better reflects the correct price reaction to the information in the 10-K. Thus, higher values of *DRIFT_REV* reflect more price discovery that occurs after the first two trading days following the 10-K filing, or a less efficient price discovery process.

We control for other variables that may influence the speed with which information from the 10-K becomes impounded into stock price. We first control for the number of analysts who follow the firm (*ANALYST_FOLLOWING*). We control for the dispersion of analyst forecasts with *ANALYST_DISPERSION*, which is the standard deviation of analyst forecasts for the last I/B/E/S statistical period preceding the 10-K filing. We control for the total disclosure volume of the 10-K using the natural logarithm of the number of words in the 10-K (*LOG_WORD_COUNT*), for the annual earnings change (*DELTA_EARNINGS*), the fog of the entire 10-K (*FOG*), and the initial filing date return (*ANN_DATE_CAR*), measured as the short-window cumulative abnormal return for days 0 and 1 relative to the filing date. We also retain all other control variables used in prior tables.

The first column of Table 8 indicates that the coefficient estimate on *REPETITION* is positive but not statistically significant. Because we expect firms are more likely to use repetition for obfuscation purposes when performance is poor, we partition our sample based on the within-firm median of annual changes in performance. Results in the second column of Table 8 indicate that coefficient estimate on *REPETITION* is negative but insignificant when earnings changes are above the median, indicating disclosure repetition does not adversely affect price discovery when performance is positive. The third column of Table 8 indicates the coefficient estimate on *REPETITION* is positive but not quite statistically significant at standard levels of significance ($p < 0.12$). We find the difference between the coefficient estimates on *REPETITION* in the third and second columns of Table 8 is marginally significant ($p < 0.07$, one-tailed test). These results provide modest evidence that disclosure repetition is associated with slower price discovery when performance is poor.

It is possible that within-firm effects are muted due to investors being relatively familiar with the content of 10-K filings from examining the same firm's filings in prior years. We next consider whether within-industry variation in 10-K disclosure repetition is associated with the length of price discovery following 10-K filings by replacing firm-fixed effects with industry-fixed effects. Results in the last three columns of Table 8 suggest 10-K disclosure repetition is significantly associated with slower price discovery after 10-K filings, and that this result appears to be driven by firm-years in which performance is poor. Overall, the results in Table 8 are consistent with the notion that disclosure repetition by poorly performing firms does not help, and may actually hinder, investors' processing of 10-K information.¹¹

Where Is Information Repeated within the 10-K?

Standard setters seeking to improve disclosure requirements are likely to benefit from evidence regarding where repetition tends to appear within the 10-K. To provide some insight into this issue, we measure the amount of repeated 10-grams that exist within the three sections of the 10-K that comprise the greatest volume of disclosure, on average, over our sample period: Item 1 (Business Description), Item 7 (Management's Discussion and Analysis), and Item 8 (Financial Statement Notes). For purposes of Table 9, *REPETITION* is measured as the percentage of 10-grams in each item that are repeated at least once anywhere in the 10-K. Similarly, *REPEATED_N_GRAMS* is the total number of 10-grams in each item that are repeated at least once anywhere in the 10-K.

Panel A of Table 9 provides distribution statistics for *REPETITION* and *REPEATED_N_GRAMS* in each 10-K item. Panel A indicates that Management's Discussion and Analysis is the 10-K section containing the most text that is repeated at least once in the 10-K, with mean and median values of *REPETITION* of approximately 0.25. This suggests that, on average, approximately one quarter of the content within the MD&A is repeated at least once elsewhere in the 10-K. The content of the financial statement notes is repeated somewhat less, with roughly 20 percent of the content appearing elsewhere in the 10-K, on average. We find Item 1 to contain relatively unique content, with less than 10 percent of Item 1 content appearing elsewhere in

¹¹ As an alternative test of the effects of disclosure repetition on the post-10-K information environment, we also examine the association between disclosure repetition and changes in bid-ask spreads around the 10-K filing date. We find no association between disclosure repetition and changes in the average daily bid-ask spread using a pre/post window of four weeks on either side of the 10-K filing date. However, we find weak evidence ($p = 0.088$) that disclosure repetition is associated with lower bid-ask spreads among relatively more strongly performing firms, but not among weakly performing firms. Although this result is weak, the tenor of this result is consistent with that of Table 8, suggesting that repetition does not aid information processing among weakly performing firms.

TABLE 8
Disclosure Repetition and Price Discovery

	<i>DRIFT_REV</i> Parameter Estimate Standard Error					
	Firm-Fixed Effects Included			Industry-Fixed Effects Included		
	All Firms	Earnings Change > Median	Earnings Change < Median	All Firms	Earnings Change > Median	Earnings Change < Median
Independent Variables						
<i>REPETITION</i>	0.4469	-0.2156	0.7168	0.5247***	0.3371	0.7747***
	0.2905	0.4357	0.4573	0.1884	0.2588	0.2751
<i>ANALYST_FOLLOWING</i>	0.0498	0.0686	0.0261	0.0003	0.2296	-0.0268
	0.0393	0.0600	0.0616	0.0263	0.0366	0.0377
<i>ANALYST_DISPERSION</i>	0.0351	0.0334	0.1725	0.0202	-0.0110	0.0461
	0.1256	0.2174	0.1904	0.0790	0.1149	0.1086
<i>LOG_WORD_COUNT</i>	-0.0348	0.0590	-0.1109	-0.0445	0.0343	-0.1404***
	0.0556	0.0827	0.0879	0.0321	0.0440	0.0476
<i>DELTA_EARNINGS</i>	-0.0877	0.2083	0.0871	-0.2148	-0.1792	-0.3841
	0.1809	0.3581	0.4534	0.1745	0.2349	0.3286
<i>FOG</i>	-0.0219	-0.0093	-0.0395	-0.0142	-0.0255	0.0008
	0.0205	0.0303	0.0325	0.0111	0.0153	0.0162
<i>ANN_DATE_CAR</i>	-1.1388***	0.0062	-1.8482	-0.9419	-0.3674	-1.5583***
	0.2971	0.4820	0.4705	0.2623	0.3707	0.3709
Table 4 control variables included	Yes	Yes	Yes	Yes	Yes	Yes
Year-fixed effects included	Yes	Yes	Yes	Yes	Yes	Yes
n	40,536	22,246	18,290	40,354	22,137	18,217
R ²	0.177	0.320	0.312	0.005	0.006	0.008

*, **, *** Denote statistical significance at the 0.10, 0.05, and 0.01 levels of significance, respectively.

Table 8 presents results from regressing a measure of price efficiency following 10-K filing dates on *REPETITION* and control variables. Standard errors are clustered by firm. All continuous variables are winsorized at the 1st and 99th percentiles.

Variable Definitions:

DRIFT_REV = the dependent variable = the abnormal drift or reversal in price following the first two days after the 10-K filing date. It is computed as the absolute value of the difference between the two-day cumulative abnormal return and the 20-day cumulative abnormal return following the 10-K filing, scaled by the 20-day cumulative abnormal return;

ANALYST_FOLLOWING = the number of analysts following the firm;

ANALYST_DISPERSION = the standard deviation of the analyst forecasts immediately prior to the earnings announcement;

LOG_WORD_COUNT = the natural logarithm of the total words in the 10-K;

DELTA_EARNINGS = the annual change in *EARNINGS*;

FOG = the Gunning Fog Index of readability measured over the entire 10-K; and

ANN_DATE_CAR = the cumulative abnormal return over trading days 0 and 1.

the 10-K. Median values of *REPEATED_N_GRAMS* within each section suggest that, on average, about 3,400 n-grams across these three sections of the 10-K are repeated at least once.

Panel B of Table 9 examines how these item-specific measures of repetition correlate with *EARNINGS* and *LITIGATION_RISK* in multivariate analyses. Panel B indicates lower values of *EARNINGS* are associated with more repeated content appearing in each 10-K item. Greater values of litigation risk increase both the proportion (*REPETITION*) and the absolute levels (*REPEATED_N_GRAMS*) of repeated content within Item 1, but neither measure of repetition in the financial statement notes. Litigation risk increases the total level of repeated content in the MD&A ($p < 0.01$), but not the proportion of repeated content, possibly due to litigation risk increasing the volume of both repeated and non-repeated content within the MD&A.

Inter-Temporal Disclosure Repetition versus Intra-Temporal Disclosure Repetition

Our analyses focus on 10-K disclosure repetition, which is a measure of intra-temporal repetition within a single document. Related research has measured inter-temporal repetition, or the similarity of two documents filed in subsequent years by the

TABLE 9
Repetitious Disclosure by 10-K Section

Panel A: Distribution of Repetitious Disclosures across 10-K Sections

10-K Section	25th Pctl.	50th Pctl.	Mean	75th Pctl.
Item 1 (Business Description)				
<i>REPETITION</i>	0.032	0.069	0.093	0.126
<i>REPEATED_N_GRAMS</i>	139	368	290	805
Item 7 (Management's Discussion and Analysis)				
<i>REPETITION</i>	0.163	0.247	0.255	0.338
<i>REPEATED_N_GRAMS</i>	825	1,727	1,384	2,982
Item 8 (Financial Statement Notes)				
<i>REPETITION</i>	0.114	0.187	0.200	0.268
<i>REPEATED_N_GRAMS</i>	459	1,268	904	2,602

Panel B: Earnings, Litigation Risk, and Repeated Disclosures by 10-K Section

	Item 1 (Business Description)		Item 7 (Management's Discussion and Analysis)		Item 8 (Financial Statement Notes)	
	<i>REPETITION</i>	<i>REPEATED_N_GRAMS</i>	<i>REPETITION</i>	<i>REPEATED_N_GRAMS</i>	<i>REPETITION</i>	<i>REPEATED_N_GRAMS</i>
Independent Variables						
<i>EARNINGS</i>	−0.1215*** 0.0043	−0.3297*** 0.0404	−0.0074** 0.0036	−0.2282*** 0.0304	−0.0213*** 0.0051	−0.2666*** 0.0708
<i>LITIGATION_RISK</i>	0.0234*** 0.0045	0.6501*** 0.1458	−0.0020 0.0126	0.3620*** 0.1123	0.0110 0.0186	0.1149 0.2564
Table 4 controls included	Yes	Yes	Yes	Yes	Yes	Yes
Year-fixed effects included	Yes	Yes	Yes	Yes	Yes	Yes
Firm-fixed effects included	Yes	Yes	Yes	Yes	Yes	Yes
n	57,126	57,126	51,290	51,290	33,923	33,923
R ²	0.664	0.772	0.669	0.795	0.703	0.703

*, **, *** Denote statistical significance at the 0.10, 0.05, and 0.01 levels of significance, respectively.

Panel A of Table 9 presents distributional statistics for measures of repeated disclosures in Items 1, 7, and 8 of the 10-K. For each 10-K item, *REPETITION* is measured as the percentage of ten-word sequences in that item that are duplicated at least once within the 10-K, scaled by the total number of ten-word sequences within that item. *REPEATED_N_GRAMS* is the raw number of repeated ten-word sequences within each item of the 10-K. Panel B of Table 9 presents results from estimating the regression from Table 4 on each of these three sections of the 10-K separately.

same firm (e.g., [Brown and Tucker 2011](#); [Brown, Tian, and Tucker 2015](#); [Peterson, Schmardebeck, and Wilks 2015](#)). This research typically relies on the edit distance between two documents, specifically the cosine similarity measure, to assess the comparability of two filings. One interesting question is whether firms' tendency to repeat similar information in the 10-K from one year to the next is related to their tendency to repeat the same information within a single 10-K. We find the Pearson correlation between *REPETITION* for a given 10-K and the cosine similarity that 10-K has with the 10-K filed by the same firm in the previous year is 0.058 ($p < 0.01$). However, in untabulated analyses, we find that controlling for 10-Ks' cosine similarity with the prior year's 10-K has very little effect on the relationship between earnings, litigation risk, and disclosure repetition. In addition, we find that earnings and litigation risk do not predict 10-K cosine similarity scores. Our results suggest disclosure repetition within a given 10-K is largely distinct from the stickiness of 10-K disclosure from one year to the next.

Additional Analyses

Our main results suggest that, on average, 10-K disclosure repetition is inversely related to earnings performance. We also consider whether 10-K repetition is inversely related to both good and bad extremes in performance, which one might expect if managers use repetition to emphasize important information regardless of its favorability. To test whether 10-K repetition is

associated with both performance extremes, we re-estimate our primary regression in Table 4 after replacing *EARNINGS* with its absolute value and interacting it with a loss indicator equal to 1 if earnings are negative. Regression results (untabulated) indicate that 10-K repetition is increasing in the magnitude of *EARNINGS* when performance is negative, but not when earnings are positive. These results do not support the notion that 10-K repetition is significantly driven by managers choosing to repeat disclosures based on their importance.

We also rerun our main results after replacing our measure of litigation risk, which is estimated as a function of firm-performance characteristics and industry affiliation, with an alternative measure of litigation risk recently proposed in the literature (Huang, Hui, and Li 2016). This alternative measure is based on the ideology of the federal judges appointed within each U.S. circuit, and is computed as the probability that a securities lawsuit will be assigned to a liberal judge majority panel, estimated as of the end of each fiscal year. Litigation risk for each firm-year is computed as of the beginning of the firm's fiscal year and is based on the circuit in which each firm is headquartered. In untabulated tests, we find that the coefficient estimate on this alternative measure of litigation risk is positive and strongly significant ($p < 0.01$). These results indicate that the positive association between 10-K repetition and litigation risk is robust to alternative measures of litigation risk.

Finally, we examine whether our two incentive variables of interest, litigation risk and earnings performance, explain 10-K length after controlling for 10-K disclosure repetition. We re-estimate the regression in Table 4 after inserting the log number of words in the 10-K as the dependent variable and controlling for the log number of repeated 10-grams in the 10-K. We find the coefficient estimate on *EARNINGS* continues to be strongly negative ($p < 0.001$) and the coefficient estimate on *LITIGATION_RISK* continues to be strongly positive ($p < 0.001$). The coefficient estimate on the log value of repeated 10-grams is also positive and highly significant ($p < 0.001$). These results suggest repetition of disclosures is not the only mechanism by which performance and litigation risk affect 10-K disclosure length.

DISCUSSION AND CONCLUSION

Standard setters, preparers, and investors have all expressed concern about the current state of financial disclosure quality in firms' 10-K reports and, particularly, the volume of these disclosures. A significant component of this volume stems from repetition found within firms' 10-K filings. Current financial reporting requirements in the United States reflect an accumulation of reporting standards that have been set at different times by different standard setters focused on different objectives. The evolutionary process that has led to current standards has produced similar or overlapping disclosure requirements for different parts of the 10-K. Regulators provide significant discretion to firms in choosing whether to respond to those overlapping disclosure requirements by providing duplicate information or by cross-referencing to the location of that information elsewhere in the 10-K.

In this study, we have conducted a series of analyses to shed light on how managers exercise their discretion in responding to these overlapping disclosure requirements. The results of our tests suggest that managers are more likely to repeat information in response to litigation concerns and to performance-related reporting incentives. On average, we find that firms repeat more information in the 10-K when performance is poor, consistent with the obfuscation hypothesis of 10-K repetition. We also find, however, that strong performers are more likely to repeat disclosures related to earnings, consistent with the salience hypothesis, explaining at least some repetition within the 10-K. We issue the caveat that our proxy for litigation risk, while commonly used in recent research, is a composite of primarily market-based variables and may thus capture market-based reporting incentives beyond litigation risk. However, our results indicate our litigation risk proxy influences disclosure repetition in ways distinct from earnings performance, suggesting both litigation risk and performance are determinants of disclosure repetition.

We find that firms that repeat more information within the 10-K are associated with slower price discovery after 10-K filing dates, particularly when firm performance is poor. Thus, the effect of repetitious disclosure in the 10-K may be to reduce information quality when performance is low and when investors' demand for quality information may be highest. Our results suggest that overlapping disclosure requirements may provide managers with an opportunity to report less transparently when their performance-related incentives to provide clear and concise disclosures is low. These findings support recent calls for the SEC and the FASB to work together to eliminate overlap in reporting requirements, with the aim of eliminating redundant reporting within the 10-K.

REFERENCES

- Ahmed, A., and S. Duellman. 2012. Managerial overconfidence and accounting conservatism. *Journal of Accounting Research* 51 (1): 1–30.
- Belch, G. 1981. An examination of comparative and noncomparative television commercials: The effects of claim variation and repetition on cognitive response and message acceptance. *Journal of Marketing Research* 18 (3): 333–349. doi:10.2307/3150974

- Bergerbest, D., D. Ghahremani, and J. Gabrieli. 2004. Neural correlates of auditory repetition priming: Reduced fMRI activation in the auditory cortex. *Journal of Cognitive Neuroscience* 16 (6): 966–977. doi:10.1162/0898929041502760
- Billings, M., and M. Cedergrén. 2015. Strategic silence, insider selling and litigation risk. *Journal of Accounting and Economics* 59: 119–142.
- Bloomfield, R. 2008. Discussion of “Annual report readability, current earnings, and earnings persistence.” *Journal of Accounting & Economics* 45 (2/3): 248–252. doi:10.1016/j.jacceco.2008.04.002
- Bonner, S. 2007. *Judgment and Decision Making in Accounting*. Upper Saddle River, NJ: Prentice Hall.
- Brown, S., and J. Tucker. 2011. Large-sample evidence on firms’ year-over-year MD&A modifications. *Journal of Accounting Research* 49 (2): 309–346. doi:10.1111/j.1475-679X.2010.00396.x
- Brown, S., X. Tian, and J. Tucker. 2015. *The Spillover Effect of SEC Comment Letters on Qualitative Corporate Disclosure: Evidence from the Risk Factor Disclosure*. Working paper, Arizona State University.
- Cazier, R., and R. Pfeiffer. 2016. Why are 10-K filings so long? *Accounting Horizons* 30 (1): 1–21. doi:10.2308/acch-51240
- Center for Audit Quality. 2012. *Financial Statement Disclosure Effectiveness: Forum Observations Summary*. Available at: http://www.thecaq.org/sites/default/files/caq_fasb_fsde.pdf
- Dardis, F. E. 2009. Attenuating the negative effects of perceived incongruence in sponsorship: How message repetition can enhance evaluations of an incongruent sponsor. *Journal of Promotion Management* 15: 36–56.
- Darley, C., and A. Glass. 1975. Effects of rehearsal and serial list position on recall. *Journal of Experimental Psychology. Human Learning and Memory* 1 (4): 453–458. doi:10.1037/0278-7393.1.4.453
- Davis, F. D., G. L. Lohse, and J. E. Kottemann. 1994. Harmful effects of seemingly helpful information on forecasts of stock earnings. *Journal of Economic Psychology* 15 (2): 253–267. doi:10.1016/0167-4870(94)90003-5
- Dyer, T., M. Lang, and L. Stice-Lawrence. 2016. *The Ever-Expanding 10-K: Why Are 10-Ks Getting So Much Longer (and Does It Matter)?* Working paper, The University of North Carolina.
- EY. 2015. *SEC Comments and Trends: An Analysis of Current Reporting Issues*. Available at: [http://www.ey.com/Publication/vwLUAssets/SECCommentsTrends_CC0421_24September2015/\\$FILE/SECCommentsTrends_CC0421_24September2015.pdf](http://www.ey.com/Publication/vwLUAssets/SECCommentsTrends_CC0421_24September2015/$FILE/SECCommentsTrends_CC0421_24September2015.pdf)
- Financial Accounting Standards Board (FASB). 2001. *GAAP-SEC Disclosure Requirements. Steering Committee Report Series*. Norwalk, CT: FASB.
- Financial Accounting Standards Board (FASB). 2012. *Invitation to Comment on Disclosure Framework*. Norwalk, CT: FASB.
- Financial Accounting Standards Board (FASB). 2013. *Q&A: FASB’s Disclosure Framework Project*. Norwalk, CT: FASB.
- Financial Accounting Standards Board (FASB). 2014. *Exposure Draft: Chapter 8: Notes to Financial Statements*. Norwalk, CT: FASB.
- Gallagher, D. 2013. *Remarks at the 2nd Annual Institute for Corporate Counsel*. Available at: <https://www.sec.gov/News/Speech/Detail/Speech/1370540462287#U-i4qKPD-Uk>
- Gibbon, D., R. Moore, and R. Winski. 1997. *Handbook of Standards and Resources for Spoken Language Systems*. Berlin, Germany: Mouton de Gruyter.
- Griffin, P. 2003. Got information? Investor response to Form 10-K and Form 10-Q EDGAR filings. *Review of Accounting Studies* 8 (4): 433–460. doi:10.1023/A:1027351630866
- Hargreaves, D. J. 1984. The effects of repetition on linking for music. *Journal of Music Education* 32 (1): 35–47.
- Higgins, K. 2014. *Keynote Address at the 2014 Angel Capital Association Summit*. Available at: <https://www.sec.gov/news/speech/2014-spch032814kfh#U3zloDJOWig>
- Hong, H., and J. Stein. 1999. A unified theory of underreaction, momentum trading and overreaction in asset markets. *The Journal of Finance* 54 (6): 2143–2184. doi:10.1111/0022-1082.00184
- Huang, A., K. Hui, and R. Li. 2016. *Federal Judge Ideology: A New Measure of ex-ante Litigation Risk*. Working paper, Hong Kong University of Science and Technology.
- Huddart, S., B. Ke, and C. Shi. 2007. Jeopardy, non-public information, and insider trading around SEC 10-K and 10-Q filings. *Journal of Accounting & Economics* 43 (1): 3–36. doi:10.1016/j.jacceco.2006.06.003
- Iliev, P., D. Miller, and L. Roth. 2014. Uninvited U.S. investors? Economic consequences of involuntary cross-listings. *Journal of Accounting Research* 52 (2): 473–519. doi:10.1111/0022-1082.00184
- Investors Technical Advisory Committee. 2007. *Disclosure Framework Proposal. Letter to FASB Chairman Robert Herz, December 11*. Norwalk, CT: FASB.
- Joe, J. R. 2003. Why press coverage of a client influences the audit opinion. *Journal of Accounting Research* 41 (1): 109–133. doi:10.1111/1475-679X.00098
- Kim, I., and D. Skinner. 2012. Measuring securities litigation risk. *Journal of Accounting & Economics* 53 (1/2): 290–310. doi:10.1016/j.jacceco.2011.09.005
- Kothari, S., S. Shu, and P. Wysocki. 2009. Do managers withhold bad news? *Journal of Accounting Research* 47 (1): 241–276. doi:10.1111/j.1475-679X.2008.00318.x
- KPMG. 2011. *Disclosure Overload and Complexity: Hidden in Plain Sight*. Available at: <http://www.kpmg.com/US/en/IssuesAndInsights/ArticlesPublications/Documents/disclosure-overload-complexity.pdf>
- Lang, M. H., and L. Stice-Lawrence. 2015. Textual analysis and international financial reporting: Large sample evidence. *Journal of Accounting & Economics* 60 (2/3): 110–135. doi:10.1016/j.jacceco.2015.09.002

- Lawrence, A. 2013. Individual investors and financial disclosure. *Journal of Accounting & Economics* 56 (1): 130–147. doi:10.1016/j.jacceco.2013.05.001
- Lehavy, R., F. Li, and K. Merkley. 2011. The effect of annual report readability on analyst following and the properties of their earnings forecasts. *The Accounting Review* 86 (3): 1087–1115. doi:10.2308/accr.00000043
- Li, F. 2008. Annual report readability, current earnings, and earnings persistence. *Journal of Accounting & Economics* 45 (2/3): 221–247. doi:10.1016/j.jacceco.2008.02.003
- Li, F. 2010. The information content of forward-looking statements in corporate filings—A naïve Bayesian machine learning approach. *Journal of Accounting Research* 48 (5): 1049–1102. doi:10.1111/j.1475-679X.2010.00382.x
- Li, H. 2014. *Repetitive Disclosures in the MD&A*. Working paper, University of Toronto.
- Loughran, T., and B. McDonald. 2010. *Measuring Readability in Financial Text*. Working paper, University of Notre Dame.
- Loughran, T., and B. McDonald. 2011. When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance* 66 (1): 35–65.
- Loughran, T., and B. McDonald. 2014. Measuring readability in financial disclosures. *The Journal of Finance* 69 (4): 1643–1671. doi:10.1111/jofi.12162
- Luotoniemi, E., E. Service, S. Maury. 2007. Good and bad effects of phonological similarity on word and nonword recall: The role of beginnings and ends. *The European Journal of Cognitive Psychology* 19 (1): 80–102. doi:10.1080/09541440600758747
- Lyon, C., J. Malcolm, and B. Dickerson. 2001. *Detecting Short Passages of Similar Text in Large Document Collections*. Working paper, University of Hertfordshire.
- McMullin, J. 2014. *Can I Borrow Your Footnotes? Learning and Network Benefits of Footnotes Similarity*. Working paper, Indiana University.
- Miller, G. 2002. Earnings performance and discretionary disclosure. *Journal of Accounting Research* 40 (1): 173–204. doi:10.1111/1475-679X.00043
- Miller, B. 2010. The effects of reporting complexity on small and large investor trading. *The Accounting Review* 85 (6): 2107–2143. doi:10.2308/accr.00000001
- Nelson, K., and A. Pritchard. 2014. *Carrot or Stick? The Shift from Voluntary to Mandatory Disclosure of Risk Factors*. Available at: <http://ssrn.com/abstract=2447066>
- Peterson, K., R. Schmardebeck, and J. Wilks. 2015. The earnings quality and information processing effects of accounting consistency. *The Accounting Review* 90 (6): 2483–2514.
- PricewaterhouseCoopers (PwC). 2014. *Public Company Disclosure Model: Is a New Approach Needed?* Available at: <https://www.pwc.com/us/en/cfodirect/assets/pdf/point-of-view-public-company-disclosure-model.pdf>
- Rennekamp, K. 2012. Processing fluency and investors' reactions to disclosure readability. *Journal of Accounting Research* 50 (5): 1319–1354. doi:10.1111/j.1475-679X.2012.00460.x
- Rock, I. 1957. The role of repetition in associative learning. *The American Journal of Psychology* 70 (2): 186–193. doi:10.2307/1419320
- Schrand, C. M., and B. R. Walther. 2000. Strategic benchmarks in earnings announcements: The selective disclosure of prior-period earnings components. *The Accounting Review* 75 (2): 151–177. doi:10.2308/accr.2000.75.2.151
- Securities and Exchange Commission (SEC). 2003. *Interpretation: Commission Guidance Regarding Management's Discussion and Analysis of Financial Condition and Results of Operations, 17 CFR Parts 211, 231 and 241 [Release Nos. 33-8350; 34-48960; FR-72]*. Available at: <https://www.sec.gov/rules/interp/33-8350.htm>
- Securities and Exchange Commission (SEC). 2014. *Standard Instructions for Filing Forms under Securities Act of 1933, Securities Exchange Act of 1934 and Energy Policy and Conservation Act of 1975—Regulation S-K. Code of Federal Regulations. Title 17, Pt 229, 2014 edition*. Available at: <https://www.ecfr.gov/cgi-bin/text-idx?node=17:3.0.1.1.11&rgn=div5action?collectionCode=CFR&browsePath=Title+17%2FChapter+II%2FPart+229&granuleId=CFR-2005-title17-vol2-part229&packageId=CFR-2005-title17-vol2&collapse=true&fromBrowse=true>
- Standard & Poor's. 2012. *Comment Letter to the FASB in Response to FASB's Invitation to Comment—Disclosure Framework, November 30, Comment Letter No. 46*. Available at: <http://www.fasb.org>
- White, M. 2013. *The Path Forward on Disclosure. Speech Presented at the National Association of Corporate Directors, National Harbor, MD*. Available at: <https://www.sec.gov/news/speech/spch101513mjw#.Uzx2VPIdV8E>
- You, H., and X. Zhang. 2009. Financial reporting complexity and investor underreaction to 10-K Information. *Review of Accounting Studies* 14 (4): 559–586. doi:10.1007/s11142-008-9083-2

APPENDIX A

Sample Redundancy in 10-K Filings

The first two excerpts illustrate the contrast in 10-K disclosure repetition between two sample companies, LSB Industries and Ford Motor Company, for fiscal year 2012. These first two excerpts illustrate variation in repetitious disclosure relating to each company's accounting policies regarding property, plant, and equipment. Underlined text indicates sequences of ten or more words repeated verbatim in prior sections of the 10-K. The last excerpt comes from American States Water Company's

MD&A portion of their 10-K for fiscal year 2009. Potentially relevant information that is not mentioned elsewhere in the 10-K (emphasized in bold italic) is sandwiched between two passages that are repeated verbatim in the financial statement notes.

LSB Industries

Property, plant, and equipment (PP&E) are stated at cost, net of accumulated depreciation, depletion, and amortization (DD&A). Leases meeting capital lease criteria are capitalized in PP&E. Major renewals and improvements that increase the life, value, or productive capacity of assets are capitalized in PP&E while maintenance, repairs, and minor renewals are expensed as incurred. In addition, maintenance, repairs, and minor renewal costs relating to planned major maintenance activities (Turnarounds) in our Chemical Business are expensed as they are incurred. As it relates to natural gas properties, leasehold costs, intangible drilling, and other costs of successful wells and development dry holes are capitalized in PP&E based on successful efforts accounting. The costs of exploratory wells are initially capitalized in PP&E, but expensed if and when the well is determined to be nonproductive. Interest cost on borrowings incurred during a significant construction or development project is capitalized in PP&E. Capitalized interest is added to the underlying asset and amortized over the estimated useful lives of the assets. Fully depreciated assets are retained in PP&E and accumulated DD&A accounts until disposal. When PP&E are retired, sold, or otherwise disposed, the assets carrying amount and related accumulated DD&A are removed from the accounts and any gain or loss is included in other income or expense.

For financial reporting purposes, depreciation of the costs of PP&E is primarily computed using the straight-line method over the estimated useful lives of the assets. DD&A of the costs of producing natural gas properties are computed using the units of production method primarily on a field-by-field basis using proved or proved developed reserves, as applicable, as estimated by our independent consulting petroleum engineer. No provision for depreciation is made on construction in progress or capital spare parts until such time as the relevant assets are put into service. No provision for DD&A is made on nonproducing leasehold costs and exploratory wells in progress until such time as the relevant assets relate to proven resources.

Our natural gas reserves are based on estimates and assumptions, which affect our DD&A calculations. Our independent consulting petroleum engineer, with our assistance, prepares estimates of natural gas reserves based on available relevant data and information. For DD&A purposes, and as required by the guidelines and definitions established by the Securities and Exchange Commission (SEC), the reserve estimates are based on average natural gas prices during the 12-month period, determined as an unweighted arithmetic average of the first-day-of-the-month price for each month.

Ford Motor Company

Net property includes land, buildings and land improvements, machinery and equipment, special tools, and other assets that we use in our normal operations. These assets are recorded at cost, net of accumulated depreciation and impairments. We capitalize new assets when we expect to use the asset for more than one year. Routine maintenance and repair costs are expensed when incurred.

Property and equipment are depreciated primarily using the straight-line method over the estimated useful life of the asset. Useful lives range from 3 years to 36 years. The estimated useful lives generally are 14.5 years for machinery and equipment, 3 years for software (8 years for mainframe and client based software), 30 years for land improvements, and 36 years for buildings. Special tools generally are amortized over the expected life of a product program using a straight-line method. If the expected production volumes for major product programs associated with the tools decline significantly, we accelerate the amortization reflecting the rate of decline.

American States Water Company

In addition, changes in the fair value of plan assets will impact future pension cost and the Plan's funded status. Volatile market conditions have affected the value of AWR's trust established to fund its future long-term pension benefits. The market value of the investments within the plan trust declined by approximately 25 percent during the year ended December 31, 2008. Unless the market recovers, reductions in the value of plan assets will result in increased future expense, an increase in the underfunded position and increased future contributions. As of February 28, 2009, the fair value of the pension plan assets decreased to \$48.6 million due to changes in market conditions. In March 2009, Registrant filed an advice letter with the CPUC requesting authorization to establish a Pension Costs Memorandum Account. If this account is approved, Registrant will track the difference between the pension costs authorized by the CPUC and included in customer rates, and actual pension costs. Registrant will not record the amounts in this account as a regulatory asset until they are reviewed and approved by the CPUC. If approved by the CPUC, Registrant will then establish a regulatory asset with a corresponding increase to earnings. Until then, Registrant expects that its earnings will be negatively affected by increasing pension costs.

APPENDIX B

Accounting Policies Word List

We constructed the following list of words based on our review of a large sample of accounting policy disclosures from firms' financial statement notes. This list is used in our study to provide descriptive evidence on the extent to which accounting policies comprise repetitious content in 10-K filings.

ACCOUNTING	DEPRECIATED	METHOD
ADOPT	DERECOGNITION	PCAOB
ADOPTED	DETERMINE	PRESENTATION
ADOPTING	DETERMINED	PRINCIPLE
AICPA	DETERMINING	PRINCIPLES
ALLOCATE	DISCLOSE	PRONOUNCEMENT
ALLOCATED	DISCLOSED	PRONOUNCEMENTS
ALLOCATING	DISCLOSING	REALIZE
ALLOWANCE	EFFECTIVE	REALIZED
AMORTIZE	EITF	RECOGNIZE
AMORTIZED	ESTIMATE	RECOGNIZED
AMORTIZING	ESTIMATED	RECOVERABILITY
ARB	EVALUATE	RESERVE
ASC	EVALUATED	RESERVES
ASSESS	FAS	RESERVED
ASSUMPTIONS	FASB	RULES
ASU	FIFO	SAB
AUTHORITATIVE	FIN	SFAC
BALANCE	FSP	SFAS
BESP	GAAP	SOP
CARRYING	GOODWILL	STAFF
CLASSIFY	GUIDANCE	STANDARD
CLASSIFICATION	IDENTIFIABLE	STRAIGHTLINE
CLASSIFIED	IMPAIRMENT	SUBTOPIC
CODIFY	IMPLIED	TEST
CODIFICATION	IMPUTE	TESTED
CODIFIED	IMPUTED	TESTING
COLLECTABILITY	INDEFINITE	TOPIC
COMPLEXITY	INTERPRETATION	TREATMENT
COMPREHENSIVE	JUDGMENT	UNREALIZED
CONSOLIDATE	LIABILITY	VALUATION
CONSOLIDATED	LIABILITIES	VIE
DEFER	LIFO	VSOE
DEFERRED	MATERIAL	WEIGHTED
DELIVERABLE	MEASURE	WRITE
DEPRECIATE	MEASURED	

Copyright of Journal of Financial Reporting is the property of American Accounting Association and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.