Journal of Accounting Research

DOI: 10.1111/1475-679X.12203
Journal of Accounting Research
Vol. 56 No. 2 May 2018
Printed in U.S.A.

CHICAGO BOOTH

Disclosure "Scriptability"

KRISTIAN D. ALLEE,* MATTHEW D. DEANGELIS,† AND JAMES R. MOON, JR.†

Received 31 October 2015; accepted 20 December 2017

ABSTRACT

In response to the increasing use of computer programs to process firm disclosures, this registered report develops a new measure of "scriptability" that reflects computerized, rather than human, information processing costs. We validate our measure using SEC filing-derived data from prior research and identify firm and disclosure characteristics related to it. In our planned hypothesis tests, we find some evidence that the speed of the market response to filings increases with scriptability, but find little evidence that scriptability affects the incidence and speed of news dissemination by Dow Jones. In additional analyses, we find that scriptability exhibits both positive and negative associations with changes in information asymmetry between market participants, depending on the filing, trading window, and measure examined. We also find little evidence that XBRL interacts with scriptability in a meaningful way. Overall, our study broadens our understanding of information processing costs and provides opportunities for new avenues of research.

*University of Arkansas; †Georgia State University.

Accepted by Christian Leuz. This paper is the final Registered Report resulting from the Registration-based Editorial Process (REP) implemented by *JAR* for its 2017 conference; details of the process are available here: https://research.chicagobooth.edu/arc/journal-of-accounting-research/2017-registered-reports. The accepted proposal and an online appendix for this report are available here: https://research.chicagobooth.edu/arc/journal-of-accounting-research/online-supplements. We thank Patrick Caldon (Morningstar), Fran Finnegan (SECInfo.com), J.W. Mike Starr (Workiva), Alex Zervakos (Securex) for excellent industry perspectives on preparation and processing of SEC filings. We thank Semir Sarajlic, Suranga Edirisinghe, and Kwai Wong for technical assistance. We also gratefully acknowledge John Campbell, Scott Dyreng, Kurt Gee, Craig Holden, Stacey Jacobsen, Todd Kravet, Bill McDonald, Jeff McMullin, Volkan Muslu, Elizabeth Odders-White, Kyle Peterson, Mark Ready, Bei Shi, Logan Steele, Brady Twedt, Ben Whipple, Ke Xu, and our anonymous reviewer for help with this project, and conference participants at the 2017 *Journal of Accounting Research* Conference for valuable insights and suggestions.

JEL codes: G10; G12; G14; G24; G30; G39; G41; M40; M41; M49

Keywords: disclosure; textual analysis; financial reporting; information processing costs; high-frequency trading; XBRL

1. Introduction

Over the last several decades, the Internet has significantly increased both the volume of information and the speed of information dissemination. In response to these changes and increasing competition for returns, market traders now utilize computing power to assist or replace human effort in the acquisition and analysis of financial information and the execution of trading strategies (Foxman [2016]). For instance, journalists estimate that 50% to 60% of all trades in the U.S. stock market are now made by algorithms (Chilton [2014], Philips [2015]), often occurring within seconds of information release (Lewis [2015]). As a result, the analysis of financial information in SEC filings has shifted away from manual processing by investors and analysts toward programmers and computers in ways that research has only begun to acknowledge.

In this paper, we examine corporate disclosure from a programmer's, or "scripter's," perspective.[1] We start by identifying two basic tasks that a scripter, whether an investor doing her own analysis or a computer scientist working at an investment bank, is likely to perform when analyzing firm disclosures. First, the scripter searches for data of interest, such as the Management's Discussion and Analysis (MD&A) section of the 10-K or a table of executive compensation figures. Second, the scripter processes identified data into decision-relevant information. We then formulate two measures based on these tasks and one composite measure of "scriptability," or the ease with which a scripter can perform these tasks on a given disclosure. In creating and analyzing our measure, we provide comprehensive descriptive statistics and empirical evidence regarding several previously unexamined dimensions of disclosure quality, such as the amount of disclosure that is machine-readable, the ease of separating text from tables, and the presence of table metadata that facilitates the extraction of key financial metrics.[2] We

---

[1] This research was conducted as a two-stage editorial process based on what is known in other fields as "Registered Reports." In stage one, authors submit a proposal that describes the hypotheses they will test, the data they will gather, and (in considerable detail) the research design and analyses they will use to interpret their results. In stage two, authors submit the "Registered Report," which describes the original intent and actual execution of the study approved in stage one, along with the results and interpretation of planned and unplanned analyses. Consistent with the registered report process, the only changes to our study from the accepted stage one proposal are reported in online appendix F.

[2] The proper use of tables facilitates the identification of numeric data and its separation from textual data. We evaluate each table in a filing and apply a series of conditions to determine whether the table contains properly formatted numerical data or textual data that would be better presented in a bulleted list or paragraph. For those tables that contain

also offer insight into dimensions of disclosure quality that affect both programmers and readers, such as the usefulness of section headings and the extent to which the disclosure references external content.[3] As such, our study highlights both the opportunities and challenges of using computer programs to analyze financial information and expands prior research on information processing costs and frictions in capital markets.

We conduct three validation tests on our scriptability measures and find evidence suggesting that our measures capture the costs of computerized information processing in firm disclosures (see online appendix E). We then examine this previously unexplored dimension of disclosure quality by identifying a set of firm and disclosure characteristics that we expect to relate to scriptability. Prior research suggests that disclosure quality increases with investment in the financial reporting function, so we expect improvements in scriptability when managers make these investments. We also expect that both technologically sophisticated managers and managers relying on external providers ("filing agents") for filing preparation are more likely to have more scriptable disclosures. While we find some evidence from these tests that investment in the financial reporting function is related to higher scriptability, our results do not suggest that technological sophistication in general translates into more scriptable filings. Overall, our most notable finding in this test is that the use of a top filing agent, a previously unexplored information intermediary, is associated with significantly better scriptability. This finding provides additional validation of our measure, as filing agents appear to improve filings along the dimensions we measure.

Controlling for the above factors, we proceed to test two hypotheses on the possible benefits of scriptability. First, market efficiency requires complete and rapid price formation (O'Hara [2003]). Investors utilizing programmatic processing techniques should be able to identify, process, and act upon disclosed information more quickly when disclosures are more scriptable. Using an intraperiod timeliness (IPT) measure from prior research (e.g., Butler, Kraft, and Weiss [2007], Bushman, Smith, and Wittenberg-Moerman [2010], McMullin, Miller, and Twedt [2015]) as well as a newly designed IPT measure based on volume, we hypothesize that disclosure scriptability positively relates to the speed of the market's response to the disclosure. Our strongest evidence is consistent with scriptability increasing the timeliness of trading (volume IPT), particularly in the 8-K filing group. These results suggest that at least some traders are impeded by poor scriptability. With regard to the timeliness of price formation, most of our evidence fails to detect a significant association between scriptability

numeric data, we further evaluate numeric tables for the presence of HTML markup, row and column consistency, and descriptive row names, all of which aid the processing of numeric data into information.

[3] Figure 1 in section 2 illustrates how our components relate to our two task-based measures and a composite measure.

and price IPT. To better understand this finding, we conduct an additional unplanned analysis that suggests that individual components of our task-based scriptability measures could be relevant for different form-types and trading windows.[4] Specifically, characteristics like external references and exhibits that are not machine-readable seem to matter more with longer documents like the 10-K and the 10-Q, while the ability to separate sections and the tabulation of numbers seem most relevant to 8-Ks, perhaps because separate sections and tables in 8-Ks often pertain to separate news items.

To further explore potential capital market consequences of scriptability, we conduct an additional, planned analysis to examine whether scriptability relates to changes in information asymmetry surrounding disclosure. Technologically sophisticated investors, particularly those implementing algorithmic strategies without human intervention, could obtain an information advantage immediately following the filing of a highly scriptable disclosure. However, more scriptable disclosures combined with low-cost computing power should lead to larger, more rapid reductions in information asymmetry following SEC filings. Thus, we explore how scriptability relates to changes in the bid-ask spread surrounding filing of disclosures and expect any adverse (beneficial) effects of scriptability to be more pronounced in shorter (longer) windows. Our results indicate that scriptability increases information asymmetry in the five-minute window following the filing, but that the information advantage of "early" scripters diminishes as the window lengthens. This pattern of results is intuitive, as it suggests that early, sophisticated traders impound the information from the filing into price, and as time goes by less sophisticated investors "catch up." In an unanticipated result, we find that the component of our measure related to identifying data relates negatively to information asymmetry, but only in longer measurement windows (60 minute and 24 hours). This evidence suggests that "late" scripters likely rely more heavily on computer programs to identify data than to process that data, perhaps suggesting that they manually process (i.e., read) the information that earlier scripters have previously processed programmatically.

For our second hypothesis tests, we build on research suggesting that the business press plays an important role in the dissemination and processing of news (Drake, Guest, and Twedt [2014], Twedt [2016]). We hypothesize that information intermediaries, such as the Dow Jones (DJ) newswire service, are more likely to locate and process information in more scriptable disclosures, leading to a higher likelihood and speed of newswire dissemination. Perhaps because this test depends on the type and degree of automated processing used by the intermediary, we find little evidence in support of these predictions. However, in an unplanned analysis, we

[4] The registered report process requires us to clearly distinguish between planned analyses that were included in the original report and unplanned analyses that were added after observing the data. See footnote 1.

demonstrate that roughly half of the components of scriptability show the predicted association with the likelihood and speed of dissemination, suggesting that portions of our measure may capture programmatic impediments to the specific processes employed by DJ.

Given the SEC's recent focus on XBRL, we examine the effects of XBRL on scriptability in an additional planned analysis. We exclude XBRL from our main measure of scriptability since XBRL information exists for only a small percentage of filings during our sample period and affects only a subset of the information in those filings (i.e., financial data). However, XBRL makes key financial data more machine-readable (Cox [2005]), so we consider whether the presence and quality of XBRL metadata serves as an alternative mechanism to improve scriptability. We find little evidence that XBRL interacts with scriptability in a meaningful way. We note, however, that our measures of scriptability are designed to be holistic to the disclosure, while XBRL metadata is limited to certain aspects such as the tagging of specific numbers for easier retrieval. As such, XBRL might only substitute for a limited set of our scriptability components. It is also possible that XBRL metadata is little utilized by investors because of frequent errors in XBRL markup (Debreceny et al. [2010]), leading to no variation based on the presence or quality of XBRL.

Our study provides several important contributions to the accounting and finance literature. First, our study contributes to the disclosure literature by providing new, detailed insights into both the data that underlie and the metadata that accompanies firm disclosures.[5] While prior capital markets research has taken full advantage of both the proliferation of data (Li [2008], Drake, Roulstone, and Thornock [2012], Blankespoor, Miller, and White [2014], Chen et al. [2014], Drake, Guest, and Twedt [2014], Jame et al. [2016], Hales, Moon, and Swenson [2018]) and expanded methods for analyzing it (Allee and DeAngelis [2015], Chen, Miao, and Shevlin [2015]), this research generally relegates discussions of the properties of disclosure that facilitate or hinder electronic analysis to footnotes or appendices (Li [2008], McDonald [2016]). We explicitly measure these properties and provide initial evidence on factors associated with disclosure scriptability. As the first study to take a holistic view of the structure, function, and quality of text and HTML in regulatory filings, our findings inform a broad range of computerized analyses and likely extend to other information sources, such as news outlets, social media, and investor relations Web sites. Additionally, our unplanned, descriptive evidence suggests that different scriptability components appear relevant to different contexts. We provide our data in order to facilitate future research, which can build on

[5] Metadata is "data about data" that can be used to describe the contents and context of data or data files and can increase the usefulness of the underlying data. For example, a Web page could include metadata specifying in what language the page is written, what tools were used to create it, and where to find more information about the subject.

our evidence to customize measures of scriptability for specific settings and purposes.

Second, prior literature suggests that lower information processing costs facilitate efficient capital allocation (e.g., Bloomfield [2002], Lundholm, Rogo, and Zhang [2014]) and reduce information asymmetry between market participants (Miller [2010], Flannery [2015]). Research also finds that firms recognize these costs and either respond with clearer disclosures (Waymire [1985], Tasker [1998], Guay, Samuels, and Taylor [2016]) or exploit them with selective and strategic disclosure (Schrand and Walther [2000], Rogers and Stocken [2005], Cheng and Lo [2006]). Prior research has primarily used disclosure readability as a proxy for information processing costs, as it reflects the level of effort required for a person to read and understand a disclosure. For example, Bloomfield [2008] presents a disclosure in which a company seemingly hides bad news by burying it under a large amount of unreadable, boilerplate text in its disclosure, thus forcing an investor to separate the useful data from the useless. Our analysis broadens the scope of this literature by recognizing that a modern investor can use a wide range of technological solutions to alleviate these costs. Simple word searches and XBRL metadata can identify information of interest no matter where it is located in a report (Hodge, Kennedy, and Maines [2004], Arnold et al. [2012]), and sophisticated algorithms can efficiently summarize large numbers of documents (Carbonell and Goldstein [1998], Barzilay, McKeown, and Elhadad [1999], Haghighi and Vanderwende [2009]). Use of these tools should reduce the effect of document length (Li [2008], You and Zhang [2009]) and the location of information (Maines and McDaniel [2000], Lee, Petroni, and Shen [2006], Bloomfield et al. [2010]) on investor processing. Future research can examine scriptability in a wide range of contexts and further distinguish the roles of manual versus computerized processing of financial information.

Third, given the SEC's interest in machine learning to identify fraud (Carney and Harker [2015]) and the difficulties that researchers often face when analyzing these disclosures (see Loughran and McDonald [2016] for a discussion of document parsing as a key "tripwire" in qualitative research),[6] our study informs regulators' ongoing efforts to improve disclosure access and quality. Starting in 1993 the SEC began the "Edgarization" process, which moved regulatory filings onto the digital EDGAR platform (Cuff [1993]). More recently, the FASB and the SEC have committed to improve the quality of the financial statements and related notes through the Financial Statement Presentation and Disclosure Framework projects (Financial Accounting Standards Board [2016]) and the implementation

[6] For instance, in our experience approximately 25% of MD&As in 10-Ks cannot be easily extracted, leading to their exclusion from some prior studies due to missing data. Our analysis includes these firms and helps to identify characteristics of fillings with "missing" MD&As.

of XBRL (Securities and Exchange Commission [2009a]) for selected financial information. Our study takes a broader view of the data and metadata throughout the electronic version of the filing, providing comprehensive details on filing characteristics that can help or hinder computerized processing. Our evidence that scriptability facilitates more efficient trading (volume IPT) and corresponds to reduced information asymmetry in longer windows also suggests that establishing best practices for data presentation and quality across many dimensions of SEC filings can improve their usefulness for less sophisticated users.

2. *Measure of Scriptability*

2.1 MEASURING COMPONENTS OF SCRIPTABILITY

We expect that the purposes of individual scripts vary as much as investor valuation models, so we do not attempt to measure the scriptability of firm disclosures for every possible use. Instead, we formulate a general approach based on features of disclosures that likely affect scripters' ability to programmatically extract needed information. We identify two basic tasks to which scripters apply computerized processing: (1) identifying data of interest and (2) processing that data into useful information. Although we expect most scripters to perform both of these tasks, scripters may differ in their reliance on identification versus processing. For instance, an investor wishing to review income tax footnotes for 30 or so firms could rely exclusively on the first task, programmatically identifying and extracting the footnote he needs from each filing using heading formatting and textual pattern matching, and then processing it manually (i.e., by reading it). On the other hand, a highly sophisticated investor could use machine learning to process all available data and use a statistical method such as ridge regression to build a complex model of firm value from the available inputs, thus relying heavily on the second task. We explore these differences in our empirical analyses by examining measures of scriptability as they relate to each of these two tasks, in addition to a combined measure.

We define our main measure of scriptability, *CompScript*, as the average of two main aspects of scriptability that correspond to each of the above two tasks: *IdentifyData* for scripters' ease in identifying data of interest and *DataToInformation* for scripters' ease in processing data into information. For *IdentifyData*, we measure four filing characteristics: the ease with which a script can (1) separate tables from text, (2) decompose text into logical sections, (3) identify the content of logical sections based on the quality of headings, and (4) find the relevant content in the filing itself rather than following links to external documents. For *DataToInformation*, we examine four filing characteristics that pertain to data analysis in SEC filings: (1) the proportion of the filing that is machine-readable as text, (2) the proportion of numeric information in the filing that is tabulated, (3) the ease of processing textual information, and (4) the ease of processing tabular information.[7] We expect that a scripter could perform many of these tasks manually, and that some characteristics that make a filing more scriptable (such as the clear use of logical sections) may also improve readability for manual processing, while other characteristics (such as the use the HTML tabular format to present bulleted lists) may represent a tradeoff between readability and scriptability.[8,9] Although we design our measure components to conform directly to our observations of SEC filings, we have attempted to make the components sufficiently general to apply to other firm disclosures and to external sources of information such as analyst reports or press articles. We provide complete technical definitions of *CompScript* and its components in online appendices A and B, and figure 1 illustrates how we combine the various components for our empirical analyses. Additionally, online appendix C includes several detailed examples to illustrate how we construct various scriptability components. Finally, online appendix D provides descriptive statistics for all scriptability components, including correlations among the various components.[10]

To facilitate the combination of variables with different measurement bases, we percentile-rank all components (including subcomponents within each component) such that increasing values correspond to more scriptable disclosures. We rank by form type and calendar year since we expect the mean level of scriptability to vary over time and by form type, and we do not want these factors to dominate variation in our measure. We omit

[7] Conforming to the most common methods of, say, table construction might result in higher scriptability than using correct table construction according to HTML guidelines since scripts can be designed to deal with similar errors in similar ways. However, given the possible variation in the many characteristics we consider, we only consider similarity across filings for characteristics where a comparison is readily available (such as cosine similarity between section headings or table row names).

[8] To expand on this example, a clearly descriptive, bolded section heading is both clearly readable and scriptable. For instance, the heading “2017 Growth in European Sales” tells a reader what to expect in the subsequent text whereas a scripter seeking information on European operations would likely identify the bolded text as a heading candidate and the description clearly articulates the sought-after information. However, using HTML table-tagging to present *qualitative* information in a bulleted list should facilitate *manual* processing (since bulleted lists are frequently used to succinctly provide a number of distinct facts) but likely impedes scriptability since tabular information is often expected to be numeric.

[9] We note that the Spearman correlation between *CompScript* and the Fog index (the most commonly used measure of readability) is only 1.4%, providing further evidence that these are different constructs, with scriptability focused on computerized processing and readability focused on manual processing of qualitative disclosures.

[10] To supplement these correlations, we performed an exploratory factor analysis and observed that the number of significant factors that are produced from a principal components analysis (PCA) varies based on the form type. Based on all 12 components, it appears that two factors explain the majority of the variance for Form 10-Q and 8-K filings, while Form DEF 14A filings would require up to four factors and Form 10-K filings appear to require three factors. Additionally, the results generally support our hypothesized grouping of the components into our two task-based measures.

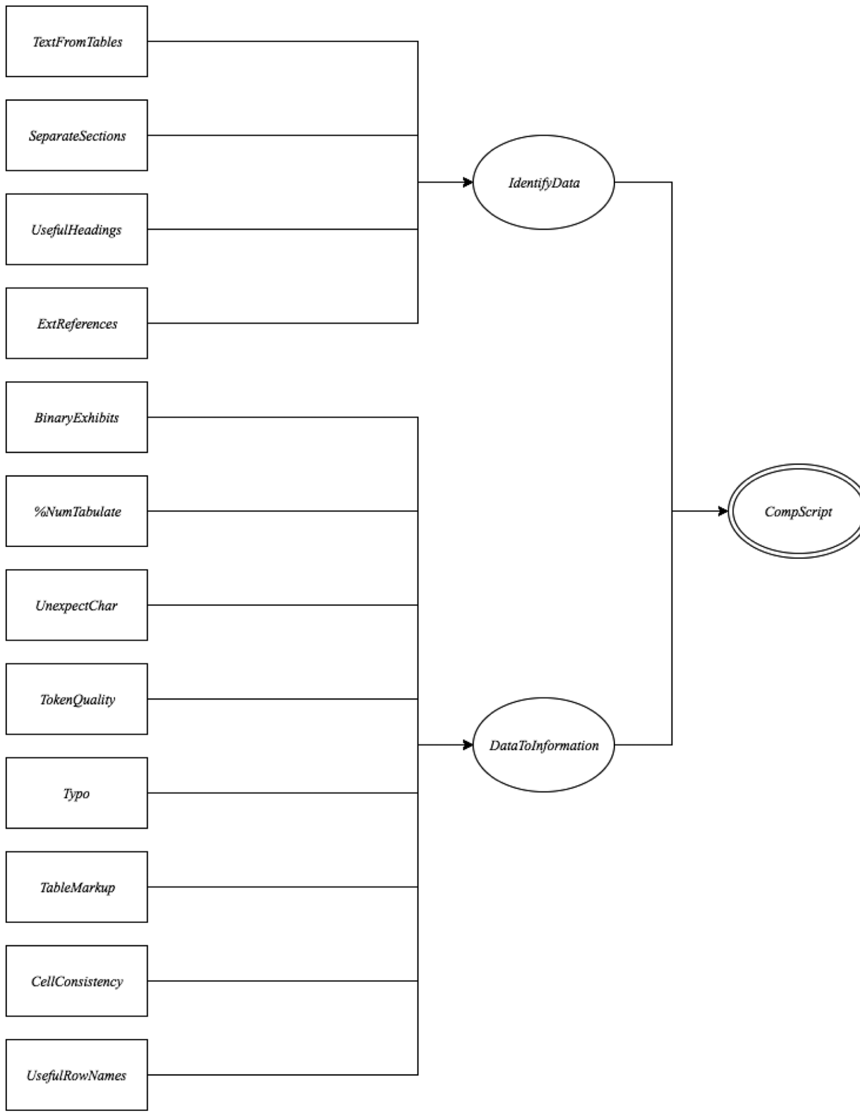

FIG. 1.—Overview of Scriptability. This figure illustrates how components of scriptability are combined into the two task-based measures, *IdentifyData* and *DataToInformation*, and ultimately *CompScript*. See online appendix A for detailed component descriptions.

components from the aggregate measure that are not possible to calculate for a filing.

Like Li [2008], we focus on measuring the scriptability of SEC filings because they are consequential to capital markets (e.g., Loughran and McDonald [2011]), publicly available, frequently accessed by investors (Drake, Roulstone, and Thornock [2015]), and generally follow a homogeneous

TABLE 1
Filing Counts for Various Samples

Form Type	Total Filings (1995–2015)	Filings with Values for *CompScript*	Filings Used in Table 3 Analysis	Max Filings Used to Test H1 (Tables 4 and 5)	Max Filings Used to Test H2 (Table 6)
10-K	263,319	104,614	71,419	12,850	15,263
10-Q	717,521	312,589	224,746	39,768	49,268
8-K	1,337,662	777,416	576,707	171,349	212,420
DEF 14A	147,960	100,073	71,479	13,348	12,929
Total	2,466,462	1,294,692	944,351	237,315	289,900

structure, especially within form type, making them prime candidates for computerized processing. Recent evidence in Rogers, Skinner, and Zechman [2017] suggests investors react to Form 4s within seconds (or even milliseconds) of filing, which must involve programmatic processing of disclosures. We note, however, that more timely disclosure outlets often supersede SEC filings. Since large institutional investors employ programmers to collect data from a range of sources (Foxman [2016]), this lack of timeliness could affect inferences from our tests.[11] As a result, although we focus on SEC filings, we design our measure so that most of our components can be adapted to other settings.

We compute scriptability on filings for four of the most accessed form types according to Drake, Roulstone, and Thornock [2015] from 1995 to 2015. These forms include (in descending order of most accessed): 10-Ks, 8-Ks, 10-Qs, and DEF 14As (definitive proxy statements). We exclude all amended filings (ending with "/A") because the scriptability of these documents likely mirrors the original filing and Form 4s, which rank fourth in terms of mean downloads per day but contain minimal variation in presentation format, making scriptability essentially constant. We classify all forms with "10-K" or "10K" ("10-Q" or "10Q") in the form type as 10-Ks (10-Qs). Table 1 presents the number of filings we identify in the EDGAR indices for each of the four form types we consider. We identify approximately 263 thousand 10-Ks, 718 thousand 10-Qs, 1.3 million 8-Ks, and 148 thousand DEF 14As. These values constitute our beginning sample for measuring scriptability. Sample sizes for selected analyses are significantly smaller due to other data requirements and restrictions as described below.[12]

[11] We discussed our study with a Senior Analyst at an investment research firm who specializes in programmatically analyzing firm disclosures. He explained that his firm relies on computer processing for several types of analysis from various sources, including table normalization, topic modeling, and proprietary sentiment measures.

[12] Upon observing the data, we noted a large number of duplicate filings pertaining to subsidiaries. In order to remove these duplicates from our sample, we require a match in the CRSP-Compustat linking file for all SEC filings in all of our samples. The attrition from column 1 to column 2 in table 1 is primarily due to this restriction. See online appendix F for details.

Before measuring each component of scriptability on a filing, we remove all SEC header information (i.e., all information preceding the first "<DOCUMENT>" tag and front matter such as the company name, state of jurisdiction, etc.), footer information such as signatures, and, where applicable, XBRL documents.[13] We conduct validation tests on both the composite measure and the task-based measures of disclosure scriptability. For brevity, we only briefly describe the results of the validation tests here, but complete results are reported in online appendix E. We first assess whether scriptability relates positively to the likelihood that a given filing is included in samples constructed by *other* researchers (Kravet and Muslu [2013], Campbell et al. [2014], Peterson, Schmardebeck, and Wilks [2015], and McMullin [2016]) and generally find this is the case. Our second two tests rely on readability and sentiment measures prepared by other researchers (Li [2008], Loughran and McDonald [2014]). We find some evidence that *DataToInformation* correlates negatively with noise in 10-K Fog and, in an unplanned test, that *CompScript* relates negatively to error in measuring 10-K tone. Overall, this evidence suggests that our measures appear to capture the costs of computerized information processing costs in firm disclosures.

2.2 FACTORS ASSOCIATED WITH SCRIPTABILITY

We examine three categories of firm and disclosure characteristics that we expect to relate to scriptability: investment in the financial reporting function, technological sophistication, and the use of a filing provider (or "agent").

Our discussions with industry representatives suggest that most managers focus on the presentation of the filing to a human reader and largely ignore the underlying text or HTML of the electronic filing. As such, firms likely achieve better scriptability indirectly through greater overall investment in the financial reporting function. In other words, we expect that firms placing a higher value on financial reporting quality invest in all aspects of disclosure quality, including technical support for the filing process. Doyle, Ge, and McVay [2007] and Ashbaugh-Skaife, Collins, and Kinney [2007] find evidence that firms invest in financial reporting quality when they have more available resources, when quality is less costly to achieve, and when the firm faces external pressures to improve quality. We therefore include firm size (*Size*) and age (*Age*), which we expect relate positively to scriptability, and financial distress (*Distress*), complexity

[13] For the purposes of measuring scriptability, we assume that the scripter uses the full text version of the filing (available as a .txt file either through EDGAR's FTP site or the Web index for each filing). The full text version clearly delineates documents, making it easy for the scripter to split the document into its components, if necessary, and saves the effort of identifying, downloading, storing, and analyzing multiple files. Using the full-text filing also allows our scriptability measures to reflect the full disclosure (i.e., the SEC Form and attached exhibits).

(*OpComplex, FinComplex*), and sales growth (*Growth*), which we expect relate negatively to scriptability. Similarly, firms facing greater scrutiny from institutional investors invest more heavily in financial reporting, so we also include the level of institutional holdings (*%IH*), which we expect to relate positively to scriptability. We provide detailed definitions of all variables in the appendix.

While we posit that investment in financial reporting has an indirect effect on scriptability, we also expect that managers will choose to directly take action to improve scriptability, and thus reduce information processing costs, when they are able to do so. Management teams with a higher level of technological sophistication more likely understand the benefits of scriptability to meet users' information processing needs. Further, these managers bear lower costs to identify and implement technological best practices. Consistent with these expectations, prior research finds that executive innovativeness and the level of technical knowledge for both managers and employees relates positively to the adoption of more sophisticated information systems (Thong [1999]). Following Kobelsky et al. [2008], who find that high-tech industries have higher IT budgets and likely greater technological sophistication, we proxy for technological sophistication with *HTI*, a dummy variable for high-tech industries, and expect it to relate positively to scriptability. We also include a proxy for technological sophistication using the 10-K filing lag (*FileLag*), as the SEC argues that technological advancement facilitates firms' ability to provide more timely filings (Securities and Exchange Commission [2002], Choudhary, Merkley, and Schloetzer [2015]).[14] We expect *FileLag* to relate positively to scriptability.

Finally, even technologically unsophisticated managers could choose to improve the quality of the filing and outsource some or all of the preparation and filing process to a professional provider. A preliminary inspection of filing metadata for recent disclosures confirms that a large number of public companies use external filing agents to "Edgarize" their filings, with filing agent Workiva currently boasting 65% of the Fortune 500 among their clients despite only being founded in 2010 (Workiva [2015]). We expect that the use of a filing provider will result in higher scriptability. However, some firms rely on the publisher of their physical annual report to also prepare the electronic EDGAR filing, and we do not expect these publishers to pay much attention to the quality of disclosure metadata. Consequently, we use two indicator variables to measure a firm's use of a filing provider for preparation and/or submission of the file. First, we identify whether firms use one of the top 25 filing agents (*Top25FA*)

[14] Over a long time series, this variable likely captures increased regulation over maximum time-to-file allowances. However, our regressions include year-month fixed effects, which limits its variation in scriptability explained by independent variables to cross-sectional differences in filing delays. Further, we control for firm size, which accounts for differences in filing deadlines for small vs. large firms ("accelerated filers").

for the form they are filing in a given year, with the expectation that top agents are more likely to improve scriptability. Second, we set an indicator variable, *Watermark*, equal to 1 if the filing provider leaves an "electronic watermark," as this is evidence of a conscious design choice in the electronic filing.[15] We expect both of these variables to relate positively to scriptability.

Table 2 presents descriptive statistics on the variables in our study. Panel A (panel B) presents statistics measured at the disclosure (firm-year or firm-quarter) level, and panel C reports correlations between each variable and scriptability.[16] To control for outliers, we replace the lower (upper) 1% of all continuous variables with the first (99th) percentile of their respective distributions. Panel A indicates that most forms are filed by a top 25 filing agent and only 9% of DEF 14As have a *Watermark* while 23% of 8-Ks are filed with a *Watermark*. Panel B shows that the median firm in our sample has \$252 million in total assets and is 13 years old. Our sample is comparable to, but somewhat different from, Li's (2008) sample, in which the median firm has \$271 million in assets and is eight years old; the differences could be due to the inclusion of form types other than the 10-K and the use of a more recent time period.

We model the association between scriptability and the constructs described above using the following multiple regression:

$$\begin{aligned} Script = {} & \alpha_0 + \Sigma\alpha_i(\textit{Investment in Financial Reporting}) \\ & + \Sigma\gamma_i(\textit{Technological Sophistication}) \\ & + \Sigma\delta_i(\textit{Use of a Filing Provider}) + \beta_0 HTML \\ & + \Sigma\,\textit{Year-MonthFEs} + \Sigma\,\textit{IndustryFEs} + e. \end{aligned} \tag{1}$$

Script equals *CompScript* or our task-based measures (*IdentifyData* or *DataToInformation*). Variable references within the summations refer to the three sets of variables described in the preceding paragraphs. We include year-month and industry fixed effects in all regressions where industries are defined using the Fama-French 48-industry designation. Further, we control for the presence of *HTML* since our measure of scriptability varies depending on whether the filing is prepared using HTML encoding. We expect *HTML* to relate positively to scriptability.

As reported in table 3, we find some support for an indirect improvement in scriptability from investments in financial reporting quality. Specifically, the coefficients on *Distress* and *Growth*, and, to a lesser extent, *OpComplex* and *FinComplex*, are generally negative and significant across most specifications, suggesting that constraints on investments in financial reporting quality reduce scriptability. We also find significantly positive coefficients for institutional ownership for *CompScript* in all form types, suggesting that

[15] Our *Watermark* measure comes from a discussion we had with an executive at Workiva.

[16] For variables measured using quarterly data, we report statistics for the average values over quarters in a given year in panel B to facilitate presentation.

TABLE 2
Descriptive Statistics

Panel A: Descriptive statistics for filing level variables

	10-K Sample				10-Q Sample				8-K Sample				DEF 14A Sample			
Variable	*n*	Mean	SD	Median	*n*	Mean	SD	Median	*n*	Mean	SD	Median	*n*	Mean	SD	Median
$IPTP^{5m}$	5,197	4.492	2.857	5.000	17,720	4.512	2.874	5.000	68,156	4.515	2.860	5.000	5,531	4.531	2.875	5.000
$IPTP^{60m}$	8,870	4.500	2.871	4.000	28,652	4.500	2.872	5.000	119,614	4.500	2.872	4.000	9,536	4.501	2.871	5.000
$IPTP^{24h}$	12,850	4.500	2.871	5.000	39,768	4.500	2.872	5.000	171,349	4.500	2.872	5.000	13,348	4.501	2.872	5.000
$IPTV^{5m}$	4,533	4.501	2.873	5.000	16,188	4.501	2.874	5.000	57,383	4.498	2.870	4.000	4,697	4.500	2.869	5.000
$IPTV^{60m}$	7,577	4.500	2.870	5.000	25,707	4.500	2.872	5.000	105,217	4.500	2.872	5.000	8,246	4.500	2.871	5.000
$IPTV^{24h}$	12,393	4.501	2.871	5.000	38,736	4.500	2.872	5.000	168,416	4.500	2.872	5.000	12,923	4.500	2.871	5.000
Dissem	16,261	0.038	0.192	0.000	51,001	0.021	0.142	0.000	217,993	0.402	0.490	0.000	17,187	0.018	0.133	0.000
NewsDelay	622	391.383	306.370	301.167	1,046	355.247	274.735	266.925	87,643	278.164	225.708	234.700	308	438.936	290.316	323.667
Top25FA	71,419	0.557	0.497	1.000	224,746	0.525	0.499	1.000	576,707	0.604	0.489	1.000	71,479	0.617	0.486	1.000
Watermark	71,419	0.133	0.340	0.000	224,746	0.145	0.352	0.000	576,707	0.228	0.419	0.000	71,479	0.087	0.282	0.000
HTML	71,419	0.566	0.496	1.000	224,746	0.548	0.498	1.000	576,707	0.740	0.438	1.000	71,479	0.510	0.500	1.000
News	16,265	2.930	1.405	3.178	51,015	2.994	1.364	3.219	218,085	3.098	1.449	3.332	17,191	3.167	1.365	3.367
FileCluster	13,038	8.983	6.181	8.000	40,257	8.656	6.495	7.000	173,013	6.965	5.559	5.000	13,527	6.704	5.338	5.000
Volume	16,261	12.159	2.170	12.389	51,001	12.212	2.154	12.447	217,993	12.638	2.095	12.852	17,187	12.258	2.148	12.500
#EAs	16,261	404.814	215.113	382.000	51,001	843.777	368.276	848.000	217,993	476.297	391.940	354.000	17,187	370.164	321.387	263.000
#Filings	16,261	6.475	0.783	6.678	51,001	6.420	0.730	6.592	217,993	6.037	0.909	6.288	17,187	6.016	0.894	6.194
DJLength	622	5.019	1.380	5.308	1,046	5.425	1.077	5.640	87,643	5.047	1.722	5.687	308	5.268	1.054	5.457
Volatility	16,261	0.418	0.264	0.343	51,001	0.419	0.264	0.346	217,993	0.420	0.277	0.340	17,187	0.409	0.255	0.337

(*Continued*)

TABLE 2—*Continued*

Panel B: Descriptive statistics for firm-year level variables

Variable	*n*	Mean	SD	25%	50%	75%
Size	81,414	5.642	2.128	4.026	5.532	7.110
Age	81,414	2.590	0.815	1.952	2.598	3.209
Distress	81,414	4.426	2.821	2.000	4.250	7.000
OpComplex	81,414	0.751	0.687	0.000	0.693	1.099
FinComplex	81,414	5.405	0.220	5.300	5.465	5.568
Growth	81,414	0.212	0.571	−0.011	0.088	0.240
%IH	81,414	0.378	0.340	0.015	0.320	0.686
HTI	81,414	0.504	0.457	0.000	0.500	1.000
FileLag	81,414	77.624	14.792	67.000	80.000	89.000
Follow	19,093	1.591	1.065	0.693	1.731	2.476
ROA	19,093	−0.003	0.045	−0.001	0.005	0.016
Loss	19,093	0.271	0.392	0.000	0.000	0.500
BTM	19,093	0.674	0.608	0.295	0.556	0.913
CapEx	19,093	0.039	0.054	0.005	0.022	0.050
Lev	19,093	0.238	0.205	0.078	0.190	0.346
Opinion	22,058	0.192	0.371	0.000	0.000	0.000
Lit	22,058	0.227	0.418	0.000	0.000	0.000

(*Continued*)

TABLE 2—*Continued*

Panel C: Correlations with *CompScript*

Form Type: Variable:	10-K *CompScript*	10-Q *CompScript*	8-K *CompScript*	DEF 14A *CompScript*
$IPTP^{5m}$	0.014	0.008	0.000	0.004
$IPTP^{60m}$	0.000	−0.005	**0.011**	0.000
$IPTP^{24h}$	**0.030**	**0.029**	−0.001	0.002
$IPTV^{5m}$	**0.074**	**0.040**	**0.016**	**0.040**
$IPTV^{60m}$	**0.048**	**0.014**	**0.020**	0.021
$IPTV^{24h}$	**0.019**	**0.017**	0.004	0.010
Dissem	**−0.031**	**−0.023**	**−0.026**	**−0.054**
NewsDelay	−0.047	0.049	0.001	0.044
#EAs	**−0.136**	**0.076**	**0.040**	**0.094**
#Filings	**−0.060**	**−0.021**	−0.002	**−0.023**
%IH	**−0.208**	**−0.151**	**−0.029**	**−0.116**
Age	**−0.100**	**−0.120**	**−0.009**	**−0.062**
BTM	**0.207**	**0.184**	**0.049**	**0.099**
CapEx	**−0.101**	**−0.086**	**−0.023**	**−0.025**
Distress	**0.063**	**0.058**	**−0.007**	**−0.025**
DJLength	0.004	**0.064**	**−0.055**	0.054

(*Continued*)

TABLE 2—*Continued*

Panel C: Correlations with *CompScript*

Form Type: Variable:	10-K *CompScript*	10-Q *CompScript*	8-K *CompScript*	DEF 14A *CompScript*
FileCluster	**−0.072**	**−0.023**	**−0.016**	**−0.021**
FileLag	**0.305**	**0.246**	**0.057**	**0.214**
FinComplex	**−0.271**	**−0.263**	**−0.036**	**−0.060**
Follow	**−0.279**	**−0.224**	**−0.057**	**−0.186**
Growth	**−0.050**	**−0.024**	**−0.018**	0.006
HTI	**0.029**	**0.023**	**0.036**	**0.050**
HTML	−0.003	**0.051**	**−0.176**	**0.208**
Lev	**−0.150**	**−0.138**	**−0.036**	**−0.077**
Lit	−0.009	**−0.031**	**0.007**	**0.054**
Loss	**0.024**	−0.004	**−0.017**	**0.045**
News	**−0.202**	**−0.200**	**−0.045**	**−0.143**
OpComplex	**−0.192**	**−0.200**	**−0.013**	**−0.093**
Opinion	**−0.042**	**−0.083**	**−0.019**	**−0.050**
ROA	−0.012	−0.006	**0.015**	**−0.028**
Size	**−0.352**	**−0.282**	**−0.069**	**−0.237**
Top25FA	**0.073**	**0.126**	**0.022**	**0.021**
Volatility	**0.125**	**0.084**	**0.040**	**0.121**
Volume	**−0.366**	**−0.303**	**−0.087**	**−0.207**
Watermark	−0.003	**−0.111**	**0.018**	**0.062**

This table presents descriptive statistics for all variables other than scriptability used in this study. Panel A presents means and standard deviations for variables measured at the filing level, by filing type. Panel B presents descriptive statistics for firm characteristics used in empirical models. For observations measured multiple times per year (e.g., *%IH*, quarterly Compustat items), we use the mean of values during the fiscal year for the purposes of presenting descriptive statistics. Panel C presents the Pearson correlation between each variable and *CompScript*. We report correlations using each of the four form groups examined. Correlations in **bold** are significant at the $p < 0.05$ level. All variables are defined in the appendix.

TABLE 3

Firm and Filing Characteristics Associated with Scriptability

Form Type:		10-K			10-Q			8-K			DEF 14A		
Dependent Variable:	*Prediction*	*CompScript*	*IdentifyData*	*DataToInfo.*	*CompScript*	*IdentifyData*	*DataToInfo.*	*CompScript*	*IdentifyData*	*DataToInfo.*	*CompScript*	*IdentifyData*	*DataToInfo.*
Size	+	−0.280***	−0.139**	−0.421***	−0.482***	−0.318***	−0.645***	−0.197***	−0.029	−0.365***	−0.243***	−0.202***	−0.283**
		(−3.23)	(−2.14)	(−3.09)	(−8.15)	(−6.38)	(−5.13)	(−5.85)	(−0.79)	(−6.54)	(−3.92)	(−4.98)	(−2.31)
Age	+	−0.600***	−1.025***	−0.176	−0.202*	−0.570***	0.166	0.019	0.062	−0.023	0.084	−0.269**	0.436**
		(−3.24)	(−4.80)	(−0.96)	(−1.71)	(−6.63)	(0.93)	(0.29)	(0.77)	(−0.24)	(0.60)	(−2.11)	(2.52)
Distress	−	−0.149***	−0.035*	−0.263***	−0.240***	−0.120***	−0.359***	−0.066***	0.042**	−0.173***	−0.108***	−0.080***	−0.135***
		(−7.99)	(−1.69)	(−8.98)	(−13.39)	(−4.93)	(−10.86)	(−3.68)	(2.30)	(−6.30)	(−4.91)	(−3.87)	(−4.27)
OpComplex	−	−0.083	0.223**	−0.390***	−0.257***	0.011	−0.524***	0.114*	−0.045	0.272**	−0.282***	−0.228**	−0.336***
		(−0.93)	(2.25)	(−2.85)	(−2.90)	(0.13)	(−3.67)	(1.72)	(−0.63)	(2.53)	(−3.90)	(−2.42)	(−2.76)
FinComplex	−	−2.358***	0.027	−4.742***	−2.409***	−0.812	−4.006***	−1.825***	−0.898	−2.753***	0.820	1.789***	−0.150
		(−3.55)	(0.04)	(−4.01)	(−3.70)	(−1.38)	(−4.02)	(−4.03)	(−1.58)	(−3.80)	(1.54)	(2.74)	(−0.19)
Growth	−	−0.218***	0.193**	−0.629***	−0.144***	0.054	−0.342***	−0.226***	−0.088*	−0.363***	0.005	0.089	−0.078
		(−3.38)	(2.56)	(−6.89)	(−4.13)	(1.13)	(−5.66)	(−5.09)	(−1.82)	(−5.10)	(0.10)	(1.48)	(−1.09)
%IH	+	0.564***	−0.205	1.333***	0.867***	0.369**	1.364***	0.313**	0.259	0.367	0.956***	0.213	1.700***
		(3.07)	(−0.90)	(5.44)	(5.34)	(2.12)	(5.83)	(2.11)	(1.51)	(1.46)	(4.82)	(1.00)	(6.23)
HTI	+	0.439*	0.964***	−0.085	0.226	0.326	0.126	0.394*	0.641***	0.147	0.092	−0.023	0.207
		(1.80)	(3.49)	(−0.26)	(1.09)	(1.40)	(0.40)	(1.67)	(2.71)	(0.43)	(0.44)	(−0.10)	(0.75)
FileLag	−	0.035***	0.062***	0.007	0.029***	0.038***	0.019**	0.004	−0.001	0.010*	0.017***	0.021***	0.012
		(5.49)	(10.34)	(0.80)	(4.86)	(7.97)	(2.05)	(1.30)	(−0.33)	(1.84)	(3.07)	(4.25)	(1.58)
Top25FA	+	4.793***	4.561***	5.025***	4.749***	4.422***	5.076***	2.303***	2.446***	2.159***	3.224***	2.761***	3.686***
		(12.87)	(8.77)	(14.72)	(12.47)	(9.86)	(12.51)	(7.49)	(6.20)	(6.77)	(9.18)	(5.04)	(13.51)
Watermark	+	2.369***	3.209***	1.529**	1.096***	0.597	1.596**	2.298***	1.586*	3.009***	1.830***	1.166**	2.495***
		(5.82)	(5.03)	(2.43)	(2.58)	(1.30)	(1.96)	(5.08)	(1.83)	(9.49)	(6.12)	(2.51)	(6.03)
HTML	+	10.693***	15.844***	5.542***	11.429***	16.147***	6.712***	−0.151	5.327***	−5.629***	8.500***	17.098***	−0.098
		(16.92)	(25.94)	(6.27)	(16.63)	(33.20)	(7.32)	(−0.12)	(6.07)	(−3.24)	(18.70)	(40.97)	(−0.14)
Year-month FEs?		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FEs?		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations		71,419	71,419	71,419	224,746	224,746	224,746	576,707	576,707	576,707	71,479	71,479	71,479
Adjusted-R^2		0.418	0.499	0.194	0.529	0.633	0.228	0.411	0.378	0.241	0.467	0.612	0.113

This table presents results from estimating equation (1) by form type for each measure of scriptability (*CompScript*, *IdentifyData*, and *DataToInformation*). All variables are defined in the appendix. Industries are defined using the Fama-French 48-industry designations. Reported *t*-statistics in parentheses are derived from standard errors clustered by firm and year. ***, **, and * denote two-tailed significance at the $p < 0.01$, $p < 0.05$, and $p < 0.10$ level, respectively. To facilitate presentation of coefficients, the dependent variable in each regression is multiplied by 100.

institutional owners value and encourage higher filing scriptability along with other improvements in financial reporting quality. In an untabulated analysis, we repeat estimation of equation (1) for disclosures submitted between 2010 and 2015, consistent with the sample period used in our hypothesis tests described in section 3. In these latter years, *%IH* becomes less consistently positive across specifications, but otherwise our results hold in the post-2009 period.

Contrary to expectations, we find that *Size* and *Age* are significantly negatively associated with scriptability. One possible explanation for this result is that larger and older firms invest more in financial reporting quality, but also have more complex filings. In addition, we have observed that many firms are slow to update disclosures once they have prepared them, creating higher quality markup for new tables but not updating existing tables.

Regarding intentional actions by managers to improve scriptability, we find a positive and weakly significant coefficient on *HTI* in 10-K and 8-K specifications using *CompScript*, and these effects appear to be driven by *IdentifyData*. In the post-2009 period, *HTI* becomes positive and significant for 10-Qs as well. Since technical expertise would likely have a greater effect on *DataToInformation*, with components such as binary exhibits and table markup, than *IdentifyData*, we do not interpret our result to constitute strong evidence that technological expertise has a direct effect on filing scriptability.[17] Contrary to expectations, we also find that *FileLag* is positively, rather than negatively, associated with scriptability. One possible explanation for this result is that the additional time taken to prepare a filing improves its quality.

Overall, the use of a high-quality filing agent shows the most significant association with higher filing scriptability. *Top25FA*, which indicates that the firm employs one of the top 25 filing agents, loads positively and significantly on all measures of scriptability for all form types, at less than the 1% level. *Watermark*, which we posit to indicate a greater focus by the filing agent on the markup of the filing, also has a positive and significant association with *CompScript* and *DataToInformation* at less than the 5% level for all form types, and a significantly positive association with *IdentifyData* for all forms but 10-Qs. Both *Top25FA* and *Watermark* are less significantly associated with scriptability after 2009 (untabulated), which implies recent convergence in methods used to prepare electronic filings.

Together, our results suggest that managers rarely directly intervene to improve scriptability, but may improve it indirectly through investments in the financial reporting function and engaging outside consultants. A lack of familiarity with the specific technologies used in financial reporting by top managers could explain this outcome.

[17] We note, however, that including industry fixed effects alongside an indicator for high-tech industries could affect the power of this test.

3. Hypothesis Development and Research Design

3.1 HYPOTHESIS DEVELOPMENT

In this section, we develop predictions to examine possible benefits of scriptability. Finance research characterizes market efficiency as the ability of markets to quickly and accurately impound information for asset allocation decisions (Fama et al. [1969], O'Hara [2003]). High-performance computing, when combined with the increasing availability of information on the World Wide Web, allows investors to analyze large amounts of data with unprecedented speed and accuracy while limiting or, in the specific case of algorithmic trading, eliminating costly human analysis (Brogaard, Hendershott, and Riordan [2014], Rogers, Skinner and Zechman [2017]). In addition to the higher speed of information collection, experimental research finds that programmatic analysis lends itself to more directed information search strategies that improve decision making (Hodge, Kennedy, and Maines [2004], Arnold et al. [2012]).

Bloomfield's [2002, p. 234] Incomplete Revelation Hypothesis (IRH) "asserts that statistics that are more costly to extract from public data are less completely revealed in market prices." This builds on prior theoretical research in economics (e.g., Grossman and Stiglitz [1980]) suggesting that market participants are constrained by information processing costs. Consistent with the IRH, empirical research associates disclosure readability with a more rapid response to the information content of 10-K filings (You and Zhang [2009]), particularly for bad news (Kim, Wang, and Zhang [2018]) and earnings surprises (Lee [2012]).

We expect that scriptability, like readability, will also reduce the time and resources necessary to process and analyze financial information, leading to a faster and more complete impounding of information into stock prices for more highly scriptable disclosures:

H1: Disclosure scriptability is positively associated with the speed of price formation.

Despite our prediction, scriptability may not affect price formation for several reasons. First, users may focus on a small group of filings and tailor their scripts to the idiosyncrasies in this group. In this case, only changes in scriptability may impede information processing. The algorithms of some traders may also be so sophisticated and endowed with resources that both scriptability levels and changes result in a negligible processing delay. Conversely, routine hand-collection of information by experts would enable rapid trading that is not sensitive to scriptability.

For our second hypothesis, we examine the relation between disclosure scriptability and intermediary dissemination. We focus on dissemination because prior research suggests that investors rely on intermediaries for "alerts" regarding regulatory filings (Li, Ramesh, and Shen [2011], Blankespoor, deHaan, and Zhu [2018]) and that news dissemination facilitates efficient investor pricing of earnings information (Drake, Guest, and Twedt [2014], Twedt [2016]). To the extent that scriptability increases the ability of business journalists and newswire providers such as DJ to quickly analyze and disseminate information in disclosures, we expect the likelihood of dissemination to increase and the delay between filing and dissemination to decrease for more scriptable disclosures:

H2: Scriptability is positively associated with the likelihood and speed of newswire dissemination.

If DJ has advance access to planned firm disclosures such as press releases and annual reports, however, any delays would occur prior to public availability. DJ could also dedicate staff to hand-collect firm information instead of relying on algorithms. As above, this would result in dissemination delays sensitive to considerations other than scriptability. Finally, Blankespoor, deHaan, and Zhu [2018] report that automated processing and dissemination of earnings press releases by the Associated Press did not occur until late 2014. While our discussions with DJ suggest some level of automation occurs before this date, we may fail to find support for H2 if fully automated dissemination is a relatively recent practice.

3.2 RESEARCH DESIGN

3.2.1. H1 Research Design. To empirically measure the speed of price formation following the filing, we follow recent studies and use the IPT of the market response (Butler, Kraft, and Weiss [2007], Bushman, Smith, and Wittenberg-Moerman [2010], McMullin, Miller, and Twedt [2015], Twedt [2016]). Using the "Accepted" timestamp from the filing's EDGAR index page, we examine IPT following the disclosure over three separate windows, five minutes, 60 minutes, and 24 hours in order to capture the response of both algorithmic scripters and traditional investors.[18] Since these short windows require intraday trading information, we modify the measure of IPT for use with Trade and Quote (TAQ) data by changing the unit of measure over which the response is measured from days or months to minutes and by using midpoints of quotes to compute effective, rather than realized, buy-and-hold returns. Following Holden and Jacobsen [2014], we eliminate nonnormal and missing quotes, quotes where the bid exceeds the offer, and extreme quotes. We then measure IPT as the scaled area under the cumulative price change curve from time 0 to a chosen end point (McMullin, Miller, and Twedt [2015]). We measure cumulative price

[18] According to Rogers, Skinner, and Zechman [2017], this time stamp precedes actual disclosure availability on EDGAR by an average of 30 seconds. However, no better time stamp is publicly available. Further, our shortest IPT measurement window is 10 times longer than this delay, which should minimize the effect of any bias due to misspecifying the start-time of each window.

change using effective returns (*ER*) for a window lasting w minutes as follows:

$$IPT = \frac{1}{2}\sum_{m=1}^{w} (ER_{m-1} + ER_m) / ER_w = \sum_{m=1}^{w-1} (ER_m/ER_w) + 0.5. \quad (2)$$

ER_m equals the ER from minute 0 to minute m. Intuitively, a faster (or "steeper") response yields a larger area under the response curve. To illustrate, if the full window return occurred in the first minute, then ER_m equals ER_w for all $m \geq 1$, so *IPT* equals $w - 0.5$. Conversely, if the full window return happened in the final minute, ER_m equals 0 for all m less than w, and *IPT* equals 0.5.

We also examine the association between scriptability and trading volume as an alternative measure of the speed of price formation (Rogers, Skinner, and Zechman [2017]). Since Beaver [1968], researchers have acknowledged that new information can affect both market price and the portfolio decisions of individual investors. Therefore, "[t]rading volume reflects investors' activity" (Bamber [1986, p. 40]) even in the absence of price movement, and we expect portfolio decisions to be made more rapidly in the presence of higher scriptability. We calculate the IPT of the volume response by replacing the *ER* terms with cumulative volume (*CV*) in equation (2) using the windows and methodology previously described.

Because finance research highlights the importance of the time at which returns are measured (e.g., Foster and Viswanathan [1993], Busse and Green [2002]), we adjust our price and volume IPT measures by a "normal" IPT, similar to Rogers, Skinner, and Zechman [2017]. Specifically, we calculate the average price and volume IPT on the same time and day as the filing over the prior 51 weeks and subtract it from the filing IPT. After calculating this adjusted IPT metric for all filings in our sample, we create decile-ranked variables by year and form type, which we use as the dependent variable in our models. *IPTP* (*IPTV*) refers to the ranked abnormal price-based (volume-based) IPT measure. The ranking procedure limits the impact of the inherent noise in returns-based area-under-the-curve measures (McMullin, Miller, and Twedt [2015], Drake, Thornock, and Twedt [2017]).

We test H1 using the following Ordinary Least Squares (OLS) regression model estimated separately for each form type:[19]

$$IPT = \alpha + \beta\ Script + \Sigma\gamma_i\ (Controls) + e. \quad (3)$$

Script equals *CompScript* or our task-based measures (*IdentifyData* and *DataToInformation*), and *IPT* equals either *IPTP* or *IPTV* measured over one of the three windows. We use superscripts "5m" (five minutes), "60m" (60 minutes) and "24h" (24 hours) on *IPTP* and *IPTV* to denote each measurement

[19] Results are qualitatively similar if we use an ordered logistic regression.

window. H1 predicts significantly positive values for β. We control for factors associated with scriptability described in subsection 2.2. In addition, we control for various aspects of firms' information environments identified to relate to IPT. Specifically, we control for analyst following (*Follow*), the number of recent news articles about the firm (*News*), performance (*ROA* and *Loss*), growth prospects (Book-to-market [*BTM*]), investments (*CapEx*), and capital structure (*Lev*). We also control for the number of filings received by the SEC in the 60 seconds prior to the filing of interest (*FileCluster*) since Rogers, Skinner, and Zechman [2017] show this variable relates positively with the time to dissemination, which could impact our measures of IPT. We exclude controls from McMullin, Miller, and Twedt [2015] related to the time of day the filing occurs since our IPT adjustment procedure controls for time-related differences in both price and volume responses. All models include year-month and industry fixed effects.

We obtain intraday trading data from TAQ. Given a large percentage of filings occur outside of trading hours, we take advantage of TAQ's coverage of pre- and postmarket sessions and consider filings made both during and outside traditional market hours. As of September 20, 2006, TAQ tracks trades made from 4:00 am to 9:30 am (premarket) and 4:00 pm to 8:00 pm (postmarket) on all three major exchanges. Our abnormal IPT measures require 52 weeks of lagged data, so the earliest we could start our data and have a constant coverage window is late 2007, which corresponds to the beginning of the financial crisis. Thus, to avoid issues with deriving expected measures of IPT during this volatile period, we begin our sample for these market tests in January of 2010, requiring data from calendar year 2009 through the end of 2015. This later sample period is also appropriate because the influence of scripters on price formation has likely increased over time with advances in technology and information availability.

3.2.2. H2 Research Design. We use DJ newswires to measure dissemination of filing information by the business press. While Li, Ramesh, and Shen [2011] describe the DJ news production process as largely manual, with reporters reviewing important disclosures and submitting news flashes with value relevant content to subscribers, a recent discussion with a representative from DJ suggests that at least a portion of the process is now highly automated. Other intermediaries such as RavenPack, which provides proprietary sentiment measures immediately following news events (Drake, Guest, and Twedt [2014]), also rely on DJ for machine-readable data. We identify dissemination by using the "Accepted" timestamp from EDGAR and searching for newswires within 24 hours of the filing, as in our design for H1.[20] To mitigate the risk of erroneously matching newswires to filings, we require newswires to either reference the type of disclosure (i.e., "10K" or "Annual Report" for form 10-K filings) or contain some combination of

[20] Twedt [2016] finds that the majority of dissemination articles occur within one day of disclosure.

"SEC" and "filing" (or "filed") in the headline or lead paragraph. *Dissem* equals 1 if we identify at least one newswire with one of these criteria in the 24 hours following filing receipt by the SEC, and *NewsDelay* equals the time between the disclosure's SEC timestamp and the newswire closest to the filing time. In the event a newswire occurs within 24 hours of two or more filings by the same firm of the same form type, we assume the disclosure closest to the newswire was the one disseminated. We then estimate the following Probit (OLS) model when *Dissem* (*NewsDelay*) is the dependent variable (*DV*) separately for each form type:

$$DV = \alpha + \beta Script + (\Sigma \gamma_k Controls_k) + e. \tag{4}$$

Script equals either *CompScript* or our task-based measures (*IdentifyData* or *DataToInformation*). H2 predicts positive (negative) values for β when *Dissem* (*NewsDelay*) is the dependent variable. We control for the factors we expect to be associated with scriptability identified in subsection 2.2 and additional controls based on the findings in Li, Ramesh, and Shen [2011]. Specifically, we include *Opinion*, which equals 1 for any disclosure issued in a year following a nonstandard audit opinion in Compustat. Second, we include *Volume,* which is the average daily trading volume over the quarter preceding the disclosure. Third, we control for the number of earnings announcements (*#EAs*) in all of Compustat occurring within 24 hours of the filing, as DJ prioritizes the dissemination of earnings announcements over other information events, potentially altering dissemination delays. For similar reasons, we also control for the total number of SEC filings occurring in the two-hour window surrounding the disclosure (*#Filings*). We also include controls for stock return volatility (*Volatility*) and litigation risk (*Lit*). When *NewsDelay* is the dependent variable, we also control for the length of the DJ article (*DJLength)* since longer articles likely take longer to prepare and disseminate. We also include *News* to control for press following. Finally, in addition to industry and year-month fixed effects, we include hour and day-of-week fixed effects to account for differences related to the time of day the filing occurs.

Footnote 7 of Li, Ramesh, and Shen [2011] indicates that DJ significantly restructured their news service in mid-2009. To avoid possible effects of this change, we limit our sample period to 2010–2015, the same period for our test of H1.

3.3 DESCRIPTIVE STATISTICS FOR HYPOTHESIS TESTS

From table 2, panel A, note that the IPT measures are uniformly distributed by definition. *Dissem* varies greatly by form type with DJ disseminating 40% of 8-K filings compared to 4% of 10-K filings and 2% of 10-Q and DEF 14A filings. Panel C documents several statistically significant positive correlations between *CompScript* and the IPT measures, especially for the IPT volume measures. Thus, we find some support for our hypotheses on a univariate basis.

4. *Empirical Results*

4.1 RESULTS FOR H1

Table 4 presents results of our test of H1 using *IPTP* to proxy for the speed of price formation. Panel A (B and C) presents results using the five-minute (60-minute and 24-hour) measurement window. H1 predicts a positive coefficient on our scriptability measures. As shown, panel A of table 4 provides no support for H1. Neither *CompScript* nor our task-based measures (*IdentifyData* and *DataToInformation*) relates significantly to $IPTP^{5m}$. It is worth noting that the adjusted-R^2 values in each column are very low and, in fact, negative in the first two columns. Inspection of the data suggests that the short-window *IPTP* measures are incredibly volatile, likely due to relatively small returns in the denominator of the IPT measures. We discuss untabulated, unplanned analyses related to this observation shortly. Panel B of table 4 provides some support for H1 using the 60-minute window. Namely, we observe a weakly positive relation between *IdentifyData* and *IPTP* for 10-Ks ($t = 1.62$). For 8-Ks, we report positive coefficients on both *CompScript* ($t = 2.26$) and *DataToInformation* ($t = 2.90$). Finally, for proxy statements, *DataToInformation* loads positively ($t = 2.35$), as predicted. Contrary to predictions, we observe a significantly *negative* coefficient ($t = -2.60$) on *CompScript* for 10-Qs. Panel C reports results using *IPTP* measured over a 24-hour window. As in panel B, we again find some support for H1 for 10-K filings. However, unlike in panel B, *CompScript* and *DataToInformation* each exhibit significantly positive coefficients ($t = 2.15$ and $t = 2.27$, respectively), whereas *IdentifyData* falls below conventional significance levels ($t = 1.18$). Also, as in panel B, we find some support for H1 in our 8-K sample, as *IdentifyData* loads positively ($t = 2.31$).

To serve as an effective proxy for the speed of price formation, *IPTP* requires a nontrivial return over the measurement window and regular changes in quotes (since we use ERs). Absent these two conditions (i.e., infrequent quotes or a total return close to 0), *IPTP* is highly volatile, which likely explains the very low explanatory power of our models. To address this issue, we conduct two unplanned, untabulated additional analyses. First, we repeat all analyses in table 4 after excluding observations that have fewer than three (31, 13) observations for the five-minute (60-minute, 24-hour) measurement window. This restriction reduces measurement error associated with infrequent changes in the quote midpoint. Using this subsample, we again fail to find any significant effect of scriptability on price formation in the five-minute window. Over the 60-minute window, *CompScript* and *DataToInformation* each exhibit positive relations with *IPTP* for 8-K filings ($t = 2.24$ for both), and *DataToInformation* also relates positively to *IPTP* for proxy statements ($t = 2.71$). Over the 24-hour window, our only support for H1 comes for 8-K filings—coefficients on both *CompScript* and *IdentifyData* are significantly positive ($t = 3.11$ and 2.68, respectively). Second, we limit the sample to filings with full-window returns

TABLE 4

Scriptability and Speed of Price Formation (IPTP)

Panel A: Price IPT for five-minute window

Form Type:	10-K		10-Q		8-K		DEF 14A	
Dependent Variable:	$IPTP^{5m}$	$IPTP^{5m}$	$IPTP^{5m}$	$IPTP^{5m}$	$IPTP^{5m}$	$IPTP^{5m}$	$IPTP^{5m}$	$IPTP^{5m}$
CompScript	0.228		0.299		−0.102		−0.250	
	(0.30)		(0.77)		(−0.87)		(−0.39)	
IdentifyData		0.186		−0.035		−0.049		−0.450
		(0.28)		(−0.19)		(−0.31)		(−0.59)
DataToInformation		0.045		0.273		−0.052		0.114
		(0.16)		(0.88)		(−0.79)		(0.26)
Follow	−0.054	−0.054	0.016	0.015	0.026	0.026	−0.019	−0.018
	(−0.95)	(−0.95)	(0.33)	(0.32)	(0.81)	(0.81)	(−0.23)	(−0.22)
ROA	−1.315	−1.305	0.334	0.327	0.130	0.130	−1.534	−1.535
	(−1.10)	(−1.11)	(0.75)	(0.73)	(1.23)	(1.27)	(−0.70)	(−0.71)
News	0.043	0.042	−0.012	−0.011	0.012	0.012	−0.017	−0.015
	(0.89)	(0.87)	(−0.52)	(−0.48)	(0.49)	(0.48)	(−0.36)	(−0.34)
BTM	−0.021	−0.022	−0.046	−0.044	0.042**	0.042*	0.190*	0.191*
	(−0.34)	(−0.34)	(−0.82)	(−0.78)	(1.96)	(1.96)	(1.69)	(1.70)
Loss	−0.054	−0.054	−0.000	0.001	0.015	0.015	−0.190	−0.188
	(−0.33)	(−0.33)	(−0.01)	(0.01)	(0.36)	(0.36)	(−1.59)	(−1.58)
CapEx	0.624	0.623	−0.630	−0.636	−0.165	−0.165	−0.584	−0.595
	(0.84)	(0.83)	(−0.96)	(−0.97)	(−1.14)	(−1.14)	(−1.00)	(−1.02)
Lev	0.264	0.269	−0.018	−0.022	0.163*	0.163*	0.886***	0.879***
	(0.90)	(0.97)	(−0.11)	(−0.14)	(1.82)	(1.80)	(3.27)	(3.25)
FileCluster	0.013**	0.013**	0.013	0.013	0.016***	0.016***	0.014***	0.015***
	(2.15)	(2.10)	(1.57)	(1.57)	(4.67)	(4.75)	(3.05)	(3.11)
Size	0.027	0.026	−0.002	0.000	0.001	0.001	−0.003	−0.001
	(0.75)	(0.79)	(−0.06)	(0.00)	(0.05)	(0.05)	(−0.06)	(−0.03)
Age	0.070	0.071	0.042	0.040	−0.040**	−0.040**	0.005	0.003
	(0.92)	(0.94)	(1.04)	(1.01)	(−2.10)	(−2.09)	(0.09)	(0.05)

(Continued)

TABLE 4—*Continued*

Panel A: Price IPT for five-minute window

Form Type:	10-K		10-Q		8-K		DEF 14A	
Dependent Variable:	$IPTP^{5m}$	$IPTP^{5m}$	$IPTP^{5m}$	$IPTP^{5m}$	$IPTP^{5m}$	$IPTP^{5m}$	$IPTP^{5m}$	$IPTP^{5m}$
Distress	−0.060**	−0.061**	−0.013	−0.012	−0.009	−0.009	−0.054***	−0.053***
	(−2.44)	(−2.48)	(−1.14)	(−1.11)	(−0.96)	(−0.95)	(−2.95)	(−2.98)
OpComplex	−0.009	−0.010	0.003	0.004	−0.001	−0.001	0.063	0.063
	(−0.12)	(−0.13)	(0.09)	(0.12)	(−0.05)	(−0.05)	(1.21)	(1.21)
FinComplex	−0.495	−0.506	−0.206	−0.197	0.005	0.005	−1.460***	−1.451***
	(−0.83)	(−0.82)	(−1.01)	(−1.00)	(0.03)	(0.02)	(−8.08)	(−7.37)
Growth	−0.094	−0.095	0.024	0.025	−0.017	−0.017	−0.003	−0.002
	(−0.77)	(−0.77)	(0.51)	(0.54)	(−0.82)	(−0.82)	(−0.03)	(−0.02)
%IH	0.322	0.322	0.025	0.025	0.091**	0.091**	0.114	0.113
	(1.38)	(1.38)	(0.37)	(0.37)	(2.42)	(2.42)	(0.67)	(0.67)
HTI	−0.328**	−0.328**	0.206***	0.205***	−0.127***	−0.127***	−0.117	−0.119
	(−2.37)	(−2.38)	(2.76)	(2.77)	(−3.87)	(−3.85)	(−0.66)	(−0.67)
FileLag	−0.000	−0.000	−0.001	−0.001	−0.000	−0.000	0.002	0.002
	(−0.07)	(−0.08)	(−0.19)	(−0.18)	(−0.37)	(−0.37)	(0.28)	(0.29)
Top25FA	−0.080	−0.078	−0.056*	−0.059*	−0.016	−0.016	−0.042	−0.050
	(−0.49)	(−0.49)	(−1.70)	(−1.70)	(−0.34)	(−0.33)	(−0.34)	(−0.40)
Watermark	−0.012	−0.012	−0.058	−0.063	0.013	0.013	−0.107	−0.112
	(−0.07)	(−0.07)	(−0.81)	(−0.87)	(0.48)	(0.49)	(−1.14)	(−1.21)
HTML	0.508**	0.492*	0.294	0.319	−0.065***	−0.065***	0.230	0.273
	(2.17)	(1.93)	(0.83)	(0.92)	(−6.64)	(−4.56)	(1.00)	(1.13)
Year-month FEs?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FEs?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,197	5,197	17,720	17,720	68,155	68,155	5,531	5,531
Adjusted-R^2	−0.003	−0.003	0.001	0.001	0.002	0.002	0.000	0.000

(*Continued*)

TABLE 4—*Continued*

Panel B: Price IPT for 60-minute window

Form Type:	10-K		10-Q		8-K		DEF 14A	
Dependent Variable:	$IPTP^{60m}$	$IPTP^{60m}$	$IPTP^{60m}$	$IPTP^{60m}$	$IPTP^{60m}$	$IPTP^{60m}$	$IPTP^{60m}$	$IPTP^{60m}$
CompScript	−0.117		−0.309***		0.283**		−0.020	
	(−0.21)		(−2.60)		(2.26)		(−0.04)	
IdentifyData		0.311		−0.073		−0.076		−0.635
		(1.62)		(−0.38)		(−0.76)		(−1.08)
DataToInformation		−0.417		−0.208		0.270***		0.462**
		(−0.84)		(−1.57)		(2.90)		(2.35)
Follow	0.022	0.023	0.033	0.034	0.028	0.029	−0.037	−0.034
	(0.38)	(0.41)	(0.91)	(0.91)	(1.55)	(1.56)	(−1.25)	(−1.19)
ROA	0.048	0.067	1.206	1.211	−0.202	−0.207	0.278	0.285
	(0.06)	(0.08)	(1.60)	(1.60)	(−0.74)	(−0.75)	(0.27)	(0.27)
News	0.006	0.002	−0.050**	−0.051**	−0.007	−0.007	0.010	0.013
	(0.09)	(0.04)	(−2.34)	(−2.34)	(−0.45)	(−0.41)	(0.18)	(0.23)
BTM	0.017	0.015	0.039	0.038	0.047*	0.046*	−0.084**	−0.082**
	(0.23)	(0.20)	(1.10)	(1.09)	(1.92)	(1.87)	(−2.07)	(−2.01)
Loss	0.032	0.029	0.034	0.034	−0.009	−0.009	0.023	0.025
	(0.38)	(0.35)	(0.47)	(0.47)	(−0.40)	(−0.39)	(0.46)	(0.51)
CapEx	0.124	0.140	−0.869***	−0.866***	0.488**	0.491**	0.472	0.455
	(0.17)	(0.19)	(−2.65)	(−2.64)	(2.19)	(2.19)	(1.04)	(1.00)
Lev	0.102	0.119	−0.166	−0.165	0.069	0.066	−0.199	−0.210
	(0.42)	(0.47)	(−0.71)	(−0.71)	(0.84)	(0.79)	(−1.16)	(−1.22)
FileCluster	0.046***	0.046***	0.047***	0.047***	0.063***	0.063***	0.049***	0.050***
	(8.20)	(8.13)	(7.59)	(7.58)	(12.80)	(12.65)	(5.15)	(5.13)
Size	−0.007	−0.013	−0.015	−0.016	−0.032	−0.031	−0.031	−0.029
	(−0.18)	(−0.31)	(−0.66)	(−0.66)	(−1.63)	(−1.62)	(−0.68)	(−0.63)

(*Continued*)

TABLE 4—*Continued*

Panel B: Price IPT for 60-minute window

Form Type:	10-K		10-Q		8-K		DEF 14A	
Dependent Variable:	$IPTP^{60m}$	$IPTP^{60m}$	$IPTP^{60m}$	$IPTP^{60m}$	$IPTP^{60m}$	$IPTP^{60m}$	$IPTP^{60m}$	$IPTP^{60m}$
Age	0.041*	0.043*	0.019	0.020	0.010	0.011	−0.092***	−0.095***
	(1.75)	(1.93)	(1.13)	(1.17)	(1.11)	(1.20)	(−2.98)	(−2.96)
Distress	−0.017	−0.019	0.010	0.009	−0.009**	−0.009**	0.028	0.028
	(−0.75)	(−0.85)	(0.71)	(0.70)	(−2.18)	(−2.00)	(1.25)	(1.26)
OpComplex	−0.078	−0.080	0.000	−0.000	0.007	0.006	−0.039	−0.038
	(−1.20)	(−1.24)	(0.01)	(−0.00)	(0.37)	(0.32)	(−1.03)	(−1.01)
FinComplex	0.214**	0.163	−0.837***	−0.841***	0.201	0.209	0.877*	0.918**
	(2.20)	(1.24)	(−2.78)	(−2.79)	(1.03)	(1.08)	(1.88)	(2.06)
Growth	0.079	0.079	−0.033	−0.034	−0.006	−0.006	−0.103**	−0.106**
	(1.38)	(1.39)	(−0.55)	(−0.56)	(−0.41)	(−0.38)	(−2.47)	(−2.49)
%IH	−0.052	−0.050	0.201***	0.201***	0.061*	0.062*	0.321*	0.316*
	(−0.31)	(−0.30)	(3.20)	(3.20)	(1.67)	(1.74)	(1.77)	(1.74)
HTI	−0.174	−0.178	0.054	0.054	0.065*	0.065*	−0.062	−0.065
	(−1.04)	(−1.06)	(0.73)	(0.74)	(1.69)	(1.72)	(−0.40)	(−0.42)
FileLag	0.003	0.003	−0.003	−0.003	−0.001	−0.001	0.000	0.000
	(0.93)	(0.80)	(−1.63)	(−1.63)	(−1.13)	(−1.14)	(0.00)	(0.10)
Top25FA	−0.092*	−0.083*	0.001	0.002	−0.020	−0.017	−0.013	−0.028
	(−1.80)	(−1.65)	(0.02)	(0.04)	(−0.86)	(−0.78)	(−0.13)	(−0.30)
Watermark	−0.031	−0.033	−0.052	−0.050	0.020	0.015	−0.104	−0.115
	(−0.90)	(−0.89)	(−1.20)	(−1.17)	(0.95)	(0.68)	(−1.31)	(−1.48)
HTML	−0.042	−0.124	0.127	0.117	−0.055	−0.034	−0.082	−0.001
	(−0.12)	(−0.36)	(0.75)	(0.78)	(−1.42)	(−0.74)	(−0.46)	(−0.00)
Year-month fixed effects?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FEs?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8,870	8,870	28,652	28,652	119,613	119,613	9,536	9,536
Adjusted-R^2	0.010	0.010	0.010	0.010	0.015	0.015	0.008	0.008

(*Continued*)

TABLE 4—*Continued*

Panel C: Price IPT for 24-hour window

Form Type:	10-K		10-Q		8-K		DEF 14A	
Dependent Variable:	$IPTP^{24h}$	$IPTP^{24h}$	$IPTP^{24h}$	$IPTP^{24h}$	$IPTP^{24h}$	$IPTP^{24h}$	$IPTP^{24h}$	$IPTP^{24h}$
CompScript	0.978**		0.105		0.208		−0.318	
	(2.15)		(0.36)		(1.63)		(−0.70)	
IdentifyData		0.508		0.052		0.233**		−0.069
		(1.18)		(0.29)		(2.31)		(−0.32)
DataToInformation		0.471**		0.053		0.028		−0.227
		(2.27)		(0.27)		(0.27)		(−0.59)
Follow	−0.040	−0.040	0.008	0.008	−0.008	−0.008	0.057	0.056
	(−0.72)	(−0.71)	(0.28)	(0.28)	(−0.59)	(−0.61)	(0.98)	(0.95)
ROA	0.286	0.287	−0.389	−0.389	−0.104	−0.102	1.129	1.125
	(0.58)	(0.58)	(−0.67)	(−0.66)	(−0.54)	(−0.53)	(1.38)	(1.38)
News	0.048**	0.048**	−0.020	−0.020	−0.067***	−0.067***	−0.010	−0.010
	(2.09)	(2.11)	(−0.65)	(−0.65)	(−3.31)	(−3.32)	(−0.17)	(−0.18)
BTM	0.047	0.047	−0.024	−0.024	0.034	0.034	−0.029	−0.029
	(1.03)	(1.01)	(−0.65)	(−0.66)	(1.37)	(1.39)	(−0.57)	(−0.57)
Loss	−0.111**	−0.111**	−0.070	−0.070	−0.010	−0.010	−0.060	−0.061
	(−2.25)	(−2.28)	(−0.86)	(−0.86)	(−0.30)	(−0.31)	(−0.72)	(−0.73)
CapEx	−0.618	−0.617	−0.245	−0.245	−0.275	−0.276	−0.327	−0.326
	(−0.86)	(−0.85)	(−0.85)	(−0.84)	(−0.98)	(−0.99)	(−0.70)	(−0.69)
Lev	−0.079	−0.078	0.050	0.050	0.095*	0.097*	−0.392	−0.391
	(−0.70)	(−0.71)	(0.43)	(0.43)	(1.76)	(1.81)	(−1.54)	(−1.55)
FileCluster	−0.032***	−0.032***	−0.048***	−0.048***	−0.063***	−0.063***	−0.043**	−0.043**
	(−5.15)	(−5.13)	(−6.11)	(−6.11)	(−10.78)	(−10.84)	(−2.49)	(−2.46)
Size	−0.075***	−0.075**	−0.094***	−0.094***	−0.051***	−0.051***	−0.115***	−0.116***
	(−2.77)	(−2.53)	(−8.01)	(−8.22)	(−3.82)	(−3.84)	(−2.94)	(−2.96)

(*Continued*)

TABLE 4—*Continued*

Panel C: Price IPT for 24-hour window

Form Type:	10-K		10-Q		8-K		DEF 14A	
Dependent Variable:	$IPTP^{24h}$	$IPTP^{24h}$	$IPTP^{24h}$	$IPTP^{24h}$	$IPTP^{24h}$	$IPTP^{24h}$	$IPTP^{24h}$	$IPTP^{24h}$
Age	−0.087*	−0.087*	−0.017	−0.017	0.006	0.006	−0.022	−0.021
	(−1.84)	(−1.81)	(−0.74)	(−0.73)	(0.40)	(0.38)	(−0.33)	(−0.31)
Distress	0.009	0.009	0.001	0.001	−0.006	−0.006	0.032**	0.032**
	(0.66)	(0.68)	(0.05)	(0.05)	(−0.92)	(−0.97)	(2.31)	(2.32)
OpComplex	0.008	0.007	−0.007	−0.007	0.007	0.007	−0.047**	−0.047**
	(0.18)	(0.17)	(−0.36)	(−0.36)	(0.41)	(0.44)	(−2.34)	(−2.34)
FinComplex	0.570	0.568	−0.737***	−0.737***	−0.036	−0.041	0.397	0.390
	(1.09)	(1.12)	(−2.62)	(−2.60)	(−0.30)	(−0.34)	(1.05)	(1.02)
Growth	−0.060	−0.060	0.041***	0.041***	−0.006	−0.006	−0.016	−0.016
	(−0.83)	(−0.82)	(3.72)	(3.67)	(−0.24)	(−0.24)	(−0.23)	(−0.23)
%IH	0.158	0.159	0.321***	0.321***	0.110**	0.109**	0.260***	0.261***
	(1.46)	(1.49)	(5.67)	(5.67)	(2.17)	(2.14)	(2.86)	(2.84)
HTI	0.087	0.087	−0.080	−0.080	−0.009	−0.009	0.099	0.100
	(0.64)	(0.66)	(−0.76)	(−0.76)	(−0.18)	(−0.19)	(1.19)	(1.19)
FileLag	−0.000	−0.000	0.001	0.001	−0.002	−0.002	−0.000	−0.000
	(−0.14)	(−0.15)	(0.88)	(0.89)	(−1.59)	(−1.58)	(−0.00)	(−0.02)
Top25FA	−0.019	−0.018	−0.003	−0.003	0.054*	0.053*	0.018	0.020
	(−0.27)	(−0.27)	(−0.11)	(−0.11)	(1.69)	(1.67)	(0.29)	(0.31)
Watermark	−0.075	−0.075	−0.021	−0.021	0.015	0.018	0.030	0.032
	(−0.88)	(−0.87)	(−0.45)	(−0.45)	(0.53)	(0.62)	(0.41)	(0.42)
HTML	−0.234	−0.238	−0.198	−0.198	0.045*	0.032	0.152*	0.140*
	(−1.00)	(−1.01)	(−1.30)	(−1.23)	(1.87)	(1.31)	(1.81)	(1.95)
Year-month FEs?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FEs?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,850	12,850	39,768	39,768	171,348	171,348	13,348	13,348
Adjusted-R^2	0.013	0.013	0.018	0.018	0.018	0.018	0.012	0.012

(*Continued*)

TABLE 4—*Continued*

Panel D: Unplanned analysis: Correlations between scriptability components and IPTP measures

	10-K			10-Q			8-K			DEF 14A		
Component	IPTP5	IPTP60	IPTP24	IPTP5	IPTP60	IPTP24	IPTP5	IPTP60	IPTP24	IPTP5	IPTP60	IPTP24
TextFromTables	0.015	−0.006	0.005	0.000	−0.010	**0.022**	−0.007	0.000	**0.015**	0.011	0.008	0.006
SeparateSections	−0.008	−0.001	0.013	0.009	−0.006	0.007	−0.002	**−0.009**	0.002	−0.023	−0.016	0.000
UsefulHeadings	0.027	0.008	0.012	−0.004	0.007	0.008	0.000	0.000	0.005	0.009	0.005	−0.014
ExtReferences	0.004	0.018	0.009	0.004	−0.006	**0.011**	0.003	−0.001	**−0.006**	0.012	0.010	−0.002
BinaryExhibits	−0.008	0.006	0.003	0.009	**0.013**	0.007	0.005	**0.021**	**−0.007**	−0.003	0.009	0.002
%NumTabulate	0.022	−0.012	0.008	0.010	−0.002	0.002	**−0.012**	0.005	**0.025**	−0.006	−0.007	−0.005
UnexpectChar	0.008	−0.002	0.007	−0.003	−0.007	**0.029**	0.001	0.005	**−0.014**	0.012	−0.002	0.009
TokenQuality	0.004	−0.008	0.017	−0.006	0.002	**0.016**	0.003	**0.012**	**−0.006**	0.007	0.001	0.000
Typo	0.000	−0.009	**0.018**	−0.004	−0.003	0.006	0.006	0.004	**−0.010**	0.020	−0.012	0.011
TableMarkup	0.003	−0.016	0.014	0.011	0.000	−0.004	−0.001	**0.015**	**0.011**	−0.003	0.017	−0.002
CellConsistency	−0.007	0.019	0.001	**0.017**	0.003	−0.003	0.009	**0.009**	**−0.020**	−0.013	−0.006	−0.005
UsefulRowNames	0.000	−0.009	**0.029**	−0.008	−0.011	**0.027**	**−0.016**	**0.010**	**0.012**	−0.011	0.001	0.015
Average across components	0.005	−0.001	0.011	0.003	−0.002	0.011	−0.001	0.006	0.001	0.001	0.000	0.001

This table presents results from estimating equation (3) with *IPTP* as the dependent variable. Panel A (B and C) presents results using the *IPTP* measure computed over the five-minute (60-minute and 24-hour) window. We estimate equation (3) by form type in all panels. All variables are defined in the appendix. Industries are defined using the Fama-French 48-industry designations. Reported *t*-statistics in parentheses are derived from standard errors clustered by firm and year. ***, **, and * denote two-tailed significance at the $p < 0.01$, $p < 0.05$, and $p < 0.10$ level, respectively. Panel D presents Spearman correlation coefficients between scriptability components and IPTP measures. Bolded correlations are significant at the $p < 0.05$ level, and shaded cells in the table denote a significant correlation consistent with H1. Note that the analysis in panel D is unplanned and not part of the original registered report proposal.

exceeding 1% in magnitude to avoid issues with very small IPT denominators, and we find support for H1 within our 8-K sample for all three windows. Specifically, *CompScript* relates positively to *IPTP* in the five-minute ($t = 1.98$) and 60-minute ($t = 2.60$) windows, *IdentifyData* relates positively to *IPTP* in the five-minute ($t = 1.89$) and 24-hour window ($t = 2.61$), and *DataToInformation* relates positively to *IPTP* in the 60-minute window ($t = 2.02$). As for other form groups, we find a positive coefficient on *CompScript* for proxy statements in the five-minute window ($t = 1.85$) and a weakly positive coefficient on *CompScript* for Form 10-Ks in the 24-hour window ($t = 1.67$).

While noise in the *IPTP* measures could explain the mixed results in table 4, it is also possible that the summary and task-based measures fail to adequately capture the components of scriptability most relevant to the speed of price formation. In panel D of table 4, we present results from an unplanned additional analysis. Specifically, we report Spearman correlations between each component of scriptability and *IPTP* measures. Shaded rows indicate that correlations are statistically significant in the predicted direction. As shown, we fail to observe much consistency in component-level correlations.

Table 5 parallels the results presented in table 4, instead using *IPTV* (a volume-based intraperiod timeliness measure) to proxy for how quickly new information is traded on. As before, panel A presents results using *IPTV* measured over a five-minute window. Our results provide some support for H1 in the five-minute window, though only for 8-Ks. Specifically, we find highly significant, positive coefficients on both *CompScript* ($t = 5.06$) and *IdentifyData* ($t = 6.10$). In panel B, we continue to find strong support for H1 in the form 8-K sample. All three scriptability measures relate positively to *IPTV* (t-statistics between 2.62 and 6.48). We find similar evidence for form 8-Ks in panel C: *CompScript* ($t = 2.24$) and *IdentifyData* ($t = 2.86$) each relate positively to *IPTV* in the 24-hour window. Similar to findings in panel B of table 4, we once again observe relations opposite those predicted in the form 10-Q sample in panels B and C. Namely, *CompScript* and *DataToInformation* relate negatively to *IPTV* in panels B and C. Since trading volume is often used as a proxy for investor disagreement (Bamber [1986]), this result is consistent with scriptability affecting different users differently.

While we find fairly consistent support for H1 using *IPTV* for 8-K filings, we fail to find that our summary measures of scriptability affect the speed of investors' response to other types of filings. One potential explanation for this is that scriptability components that affect the speed of 8-K response differ from those that facilitate rapid price formation following 10-K, 10-Q, or proxy statement filings. To investigate this possibility, we present an unplanned, descriptive analysis in panel D of table 5 similar to panel D in table 4. Shaded cells indicate components that exhibit the predicted signed correlation. We observe numerous correlation patterns consistent with our results, and the pattern of shading suggests that scriptability is not a

TABLE 5

Scriptability and Speed of Price Formation (IPTV)

Panel A: Volume IPT for five-minute window

Form Type:	10-K		10-Q		8-K		DEF 14A	
Dependent Variable:	$IPTV^{5m}$	$IPTV^{5m}$	$IPTV^{5m}$	$IPTV^{5m}$	$IPTV^{5m}$	$IPTV^{5m}$	$IPTV^{5m}$	$IPTV^{5m}$
CompScript	0.107		−0.386		0.577***		−0.718	
	(0.24)		(−1.54)		(5.06)		(−0.81)	
IdentifyData		0.309		−0.323		0.780***		0.128
		(0.80)		(−1.36)		(6.10)		(0.26)
DataToInformation		−0.182		−0.106		−0.005		−0.722
		(−0.40)		(−0.43)		(−0.06)		(−1.13)
Follow	−0.241***	−0.240***	−0.032	−0.032	0.035	0.035	−0.191*	−0.195*
	(−4.96)	(−4.98)	(−0.69)	(−0.69)	(0.74)	(0.73)	(−1.90)	(−1.91)
ROA	−0.225	−0.207	1.000**	0.992**	1.332**	1.349**	−0.451	−0.464
	(−0.17)	(−0.16)	(2.36)	(2.35)	(2.22)	(2.25)	(−0.34)	(−0.34)
News	−0.131	−0.133	−0.012	−0.011	0.009	0.007	−0.119	−0.121
	(−1.53)	(−1.58)	(−0.37)	(−0.33)	(0.32)	(0.23)	(−1.16)	(−1.18)
BTM	−0.074	−0.076	−0.164***	−0.162***	−0.091**	−0.090**	−0.007	−0.007
	(−0.64)	(−0.66)	(−2.83)	(−2.88)	(−2.09)	(−2.06)	(−0.08)	(−0.08)
Loss	−0.131	−0.132	0.034	0.035	0.133***	0.132***	−0.240*	−0.244*
	(−0.99)	(−0.99)	(0.62)	(0.63)	(2.69)	(2.68)	(−1.88)	(−1.90)
CapEx	0.147	0.158	0.169	0.166	−1.963***	−1.981***	0.323	0.338
	(0.14)	(0.15)	(0.20)	(0.19)	(−5.01)	(−5.06)	(0.36)	(0.38)
Lev	−0.473	−0.458	−0.528**	−0.531**	0.025	0.032	−0.254	−0.242
	(−1.06)	(−1.02)	(−2.48)	(−2.44)	(0.17)	(0.23)	(−1.01)	(−0.97)
FileCluster	0.119***	0.119***	0.134***	0.134***	0.102***	0.102***	0.138***	0.137***
	(15.87)	(15.74)	(16.13)	(16.13)	(17.14)	(17.18)	(13.65)	(13.27)
Size	−0.106*	−0.109*	−0.219***	−0.218***	−0.124***	−0.124***	−0.098	−0.099
	(−1.70)	(−1.73)	(−4.87)	(−4.93)	(−4.49)	(−4.47)	(−1.44)	(−1.47)

(*Continued*)

TABLE 5—*Continued*

Panel A: Volume IPT for five-minute window

Form Type:	10-K		10-Q		8-K		DEF 14A	
Dependent Variable:	$IPTV^{5m}$	$IPTV^{5m}$	$IPTV^{5m}$	$IPTV^{5m}$	$IPTV^{5m}$	$IPTV^{5m}$	$IPTV^{5m}$	$IPTV^{5m}$
Age	0.016	0.017	−0.020	−0.021	−0.086***	−0.087***	−0.003	−0.001
	(0.24)	(0.26)	(−0.86)	(−0.95)	(−3.56)	(−3.57)	(−0.04)	(−0.01)
Distress	0.003	0.001	0.006	0.006	0.010	0.009	0.002	0.001
	(0.10)	(0.03)	(0.50)	(0.52)	(0.91)	(0.79)	(0.08)	(0.05)
OpComplex	−0.146*	−0.147*	0.014	0.015	−0.001	0.001	−0.068	−0.067
	(−1.67)	(−1.67)	(0.54)	(0.55)	(−0.04)	(0.03)	(−1.39)	(−1.39)
FinComplex	0.465	0.434	−1.361***	−1.354***	−0.639***	−0.644***	−1.045	−1.065
	(1.02)	(0.95)	(−3.02)	(−3.03)	(−2.76)	(−2.81)	(−1.07)	(−1.09)
Growth	−0.055	−0.056	−0.006	−0.006	0.065	0.064	−0.161	−0.158
	(−0.38)	(−0.39)	(−0.10)	(−0.09)	(1.63)	(1.62)	(−1.17)	(−1.16)
%IH	−0.570***	−0.570***	−0.578***	−0.579***	−0.734***	−0.738***	−0.420***	−0.413***
	(−3.25)	(−3.24)	(−3.86)	(−3.88)	(−8.19)	(−8.25)	(−3.14)	(−3.09)
HTI	−0.220	−0.220	−0.166*	−0.166*	0.057	0.057	−0.334**	−0.328**
	(−0.99)	(−0.98)	(−1.91)	(−1.91)	(0.55)	(0.55)	(−2.34)	(−2.27)
FileLag	0.012***	0.012***	0.012***	0.012***	0.014***	0.014***	0.023***	0.023***
	(3.30)	(3.13)	(3.81)	(3.80)	(5.82)	(5.80)	(4.70)	(4.69)
Top25FA	−0.026	−0.019	−0.029	−0.031	0.030	0.025	0.109	0.121
	(−0.27)	(−0.19)	(−0.36)	(−0.40)	(0.72)	(0.59)	(1.05)	(1.19)
Watermark	0.129*	0.129*	0.021	0.018	0.045	0.055	−0.111	−0.105
	(1.79)	(1.74)	(0.36)	(0.30)	(0.96)	(1.15)	(−0.83)	(−0.78)
HTML	−0.033	−0.093	−0.776**	−0.758**	0.152***	0.103***	−0.189	−0.253
	(−0.06)	(−0.14)	(−2.57)	(−2.41)	(3.83)	(2.62)	(−1.02)	(−1.10)
Year-month FEs?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FEs?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,533	4,533	16,188	16,188	57,383	57,383	4,697	4,697
Adjusted-R^2	0.108	0.108	0.108	0.108	0.073	0.074	0.092	0.092

(*Continued*)

TABLE 5—*Continued*

Panel B: Volume IPT for 60-minute window

Form Type:	10-K		10-Q		8-K		DEF 14A	
Dependent Variable:	$IPTV^{60m}$	$IPTV^{60m}$	$IPTV^{60m}$	$IPTV^{60m}$	$IPTV^{60m}$	$IPTV^{60m}$	$IPTV^{60m}$	$IPTV^{60m}$
CompScript	−0.390		−0.825**		0.642***		−0.258	
	(−0.84)		(−2.52)		(6.48)		(−0.88)	
IdentifyData		0.311		−0.049		0.315***		0.082
		(0.91)		(−0.17)		(2.62)		(0.23)
DataToInformation		−0.679		−0.651***		0.325***		−0.287
		(−1.07)		(−4.97)		(3.93)		(−0.95)
Follow	−0.215***	−0.213***	−0.004	−0.004	−0.027	−0.027	−0.104**	−0.105**
	(−5.76)	(−5.61)	(−0.11)	(−0.10)	(−0.86)	(−0.86)	(−2.54)	(−2.57)
ROA	−0.082	−0.039	2.638***	2.655***	0.375	0.375	2.185*	2.180*
	(−0.11)	(−0.05)	(5.03)	(5.09)	(1.10)	(1.10)	(1.94)	(1.94)
News	−0.118***	−0.122***	−0.050	−0.052	−0.062***	−0.062***	−0.155*	−0.156*
	(−3.45)	(−3.59)	(−1.17)	(−1.22)	(−3.16)	(−3.12)	(−1.86)	(−1.85)
BTM	−0.001	−0.003	−0.106***	−0.109***	−0.032	−0.032	0.093	0.092
	(−0.06)	(−0.15)	(−2.80)	(−2.91)	(−1.01)	(−1.01)	(1.36)	(1.36)
Loss	−0.087	−0.090	0.019	0.017	0.029	0.029	−0.011	−0.012
	(−0.63)	(−0.64)	(0.26)	(0.22)	(0.71)	(0.71)	(−0.11)	(−0.12)
CapEx	0.761	0.776	−0.198	−0.188	−0.836***	−0.836***	−0.753	−0.750
	(0.75)	(0.77)	(−0.50)	(−0.48)	(−3.10)	(−3.10)	(−0.96)	(−0.96)
Lev	−0.497	−0.468	−0.197*	−0.189	0.099	0.099	−0.124	−0.119
	(−1.40)	(−1.33)	(−1.66)	(−1.64)	(1.07)	(1.07)	(−0.78)	(−0.73)
FileCluster	0.121***	0.121***	0.101***	0.101***	0.082***	0.082***	0.130***	0.129***
	(14.06)	(13.88)	(15.00)	(15.00)	(15.46)	(15.59)	(19.02)	(18.79)
Size	−0.042*	−0.050**	−0.115***	−0.118***	−0.065***	−0.065***	0.006	0.005
	(−1.88)	(−1.99)	(−2.90)	(−3.05)	(−3.19)	(−3.20)	(0.13)	(0.11)

(*Continued*)

TABLE 5—*Continued*

Panel B: Volume IPT for 60-minute window

Form Type:	10-K		10-Q		8-K		DEF 14A	
Dependent Variable:	$IPTV^{60m}$	$IPTV^{60m}$	$IPTV^{60m}$	$IPTV^{60m}$	$IPTV^{60m}$	$IPTV^{60m}$	$IPTV^{60m}$	$IPTV^{60m}$
Age	−0.075*	−0.073**	−0.078***	−0.074***	−0.086***	−0.086***	−0.209***	−0.208***
	(−1.92)	(−1.97)	(−3.08)	(−2.84)	(−3.51)	(−3.49)	(−3.47)	(−3.36)
Distress	0.007	0.003	0.022*	0.021*	0.008	0.008	0.025	0.025
	(0.27)	(0.14)	(1.74)	(1.69)	(1.12)	(1.11)	(1.13)	(1.12)
OpComplex	0.038*	0.036	−0.020	−0.022	0.030	0.030	−0.037	−0.037
	(1.69)	(1.52)	(−0.59)	(−0.66)	(1.41)	(1.41)	(−0.69)	(−0.70)
FinComplex	0.012	−0.060	−0.742**	−0.758**	0.152	0.152	0.471	0.459
	(0.02)	(−0.11)	(−2.06)	(−2.10)	(0.57)	(0.58)	(0.48)	(0.47)
Growth	0.056	0.054	0.041	0.040	0.043**	0.043**	0.125***	0.126***
	(0.60)	(0.59)	(0.94)	(0.91)	(2.10)	(2.10)	(3.54)	(3.75)
%IH	−0.285**	−0.283**	−0.157*	−0.156*	−0.248***	−0.248***	−0.222	−0.220
	(−2.13)	(−2.14)	(−1.77)	(−1.77)	(−2.76)	(−2.77)	(−1.29)	(−1.27)
HTI	0.017	0.011	−0.031	−0.031	−0.057	−0.057	−0.174	−0.172
	(0.09)	(0.06)	(−0.25)	(−0.24)	(−0.60)	(−0.59)	(−1.01)	(−1.00)
FileLag	0.009***	0.009***	0.015***	0.014***	0.007***	0.007***	0.004*	0.003*
	(3.59)	(3.19)	(4.30)	(4.30)	(4.86)	(4.86)	(1.96)	(1.90)
Top25FA	0.194*	0.209*	−0.077	−0.070	0.017	0.017	−0.082	−0.077
	(1.86)	(1.91)	(−1.55)	(−1.39)	(0.38)	(0.38)	(−1.04)	(−1.05)
Watermark	0.189*	0.189*	−0.016	−0.007	−0.035	−0.035	−0.224**	−0.220**
	(1.89)	(1.89)	(−0.43)	(−0.17)	(−0.86)	(−0.86)	(−2.55)	(−2.57)
HTML	−0.205	−0.320	−0.393	−0.437	0.046	0.047	0.046	0.020
	(−0.54)	(−0.80)	(−1.35)	(−1.53)	(1.13)	(1.14)	(0.23)	(0.12)
Year-month fixed effects?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FEs?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,577	7,577	25,707	25,707	105,216	105,216	8,246	8,246
Adjusted-R^2	0.087	0.087	0.066	0.066	0.037	0.037	0.067	0.067

(*Continued*)

TABLE 5—*Continued*

Panel C: Volume IPT for 24-hour window

Form Type:	10–K		10–Q		8-K		DEF 14A	
Dependent Variable:	$IPTV^{24h}$	$IPTV^{24h}$	$IPTV^{24h}$	$IPTV^{24h}$	$IPTV^{24h}$	$IPTV^{24h}$	$IPTV^{24h}$	$IPTV^{24h}$
CompScript	−0.048		−0.262**		0.357**		−0.616	
	(−0.18)		(−2.01)		(2.24)		(−1.04)	
IdentifyData		0.057		0.113		0.218***		−0.382
		(0.13)		(0.54)		(2.86)		(−0.84)
DataToInformation		−0.100		−0.292**		0.156		−0.252
		(−0.33)		(−1.97)		(1.17)		(−0.70)
Follow	−0.005	−0.005	−0.017	−0.016	0.019	0.019	−0.006	−0.006
	(−0.16)	(−0.15)	(−0.47)	(−0.45)	(1.16)	(1.16)	(−0.22)	(−0.20)
ROA	−1.361***	−1.356***	0.492	0.510	0.193	0.193	−0.183	−0.181
	(−2.96)	(−3.04)	(1.31)	(1.33)	(0.80)	(0.80)	(−0.33)	(−0.32)
News	−0.102***	−0.102**	−0.030	−0.032	−0.075***	−0.075***	−0.062	−0.062
	(−2.59)	(−2.47)	(−1.53)	(−1.56)	(−2.63)	(−2.63)	(−0.73)	(−0.73)
BTM	0.156***	0.155***	0.029	0.027	0.015	0.015	0.080	0.080
	(4.34)	(4.37)	(1.25)	(1.18)	(0.78)	(0.78)	(1.42)	(1.42)
Loss	−0.129*	−0.129*	−0.036	−0.037	0.057*	0.057*	0.012	0.013
	(−1.71)	(−1.68)	(−0.74)	(−0.78)	(1.70)	(1.70)	(0.19)	(0.20)
CapEx	0.470	0.471	−0.127	−0.120	−0.061	−0.062	−0.150	−0.152
	(0.63)	(0.63)	(−0.31)	(−0.29)	(−0.20)	(−0.20)	(−0.18)	(−0.18)
Lev	−0.111	−0.108	−0.166	−0.163	−0.048	−0.048	0.031	0.030
	(−0.62)	(−0.60)	(−1.60)	(−1.54)	(−0.80)	(−0.79)	(0.09)	(0.09)
FileCluster	−0.005	−0.005	−0.021***	−0.021***	−0.025***	−0.025***	−0.004	−0.004
	(−1.46)	(−1.46)	(−7.00)	(−7.01)	(−18.07)	(−17.43)	(−0.39)	(−0.38)
Size	−0.047**	−0.048*	−0.047*	−0.050**	0.028**	0.028**	0.012	0.012
	(−2.00)	(−1.87)	(−1.90)	(−1.97)	(2.10)	(2.08)	(0.50)	(0.51)

(*Continued*)

TABLE 5—*Continued*

Panel C: Volume IPT for 24-hour window								
Form Type:	10−K		10−Q		8-K		DEF 14A	
Dependent Variable:	$IPTV^{24h}$	$IPTV^{24h}$	$IPTV^{24h}$	$IPTV^{24h}$	$IPTV^{24h}$	$IPTV^{24h}$	$IPTV^{24h}$	$IPTV^{24h}$
Age	−0.025	−0.025	−0.184***	−0.182***	−0.112***	−0.112***	−0.131***	−0.131***
	(−0.45)	(−0.44)	(−8.52)	(−8.17)	(−8.03)	(−8.07)	(−3.69)	(−3.71)
Distress	−0.022	−0.023	−0.004	−0.004	0.004	0.004	0.001	0.001
	(−1.49)	(−1.59)	(−0.40)	(−0.48)	(1.06)	(1.06)	(0.05)	(0.05)
OpComplex	−0.004	−0.005	0.075***	0.074***	0.015	0.016	0.031	0.031
	(−0.08)	(−0.09)	(2.89)	(2.80)	(1.20)	(1.19)	(0.68)	(0.68)
FinComplex	−0.259	−0.269	−0.743**	−0.752**	−0.088	−0.090	−0.125	−0.120
	(−0.65)	(−0.74)	(−2.34)	(−2.40)	(−0.42)	(−0.43)	(−0.33)	(−0.31)
Growth	0.132*	0.132*	0.064**	0.063**	0.040*	0.040*	0.093	0.092
	(1.80)	(1.80)	(2.15)	(2.13)	(1.79)	(1.79)	(1.36)	(1.35)
%IH	−0.328***	−0.327***	−0.151	−0.150	−0.332***	−0.332***	−0.121	−0.122
	(−3.07)	(−3.08)	(−1.37)	(−1.36)	(−7.78)	(−7.79)	(−1.09)	(−1.09)
HTI	0.021	0.020	0.048	0.047	0.166***	0.165***	−0.071	−0.071
	(0.16)	(0.15)	(0.49)	(0.48)	(3.12)	(3.12)	(−0.38)	(−0.39)
FileLag	−0.001	−0.001	−0.000	−0.000	0.004***	0.004***	0.005	0.005
	(−0.27)	(−0.29)	(−0.12)	(−0.15)	(3.96)	(3.96)	(1.23)	(1.25)
Top25FA	0.042	0.044	−0.077**	−0.074**	−0.014	−0.015	0.025	0.023
	(0.57)	(0.64)	(−2.30)	(−2.18)	(−0.51)	(−0.52)	(0.27)	(0.26)
Watermark	−0.006	−0.006	−0.064	−0.059	−0.023	−0.022	−0.081	−0.083
	(−0.08)	(−0.09)	(−1.56)	(−1.43)	(−1.24)	(−1.15)	(−0.92)	(−0.91)
HTML	−0.229***	−0.247	−0.202**	−0.231***	0.270***	0.266***	0.154	0.164
	(−4.02)	(.)	(−2.29)	(−3.09)	(5.67)	(5.49)	(0.93)	(0.95)
Year-month FEs?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FEs?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,393	12,393	38,736	38,736	168,415	168,415	12,923	12,923
Adjusted-R^2	0.012	0.011	0.011	0.011	0.011	0.011	0.007	0.007

(*Continued*)

TABLE 5—*Continued*

Panel D: Unplanned analysis: Correlations between scriptability components and IPTV measures

	10-K			10-Q			8-K			DEF 14A		
Component	IPTV5	IPTV60	IPTV24	IPTV5	IPTV60	IPTV24	IPTV5	IPTV60	IPTV24	IPTV5	IPTV60	IPTV24
TextFromTables	**0.051**	**0.024**	0.013	−0.014	**−0.021**	0.008	**0.033**	**0.009**	**0.018**	**0.035**	0.017	**0.018**
SeparateSections	−0.001	0.012	−0.012	0.004	0.002	0.001	**0.012**	0.006	**0.005**	−0.002	0.005	−0.010
UsefulHeadings	**0.036**	0.007	0.005	**0.036**	**0.020**	0.000	**0.028**	**0.009**	**0.012**	**0.041**	0.003	0.005
ExtReferences	**0.070**	**0.061**	0.015	0.002	**0.018**	**0.015**	**−0.019**	**−0.009**	**−0.016**	**0.050**	**0.024**	0.012
BinaryExhibits	**0.080**	**0.067**	0.017	**0.070**	**0.043**	**0.011**	0.005	**0.031**	**−0.007**	**0.046**	**0.028**	0.007
%NumTabulate	**−0.051**	**−0.048**	−0.011	**−0.048**	**−0.057**	**−0.011**	**0.082**	**0.031**	**0.076**	**−0.101**	**−0.077**	**−0.031**
UnexpectChar	0.004	−0.001	**0.022**	**0.016**	0.003	0.006	**−0.047**	**−0.016**	**−0.040**	**0.085**	**0.039**	**0.030**
TokenQuality	**0.038**	**0.037**	0.016	**0.042**	**0.022**	**0.015**	0.007	**0.022**	**−0.014**	−0.003	0.004	0.004
Typo	0.005	0.018	0.014	**0.039**	**0.013**	**0.017**	**−0.039**	**−0.008**	**−0.029**	0.005	−0.004	0.001
TableMarkup	**0.030**	0.013	0.013	0.007	0.007	−0.006	**0.040**	**0.019**	**0.036**	−0.015	0.005	0.000
CellConsistency	−0.005	**−0.026**	−0.017	0.005	0.001	0.003	**−0.050**	**−0.009**	**−0.027**	**−0.036**	−0.021	−0.015
UsefulRowNames	**0.070**	**0.042**	0.010	**0.063**	**0.030**	0.009	**0.093**	**0.031**	**0.054**	**0.081**	**0.045**	**0.023**
Average across components	0.027	0.017	0.007	0.019	0.007	0.006	0.012	0.010	0.006	0.015	0.006	0.004

This table presents results from estimating equation (3) with *IPTV* as the dependent variable. Panel A (B and C) presents results using the *IPTV* measure computed over the five-minute (60-minute and 24-hour) window. We estimate equation (3) by form type in all panels. All variables are defined in the appendix. Industries are defined using the Fama-French 48-industry designations. Reported *t*-statistics in parentheses are derived from standard errors clustered by firm and year. ***, **, and * denote two-tailed significance at the $p < 0.01$, $p < 0.05$, and $p < 0.10$ level, respectively. Note that the standard error for HTML cannot be estimated in column 2 of panel C, likely due to a combination of clustering standard errors by both firm and year and a lack of variation in the HTML indicator during the years of our sample period. When we cluster our standard errors by firm only, the *t*-statistics is −0.91. Panel D presents Spearman correlation coefficients between scriptability components and IPTV measures. Bolded correlations are significant at the $p < 0.05$ level, and shaded cells in the table denote a significant correlation consistent with H1. Note that the analysis in panel D is unplanned and not part of the original registered report proposal.

"one-size-fits-all" construct. For instance, *ExtReferences* and *BinaryExhibits* appear to correlate with *IPTV* for 10-Ks, 10-Qs, and proxy statements, though not for 8-Ks. Conversely, *%NumTabulate* correlates in the expected direction only for 8-Ks and relates negatively to *IPTV* for the other form types. In fact, the only components that exhibit positive associations with *IPTV* across all form types are *UsefulHeadings* and *UsefulRowNames*. These findings are consistent with our discussion in section 2 that scripters perform different information identification and processing procedures in different settings.

To summarize, we find some support for H1, though the relation between scriptability and the speed of the market response appears to vary considerably depending on form type and measurement window. Our results are most pronounced with 8-Ks, which is perhaps not surprising since these filings often convey forward-looking, voluntarily disclosed information that has not already been priced prior to filing (Lerman and Livnat [2010]). However, our subsequent analysis in panel D of table 5 suggests that some components impact the speed of information acquisition in other (non-8-K) filings. We also generally find stronger results using volume to assess the market's response. Given that measures of volume IPT are less volatile (i.e., less impacted by small denominators) and volume is the *sum* of all individual investors' (or scripters') trades, or actions (whereas price changes reflect changes in the aggregate market's *average* beliefs), we think this pattern is intuitive.

4.2 RESULTS FOR H2

Table 6 presents results for H2. Panel A presents results using *CompScript*, and panel B presents results using our task-based measures *IdentifyData* and *DataToInformation*. We expect each measure of scriptability to relate positively to *Dissem* (the likelihood of DJ dissemination increases with scriptability) and negatively with *NewsDelay* (the time from SEC acceptance to dissemination decreases as scriptability increases). As shown in panel A, we find very little support for H2 using *CompScript*. For 8-K filings, we do observe a positive coefficient with *Dissem* as the dependent variable ($t = 2.66$). In other columns, we either fail to observe a significant relation between *CompScript* and our dissemination variables, or, in three cases, significant in the direction opposite our prediction (columns 3, 6, and 7). In panel B, we again find some support for H2 for Form 8-Ks. *IdentifyData* relates positively (negatively) to *Dissem* (*NewsDelay*), as predicted. However, for Form 8-Ks, *DataToInformation* exhibits a significant association opposite our predictions, relating negatively (positively) to *Dissem* (*NewsDelay*). We also observe some evidence contrary to our predictions for 10-Qs. *DataToInformation* exhibits a weakly negative relation with *Dissem* and *IdentifyData* relates positively to *NewsDelay*.

We note that scriptability might not predict dissemination likelihood or timeliness to the extent that the process of creating DJ newswires is manual instead of computerized. An additional possible explanation for our results

TABLE 6

Scriptability and Dow Jones News Dissemination

Panel A: Composite scriptability

Form Type:	10-K		10-Q		8-K		DEF 14A	
Dependent Variable:	*Dissem*	*NewsDelay*	*Dissem*	*NewsDelay*	*Dissem*	*NewsDelay*	*Dissem*	*NewsDelay*
CompScript	0.492	−134.793	−1.401**	241.823	0.422***	71.962***	−2.198**	−61.639
	(0.67)	(−0.57)	(−2.17)	(1.19)	(2.66)	(6.61)	(−2.57)	(−0.26)
Opinion	0.441***	55.633	0.171	−33.349	−0.010	3.712	−0.029	3.504
	(3.60)	(1.47)	(1.62)	(−1.42)	(−0.41)	(1.29)	(−0.14)	(0.06)
Volume	0.026	−11.544	−0.061	8.685	0.007	2.670**	0.048	−14.059
	(0.39)	(−0.82)	(−0.98)	(0.83)	(0.40)	(2.09)	(0.41)	(−0.99)
#EAs	0.001**	0.095	0.000	0.013	0.001***	−0.038***	−0.000	0.148
	(2.37)	(1.28)	(1.17)	(0.26)	(32.08)	(−8.04)	(−0.46)	(0.72)
Lit	−0.656	241.900***	−0.132	118.837**	0.111*	−5.526	0.273	−0.955
	(−1.60)	(6.71)	(−0.49)	(2.19)	(1.80)	(−0.89)	(1.15)	(−0.01)
Volatility	0.576***	17.116	0.531***	42.212	−0.442***	6.632	0.296	−233.763
	(3.61)	(0.25)	(2.65)	(0.59)	(−6.99)	(1.40)	(0.51)	(−1.44)
#Filings	−0.259***	3.572	−0.196***	18.583***	−0.331***	21.711***	−0.065	2.477
	(−4.20)	(0.29)	(−3.14)	(2.66)	(−24.23)	(15.62)	(−0.85)	(0.08)
News	0.338***	37.657	0.344***	21.381	0.430***	19.329***	1.207***	83.283**
	(6.07)	(1.40)	(7.41)	(1.15)	(10.36)	(11.00)	(5.19)	(2.55)
DJLength		23.999**		2.192		2.866***		28.964*
		(2.08)		(0.26)		(3.85)		(1.85)
Size	0.066	15.843	0.073	7.186	−0.070***	−2.903*	0.140	−25.991
	(1.03)	(1.02)	(1.07)	(0.50)	(−4.45)	(−1.74)	(1.10)	(−0.77)
Age	0.081	−12.854	0.105	−7.197	−0.085***	0.121	0.072	−0.052
	(0.64)	(−0.99)	(1.49)	(−0.80)	(−4.89)	(0.06)	(0.56)	(−0.00)

(*Continued*)

TABLE 6—*Continued*

Panel A: Composite scriptability

Form Type:	10-K		10-Q		8-K		DEF 14A	
Dependent Variable:	*Dissem*	*NewsDelay*	*Dissem*	*NewsDelay*	*Dissem*	*NewsDelay*	*Dissem*	*NewsDelay*
Distress	0.015	0.684	0.048**	0.837	−0.011**	0.017	0.029	8.065
	(0.63)	(0.19)	(2.00)	(0.14)	(−2.10)	(0.06)	(0.59)	(1.14)
OpComplex	0.023	1.184	−0.031	23.639	−0.050***	5.385***	0.152	−78.137***
	(0.23)	(0.07)	(−0.41)	(1.00)	(−2.69)	(3.60)	(1.54)	(−3.93)
FinComplex	1.825**	−71.371	0.625	−122.650	−0.204	25.210	−0.542	311.871
	(2.19)	(−0.31)	(0.41)	(−0.88)	(−1.07)	(1.41)	(−0.30)	(0.89)
Growth	−0.030	21.457	0.021	21.393	−0.043*	1.382	−0.338	−7.736
	(−0.22)	(0.70)	(0.31)	(0.54)	(−1.68)	(0.49)	(−1.14)	(−0.22)
%IH	−0.328*	19.884	−0.264	32.383**	0.260***	−5.757	−0.670**	208.230**
	(−1.86)	(0.42)	(−1.34)	(2.17)	(5.96)	(−1.31)	(−2.37)	(2.31)
HTI	0.311	−168.943**	0.001	−87.714**	0.074	11.798**	0.050	−261.423
	(0.80)	(−2.14)	(0.00)	(−2.08)	(1.13)	(2.55)	(0.11)	(−1.26)
FileLag	0.004	2.121	0.007	3.380**	−0.002	0.033	−0.002	4.153
	(0.72)	(0.64)	(1.12)	(2.06)	(−1.35)	(0.31)	(−0.20)	(1.41)
Top25FA	0.099	29.450	0.099	−61.669	−0.132***	6.270	−0.198	33.935
	(0.47)	(0.37)	(0.90)	(−1.36)	(−3.36)	(1.51)	(−1.41)	(0.34)
Watermark	0.051	11.869	0.258*	−45.114	0.010	−1.794	0.049	83.599
	(0.32)	(0.18)	(1.69)	(−1.18)	(0.28)	(−0.65)	(0.37)	(0.81)
HTML	−0.712	−28.719	−1.109**	−260.491***	1.247***	−123.351***	0.245	21.483
	(−1.58)	(−0.38)	(−2.37)	(−2.72)	(22.82)	(−18.97)	(0.52)	(0.37)
Year-month FEs?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FEs?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hour FEs?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Day of week FEs?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	15,263	622	49,268	1,046	212,420	87,643	12,929	308
Pseudo-R^2/adjusted-R^2	0.090	–0.008	0.065	0.072	0.162	0.059	0.272	0.177

(*Continued*)

TABLE 6—*Continued*

Panel B: Task-based measures

Form Type:	10-K		10-Q		8-K		DEF 14A	
Dependent Variable:	*Dissem*	*NewsDelay*	*Dissem*	*NewsDelay*	*Dissem*	*NewsDelay*	*Dissem*	*NewsDelay*
IdentifyData	0.128	−137.772	−0.603	411.476***	0.970***	−36.415***	−0.651	−158.674
	(0.21)	(−0.93)	(−0.85)	(2.99)	(8.26)	(−4.19)	(−0.79)	(−0.52)
DataToInformation	0.358	−5.024	−0.766*	−81.372	−0.226**	85.758***	−1.374	67.784
	(0.68)	(−0.04)	(−1.76)	(−0.83)	(−2.36)	(14.12)	(−1.55)	(0.56)
Opinion	0.442***	55.619	0.171	−35.322	−0.010	3.859	−0.031	2.964
	(3.59)	(1.45)	(1.62)	(−1.56)	(−0.45)	(1.28)	(−0.16)	(0.05)
Volume	0.027	−11.918	−0.061	6.864	0.004	2.981**	0.048	−14.624
	(0.40)	(−0.85)	(−0.97)	(0.63)	(0.23)	(2.36)	(0.41)	(−0.99)
#EAs	0.001**	0.101	0.000	0.007	0.001***	−0.038***	−0.000	0.161
	(2.39)	(1.38)	(1.16)	(0.15)	(31.65)	(−8.08)	(−0.52)	(0.72)
Lit	−0.653	245.891***	−0.133	100.646*	0.114*	−6.420	0.272	3.671
	(−1.61)	(6.82)	(−0.49)	(1.78)	(1.85)	(−1.04)	(1.15)	(0.04)
Volatility	0.577***	15.977	0.530***	50.312	−0.440***	6.682	0.299	−240.697
	(3.66)	(0.23)	(2.67)	(0.73)	(−7.02)	(1.43)	(0.51)	(−1.45)
#Filings	−0.259***	3.211	−0.196***	19.466***	−0.324***	21.209***	−0.063	2.515
	(−4.19)	(0.27)	(−3.15)	(2.96)	(−23.58)	(14.89)	(−0.83)	(0.08)
News	0.338***	38.701	0.344***	20.852	0.430***	19.597***	1.203***	83.990**
	(6.11)	(1.44)	(7.40)	(1.17)	(10.32)	(11.24)	(5.19)	(2.51)
DJLength		24.397**		3.453		3.044***		28.499*
		(2.05)		(0.46)		(4.08)		(1.91)
Size	0.067	16.487	0.072	8.275	−0.069***	−3.041*	0.138	−24.702
	(1.05)	(1.03)	(1.07)	(0.58)	(−4.43)	(−1.85)	(1.10)	(−0.70)
Age	0.080	−12.589	0.106	−10.473	−0.088***	0.420	0.074	1.746
	(0.64)	(−0.94)	(1.50)	(−1.06)	(−5.04)	(0.21)	(0.58)	(0.05)

(*Continued*)

TABLE 6—*Continued*

Panel B: Task-based measures								
Form Type:	10-K		10-Q		8-K		DEF 14A	
Dependent Variable:	*Dissem*	*NewsDelay*	*Dissem*	*NewsDelay*	*Dissem*	*NewsDelay*	*Dissem*	*NewsDelay*
Distress	0.016	1.142	0.048**	0.535	−0.012**	0.111	0.030	7.383
	(0.65)	(0.33)	(1.98)	(0.08)	(−2.24)	(0.37)	(0.60)	(1.07)
OpComplex	0.024	1.956	−0.032	20.215	−0.047**	5.140***	0.153	−78.804***
	(0.24)	(0.11)	(−0.42)	(0.86)	(−2.56)	(3.46)	(1.53)	(−4.25)
FinComplex	1.832**	−69.012	0.622	−161.981	−0.216	29.225	−0.534	318.409
	(2.19)	(−0.30)	(0.41)	(−1.16)	(−1.15)	(1.61)	(−0.30)	(0.89)
Growth	−0.030	21.501	0.021	19.189	−0.045*	1.479	−0.335	−8.902
	(−0.22)	(0.71)	(0.30)	(0.48)	(−1.74)	(0.52)	(−1.13)	(−0.24)
%IH	−0.329*	17.468	−0.264	27.685*	0.255***	−4.849	−0.670**	210.337**
	(−1.86)	(0.35)	(−1.34)	(1.81)	(5.77)	(−1.09)	(−2.38)	(2.31)
HTI	0.312	−171.156**	0.001	−75.842*	0.070	12.189***	0.055	−259.396
	(0.80)	(−2.15)	(0.00)	(−1.84)	(1.07)	(2.64)	(0.12)	(−1.24)
FileLag	0.005	2.175	0.007	3.147*	−0.002	0.034	−0.002	4.245
	(0.72)	(0.64)	(1.10)	(1.96)	(−1.32)	(0.32)	(−0.21)	(1.38)
Top25FA	0.095	30.332	0.101	−53.776	−0.140***	7.367*	−0.183	29.340
	(0.44)	(0.39)	(0.92)	(−1.18)	(−3.65)	(1.82)	(−1.56)	(0.31)
Watermark	0.051	12.700	0.260*	−39.864	0.026	−3.381	0.062	79.725
	(0.32)	(0.20)	(1.68)	(−1.04)	(0.77)	(−1.34)	(0.42)	(0.78)
HTML	−0.685	−8.002	−1.122**	−307.920***	1.179***	−115.855***	0.191	42.986
	(−1.38)	(−0.11)	(−2.46)	(−3.44)	(20.88)	(−18.46)	(0.50)	(0.70)
Year-month FEs?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FEs?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hour FEs?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Day of week FEs?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	15,263	622	49,268	1,046	212,420	87,643	12,929	308
Pseudo-R^2/adjusted-R^2	0.090	−0.009	0.065	0.081	0.164	0.060	0.272	0.174

(*Continued*)

TABLE 6—*Continued*

Panel C: Unplanned analysis: correlations between scriptability components and dissemination measures

Component	*Dissem*	*NewsDelay*
TextFromTables	**0.117**	**−0.105**
SeparateSections	**0.030**	**−0.017**
UsefulHeadings	**0.103**	**−0.032**
ExtReferences	**−0.094**	**−0.056**
BinaryExhibits	**−0.114**	**0.048**
%NumTabulate	**0.230**	**−0.054**
UnexpectChar	**−0.137**	**0.022**
TokenQuality	**−0.105**	**0.045**
Typo	**−0.121**	**0.017**
TableMarkup	**0.108**	**−0.022**
CellConsistency	**−0.182**	**0.009**
UsefulRowNames	**0.174**	−0.002
Average across components:	0.001	−0.012

This table presents results from estimating equation (4), estimated separately by form type. Panel A (B) presents results using the *CompScript* and panel B presents results using components of *CompScript*, *IdentifyData*, and *DataToInformation*. All variables are defined in the appendix. Industries are defined using the Fama-French 48-industry designations. Reported *t*-statistics in parentheses are derived from standard errors clustered by firm and year. ***, **, and * denote two-tailed significance at the $p < 0.01$, $p < 0.05$, and $p < 0.10$ level, respectively. Panel C presents the Spearman correlation coefficient between each component of scriptability and our two measures of news dissemination, *Dissem* and *NewsDelay*, for 8-K disclosures. Bolded correlations are significant at the $p < 0.05$ level, and shaded cells indicate a significant correlation consistent with predictions in H2. Note that the analysis in panel C is unplanned and not part of the original registered report proposal.

is that our tests lack power because DJ does not frequently disseminate SEC filings. Descriptive statistics presented earlier in panel A of table 2 provide some support for the latter. Namely, we only identify DJ dissemination of 3.8%, 2.1%, and 1.8% of Form 10-K, 10-Q, and DEF 14A disclosures, respectively. These low base rates limit the power of our tests, particularly those involving *NewsDelay*, and indicate that factors other than scriptability determine whether DJ disseminates these filings.

To shed additional light on our findings, we conduct an additional unplanned, descriptive analysis related to H2, similar to panel D of tables 4 and 5. Since we find that a relatively high percentage of 8-Ks, approximately 40%, are disseminated, we limit our unplanned analysis to this form group. Panel C of table 6 presents unconditional Spearman correlations between each component of scriptability and both *Dissem* and *NewsDelay*. As shown, several individual components exhibit significant correlations with each dependent variable, some consistent and others inconsistent with our expectations. Specifically, *TextFromTables, SeparateSections, UsefulHeadings, %NumTabulate, TableMarkup,* and *UsefulRowNames* each exhibit correlations in the predicted directions with both *Dissem* and *NewsDelay*. Conversely, *ExtReferences, BinaryExhibits, UnexpectChar, TokenQuality, Typo,* and *CellConsistency* generally correlate in directions contrary to our predictions. The resulting combination yields only weak mean correlations, reported in the bottom row of panel C, consistent with our lack of results in panels A and B. Predicting which specific components of scriptability affect the DJ dissemination process (or any process for that matter) requires intimate knowledge of the exact scripts used to monitor and process the SEC filings. Nonetheless, it is perhaps not surprising that features related to separating textual and numeric information (*TextFromTables, %NumTabulate*) and to identifying specific data in the filing (*UsefulHeadings, UsefulRowNames, TableMarkup*) exhibit relations in the expected directions, whereas measures of the "quality" or "readability" of the filing (*BinaryExhibits, TokenQuality, Typo*) exhibit unexpected relations.[21]

To summarize our results, we generally fail to find support for H2. However, our unplanned descriptive evidence suggests that certain features of filings could indeed relate to the likelihood and speed of DJ newswires dissemination of form 8-K disclosures. This result is consistent with the influence of scriptability being context specific.

[21] In untabulated, unplanned analyses, we construct an alternative summary measure of scriptability using the average of six components highlighted in panel C of table 6. We reestimate columns 5 and 6 of table 6, panel A, using this measure instead of *CompScript* and find highly significant coefficients in both columns ($p < 0.01$). Note that the only purpose of this analysis is to verify that directionally consistent correlations in panel C are not due to correlated omitted variables.

5. *Planned Additional Analyses*

5.1 XBRL AND SCRIPTABILITY

XBRL metadata is designed to make existing information easier to process without increasing the information content of regulatory filings (Securities and Exchange Commission [2005]). Among other benefits, XBRL allows "what is currently static, text-based information [to] be dynamically searched and analyzed, facilitating the comparison of financial and business performance across companies, reporting periods, and industries" (Securities and Exchange Commission [2009b]). Given the natural relation between programmatic analysis and "dynamic search and analysis," XBRL should make documents more scriptable. Despite its potential usefulness, however, research documents substantial errors in early XBRL filings, although recent evidence suggests some improvement in data quality over time (Debreceny et al. [2010], Du, Vasarhelyi, and Zheng [2013]). While prior studies cite noise in XBRL metadata as a weakness (Hoitash and Hoitash [2018]), we exploit variation in both the presence and quality of XBRL metadata to assess whether XBRL moderates the effects of disclosure scriptability by providing an alternative mechanism for extracting information.

In our first supplemental analysis, we test whether associations between the scriptability of filings and *IPTP*, *IPTV*, *Dissem*, and *NewsDelay* vary depending on whether the firm includes XBRL. We restrict our sample to 10-K filings, as they contain the largest concentration of XBRL metadata, filed between 2010 and 2012, the period where the SEC required only some filers to tag both the financial statements and the notes. We identify a filing as containing XBRL if it has the designated exhibits (i.e., EX-101) and estimate equations (3) and (4) separately for the XBRL and non-XBRL filings. Results from this analysis are presented in panel A of table 7. For brevity, we report only results using *CompScript* and do not tabulate coefficient estimates for variables other than *CompScript*. As shown in panel A of table 7, with one exception, we fail to observe significant differences across partitions in the relation between *CompScript* and each of our dependent variables. We do observe a significant difference using *NewsDelay* as the dependent variable in the direction of our expectations. However, this result is difficult to interpret given the fact that DJ appears to only rarely disseminate 10-K filings. In untabulated tests, we also analyze associations between XBRL and our task-based measures of scriptability (i.e., *IdentifyData* and *DataToInformation*). The relation between *DataToInformation* and *IPTP*[60] is significantly less positive (more negative) for the sample of XBRL filings, consistent with our expectations. However, consistent with reported results for *CompScript*, the association between *IdentifyData* and *NewsDelay* is marginally more negative in XBRL filings. In general, this evidence does not support our conjecture that XBRL substitutes for scriptability in 10-K filings.

TABLE 7

XBRL and Disclosure Scriptability

Panel A: Scriptability in non-XBRL 10-K filings

Dependent Variable:	$IPTP^{5m}$	$IPTP^{60m}$	$IPTP^{24h}$	$IPTV^{5m}$	$IPTV^{60m}$	$IPTV^{24h}$	*Dissem*	*NewsDelay*
CompScript (XBRL = 0)	−0.513	0.161	0.871	0.024	−0.224	−0.602	0.035	−641.110***
	(−0.45)	(0.21)	(1.28)	(0.02)	(−0.25)	(−0.86)	(1.01)	(−2.79)
CompScript (XBRL = 1)	−0.043	−1.339	1.338*	−0.112	−1.220	0.216	0.035	418.597
	(−0.04)	(−1.58)	(1.73)	(−0.10)	(−1.47)	(0.28)	(0.72)	(1.13)
Test of Difference	p = 0.77	p = 0.19	p = 0.65	p = 0.93	p = 0.42	p = 0.42	p = 0.99	p = 0.02
Controls and determinants?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-month FEs?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hour FEs?	No	No	No	No	No	No	Yes	Yes
Industry FEs?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Day of week FEs?	No	No	No	No	No	No	Yes	Yes
Observations (XBRL = 0/1)	1,146/1,533	2,183/2,473	3,247/3,141	896/1355	1,661/2,125	3,058/3,019	4,364/3,868	137/171
Adjusted-R^2/pseudo-R^2 (XBRL = 0/1)	−2.3%/*0.5%*	0.0%/*1.5%*	0.7%/*1.1%*	15.1%/*5.8%*	10.3%/*6.4%*	1.4%/*0.5%*	4.5%/*3.5%*	26.9%/*−1.0%*

(*Continued*)

TABLE 7—*Continued*

Panel B: Scriptability and XBRL quality in 10-K filings

	$IPTP^{5m}$	$IPTP^{60m}$	$IPTP^{24h}$	$IPTV^{5m}$	$IPTV^{60m}$	$IPTV^{24h}$	*Dissem*	*NewsDelay*
CompScript	23.521	−5.228	−12.658	23.188	2.554	0.998	0.067	5,178.097
	(1.57)	(−0.47)	(−1.42)	(1.56)	(0.21)	(0.10)	(0.12)	(0.83)
% Tagged	4.081	−2.799	0.894	5.724	1.462	1.793	−0.340**	122.998
	(0.98)	(−0.92)	(0.35)	(1.34)	(0.44)	(0.67)	(−2.06)	(0.07)
StandardTags	−0.565	1.800	1.442	1.292	2.337	−0.662	−0.224	237.355
	(−0.11)	(0.48)	(0.46)	(0.25)	(0.55)	(−0.22)	(−1.14)	(0.13)
TagQuality	8.942	−1.184	−8.293**	5.680	0.450	0.276	0.327	2,127.010
	(1.41)	(−0.26)	(−2.19)	(0.95)	(0.09)	(0.07)	(1.38)	(0.78)
CompScript × *% Tagged*	−8.086	5.314	−1.370	−13.746	−4.698	−3.748	0.608*	385.373
	(−0.95)	(0.86)	(−0.26)	(−1.56)	(−0.69)	(−0.69)	(1.93)	(0.11)
CompScript × *StandardTags*	−18.255	3.739	17.255**	−14.063	2.501	1.522	−0.823*	−5,636.499
	(−1.42)	(0.41)	(2.27)	(−1.14)	(0.25)	(0.19)	(−1.74)	(−1.05)
CompScript × *TagQuality*	0.057	−2.485	−2.951	−1.502	−3.638	0.241	0.398	−224.517
	(0.01)	(−0.34)	(−0.48)	(−0.14)	(−0.43)	(0.04)	(1.09)	(−0.06)
Controls and determinants?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-month FEs?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hour FEs	No	No	No	No	No	No	Yes	Yes
Industry FEs?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Day of week FEs?	No	No	No	No	No	No	Yes	Yes
Observations	2,426	4,031	6,119	2,212	3,640	5,974	7,581	267
Adjusted-R^2/pseudo-R^2	−0.009	0.010	0.016	0.115	0.090	0.018	0.026	−0.074

This table reports results from the planned supplemental analyses described in section 5. Panel A presents results from estimating equations (3) and (4) separately for 10-Ks filed between 2010 and 2012 including (*XBRL* = 1) and excluding (*XBRL* = 0) XBRL metadata. We report coefficients (*t*-statistics) for *CompScript* for each estimation and a test of difference. Panel B reports results from estimating equation (5) using 10-Ks filed between 2013 and 2015. We report coefficients (*t*-statistics) for variables of interest suppressing tabulation of all control variables. Reported *t*-statistics in parentheses are derived from standard errors clustered by firm. ***, **, and * denote two-tailed significance at the $p < 0.01$, $p < 0.05$, and $p < 0.10$ level, respectively.

In addition to examining whether the presence of XBRL moderates the effect of scriptability on market participants and news outlets, we also assess whether the extent and quality of XBRL metadata increases its value. To measure the extent of XBRL tagging, we divide the number of numeric tags identified in the XBRL metadata by the number of unique numbers in the document. The resulting measure, *%Tagged*, indicates the degree to which XBRL can substitute for data collection through scripting. To measure quality, we rely on data from the "EDGAR Dashboard" developed by xbrlcloud.com, which rates the quality of firms' XBRL metadata over a variety of syntactical and regulation-based dimensions (e.g., XBRL Technical Syntax Rules, Fundamental Accounting Concepts and Relations Rules, U.S. GAAP Reportability Rules, etc.).[22] For each dimension, the dashboard quantifies the percentage of "custom" tags, or tags that deviate from the standard taxonomy, and identifies errors or warnings associated with various XBRL rules. We collect ratings for 10-K filings between 2013 and 2015 and develop two measures of XBRL quality.[23] First, we set *StandardTags* equal to one minus the proportion of extended tags. Extended tags prevent users from relying on the standard taxonomy, making XBRL metadata harder to process. Second, we compute *TagQuality* as the proportion of quality dimensions classified as "OK" (i.e., no error or warnings), as we expect errors and warnings to reduce the usefulness of XBRL. We then estimate the following model:

$$\begin{aligned} DV = \alpha_0 &+ \alpha_1 CompScript + \alpha_2 \%Tagged + \alpha_3 StandardTags \\ &+ \alpha_4 TagQuality + \alpha_5 CompScript \times \%Tagged + \alpha_6 CompScript \\ &\times StandardTags + \alpha_7 CompScript \times TagQuality \\ &+ (\Sigma \gamma_k Controls_k) + \varepsilon. \end{aligned} \tag{5}$$

DV is *IPTP*, *IPTV*, *Dissem*, or *NewsDelay*. If XBRL quality substitutes for scriptability, we expect the three interaction terms to be negative for all specifications except when *NewsDelay* is the dependent variable. We employ the same controls used in equations (3) and (4) in estimating equation (5). We report results for this analysis in panel B of table 7. Beginning with the *IPTP* and *IPTV* measures, we find little evidence of XBRL substituting for scriptability. For *Dissem*, the interaction between *CompScript* and *%Tagged* is significantly positive, suggesting a complementary association between XBRL and scriptability, but the interaction between *CompScript* and *StandardTags* is significantly negative, suggesting a substitutive relationship. In an untabulated analysis, we repeat estimation of (5) after decomposing *CompScript* into *IdentifyData* and *DataToInformation*. Unlike the tabulated results, these results provide more consistent evidence of a substitutive

[22] See https://edgardashboard.xbrlcloud.com/edgar-dashboard/ for the dashboard.

[23] The EDGAR dashboard extends back to the earliest adoption of XBRL in 2009, but error and warning assessments are not available until late 2012, limiting our sample to later years of XBRL filings.

relationship between XBRL and scriptability. Coefficients on interactions between *IdentifyData* and either *TagQuality* and *StandardTags* are significantly negative in regressions using $IPTP^{24h}$, $IPTV^{5m}$, and *Dissem,* and we observe a significantly negative interaction between *DataToInformation* and *PctTagged* (*StandardTags*) in the model using $IPTV^{60m}$ ($IPTP^{60m}$). Consistent with tabulated results, the coefficient on *DataToInformation* interacted with *TagQuality* is significantly positive in the $IPTP^{24h}$ regression. In sum, results using *CompScript* suggest limited substitution, and possible complementarity, between XBRL and scriptability, and results using task-based measures suggest more consistent substitution. The inconsistency of our results likely reflects differences in scope between our measures of scriptability, which are designed to be holistic to the disclosure, and XBRL metadata that is limited to certain aspects such as the tagging of specific numbers for easier retrieval. As such, XBRL might only substitute for a limited set of our scriptability components. Additionally, we point out that many of our prior tests failed to detect significant associations between scriptability and our variables of interest for 10-K filings (i.e., most of our significant results relate to 8-K filings). Thus, it is possible that other capital market outcomes would be more appropriate for assessing the association between XBRL and scriptability.

5.2 SCRIPTABILITY AND CHANGES IN INFORMATION ASYMMETRY

The effect of scriptability on information asymmetry between market participants is ex ante unclear. Consistent with trends in business (Downes and Nunes [2013], Eule [2015]), we expect advances in technology and the availability of low-cost computing power (Kim [2015]) to level the playing field between investors by enabling data analysis at scale without significant investments in personnel. In this case, higher scriptability should lead to a further decline in information asymmetry by reducing the knowledge and resources necessary to perform computerized analysis and allowing even unsophisticated programmers to trade on scripted information. However, unsophisticated investors are unlikely to implement automated trading strategies such as those used by algorithmic traders.[24] As a result, more scriptable filings could result in a short-term information advantage for larger and more sophisticated investors, allowing them to trade on filing information more rapidly than other market participants and resulting in increases in information asymmetry immediately following disclosures. Consistent with this conjecture, Rogers, Skinner, and Zechman [2017] show that abnormal bid-ask spreads increase by 20–25% in the 30 to 60 seconds following Form 4 filings disclosing insider purchases, implying high-frequency traders that target these filings have an information advantage.

[24] For retail investors, it is not ex ante clear that technological and programming sophistication correlates with financial sophistication, so not all algorithms will necessarily trade profitably.

Given the above, we test whether scriptability increases or decreases information asymmetry over both short and long trading windows. As our proxy for information asymmetry, we estimate abnormal spreads in a manner similar to Rogers, Skinner, and Zechman [2017] by computing the quoted bid-ask spread five minutes, 60 minutes, and 24 hours following filing acceptance by the SEC and divide by the spread measured 60 seconds prior to the disclosure (*%AbSpread*). Panel C of figure 4 in Rogers, Skinner, and Zechman [2017] implies that algorithmic traders' information advantage remains significant for at least one minute following dissemination, so we expect an increase in information asymmetry during the five-minute window if scriptability facilitates automated trading. For the longer windows of 60 minutes and 24 hours, we expect scriptability to reduce this information advantage, leading to either a smaller increase or a decrease in information asymmetry. We estimate equation (6) below for each window:

$$\%AbSpread = \alpha + \beta Script + \Sigma \gamma_i \, (Controls) + e. \tag{6}$$

Script equals *CompScript* or our task-based measures (*IdentifyData* and *DataToInformation*). *Controls* refer to the same set of controls used to test H1 and described in subsection 3.2.1. Consistent with earlier tests, we estimate (6) by form type. Further, we limit our sample period to 2010–2015 to match the years used in our hypotheses tests. Results from estimating equation (6) are reported in table 8. As in tables 4 and 5, we report results using the five-minute (60-minute and 24-hour) measurement window in panel A (B and C). In panel A, we observe positive relations between *DataToInformation* and the abnormal bid-ask spread for 10-K ($t = 1.97$) and 8-K ($t = 2.08$) filings. We also observe a positive association between *CompScript* and *%AbSpread* for 8-K filings ($t = 2.56$). For proxy statements, we observe a weakly negative coefficient ($t = -1.70$) on *DataToInformation*, suggesting higher scriptability corresponds to a reduction in information asymmetry for these filings. In panels B and C, we observe that the positive coefficient in the five-minute window for 10-Ks declines to the point of insignificance for the 60-minute window. We observe a similar pattern for 8-Ks: the positive coefficient on *CompScript* weakens slightly in panel B (relative to panel A) and becomes insignificant in panel C. This shift in the significance of the coefficient occurs despite a significantly positive coefficient on *DataToInformation*, likely because *IdentifyData* loads negatively in panels B and C (increasing in significance from a *t*-statistics of -1.94 to -2.86). This could suggest, consistent with our discussion in subsection 2.1, that different types of scripters utilize different methods to reach the same information, so that unsophisticated scripters "close the gap" with sophisticated scripters through information search instead of computerized data processing. As such, it appears that scriptability is associated with information asymmetry, but that the association varies by form type and with time elapsed.

TABLE 8

Scriptability and Changes in Information Asymmetry (%AbSpread)

Panel A: *%AbSpread* for five-minute window

Form Type:	10-K		10-Q		8-K		DEF 14A	
Dependent Variable:	$\%AbSpread^{5m}$	$\%AbSpread^{5m}$	$\%AbSpread^{5m}$	$\%AbSpread^{5m}$	$\%AbSpread^{5m}$	$\%AbSpread^{5m}$	$\%AbSpread^{5m}$	$\%AbSpread^{5m}$
CompScript	1.757		0.606		0.811**		−0.737	
	(1.18)		(0.75)		(2.56)		(−0.68)	
IdentifyData		1.165		−0.220		0.111		0.284
		(0.81)		(−0.39)		(0.57)		(0.18)
DataToInformation		0.606**		0.652		0.580**		−0.853*
		(1.97)		(1.30)		(2.08)		(−1.70)
Follow	0.172*	0.172**	0.035	0.034	−0.088*	−0.087*	0.141	0.137
	(1.94)	(1.97)	(0.62)	(0.60)	(−1.80)	(−1.79)	(0.64)	(0.63)
ROA	1.216	1.240	−1.218	−1.250	−0.193	−0.203	−5.718***	−5.712***
	(0.70)	(0.71)	(−0.65)	(−0.66)	(−0.20)	(−0.21)	(−3.78)	(−3.78)
News	0.111	0.108	−0.007	−0.003	−0.130***	−0.129***	−0.018	−0.020
	(0.92)	(0.93)	(−0.23)	(−0.10)	(−4.52)	(−4.49)	(−0.21)	(−0.24)
BTM	−0.187	−0.190*	−0.025	−0.019	0.005	0.004	−0.045	−0.045
	(−1.63)	(−1.69)	(−0.19)	(−0.15)	(0.06)	(0.05)	(−0.26)	(−0.26)
Loss	−0.232	−0.233	−0.067	−0.063	−0.089	−0.089	−0.275	−0.278
	(−0.58)	(−0.58)	(−0.25)	(−0.24)	(−1.12)	(−1.12)	(−1.19)	(−1.20)
CapEx	3.519*	3.530*	−3.158***	−3.170***	−0.763	−0.758	4.423***	4.451***
	(1.92)	(1.91)	(−2.81)	(−2.82)	(−0.75)	(−0.75)	(3.27)	(3.31)
Lev	−0.181	−0.164	0.016	0.004	0.354**	0.350**	−0.059	−0.040
	(−0.56)	(−0.45)	(0.04)	(0.01)	(2.39)	(2.39)	(−0.09)	(−0.06)
FileCluster	0.093***	0.093***	0.115***	0.115***	0.138***	0.138***	0.138***	0.137***
	(4.42)	(4.44)	(5.75)	(5.76)	(5.67)	(5.67)	(4.98)	(5.10)
Size	−0.084	−0.088	−0.021	−0.016	0.045	0.045	−0.057	−0.059
	(−1.18)	(−1.27)	(−0.70)	(−0.52)	(1.10)	(1.11)	(−0.89)	(−0.95)
Age	−0.110	−0.109	−0.109	−0.115	−0.252***	−0.251***	−0.076	−0.072
	(−0.72)	(−0.71)	(−1.14)	(−1.18)	(−4.82)	(−4.81)	(−0.47)	(−0.44)

(*Continued*)

TABLE 8—*Continued*

Panel A: *%AbSpread* for five-minute window								
Form Type:	10-K		10-Q		8-K		DEF 14A	
Dependent Variable:	$\%AbSpread^{5m}$	$\%AbSpread^{5m}$	$\%AbSpread^{5m}$	$\%AbSpread^{5m}$	$\%AbSpread^{5m}$	$\%AbSpread^{5m}$	$\%AbSpread^{5m}$	$\%AbSpread^{5m}$
Distress	0.029	0.027	−0.010	−0.008	−0.022*	−0.021*	0.033	0.032
	(1.10)	(0.98)	(−0.31)	(−0.24)	(−1.87)	(−1.84)	(0.80)	(0.76)
OpComplex	0.044	0.043	−0.017	−0.014	0.039	0.037	0.123***	0.122***
	(0.38)	(0.37)	(−0.21)	(−0.17)	(0.79)	(0.77)	(2.67)	(2.65)
FinComplex	−1.062**	−1.103*	−0.248	−0.227	0.921	0.923	2.337*	2.307*
	(−2.08)	(−1.90)	(−0.21)	(−0.19)	(1.17)	(1.17)	(1.76)	(1.77)
Growth	−0.121	−0.121	−0.082	−0.080	0.038	0.039	0.055	0.057
	(−1.08)	(−1.08)	(−0.82)	(−0.79)	(0.92)	(0.94)	(0.22)	(0.23)
%IH	−0.341	−0.340	0.520**	0.518**	0.738***	0.740***	0.031	0.038
	(−1.33)	(−1.32)	(2.46)	(2.45)	(4.87)	(4.90)	(0.05)	(0.07)
HTI	0.490	0.489	0.196	0.195	0.139	0.140	0.326	0.331
	(1.11)	(1.11)	(0.63)	(0.63)	(0.58)	(0.59)	(0.81)	(0.83)
FileLag	−0.019***	−0.019***	−0.006	−0.006	−0.006*	−0.006*	−0.011	−0.011
	(−2.90)	(−2.70)	(−1.56)	(−1.52)	(−1.78)	(−1.78)	(−1.46)	(−1.52)
Top25FA	−0.355	−0.347	−0.131	−0.140	−0.029	−0.026	−0.004	0.012
	(−0.96)	(−0.98)	(−0.67)	(−0.70)	(−0.32)	(−0.28)	(−0.01)	(0.04)
Watermark	−0.022	−0.021	0.099	0.086	0.004	−0.003	0.116	0.125
	(−0.08)	(−0.07)	(0.61)	(0.51)	(0.04)	(−0.02)	(0.67)	(0.70)
HTML	−1.832	−1.900	0.053	0.122	−0.079	−0.051	−0.359	−0.449
	(−0.93)	(−1.00)	(0.08)	(0.19)	(−0.76)	(−0.43)	(−0.73)	(−0.76)
Year-month FEs?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FEs?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,108	6,108	20,312	20,312	74,207	74,207	6,753	6,753
Adjusted-R^2	0.015	0.014	0.013	0.013	0.018	0.018	0.016	0.016

(*Continued*)

TABLE 8—*Continued*

Panel B: *%AbSpread* for 60-minute window

Form Type:	10-K		10-Q		8-K		DEF 14A	
Dependent Variable:	*%AbSpread*^{60m}	*%AbSpread*^{60m}	*%AbSpread*^{60m}	*%AbSpread*^{60m}	*%AbSpread*^{60m}	*%AbSpread*^{60m}	*%AbSpread*^{60m}	*%AbSpread*^{60m}
CompScript	0.050		3.497		2.770**		3.879	
	(0.01)		(0.76)		(1.99)		(0.72)	
IdentifyData		−1.996		4.482		−2.200*		−1.027
		(−0.64)		(1.36)		(−1.94)		(−0.29)
DataToInformation		1.941		−0.084		3.511***		4.169
		(0.44)		(−0.03)		(3.77)		(1.09)
Follow	−1.444	−1.448	−0.119	−0.114	−0.721***	−0.719***	0.252	0.272
	(−1.37)	(−1.37)	(−0.35)	(−0.33)	(−4.31)	(−4.28)	(0.47)	(0.50)
ROA	6.278	6.114	−2.103	−1.941	3.830	3.752	10.394	10.426
	(1.26)	(1.26)	(−0.44)	(−0.41)	(1.44)	(1.44)	(1.43)	(1.45)
News	−0.440	−0.420	−0.395*	−0.416*	−0.514***	−0.502***	−0.858	−0.846
	(−0.95)	(−0.93)	(−1.68)	(−1.83)	(−3.30)	(−3.27)	(−1.38)	(−1.34)
BTM	−0.735	−0.717	−0.527	−0.551	−0.036	−0.047	−0.563	−0.561
	(−1.08)	(−1.04)	(−1.22)	(−1.27)	(−0.14)	(−0.18)	(−1.44)	(−1.43)
Loss	−0.546	−0.533	0.222	0.204	−0.774*	−0.771*	1.217	1.230
	(−0.68)	(−0.65)	(0.40)	(0.36)	(−1.90)	(−1.91)	(1.16)	(1.18)
CapEx	3.766	3.613	−8.012	−7.922	0.414	0.454	17.723*	17.500*
	(0.50)	(0.47)	(−1.43)	(−1.42)	(0.21)	(0.23)	(1.84)	(1.81)
Lev	2.695	2.584	1.484	1.539*	0.710	0.660	−0.155	−0.233
	(0.69)	(0.65)	(1.60)	(1.65)	(1.16)	(1.09)	(−0.07)	(−0.10)
FileCluster	0.002	0.003	0.109	0.109	0.242**	0.239**	0.456***	0.460***
	(0.02)	(0.02)	(0.83)	(0.82)	(2.53)	(2.51)	(3.79)	(3.87)
Size	0.016	0.046	−0.170	−0.196	−0.079	−0.075	−0.151	−0.141
	(0.04)	(0.10)	(−0.76)	(−0.84)	(−0.56)	(−0.53)	(−0.36)	(−0.34)
Age	−1.784*	−1.794*	−0.638	−0.605	−0.494***	−0.481***	−0.610*	−0.630*
	(−1.87)	(−1.87)	(−1.62)	(−1.54)	(−2.65)	(−2.59)	(−1.69)	(−1.70)

(*Continued*)

TABLE 8—*Continued*

Panel B: *%AbSpread* for 60-minute window								
Form Type:	10-K		10-Q		8-K		DEF 14A	
Dependent Variable:	$\%AbSpread^{60m}$	$\%AbSpread^{60m}$	$\%AbSpread^{60m}$	$\%AbSpread^{60m}$	$\%AbSpread^{60m}$	$\%AbSpread^{60m}$	$\%AbSpread^{60m}$	$\%AbSpread^{60m}$
Distress	−0.095	−0.082	−0.074	−0.084	−0.107**	−0.099**	0.213	0.218
	(−0.36)	(−0.31)	(−0.69)	(−0.77)	(−2.29)	(−2.19)	(0.70)	(0.72)
OpComplex	0.221	0.232	−0.503	−0.520	−0.201	−0.216	0.749*	0.749*
	(0.52)	(0.53)	(−1.26)	(−1.33)	(−1.31)	(−1.42)	(1.72)	(1.73)
FinComplex	17.056***	17.300***	3.865	3.771	2.913	2.985	17.972***	18.181***
	(4.57)	(4.73)	(0.62)	(0.60)	(1.07)	(1.10)	(2.86)	(2.95)
Growth	0.060	0.066	−0.705***	−0.719***	−0.348**	−0.340**	−0.234	−0.247
	(0.15)	(0.16)	(−2.71)	(−2.82)	(−2.23)	(−2.20)	(−0.66)	(−0.69)
%IH	8.846***	8.837***	9.428***	9.431***	4.637***	4.668***	4.114***	4.081***
	(2.69)	(2.68)	(5.81)	(5.81)	(6.50)	(6.50)	(2.91)	(2.89)
HTI	1.231	1.245	−2.096	−2.091	−0.812	−0.800	−0.917	−0.941
	(0.64)	(0.65)	(−1.05)	(−1.05)	(−1.36)	(−1.34)	(−0.31)	(−0.32)
FileLag	−0.095*	−0.093*	−0.017	−0.018	−0.026***	−0.026**	−0.025	−0.024
	(−1.78)	(−1.78)	(−0.46)	(−0.48)	(−2.60)	(−2.58)	(−0.55)	(−0.54)
Top25FA	−1.777***	−1.829***	−0.726	−0.691	−0.040	−0.003	1.438	1.364
	(−2.88)	(−3.00)	(−0.76)	(−0.72)	(−0.16)	(−0.01)	(0.82)	(0.76)
Watermark	−1.570***	−1.563***	0.174	0.234	0.384**	0.302**	2.273*	2.231*
	(−2.94)	(−3.04)	(0.15)	(0.20)	(2.50)	(2.22)	(1.90)	(1.87)
HTML	0.479	0.932	0.868	0.537	−0.700***	−0.368**	1.652	2.034*
	(0.13)	(0.27)	(0.29)	(0.18)	(−4.07)	(−2.44)	(1.47)	(1.80)
Year-month FEs?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FEs?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,468	7,468	24,149	24,149	94,412	94,412	8,297	8,297
Adjusted-R^2	0.015	0.015	0.010	0.010	0.012	0.013	0.018	0.018

(*Continued*)

TABLE 8—*Continued*

Panel C: *%AbSpread* for 24-hour window

Form Type:	10-K		10-Q		8-K		DEF 14A	
Dependent Variable:	$\%AbSpread^{24h}$	$\%AbSpread^{24h}$	$\%AbSpread^{24h}$	$\%AbSpread^{24h}$	$\%AbSpread^{24h}$	$\%AbSpread^{24h}$	$\%AbSpread^{24h}$	$\%AbSpread^{24h}$
CompScript	8.519		8.218		2.926		8.389	
	(1.28)		(1.25)		(1.34)		(0.89)	
IdentifyData		6.819		7.544		−3.090***		6.434
		(1.24)		(1.22)		(−2.86)		(0.94)
DataToInformation		1.855		1.821		4.172***		2.515
		(0.22)		(0.49)		(3.69)		(0.57)
Follow	1.315*	1.320*	2.187***	2.194***	−0.234	−0.229	1.811***	1.796***
	(1.85)	(1.86)	(3.31)	(3.32)	(−1.11)	(−1.07)	(2.78)	(2.82)
ROA	−9.418	−9.249	−10.690*	−10.482*	−1.662	−1.745	−18.171	−18.245
	(−0.87)	(−0.87)	(−1.73)	(−1.71)	(−0.68)	(−0.71)	(−1.51)	(−1.51)
News	0.110	0.086	0.363	0.340	0.469***	0.484***	0.181	0.173
	(0.27)	(0.21)	(0.67)	(0.62)	(3.37)	(3.54)	(0.27)	(0.26)
BTM	0.796	0.775*	0.087	0.057	0.730***	0.717***	0.663	0.663
	(1.60)	(1.67)	(0.27)	(0.18)	(2.92)	(2.83)	(0.74)	(0.74)
Loss	−2.406*	−2.424*	−0.272	−0.293	−0.570*	−0.564*	−0.243	−0.252
	(−1.72)	(−1.75)	(−0.44)	(−0.47)	(−1.73)	(−1.70)	(−0.19)	(−0.19)
CapEx	−12.637	−12.500	−13.969**	−13.848**	−1.076	−1.024	6.984	7.140
	(−0.96)	(−0.95)	(−2.06)	(−2.05)	(−0.36)	(−0.34)	(0.51)	(0.52)
Lev	6.386	6.502	−4.037	−3.986	1.768**	1.717**	−2.759	−2.702
	(1.34)	(1.30)	(−1.16)	(−1.14)	(2.04)	(1.98)	(−1.59)	(−1.55)
FileCluster	−0.602***	−0.603***	−0.539***	−0.540***	−0.103	−0.106	−0.420***	−0.423***
	(−5.93)	(−6.04)	(−5.34)	(−5.37)	(−1.52)	(−1.57)	(−3.77)	(−3.81)
Size	0.630	0.595	0.301	0.267	0.206**	0.209**	0.606	0.599
	(0.86)	(0.77)	(0.70)	(0.64)	(1.98)	(2.02)	(1.05)	(1.04)
Age	−1.194***	−1.184***	−0.677	−0.637	−0.059	−0.045	−0.803	−0.789
	(−3.22)	(−3.08)	(−0.83)	(−0.77)	(−0.23)	(−0.17)	(−1.29)	(−1.25)

(*Continued*)

TABLE 8—*Continued*

Panel C: *%AbSpread* for 24-hour window

Form Type:	10-K		10-Q		8-K		DEF 14A	
Dependent Variable:	$\%AbSpread^{24h}$	$\%AbSpread^{24h}$	$\%AbSpread^{24h}$	$\%AbSpread^{24h}$	$\%AbSpread^{24h}$	$\%AbSpread^{24h}$	$\%AbSpread^{24h}$	$\%AbSpread^{24h}$
Distress	−0.169	−0.183	0.406**	0.395**	−0.026	−0.017	0.420	0.417
	(−0.62)	(−0.62)	(2.13)	(2.05)	(−0.42)	(−0.27)	(1.47)	(1.46)
OpComplex	0.261	0.246	0.548	0.527	−0.293*	−0.310*	−0.278	−0.277
	(0.30)	(0.28)	(0.67)	(0.64)	(−1.85)	(−1.96)	(−0.42)	(−0.42)
FinComplex	−1.674	−1.956	2.003	1.886	−1.275	−1.170	9.935	9.762
	(−0.16)	(−0.18)	(0.33)	(0.31)	(−1.19)	(−1.06)	(1.26)	(1.24)
Growth	−1.061*	−1.065*	−0.946*	−0.961**	−0.273***	−0.266***	−0.001	0.008
	(−1.82)	(−1.84)	(−1.93)	(−1.97)	(−3.99)	(−3.83)	(−0.00)	(0.02)
%IH	5.009*	5.015*	5.461**	5.470**	2.019**	2.053**	−1.000	−0.977
	(1.92)	(1.91)	(2.21)	(2.21)	(2.34)	(2.37)	(−0.39)	(−0.38)
HTI	−2.674	−2.711	−3.328*	−3.328*	−0.682	−0.666	−0.164	−0.150
	(−1.54)	(−1.62)	(−1.69)	(−1.69)	(−1.05)	(−1.02)	(−0.06)	(−0.06)
FileLag	−0.036	−0.038	−0.057	−0.058	−0.041***	−0.041**	−0.100**	−0.100**
	(−0.51)	(−0.61)	(−1.46)	(−1.51)	(−2.58)	(−2.58)	(−2.38)	(−2.40)
Top25FA	−0.689	−0.630	0.545	0.587	−0.045	0.004	−0.514	−0.461
	(−0.43)	(−0.37)	(0.59)	(0.64)	(−0.14)	(0.01)	(−0.32)	(−0.29)
Watermark	−1.148	−1.158	0.228	0.301	−0.375	−0.479*	0.720	0.756
	(−1.59)	(−1.57)	(0.28)	(0.37)	(−1.50)	(−1.82)	(0.51)	(0.55)
HTML	−0.182	−0.750	−0.344	−0.737	−1.044***	−0.620**	−0.461	−0.747
	(−0.04)	(−0.23)	(−0.11)	(−0.22)	(−3.67)	(−2.10)	(−0.24)	(−0.34)
Year-month FEs?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FEs?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8,479	8,479	26,777	26,777	108,042	108,042	9,288	9,288
Adjusted-R^2	0.033	0.033	0.034	0.034	0.008	0.008	0.016	0.016

This table presents results from estimating equation (5) with *%AbSpread* as the dependent variable. This was a planned supplemental analysis. Panel A (B and C) presents results using the *%AbSpread* measure computed over the five-minute (60-minute and 24-hour) window. We estimate equation (5) by form type in all panels. All variables are defined in the appendix. Industries are defined using the Fama-French 48-industry designations. Reported *t*-statistics in parentheses are derived from standard errors clustered by firm and year. ***, **, and * denote two-tailed significance at the $p < 0.01$, $p < 0.05$, and $p < 0.10$ level, respectively.

6. *Conclusion*

Prior research examines the readability of disclosures for managers, investors, and regulators with a focus on the level of effort required for a *person* to read and understand the disclosure (Li [2008]). We examine the "scriptability" of firm disclosures, which represents the relative ease with which a *computer program* or a *computer programmer* can transform the large amounts of unstructured data contained in various firm disclosures into usable information (Bloomfield [2002]). We identify two basic tasks that a "scripter" of firm filings is likely to perform, identifying data of interest and processing that data into decision-relevant information, and measure characteristics of filings that are likely to facilitate or inhibit the automation of these tasks. We validate our measures using a series of tests relying on researcher-derived samples and measures commonly used in prior literature. We also assess whether our measure is associated with a set of firm and disclosure characteristics that we expect to relate to scriptability and find some limited evidence from these tests that investment in the financial reporting function is related to higher scriptability, although characteristics such as technological sophistication, firm size, or age do not appear to translate into more scriptable filings. Overall, the most significant predictor of filing scriptability appears to be preparation by a top filing agent. In our primary tests, we provide some evidence that scriptability relates positively to measures of the speed of the market response to SEC filings, especially as captured in our volume measures. We also predict that scriptability increases the likelihood and speed of DJ dissemination of SEC filings, but find limited evidence consistent with this prediction. In additional, planned analyses, we find evidence in some specifications that XBRL could substitute for scriptability, at least with respect to our task-based measures. However, given limited evidence in our main tests that scriptability affects 10-K filing processing, it is perhaps unsurprising that we find little evidence regarding the interaction of XBRL data with scriptability. Finally, the association between scriptability and information asymmetry varies considerably depending on the measurement window, form type, and aspect of scriptability examined.

Our proposed research provides several important contributions to the accounting and finance literature. Perhaps most importantly, our study is one of the first to examine programmatic extraction (i.e., mining) of data from firm disclosures. By providing detailed insights into the data and metadata quality of firm disclosures, we lay the groundwork for future research to explore how scriptability affects information processing. Although all of our results are based on cross-sectional analyses, and thus might be subject to omitted variable concerns, one takeaway from our study is that different aspects of scriptability likely affect information processing in different ways, making a single-measure construct, such as *CompScript*, inappropriate in some settings. We encourage future researchers examining scriptability to choose components of our measure that fit their research questions and the nature of their tests. Additionally, our analysis broadens the scope of literature suggesting that lower information processing costs facilitate efficient capital allocation (e.g., Lundholm, Rogo, and Zhang [2014]). Our examination of the role of technology in alleviating these costs informs regulators increasingly focused on information processing costs as an impediment to capital formation and as a source of information asymmetry among market participants (Flannery [2015]). Our study also informs regulators' ongoing efforts to improve disclosure quality from a "machine-readable" perspective by suggesting that the electronic quality of regulatory filings is associated with their usefulness.

APPENDIX

Description of Variables Used in the Analysis of Scriptability

Note that variable names appearing in braces refer to variable names used by data source in brackets.

Key Variables of Interest	
CompScript	Composite measure of disclosure scriptability that equals the average of task-based measures *IdentifyData* and *DataToInformation*. See online appendix A for detail [EDGAR].
IdentifyData	The ease with which a program can identify data. See online appendix A for detail [EDGAR].
DataToInformation	The ease with which a program can transform data to usable information. See online appendix A for detail [EDGAR].
%AbSpread^w	The ratio of the bid-ask spread measured *w* following the disclosure divided by the spread measured 60 seconds prior to disclosure. *w* equals five minutes, 60 minutes, or 24 hours [TAQ].
%Tagged	The number of numeric XBRL tags in the filing divided by number of unique numbers in the document [EDGAR].
Dissem	Equals 1 if DJ disseminates at least one newswire in the 24 hours following filing receipt by the SEC, denoted by the "Accepted" timestamp. If more than one newswire occurs within 24 hours of two or more filings by the same firm of the same form type, we assume the disclosure closest to the newswire was the one disseminated. To mitigate the risk of erroneously matching newswires to filings, we require newswires to contain one of the following in the headline or lead paragraph: 1. Category of form type. Specifically, we search for "10K," "10-K," or "Annual Report" for all form 10-K variants; "10Q," "10-Q," or "Quarterly Report" for all form 10-Q variants; "8K," "8-K," or "Press Release" for all form 8-Ks; and "Proxy," "DEF 14A," or "DEF14A" for all form DEF 14As. 2. The specific form type of the filing if different from the category (e.g., 10-KSB, etc.). 3. Some combination of "SEC" and "filing" or "filed" [DOWJONES].

(*Continued*)

Key Variables of Interest	
IPTP^w	Ranked abnormal price-based IPT measure. We first calculate the price IPT as the area under the price response curve over the first *w* minutes following a disclosure, where *w* equals five minutes ("5m"), 60 minutes ("60m"), or 24 hours ("24h"). We use effective returns based on quoted midpoints of bids {bid} and offers {ofr} to compute returns over each window. We exclude quotes with nonnormal quote conditions {qu_cond = A, B, H, O, R, and W}, quotes where the bid is greater than or equal to the offer for any exchange or market maker, quotes where the spread exceeds \$5.00 and the bid (offer) price is less (greater) than the previous midpoint −\$2.50 (previous midpoint +\$2.50), cases in which trading has been halted, and quotes where the bid (ask) price or bid (offer) size {bidsiz, asksiz} is zero or missing. We use the National Best Bid and Best Offer (NBBO) at each minute in our measurement window. Since Holden and Jacobsen [2014] find that NBBO data in TAQ is incomplete, we use their procedure to compute the NBBO. SAS code to perform this computation is available at http://kelley.iu.edu/cholden/ (as of April 14, 2016). For the five minute (60 minute) window, we require the disclosure to be filed by five minutes (60 minutes) prior to the end of TAQ's after-hours trading data. For the 24-hour window, we consider all filings since all include a postmarket, overnight window. We then adjust this IPT measure by a "normal" IPT measure, computed as the mean price IPT over the previous 52 weeks for the same minute, hour, and day as that of the filing. Finally, we decile-rank these IPT measures by year, by form-type [TAQ].
IPTV^w	Ranked abnormal volume-based IPT measure. We first calculate the volume IPT as the area under the cumulative volume curve over the first *w* following a disclosure, where *w* equals five minutes ("5m"), 60 minutes ("60m"), or 24 hours ("24h"). We compute CV based on executed trades from TAQ {size}. For the five-minute (60-minute) window, we require the disclosure to be filed by five minutes (60 minutes) prior to the end of TAQ's after-hours trading data. For the 24-hour window, we consider all filings since all include a postmarket, overnight window. We then adjust this IPT measure by a "normal" IPT measure, computed as the mean volume IPT over the previous 52 weeks for the same minute, hour, and day as that of the filing. Finally, we decile-rank these IPT measures by year, by form-type [TAQ].
NewsDelay	Minutes elapsed between the SEC's "Accepted" timestamp and the newswire identified for *Dissem*. *NewsDelay* is only defined for filings where *Dissem* equals 1 [DOWJONES].
StandardTags	One minus the proportion of extended XBRL tags (i.e., tags not in a standard FASB taxonomy) to total XBRL tags in the filing [EDGAR].
TagQuality	The proportion of XBRL-quality dimensions identified as "OK" by xbrlcloud.com's EDGAR Dashboard [XBRLcloud].

(*Continued*)

Other Variables	
#EAs	Number of earnings announcements occurring in the three-day window surrounding the filing. Compustat {rdq} is used to measure *#EAs* [Compustat].
#Filings	Natural log of the total number of SEC filings occurring in the two-hour window surrounding the disclosure [EDGAR].
%IH	The percentage of shares held by institutions {instown_perc} reported by Thomson Reuters Stock Ownership Summary as of the 13F reports filed closest, but prior, to the disclosure [Thomson Reuters].
Age	Firm age, measured as the natural log of firm age, measured as the number of years between the filing date and the earliest date the firm appears in Compustat [Compustat].
BTM	Firm's book-to-market ratio {ceqq/(prccq × cshoq)} as of the end of the quarter ending closest, but prior, to the filing [Compustat].
CapEx	Capital expenditures scaled by assets {capx/at} for the fiscal year ending closest, but prior, to the filing [Compustat].
Distress	The inverse of financial strength, measured as the decile rank of Altman's [1968] z-score (1.2({act-lct})/{at} + 1.4({re/at}) + 3.3({ebit/at}) + 0.6({prcc_f∗csho})/({dltt+dlc}) + 1.0({sale/at}) as of the end of the fiscal year ending closest, but prior, to the date of the filing. If act (lct) is missing, we use the difference between total assets {at} and gross PP&E {ppegt} (total liabilities {lt} and long-term debt {dltt}) instead [Compustat].
DJLength	Natural log of the number of words in the DJ news article [DOWJONES].
FileCluster	Number of filings received by the SEC in the 60 seconds prior to the filing of interest [EDGAR].
FileLag	Number of days from fiscal year-end to 10-K filing date for 10-K filed closest, but prior, to current filing [EDGAR].
FinComplex	Firm financial complexity, measured as the logarithm of the number of nonmissing items in the Compustat Annual file for the fiscal year closest, but prior, to the filing [Compustat].
Follow	The natural log of one plus the number analysts following the firm at the time of the disclosure, proxied for using the number of estimates {numest} for the current fiscal year's EPS {fpi = 1} included in the monthly IBES summary file dated closest, but prior, to the date of the filing [IBES].
Growth	Sales growth, measured as the average percentage increase in sales {revt} over the last year. For firms with missing values, we use the industry median for that year [Compustat].
HTI	Indicator equaling 1 if primary SIC code {sic} for firm is in technology industries, and 0 otherwise. Industries include SIC code ranges: 3661–3669, 3571–3579, 7370–7379, 3670, 3500, 3600, 3810–3870, and 3840–3849 [Compustat].
HTML	Indicator variable equaling 1 if disclosure is prepared using HTML encoding, and 0 otherwise [EDGAR].
Lev	Long-term debt scaled by total assets {(dltt + dlc)/at} as of the end of the quarter ending closest, but prior, to the filing [Compustat].

(*Continued*)

Other Variables	
Lit	Indicator equaling 1 if primary SIC code [sic] for firm is in high-litigation risk industries. These industries include the following SIC code ranges: 2833–2836, 3570–3577, 3600–3674, 5200–5961, 7370–7374, and 8731–8734 [Compustat].
Loss	An indicator variable equal to 1 if net income {niq} for the quarter ending closest, but prior, to the filing was less than 0, and 0 otherwise [Compustat].
News	The natural log of the number of news articles mentioning the firm appearing in the DJ News Archives during the 90 days (one quarter) ending the day prior to the filing [DOW JONES].
OpComplex	Firm operating complexity, measured as the log of the sum of the number of operating and geographic segments reported in Compustat's Annual Segments file for the fiscal year closest, but prior, to the filing [Compustat].
Opinion	Equals 1 for any disclosure issued in a year following a nonstandard audit opinion {auop not equal to 1}, and 0 otherwise [Compustat].
ROA	Return on assets for the quarter ending closest, but prior, to the filing, calculated as net income before extraordinary items {ibq} divided by total assets {atq} [Compustat].
Size	Firm size, measured as the log of the market value {prccq x cshoq} of equity as of the end of the fiscal quarter closest, but prior, to the filing [Compustat].
Top25FA	Indicator variable equaling 1 if the firm filing is prepared by one of the top 25 filing agents for that calendar year for that form, and 0 otherwise [SECFILE.COM/EDGAR].
Volatility	The standard deviation of logged daily returns (i.e., ln(1+{ret}) over the 60 trading days prior to the filing, annualized by multiplying by the square-root of 252 [CRSP].
Volume	The natural log of the average daily trading volume {vol} over the 90 days preceding the disclosure [CRSP].
Watermark	Indicator variable equaling 1 if the filing contains a preparer's electronic watermark, and 0 otherwise [SECFILE.COM/EDGAR].
XBRL	Equals 1 for filings with XBRL metadata and 0 otherwise [EDGAR].

REFERENCES

ALLEE, K. D., AND M. D. DEANGELIS. "The Structure of Voluntary Disclosure Narratives: Evidence from Tone Dispersion." *Journal of Accounting Research* 53 (2015): 241–74.

ARNOLD, V.; J. C. BEDARD; J. R. PHILLIPS; AND S. G. SUTTON. "The Impact of Tagging Qualitative Financial Information on Investor Decision Making: Implications for XBRL." *International Journal of Accounting Information Systems* 13 (2012): 2–20.

ASHBAUGH-SKAIFE, H.; D. W. COLLINS; AND W. R. KINNEY JR. "The Discovery and Reporting of Internal Control Deficiencies Prior to SOX-Mandated Audits." *Journal of Accounting and Economics*, Conference Issue on Corporate Governance: Financial Reporting, Internal Control, and Auditing, 44 (2007): 166–92.

BAMBER, L. S. "The Information Content of Annual Earnings Releases: A Trading Volume Approach." *Journal of Accounting Research* 24 (1986): 40–56.

BARZILAY, R. K.; K. R. MCKEOWN; AND M. ELHADAD. "Information Fusion in the Context of Multi-Document Summarization," in *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics on Computational Linguistics*, 1999: 550–57.

BEAVER, W. H. "The Information Content of Annual Earnings Announcements." *Journal of Accounting Research* 6 (1968): 67–92.

BLANKESPOOR, E; E. DEHAAN; AND C. ZHU. "Capital Market Effects of Media Synthesis and Dissemination: Evidence from Robo-Journalism." *Review of Accounting Studies* 23 (2018): 1–36.

BLANKESPOOR, E; G. S. MILLER; AND H. D. WHITE. "The Role of Dissemination in Market Liquidity: Evidence from Firms' Use of Twitter™." *The Accounting Review* 89 (2014): 79–112.

BLOOMFIELD, R. J. "The 'Incomplete Revelation Hypothesis' and Financial Reporting." *Accounting Horizons* 16 (2002): 233–43.

BLOOMFIELD, R. J. "Discussion of 'Annual Report Readability, Current Earnings, and Earnings Persistence.'" *Journal of Accounting and Economics* 45 (2008): 248–52.

BLOOMFIELD, R. J.; F. D. HODGE; P. E. HOPKINS; AND K. M. RENNEKAMP. "Does Coordinated Presentation Help Credit Analysts Identify Firm Characteristics?" SSRN Scholarly Paper ID 1650906, 2010.

BROGAARD, J.; T. HENDERSHOTT; AND R. RIORDAN. "High-Frequency Trading and Price Discovery." *Review of Financial Studies* 27 (2014): 2267–306.

BUSHMAN, R. M.; A. J. SMITH; AND R. WITTENBERG-MOERMAN. "Price Discovery and Dissemination of Private Information by Loan Syndicate Participants: Price Discovery and Dissemination of Private Information." *Journal of Accounting Research* 48 (2010): 921–72.

BUSSE, J. A., AND C. T. GREEN. "Market Efficiency in Real Time." *Journal of Financial Economics* 65 (2002): 415–37.

BUTLER, M.; A. KRAFT; AND I.S. WEISS. "The Effect of Reporting Frequency on the Timeliness of Earnings: The Cases of Voluntary and Mandatory Interim Reports." *Journal of Accounting and Economics* 43 (2007): 181–217.

CAMPBELL, J. L.; H. CHEN; D. S. DHALIWAL; H. M. LU; AND L. B. STEELE. "The Information Content of Mandatory Risk Factor Disclosures in Corporate Filings." *Review of Accounting Studies* 19 (2014): 396–455.

CARBONELL, J., AND J. GOLDSTEIN. "The Use of MMR, Diversity-Based Reranking for Reordering Documents and Producing Summaries," in *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1998: 335–36.

CARNEY, J., AND F. HARKER. "How SEC's New RoboCop Profiles Companies for Accounting Fraud." *Forbes*, 2015. Available at http://www.forbes.com/sites/janetnovack/2013/08/09/how-secs-new-robocop-profiles-companies-for-accounting-fraud/.

CHEN, H.; P. DE; Y. HU; AND B.H. HWANG. "Wisdom of Crowds: The Value of Stock Opinions Transmitted Through Social Media." *Review of Financial Studies* 27 (2014): 1367–403.

CHEN, S.; B. MIAO; AND T. SHEVLIN. "A New Measure of Disclosure Quality: The Level of Disaggregation of Accounting Data in Annual Reports." *Journal of Accounting Research* 53 (2015): 1017–54.

CHENG, Q., AND K. LO. "Insider Trading and Voluntary Disclosures." *Journal of Accounting Research* 44 (2006): 815–48.

CHILTON, B. "No Need to Demonize High-Frequency Trading." *New York Times*, July 7, 2014. Available at http://dealbook.nytimes.com/2014/07/07/no-need-to-demonize-high-frequency-trading/.

CHOUDHARY, P.; K. J. MERKLEY; AND J. D. SCHLOETZER. "Early Annual Reports." SSRN Scholarly Paper ID 1436538, 2015.

COX, C. "Speech by SEC Chairman: Remarks at the 12th XBRL International Conference," 2005. Available at http://www.sec.gov/news/speech/spch110705cc.htm.

CUFF, D. C. "Making a Difference; Edgarizing the S.E.C." *The New York Times*, February 21, 1993. Available at http://www.nytimes.com/1993/02/21/business/making-a-difference-edgarizing-the-sec.html.

DEBRECENY, R.; S. FAREWELL; M. PIECHOCKI; C. FELDEN; AND A. GRÄNING. "Does It Add Up? Early Evidence on the Data Quality of XBRL Filings to the SEC." *Journal of Accounting and Public Policy* 29 (2010): 296–306.

DOWNES, L., AND P. NUNES. "Big-Bang Disruption." *Harvard Business Review*, March 1, 2013.

DOYLE, J.; W. GE; AND S. MCVAY. "Determinants of Weaknesses in Internal Control over Financial Reporting." *Journal of Accounting and Economics*, Conference Issue on Corporate Governance: Financial Reporting, Internal Control, and Auditing 44 (2007): 193–223.

DRAKE, M. S.; N. M. GUEST; AND B.J. TWEDT. "The Media and Mispricing: The Role of the Business Press in the Pricing of Accounting Information." *The Accounting Review* 89 (2014): 1673–701.

DRAKE, M. S.; D. T. ROULSTONE; AND J. R. THORNOCK. "Investor Information Demand: Evidence from Google Searches Around Earnings Announcements." *Journal of Accounting Research* 50 (2012): 1001–40.

DRAKE, M. S.; D. T. ROULSTONE; AND J. R. THORNOCK. "The Determinants and Consequences of Information Acquisition via EDGAR." *Contemporary Accounting Research* 32 (2015): 1128–61.

DRAKE, M. S.; J. R. THORNOCK; AND B. J. TWEDT. "The Internet as an Information Intermediary." *Review of Accounting Studies* 22 (2017): 543–76.

DU, H.; M. A. VASARHELYI; AND X. ZHENG. "XBRL Mandate: Thousands of Filing Errors and So What?" *Journal of Information Systems* 27 (2013): 61–78.

EULE, A. "Robo Advisors Take on Wall Street." *Barron's Online*, May 23, 2015.

FAMA, E. F.; L. FISHER; M. C. JENSEN; AND R. ROLL. "The Adjustment of Stock Prices to New Information." *International Economic Review* 10 (1969): 1–21.

FINANCIAL ACCOUNTING STANDARDS BOARD. "Disclosure Framework." 2016. Available at http://www.fasb.org/jsp/FASB/FASBContent_C/ProjectUpdatePage&cid=1176156344894.

FLANNERY, M. J. "Economic Analysis: Providing Insight to Advance the Missions of the SEC and the PCAOB," 2015. Available at http://www.sec.gov/news/speech/keynote-address-pcaob-missions-of-sec-and-pcaob.html

FOSTER, F. D., AND S. VISWANATHAN. "Variations in Trading Volume, Return Volatility, and Trading Costs; Evidence on Recent Price Formation Models." *The Journal of Finance* 48 (1993): 187–211.

FOXMAN, S. "Cohen's Point72 Hires 30 People for Big Data Investing." Bloomberg.com., 2016. Available at http://www.bloomberg.com/news/articles/2015-03-10/cohen-s-firm-said-to-hire-30-seeking-edge-in-public-data.

GROSSMAN, S. J., AND J. E. STIGLITZ. "On the Impossibility of Informationally Efficient Markets." *The American Economic Review* 70 (1980): 393–408.

GUAY, W. R.; D. SAMUELS; AND D. J. TAYLOR. "Guiding Through the Fog: Financial Statement Complexity and Voluntary Disclosure." *Journal of Accounting and Economics* 62 (2016): 234–69.

HAGHIGHI, A., AND L. VANDERWENDE. "Exploring Content Models for Multi-Document Summarization," in *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, 2009: 362–70. Available at http://dl.acm.org/citation.cfm?id=1620754.1620807.

HALES, J.; J. R. MOON JR; AND L. A. SWENSON. "A New Era of Voluntary Disclosure? Empirical Evidence on How Employee Postings on Social Media Relate to Future Corporate Disclosures." Working paper, Georgia Tech Scheller College of Business Research, 2018.

HODGE, F. D.; J. J. KENNEDY; AND L. A. MAINES. "Does Search-Facilitating Technology Improve the Transparency of Financial Reporting?" *The Accounting Review* 79 (2004): 687–703.

HOITASH, R., AND U. HOITASH. "Measuring Accounting Reporting Complexity with XBRL." *The Accounting Review* 93 (2018): 259–87. https://doi.org/10.2308/accr-51762.

HOLDEN, C. W., AND S. JACOBSEN. "Liquidity Measurement Problems in Fast, Competitive Markets: Expensive and Cheap Solutions." *The Journal of Finance* 69 (2014): 1747–85.

JAME, R.; R. JOHNSTON; S. MARKOV; AND M. C. WOLFE. "The Value of Crowdsourced Earnings Forecasts." *Journal of Accounting Research* 54 (2016): 1077–110.

KIM, C.; K. WANG; AND L. ZHANG. "Readability of 10-K Reports and Stock Price Crash Risk." Working paper, University of Alberta School of Business, 2018.

KIM, E. "This One Chart Shows the Vicious Price War Going on in Cloud Computing." *Business Insider*, January 14, 2015. Available at http://www.businessinsider.com/cloud-computing-price-war-in-one-chart-2015-1.

KOBELSKY, K. W.; V. J. RICHARDSON; R. E. SMITH; AND R. W. ZMUD. "Determinants and Consequences of Firm Information Technology Budgets." *Accounting Review* 83 (2008): 957–95.

KRAVET, T., AND V. MUSLU. "Textual Risk Disclosures and Investors' Risk Perceptions." *Review of Accounting Studies* 18 (2013): 1088–122.

LEE, Y. J. "The Effect of Quarterly Report Readability on Information Efficiency of Stock Prices." *Contemporary Accounting Research* 29 (2012): 1137–70.

LEE, Y. J.; K. R. PETRONI; AND M. SHEN. "Cherry Picking, Disclosure Quality, and Comprehensive Income Reporting Choices: The Case of Property-Liability Insurers." *Contemporary Accounting Research* 23 (2006): 655–700.

LERMAN, A., AND J. LIVNAT. "The New Form 8-K Disclosures." *Review of Accounting Studies* 15 (2010): 752–78.

LEWIS, M. *Flash Boys: A Wall Street Revolt*, Third edition. New York: W. W. Norton & Company, 2015.

LI, E. X.; K. RAMESH; AND M. SHEN. "The Role of Newswires in Screening and Disseminating Value-Relevant Information in Periodic SEC Reports." *Accounting Review* 86 (2011): 669–701.

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

LOUGHRAN, T., AND B. MCDONALD. "When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks." *The Journal of Finance* 66 (2011): 35–65.

LOUGHRAN, T., AND B. MCDONALD. "Measuring Readability in Financial Disclosures: Measuring Readability in Financial Disclosures." *The Journal of Finance* 69 (2014): 1643–71.

LOUGHRAN, T., AND B. MCDONALD. "Textual Analysis in Accounting and Finance: A Survey." *Journal of Accounting Research* 54 (2016): 1187–230.

LUNDHOLM, R. J.; R. ROGO; AND J. ZHANG. "Restoring the Tower of Babel: How Foreign Firms Communicate with U.S. Investors." *The Accounting Review* 89 (2014): 1453–85.

MAINES, L. A., AND L. S. MCDANIEL. "Effects of Comprehensive-Income Characteristics on Nonprofessional Investors' Judgments: The Role of Financial-Statement Presentation Format." *The Accounting Review* 75 (2000): 179–207.

MCDONALD, B. "Bill McDonald's Word Lists Page." 2016. Available at http://www3.nd.edu/˜mcdonald/Word_Lists.html.

MCMULLIN, J. L. "Can I Borrow Your Footnotes? Learning and Network Benefits of Footnote Similarity." SSRN Scholarly Paper ID 2381950, 2016.

MCMULLIN, J. L.; B. P. MILLER; AND B. J. TWEDT. "Increased Mandated Disclosure Frequency and Price Formation: Evidence from the 8-K Expansion Regulation," 2015. Available at http://mays.tamu.edu/phd-accounting/wp-content/uploads/sites/31/2015/10/MMT-2015_09_29_Final.pdf.

MILLER, B. P. "The Effects of Reporting Complexity on Small and Large Investor Trading." *Accounting Review* 85 (2010): 2107–43.

O'HARA, M. "Presidential Address: Liquidity and Price Discovery." *The Journal of Finance* 58 (2003): 1335–54.

PETERSON, K.; R. SCHMARDEBECK; AND T. J. WILKS. "The Earnings Quality and Information Processing Effects of Accounting Consistency." *The Accounting Review* 90 (2015): 2483–514.

PHILIPS, M. "How the Robots Lost: High-Frequency Trading's Rise and Fall." Bloomberg, 2015. Available at http://www.bloomberg.com/bw/articles/2013-06-06/how-the-robots-lost-high-frequency-tradings-rise-and-fall.

ROGERS, J. L.; D. J. SKINNER; AND S. L. ZECHMAN. "Run EDGAR Run: SEC Dissemination in a High-Frequency World." *Journal of Accounting Research* 55 (2017): 459–505.

ROGERS, J. L., AND P. C. STOCKEN. "Credibility of Management Forecasts." *The Accounting Review* 80 (2005): 1233–60.

SCHRAND, C. M., AND B. R. WALTHER. "Strategic Benchmarks in Earnings Announcements: The Selective Disclosure of Prior Period Earnings Components." *The Accounting Review* 75 (2000): 151–77.

SECURITIES AND EXCHANGE COMMISSION. "Final Rule: Acceleration of Periodic Report Filing Dates and Disclosure Concerning Website Access to Reports." 2002. Available at https://www.sec.gov/rules/final/33-8128.htm.

SECURITIES AND EXCHANGE COMMISSION. "Final Rule: XBRL Voluntary Financial Reporting Program on the EDGAR System," 2005. Available at http://www.sec.gov/rules/final/33-8529.htm.

SECURITIES AND EXCHANGE COMMISSION. "Small Business Compliance Guide: Interactive Data for Financial Reporting," 2009a. Available at https://www.sec.gov/info/smallbus/secg/interactivedata-secg.htm.

SECURITIES AND EXCHANGE COMMISSION. "Final Rule: Interactive Data to Improve Financial Reporting—33–9002.pdf," 2009b. Available at http://www.sec.gov/rules/final/2009/33-9002.pdf.

TASKER, S. C. "Bridging the Information Gap: Quarterly Conference Calls as a Medium for Voluntary Disclosure." *Review of Accounting Studies* 3 (1998): 137–67.

THONG, Y. L. "An Integrated Model of Information Systems Adoption in Small Businesses." *Journal of Management Information Systems* 15 (1999): 187–214.

TWEDT, B. "Spreading the Word: Price Discovery and Newswire Dissemination of Management Earnings Guidance." *The Accounting Review* 91 (2016): 317–46.

WAYMIRE, G. "Earnings Volatility and Voluntary Management Forecast Disclosure." *Journal of Accounting Research* 23 (1985): 268–95.

WORKIVA. "Workiva Ranked as One of the Fastest Growing Companies in North America on Deloitte's 2015 Technology Fast 500™ List." News release, 2015. Available at http://www.prnewswire.com/news-releases/workiva-ranked-as-one-of-the-fastest-growing-companies-in-north-america-on-deloittes-2015-technology-fast-500-list-300178399.html.

YOU, H., AND X. J. ZHANG. "Financial Reporting Complexity and Investor Underreaction to 10-K Information." *Review of Accounting Studies* 14 (2009): 559–86.