The Information Content of Corporate Websites

Bradford Lynch
lynchb@wharton.upenn.edu
The Wharton School
University of Pennsylvania

Daniel Taylor*
dtayl@wharton.upenn.edu
The Wharton School
University of Pennsylvania

May 2021

Abstract

In 2008, the SEC published guidance allowing firms to use corporate websites as an alternative disclosure channel to EDGAR. While the information content and market reaction to traditional disclosure channels such as EDGAR filings and press releases are well-documented, evidence on corporate websites as a disclosure channel is scarce. In this paper, we take the first step toward shedding light on corporate websites as an important source of information to investors. Employing standard event study methods, we develop a novel measure of corporate website content and find that increases in website content provide significant value-relevant information to investors incremental to that contained in traditional disclosure channels. In addition, we find a negative relation between website content and information asymmetry, and that this negative relation is most pronounced after the SEC's 2008 guidance. Comparing the content of websites to EDGAR, we find that websites contain *more* content related to business operations and that the market reaction to such information is heightened. Collectively, our findings indicate that corporate websites are an economically significant source of new information that supplements traditional disclosure channels considered in prior literature.

*Corresponding author. We appreciate helpful comments and suggestions from Brian Bushee, Paul Fischer, Wayne Guay, Bob Holthausen, Rick Lambert, Catherine Schrand, Christina Zhu and seminar participants at The Wharton School. This paper is based on Bradford Lynch's second year summer paper. Data and code used in this paper will be made publicly available upon publication.

1. Introduction

Over the past 25 years, corporate websites have evolved to become a cornerstone of modern business communications between firms and their stakeholders. Recognizing the value of internet-based dissemination, the SEC first recommended firms use corporate websites to communicate with investors as far back as 1995 (SEC Release 33-7233). While numerous academic surveys catalog how investors respond to traditional forms of disclosure such as EDGAR filings, management forecasts, and press releases, academic study of whether and how firms use corporate websites to communicate to stakeholders is virtually non-existent. Indeed, the word "website" does not appear in any of the academic surveys of the disclosure literature (e.g., Beyer et al., 2010; Leuz and Wysocki, 2016; Blankespoor, deHaan, and Marinovic, 2020).

The absence of empirical evidence on whether and how firms use corporate websites as a disclosure channel is striking when one considers that the SEC allows firms to use corporate websites as an alternative to EDGAR, thereby enabling firms to post material information to their websites but *not* to EDGAR (SEC Release 34-58288 p. 14, emphasis added):

> "*[W]e have given companies the choice and flexibility of satisfying an Exchange Act disclosure requirement either by filing the disclosure on EDGAR or by making it available on the company's web site, thereby using company web sites as an alternative to EDGAR.*"

In this paper, we take the first step toward shedding light on whether corporate websites are an important source of information to investors; whether such websites provide novel information or merely reinforce information provided through traditional disclosure channels; and whether information on websites exacerbates or mitigates the information advantage of sophisticated investors.

We begin our analysis by developing a novel measure of website content that is broadly applicable to the universe of firms appearing on standard research databases (e.g., Compustat and CRSP). For each firm, we locate and download the entire website including the homepage and interior pages on a daily basis. Similar to Bonsall, Leone, and Miller (2017), we then measure five fundamental attributes of the website—character length, tags, links, multimedia content, and dynamic scripting. We conduct a principal component analysis of these attributes, and find the first principal component explains 93% of the variation in the attributes. We use the first principal component as a composite measures of daily website content, and estimate three distinct sets of tests that represent joint tests of our measures of website content and our economic predictions.

In our first set of tests, we use an event study design to examine the market reaction to large increases in daily website content. Specifically, we study the price and volume reaction on those days with large increases in website content, but yet there is no confounding EDGAR filings or press releases in the [-1,+1] window around the change in content. Focusing on large increases in website content allows the design to cleanly separate routine small increases, e.g., addition of a stock price ticker, from more substantive changes such as the issuance of a Regulation Fair Disclosure (Reg FD) announcement. Focusing on days where there were no EDGAR filings or press releases rules out the possibility that investors are responding to some other disclosure channel.

In addition to focusing on changes in prices and volume in a tight window around an increase in website content, we also include firm-quarter fixed effects in our event study design. This feature of our design should mitigate concerns that our results are attributable to omitted firm characteristics (e.g., firm size, operating environment, etc). To the extent that an omitted variable does not vary *within a given firm-quarter* (e.g., within Intel's 2019-Q4), our analysis controls for

the omitted variable. Consistent with corporate websites providing investors with significant value-relevant information, we find a significant market reaction to increases in website content. This finding is consistent with the importance of corporate websites as a disclosure medium.

One challenge with interpreting market reaction tests is that—in a world where investors have limited attention and information processing costs—such tests do not speak to whether the market reaction is driven by the provision of novel information, or by increased dissemination of information already present in the public domain (e.g., Blankespoor, deHaan, and Marinovic, 2020). The "information or dissemination" distinction is an important one, and speaks to whether firms are using corporate websites as a true alternative information channel, as permitted by SEC guidance, or as a supplemental channel to reinforce the information provided through more traditional disclosure channels (see Blankespoor, Miller, and White, 2014 and Blankespoor, 2018 for a discussion in the context of social media).

To shed light on this question, we restrict our analysis to a setting where there are no EDGAR disclosures or press releases 15 days before or after large increases in website content. In such a setting, the website is not being used as part of an overall disclosure strategy to reinforce or disseminate information provided through other disclosure channels—because no other channels are being used. Instead, this setting isolates the cases where the website is being used as a standalone channel for disclosure. On the one hand, if the informational value of website content is derived from its ability to magnify and disseminate information provided in traditional disclosure channels, then in this setting—where such channels are silent—we expect to find a muted market reaction. On the other hand, if firms are using corporate websites to provide novel information not available from other channels, then we continue to expect a market reaction even in the absence of temporally proximate EDGAR filings and press releases. Consistent with the

latter, we find an economically and statistically significant market reaction to increases in website content—regardless of whether the firm is also providing information through other disclosure channels in the preceding or subsequent weeks.

In our second set of tests, we repeat our event study design using two market-based measures of information asymmetry (i.e., the daily bid-ask spread and the Amihud (2002) measure of illiquidity). It is well known that information asymmetry temporarily spikes around firm-initiated disclosure such as earnings announcements and EDGAR filings (e.g., Lee et al., 1993; Coller and Yohn, 1997; Amiram et al., 2016). These studies document a temporary increase in information asymmetry as potentially complex disclosure is more quickly processed by sophisticated investors than unsophisticated investors (e.g., Kim and Verrecchia, 1994).

In a similar vein, if corporate websites provide complex information that is easier for sophisticated investors to digest, we expect increases in website content are associated with a temporary increase in information asymmetry. However, if website content is easily digestible by unsophisticated investors then—in contrast to earnings announcement and EDGAR filings—we expect a reduction in information asymmetry. Consistent with the latter, we find a substantial reduction in information asymmetry around increases in website content. This finding is similar to prior work that finds corporate use of social media reduces information asymmetry (e.g., Blankespoor et al., 2014). However, unlike prior work on corporate social media, we find information asymmetry is reduced even when the website is being used as a *standalone* medium for disclosure—where there are no EDGAR disclosures or press releases in the surrounding 15 days.

Finally, in our third set of tests, we examine how the relation between corporate websites and the information environment varies over time. To address this question, we use the Wayback

Machine to reconstruct firm's historical websites every quarter, since 2004. One limitation of this analysis is that—whereas our earlier tests used daily data on website content—by virtue of relying on the Wayback Machine for website archives, these tests are necessarily at the quarterly-level. Using quarterly website archives, we compute our measure of website content for each firm-quarter as far back as 2004. We then use a *within-firm* design to estimate the relation between quarterly website content and information asymmetry in the subsequent quarter. We find that within-firm variation in website content is negatively associated with information asymmetry. This is consistent with the extant disclosure literature which finds that disclosure tends to enhance firm's information environment.

We next examine whether the importance of corporate websites for the information environment varies before and after the SEC published guidance allowing firms to use their corporate websites as an alternative to EDGAR. To do so, we estimate separate regressions relating website content to information asymmetry, before and after the 2008 SEC guidance. One can think of this design as akin to a difference-in-differences design, where the regression is estimated separately in the pre-period and post-period (rather than pooling over both periods) and inferences are drawn on the difference in the coefficients between the two regressions (see e.g., deHaan 2020). We find that the negative relation between website content and information asymmetry is roughly twice as large after the 2008 SEC guidance than before the SEC guidance.

Collectively, our findings suggest corporate websites are an important, unexamined source of information to investors, and that the information provided by such websites is often incremental to that contained in more traditional disclosure channels (e.g., EDGAR filings and press releases). These findings should be of interest to both academics and regulators. With respect to academics, our study extends a long line of research on corporate disclosure. While prior literature has

extensively studied the information content of EDGAR filings and press releases, evidence on whether and how firms use corporate websites to provide information is virtually non-existent. In this regard our findings relate to a stream of recent research examining how advances in technology have altered traditional disclosure channels (e.g., Miller and Skinner, 2015). By making our data publicly available, our study potentially opens up a new stream of research examining the use corporate websites as an information channel, how analysts and other intermediaries rely on this channel, how the cost and benefits of this channel relate to and interact with traditional channels, and how corporate websites fit in to the firm's overall disclosure strategy.[1] With respect to regulators, our study suggests corporate websites are an important unregulated information channel. To the extent that websites can provide similar information to investors as EDGAR filings—as our evidence suggests—any regulatory concerns about opportunistic disclosure and investor protection would also seem applicable to the information provided on corporate websites.

The remainder of the paper proceeds as follows. Section 2 discusses institutional features of our setting and related literature. Section 3 describes our sample and measurement choices. Section 4 describes our research design and presents results. Section 5 provides concluding remarks.

2. Background and Related Literature

Firms' use of corporate websites to communicate with stakeholders is a decades-old phenomenon. For example, internet archives show that Dow Chemical has been posting investor relations materials such as annual reports to their website since 1996 at the latest. From a regulatory

[1] We are committing to make our measures of the website content publicly available upon publication.

perspective, the SEC has acknowledged the benefits of dissemination via corporate websites since at least October 1995 (SEC Release 33-7233, p. 2):

> *"Until recently, on-line use of corporate information was generally limited to large corporations and institutional investors. The dramatic growth in personal computer ownership, however, is enabling many small investors to access on-line corporate information just as readily as institutions. Access to information through electronic means permits small investors to communicate quickly and efficiently with companies as well as with each other."*

In April 2000, the SEC openly recommended that firms use their websites to communicate with investors, even if such communication was solely for dissemination purposes, e.g., by providing hyperlinks to EDGAR from their investor relations page (SEC Release 33-7856). Two years later, in 2002, the SEC recommended that all firms disclose the address of their website in annual reports along with a statement regarding the disclosures that are made available on the website. The SEC's statements would suggest that they viewed corporate websites as an integral part of the mosaic surrounding a firm.

The SEC gradually expanded its guidance as websites and the internet progressed in importance, and explicitly acknowledged corporate websites as achieving similar status to EDGAR in terms of being a primary source of firm information. For example, on August 7, 2008, the SEC began allowing companies to substitute website disclosures for Regulation FD filings on EDGAR. The SEC stated that substitution between website disclosure and EDGAR disclosure was permitted under three conditions (SEC Release 34-58288, p. 18):

> *"1) A company web site is a recognized channel of distribution, 2) Posting of information on a company web site disseminates the information in a manner making it available to the securities marketplace in general, 3) There has been a reasonable waiting period for investors and the market to react to the posted information."*

By April 2013, the SEC responded to the rise of social media platforms built on internet technologies, and specifically Reed Hasting's use of his personal Facebook page to disclose

material information about Netflix, by issuing guidance that explicitly allows companies to use social media outlets like Facebook and Twitter to announce key information. The SEC's guidance confirmed that Regulation FD applies to social media and other emerging means of communication used by public companies the same way it applies to company websites (SEC Press Release 2013-51). Thus, understanding how corporate websites fit into firms' overall disclosure strategies, should be of interest to regulators, practitioners, and academics alike.

Despite a large and growing academic literature on corporate disclosure, we know very little about how firms use corporate websites as a disclosure channel. Perhaps most closely related, is the recent literature on corporate use of social media—primarily Twitter. Blankespoor, Miller and White (2014) examine firms' use of Twitter to disseminate and reinforce corporate financial information provided through other channels (e.g., an earnings announcement). They find a significant decrease in information asymmetry when firms use Twitter to supplement existing disclosure channels. Lee, Hutton, and Shu (2015) examine firm's use of Twitter in conjunction with a product recall announcement. They find corporate use of Twitter can reduce the negative reaction to a product recall announcement. Jung, Naughton, Tahoun, and Wang (2018) examine whether firms' strategically use Twitter to disseminate favorable news and find firms are less likely to use Twitter to disseminate bad earnings news. Finally, Campbell, Drake, Thornock, and Twedt (2020) examine the factors associated with "earnings virality"—earnings news going "viral" on Twitter, Facebook, LinkedIn, Pinterest and YouTube. They find earnings virality is associated with increased noise trading and reduced price efficiency, which they interpret as evidence that extreme social media coverage "over-broadcasting information" and acting to correlate noise trading.

While technological change in disclosure technologies has led recent papers to examine corporate use of social media (Miller and Skinner, 2015), the literature seems to have overlooked

the role of corporate websites in shaping the firm's information environment. In this paper, we take the first step toward shedding light on corporate websites as an important source of information to investors. Specifically, we examine whether increases in website content provide significant value-relevant information to investors incremental to that contained in traditional disclosure channels, and how website content affects the balance of information between sophisticated and unsophisticated investors.

3. Sample Construction and Variable Measurement

3.1. Sample

Our analysis requires data on firm websites. We identify corporate websites using firms' 10-Ks. If a firm's web address is not disclosed in their 10-K then we use the web address listed in Compustat. Using the identified web addresses, when then construct two different samples. The first sample, hereafter the *"*daily sample," is used in our daily event study tests. The second sample, hereafter the "quarterly sample," is used in our pooled regression tests that examine the relation between website content and information environment as far back as 2004.

Daily Sample. We compile the daily sample by scraping firms' websites every day prior to the market open, from July 2019 through October 2019. The window of time for the daily sample is limited because of the sheer amount of data involved with scraping and saving over 5,000 corporate websites and associated interior pages every day. The tradeoffs for this sample are clear. On the one hand, daily observations allow our event study tests to tightly identify shifts in website content and corresponding capital market effects, and to disentangle the effect of these shifts on capital markets from that of temporally proximate EDGAR filings or press releases. On the other

hand, the enormous mass of data needed to collect daily observations of corporate websites precludes an analysis of daily data over any horizon greater than a few months.[2]

Quarterly Sample. We compile this sample by using the Wayback Machine to obtain archived copies of corporate websites every quarter from 2004 to 2018. This sample provides a meaningful time-series of website content before and after various SEC guidance on the use of corporate websites. The tradeoffs for this sample are also clear. On the one hand, quarterly observations of websites over two decades maximizes sampling variation in website content, and the resulting variation generalizes to a greater span of time. This allows our tests to examine whether the relation between website content and the information environment has changed overtime, and specifically in relation to SEC rules and guidance. On the other hand, because the observations are quarterly, it will be more difficult to control for confounding firm events that occur in the same fiscal quarter (e.g., EDGAR filings and press releases).

Importantly, the quarterly sample and associated analysis is designed to complement the daily sample and associated event study. By using both samples, and presenting results across both sets of tests, our collective analysis should mitigate concerns about generalizability and specific identification that are inherent to any single sample or any one particular set of tests.

3.2. Measuring Website Content

Corporate websites consist of a combination of text, images, and/or videos, as well as tabular data and hyperlinks. Some companies rely more heavily on text-based content while others use less text and rely more on graphical content. For example, Tesla's website contains a combination of text, images, links, and dynamic elements (See Figure 1 Panel A). In contrast, Berkshire Hathaway's website contains almost exclusively text and links (See Figure 1 Panel B).

[2] Three months of daily data on all corporate websites and associated interior pages consumes about 150 terabytes of raw storage.

Accordingly, to strike a balance between the various content elements, similar to Bonsall, Leone, and Miller (2017), we measure five fundamental attributes of the website—character length, tags, links, multimedia content, and dynamic scripting. *Length* is the number of characters on the website.[3] *Formatting* is the count of the number of HTML tags that are associated with formatting the content on websites, e.g., headings, paragraphs, etc. *Links* is the count of hyperlinks on the website. *Rich* is the count of tags associated with displaying content such as audio, images, and video. Lastly, *Dynamic* is the count of HTML tags associated with enabling interactive page content such as tables with sorting and filtering functionality.

Panel A of Table 1 presents descriptive statistics for our measures for the daily sample of 198,756 unique website-days. Panel A indicates that, on average, corporate websites in our sample are 3,716 characters long, contain 397 formatting related tags, 127 links, 38 pieces of rich content, and 23 dynamic tags.

Next, we use factor analysis to analyze the common component to these five content elements. Table 1 Panel B presents results from our factor analysis. Variables are standardized to be mean zero and unit standard deviation prior to this analysis. Panel B shows that the first principal factor explains 93% of the variation in our five content proxies. The first factor also loads positively on all five content elements (loadings of 0.40, 0.88, 0.78, 0.47, and 0.24 on *Length*, *Formatting*, *Links*, *Rich*, and *Dynamic* respectively) and is the only factor with an eigenvalue greater than one. In the remainder of our analysis, we use this first principal factor as our measure of website content, *WebContent*. Table 1 Panel C presents descriptive statistics for

[3] *Length* is based purely on the actual text on the webpage and does not include any characters that are part of the website's code, e.g., HTML tags.

our measure of website content. The mean and median values of *WebContent* are 0.00 and -0.29 respectively, suggesting significant right skew in website content.[4]

4. Daily Event Study

4.1. Descriptive Statistics

We use a standard event study to examine whether corporate websites contain information useful for investors that is not conveyed through traditional disclosure channels. To conduct our event study, we focus on firm-days where there was a large increase in content on the firm's website but no other forms of disclosure (i.e., EDGAR filings or press releases) within a specific window around the website change. Specifically, we compute daily changes in website content (Δ*WebContent*), and define an "event" (i.e., day 0) as a circumstance where Δ*WebContent* is in the top tercile, and there are no press releases or EDGAR filings within the [–1,+1] window around the increase in *WebContent*.[5] This results in a sample of 11,294 unique events with large increases in *WebContent* and no contemporaneous corporate disclosures via other channels.

For each event, we then collect data for days [–30,+30] around the increase in website content. We collect stock market data from CRSP and the NYSE Trade and Quote (TAQ) databases, financial statement data from Compustat, data on press releases from RavenPack, and data on 8-K filings from SEC EDGAR. We exclude firms with missing data for total assets, net income, and stock returns. The resulting sample consists of 699,060 firm-days.

[4] *WebContent* is standardized by construction. The mean of a linear combination of standardized variables is zero.
[5] In untabulated analyses, we find our results are robust to defining events based on above median changes in *WebContent*, top quartile changes in *WebContent*, and top decile changes in *WebContent*.

Table 2 Panel A presents descriptive statistics for the variables used in our daily event study tests. Average daily trading volume is 0.88% of shares outstanding, average bid-ask spread is 0.33% of price and 12% (18%) of observations are associated with an EDGAR disclosure (press release). Table 2 Panel B presents pearson correlations between our measures of website content and the variables used in our analysis. Notably, our measures of website content and its components are only weakly correlated with other firm characteristics. In absolute terms, *WebContent* is most correlated with *Volume*, *Returns*, and *Spread* (correlation coefficients of 0.15, -0.07, and -0.06 respectively).

4.2. Information Content of Corporate Websites

In our first set of tests, we use a standard short-window event study to examine the market reaction to increases in the content of corporate websites. Following extant literature, we measure the market reaction using the absolute value of daily returns expressed as a percent, *AbsoluteReturn*, and trading volume as a percent of shares outstanding, *Volume*. Figure 4 Panel A (Panel B) plots the mean abnormal *AbsoluteReturn* (*Volume*) around the website change.[6] Both panels show a pronounced increase in price and volume reaction in the ten days after the website change.

To test whether the price and volume reactions are statistically different than the non-event day window, we estimate the following OLS regression, pooling across all firm-days in the [-30, +30] window around the increase in website content:

$$Outcome_{i,t+1} = \alpha + \beta \cdot EventDay[0,10] + \gamma \cdot Controls_{i,t} + \varepsilon_{i,t} \quad (1)$$

where *Outcome* is a measure of the market reaction, *EventDay[0,10]* is an indicator variable equal to one for days zero through 10 after the increase in website content, and following

[6] For the purposes of Figure 2, we regress each measure on the vector of controls in Eq. (1) and plot the standardized residuals.

Bushee et al. (2018), *Controls* is a vector of control variables including *DailyEDGAR*, Daily*PR*, *DailyReturn*, *DailyVolat*, *Size*, *Leverage*, *MTB*, *AdExp*, *CorpAcq*, *CapIntens*, *CapEx*, *R&D*, *Financing*, *Loss*, *Returns*, *IdioVol*, *Surprise*, *SpecItems*, and *SmallBeat*. All variables are defined in Table 2. The coefficient of interest is β. This coefficient represents the difference in the dependent variable between the non-event period (e.g., days -30, …, -1, 11, …, 30) and the event period (e.g., days 0, …, 10). Throughout our analyses, we calculate standard errors clustered by firm and date which allows for arbitrary correlations across time within a given firm and across firms within a given date.

For each outcome measure, we estimate two versions of equation (1). We estimate the first version using firm and date fixed effects and the second version using firm-quarter (e.g., Intel's 2009-Q4) and date fixed effects. The firm-quarter fixed effects subsume any variables that are measured at either an annual or quarterly frequency, e.g., corporate governance, quarterly earnings surprise, etc. The former specification focuses on within firm variation in information content, while the latter focuses exclusively on within *firm-quarter* variation.

Table 3 presents results. Across all specifications we find a statistically and economically significant increase in absolute returns and trading volume. Specifically, the *t*-statistics on *EventDay[0,10]* vary from 6.89 to 8.01 and the size of the coefficient on *EventDay[0,10]* ranges from 0.45 to 0.48 whereas the size of the coefficient on *DailyPR* ranges from 0.43 to 1.16, and the size of the coefficient on *DailyEDGAR* ranges from 3.23 to 4.47. Thus, the economic significance of the market reaction to corporate website content is of a similar order of magnitude as press releases, but is an order of magnitude smaller than that of EDGAR disclosure.

4.3. Dissemination Versus New Information

One challenge with interpreting market reaction tests is that—in a world where investors have limited attention and information processing costs—such tests do not speak to whether the market reaction is driven by the provision of novel information, or by increased dissemination of information already present in the public domain. To distinguish between these two possibilities, we examine increases in corporate website content when there is no SEC filing or press release in the surrounding period. This design isolates changes in website content independent of other disclosure channels and corporate events.

To test whether the market reaction to website content is attributable to a dissemination effect or the release of new information, we re-estimate the within-firm-quarter regressions of Table 3 restricting the sample to a change in website content where there have been no EDGAR filings or press releases within a [-*s*, +*s*] window around the event date, where s $\in$ {5, 10, 15}. Table 4 presents results. Across all specifications we find that our results are generally unchanged. Specifically, for the baseline within-firm-quarter specification in Table 3, when the dependent variable is *AbsoluteReturn* (*Volume*) the coefficient on *EventDay[0,10]* is 0.48 (0.45). The rows of Table 4 show that as *s* increases, and the dependent variable is *AbsoluteReturn* (*Volume*), the coefficient on *EventDay[0,10]* ranges from 0.53 to 0.63 (0.30 to 0.55). Across all specifications the coefficient remains highly statistically significant. Thus, finding similar, and in some cases stronger results in the absence of EDGAR disclosure and press releases in the surrounding two weeks, suggests that firms are using their websites to provide novel information not present in other disclosure channels.

4.4. Information Asymmetry around Corporate Website Changes

In our second set of tests, we use a standard short-window event study to examine how increases in the content of corporate websites affect information asymmetry. If unsophisticated

investors can readily process information on corporate websites, we expect a decrease in information asymmetry around the increase in website content. In contrast, if website information is complex and requires significant processing costs, then, similar to prior work on earnings announcements, management forecasts, and 10-K filings (Lee et al., 1993; Coller and Yohn, 1997; Amiram et al., 2016), we expect to observe an increase in information asymmetry around changes in website content.

We measure information asymmetry using *AmihudIlliquidity*, defined as the absolute value of daily return divided by the dollar volume traded multiplied by one million, and *EffectiveSpread*, defined as twice the absolute difference between the trade execution price and the midpoint, scaled by the midpoint and averaged over all trades during date *t*, and expressed as a percent. Figure 5 Panel A (Panel B) plots the mean abnormal *AmihudIlliquidity* (*EffectiveSpread*) around the website change.[7] Both panels show a pronounced decrease in information asymmetry in the ten days after the website change. To test whether the level of information asymmetry during the event day window is statistically different than the level during the non-event day window, after controlling for various firm characteristics associated with information asymmetry, we re-estimate equation using *AmihudIlliquidity* and *EffectiveSpread* as the dependent variables.

Table 5 presents results. Across all specifications we find a statistically and economically significant decrease in information asymmetry around the increase in website content. Specifically, the *t*-statistics on *EventDay[0,10]* vary from -5.18 to -4.94 and the size of the coefficient on *EventDay[0,10]* ranges from -0.12 to -0.06. The negative coefficient on *EventDay [0,10]* in Table 5 suggests that increases in corporate website content alleviate information asymmetry between investors. This stands in contrast to prior work that has found a spike in information asymmetry

[7] For the purposes of Figure 3, we regress each measure on the vector of controls in Eq. (1) and plot the standardized residuals.

leading up to and during information events related to earnings announcement and financial statements (Lee et al., 1993; Amarim et al., 2016). Consistent with these studies, we find the coefficients on *DailyEDGAR* and *DailyPR* are positive (and range from 0.03 to 0.41 for *DailyPR,* and 0.07 to 0.39 for *DailyEDGAR*).

We test whether our information asymmetry results are driven by corporate websites serving a dissemination role rather than providing novel information using the same methodology outlined previously: specifically we re-estimate the results in Table 5 while lengthening the width of the non-disclosure window from one week around the increase in website content, to two weeks around the increase in website content. Table 6 presents results. Across all specifications we find our results are quantitatively similar. Notably, in all cases the economic significance increases over the baseline specifications presented in Table 5, suggesting one rationale for using corporate websites as a standalone disclosure channel is that it magnifies the reduction in information asymmetry as compared to a circumstance where website disclosure is paired with an EDGAR disclosure or press release.

5. Corporate Websites Over Time

The results of our daily event studies suggest that corporate websites contain information content that is incremental to the disclosure channels traditionally examined in the literature, and that the provision of information on corporate websites significantly affects the firm's information environment. In our third set of tests, we attempt to answer whether the website's effect on the information environment is short or long-lived and how its effect has changed over time as website and internet technologies have advanced—specifically after August 2008 when the SEC officially allowed website disclosure as a substitute for 8-Ks.

Addressing this question requires data on website content over a long time-series. To address this question, we collect new data using the Wayback Machine. Specifically, we collect archived data on firm websites from 2004 until 2018. Unlike our earlier tests that employ daily data, archives are only available at the quarterly frequency. Based on this archived data we compute our measure of website content (*WebContent*) at the quarterly level as far back as 2004. We then merge these data with stock market data from CRSP and the NYSE TAQ databases, financial statement data from Compustat, data on press releases from RavenPack, and data on 8-K filings from SEC EDGAR. We exclude firms with missing data for total assets, net income, and stock returns. The resulting sample consists of 178,689 firm-quarter observations from 2004 to 2018.[8]

5.1. Quarterly Analysis of Website Changes and Information Asymmetry

To test whether changes in website content have a statistically significant effect on information asymmetry after controlling for a variety of firm characteristics associated with the economic activities of the firm, we estimate the following ordinary least squares (OLS) regression, pooling across all firm-quarter observations:

$$InfoAsymm_{i,q+1} = \alpha + \beta \cdot WebContent_{i,q} + \gamma \cdot Controls_{i,q} + \varepsilon_{i,q} \quad (2)$$

where *InfoAsymm* is one of our two daily measures of information asymmetry averaged over quarter q+1, *WebContent* is our measure of website content, *Controls* is a vector of control variables including *8Ks*, *PRs*, *Size*, *Leverage*, *MTB*, *AdExp*, *CorpAcq*, *CapIntens*, *CapEx*, *R&D*, *Financing*, *Loss*, *Returns*, *IdioVol*, *Surprise*, *Turnover*, *SpecItems*, *SmallBeat*, and *InvPrice*, and firm and year-quarter fixed effects (Bushee et al., 2018).[9] *8Ks* is the number of 8-Ks filed during

[8] The start date of our sample is limited by the availability of press releases through RavenPack. Our results are robust to excluding press releases and starting the sample in 1997.
[9] Our results are robust to other fixed effects structures such as industry and year-quarter.

the quarter, *PRs* is the number of press releases filed during the quarter, *Turnover* is the dollar volume of shares traded during the quarter divided by the market value of the firm at the start of the quarter, and *InvPrice* is the inverse of the firm's price at the end of the quarter. All other variables are defined in Tables 1, and standard errors are clustered by firm and quarter.

To facilitate interpretation of our results, similar to Armstrong et al. (2012) and Bushee et al. (2018), we use the scaled decile rank of the independent variables in our regression analysis (scaled decile ranks range from 0 to 1). We use the decile rank of each independent variable to ensure that all independent variables are of similar scale. This, in turn, allows us to meaningfully compare the relative economic significance of each variable. An added advantage is that the specification is robust to both outliers and nonlinearities. The coefficient of interest is β. This coefficient represents the change in information asymmetry attributable to moving from the lowest decile of website content to highest decile of website content.

We estimate three versions of Eq (2). First, we estimate Eq (2) pooling across all firm-quarters from 2004 to 2018. Next, we re-estimate equation (2) before and after the SEC began allowing firms to substitute disclosure on their websites for SEC filings in August of 2008. As part of the change in SEC policy, firms also became responsible for the accuracy of all content posted to their websites. To explore how this change in policy affected firms' use of corporate websites, we re-estimate equation (2) before and after the policy change. By re-estimating equation (2) separately for each period—pre-August 2008 and post-August 2008—we allow the coefficients on our control variables to vary between the two samples. We draw inferences on the difference in the coefficient on *WebContent* between the two periods. One can think of this design as akin to a difference-in-differences design, where the regression is estimated separately in the pre-period and post-period (rather than pooling over both periods) and inferences are

drawn on the difference in the coefficients between the two regressions (see e.g., deHaan 2020).[10]

Table 7 presents results. Panel A (Panel B) presents results for *AmihudIlliquidityQtrly* (*EffectiveSpreadQtrly*). Across both panels we find a robust negative relation between website content and information asymmetry, and that the negative relation is roughly twice as large after the 2008 SEC guidance than before the SEC guidance. Specifically, when the dependent variable is *AmihudIlliquidityQtrly* (*EffectiveSpreadQtrly*), the coefficient on *WebContent* increases from -0.92 during the pre-period to -1.95 during the post-period (-0.88 during the pre-period to -1.38 during the post-period). A two-tailed test for the difference in coefficients before and after the change in regulation is significant (*p*-values of 0.018 and 0.033 respectively).

6. Conclusion

While numerous academic surveys catalog how investors respond to traditional forms of disclosure such as EDGAR filings, management forecasts, and press releases, academic study of whether and how firms use corporate websites to communicate to stakeholders is virtually non-existent. In this paper, we take the first step toward shedding light on whether corporate websites are an important source of information to investors, whether such websites provide novel information or merely reinforce information provided through traditional disclosure channels, and whether information on corporate websites exacerbates or mitigates the information advantage of sophisticated investors.

[10] Here the coefficient on *WebContent* represents the first difference (difference in information asymmetry between firms with high and low website content) and the difference in the coefficients between the two periods represents the second difference.

We begin our analysis by developing a novel measure of website content that is broadly applicable to the universe of firms appearing on standard research databases (e.g., Compustat and CRSP). We then estimate three distinct sets of tests that represent joint tests of our measures of website content and our economic predictions.

In our first set of tests, we use an event study design to examine the market reaction to large increases in website content. We find (i) an economically and statistically significant market reaction to increases in website content; (ii) that the magnitude of the price and volume reaction is, on average, similar to that of press releases, and (iii) that the reaction is economically and statistically significant even in a circumstance when firms use their website as a standalone disclosure medium rather than as a supplemental channel to reinforce information also provided through more traditional disclosure channels. The evidence suggests that corporate websites provide significant value-relevant information to investors—information that is incremental to that contained in traditional disclosure channels.

In our second set of tests, we examine how website content affects the balance of information between sophisticated and unsophisticated investors. In contrast to the prior literature documenting a spike in information asymmetry around corporate filings and earnings announcements, we find evidence of a sharp reduction in information asymmetry around increases in website content. This finding is similar to prior work that finds corporate use of social media reduces information asymmetry. However, unlike the findings on corporate social media, we find information asymmetry is reduced even when the website is used as a *standalone* disclosure medium.

Finally, in our third set of tests, we examine how the relation between corporate websites and the information environment varies over time. We use the Wayback Machine to reconstruct

firms' historical websites every quarter, since 2004. We then estimate the relation between quarterly website content and information asymmetry. We find a robust negative relation between website content and information asymmetry, and that the negative relation is roughly twice as large after the SEC issued guidance allowing firms to use corporate websites as an alternative to EDGAR.

Collectively, our findings suggest corporate websites are an important, unexamined source of information to investors, and that the information provided by such websites is often incremental to that contained in more traditional disclosure channels (e.g., EDGAR filings and press releases). By making our data publicly available, we aim to potentially open up a new stream of research related to how corporate websites fit in to the firm's overall disclosure strategy, how the cost and benefits of this channel relate to that of traditional channels, and how analysts and other intermediaries rely on this channel.

References

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets*, *5*(1), 31–56. https://doi.org/10.1016/S1386-4181(01)00024-6

Amiram, D., Owens, E., & Rozenbaum, O. (2016). Do information releases increase or decrease information asymmetry? New evidence from analyst forecast announcements. *Journal of Accounting and Economics*, *62*(1), 121–138. https://doi.org/10.1016/j.jacceco.2016.06.001

Armstrong, C. S., Blouin, J. L., & Larcker, D. F. (2012). The incentives for tax planning. *Journal of Accounting and Economics*, *53*(1), 391–411. Retrieved from https://econpapers.repec.org/RePEc:eee:jaecon:v:53:y:2012:i:1:p:391-411

Beyer, A., Cohen, D. A., Lys, T. Z., & Walther, B. R. (2010). The financial reporting environment: Review of the recent literature. *Journal of Accounting and Economics*, *50*(2–3), 296–343. https://doi.org/10.1016/j.jacceco.2010.10.003

Blankespoor, E. (2018). Firm communication and investor response: A framework and discussion integrating social media. *Accounting, Organizations and Society*, *68–69*, 80–87. https://doi.org/10.1016/j.aos.2018.03.009

Blankespoor, E., Ed deHaan, & Marinovic, I. (2020). Disclosure Processing Costs, Investors' Information Choice, and Equity Market Outcomes: A Review. *Journal of Accounting and Economics*, 101344. https://doi.org/10.1016/j.jacceco.2020.101344

Blankespoor, E., Miller, G. S., & White, H. D. (2014). The role of dissemination in market liquidity: Evidence from firms' use of Twitter™. *Accounting Review*, *89*(1), 79–112. https://doi.org/10.2308/accr-50576

Bushee, B. J., Gow, I. D., & Taylor, D. J. (2018). Linguistic Complexity in Firm Disclosures: Obfuscation or Information? *Journal of Accounting Research*, *56*(1), 85–121. https://doi.org/10.1111/1475-679X.12179

Coller, M., & Yohn, T. L. (1997). Management Forecasts and Information Asymmetry: an Examination of Bid-Ask Spreads. *Journal of Accounting Research*, *35*(2), 181–191.

deHaan, E. (2020). Practical Guidance on Using and Interpreting Fixed Effects Models. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3699777

Drake, M. S., Thornock, J. R., & Twedt, B. J. (2020). *Earnings Virality*.

Dyer, T., Lang, M., & Stice-Lawrence, L. (2017). The evolution of 10-K textual disclosure: Evidence from Latent Dirichlet Allocation. *Journal of Accounting and Economics*, *64*(2–3), 246–252. https://doi.org/10.1016/j.jacceco.2017.07.004

Jung, M. J., Naughton, J. P., Tahoun, A., & Wang, C. (2018). Do firms strategically disseminate? evidence from corporate use of social media. *Accounting Review*, *93*(4), 225–252. https://doi.org/10.2308/accr-51906

Kim, O., & Verrecchia, R. E. (1994). Market liquidity and volume around earnings announcements. *Journal of Accounting and Economics*, *17*(1–2), 41–67. https://doi.org/10.1016/0165-4101(94)90004-3

Lee, C. M. C., Mucklow, B., & Ready, M. J. (1993). Spreads, Depths , and the Impact of Earnings Information : An Intraday Analysis Author (s): Charles M . C . Lee , Belinda Mucklow and Mark J . Ready Source : The Review of Financial Studies , Vol . 6 , No . 2 (1993),. *The Review of Financial Studies*, *6*(2), 345–374.

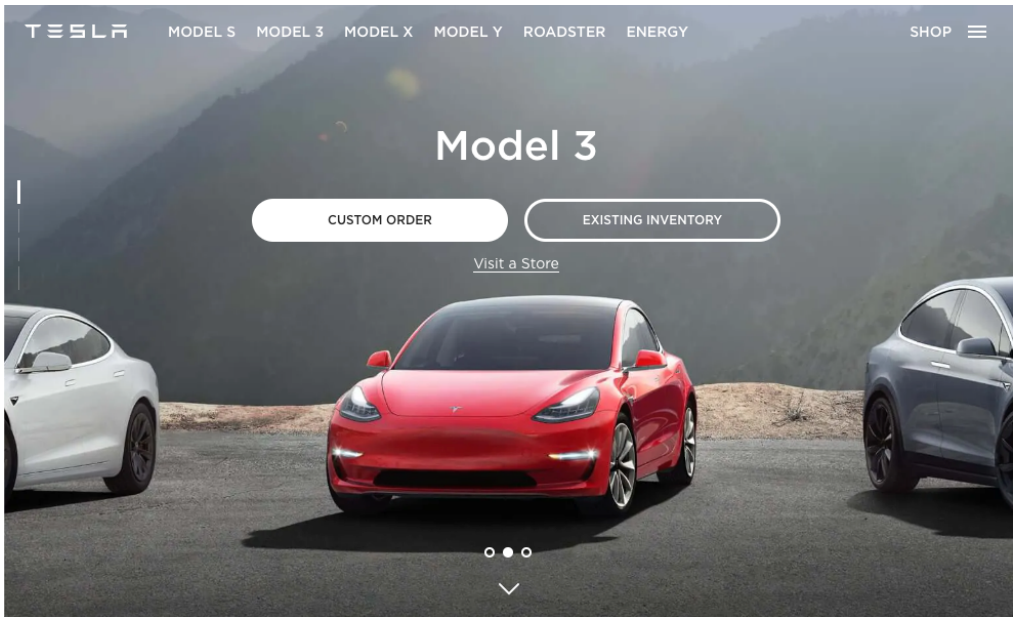
Lee, L. F., Hutton, A. P., & Shu, S. (2015). The Role of Social Media in the Capital Market: Evidence from Consumer Product Recalls. *Journal of Accounting Research*, *53*(2), 367–404. https://doi.org/10.1111/1475-679X.12074

Leuz, C., & Wysocki, P. D. (2016). The Economics of Disclosure and Financial Reporting Regulation: Evidence and Suggestions for Future Research. *Journal of Accounting Research*, *54*(2), 525–622. https://doi.org/10.1111/1475-679X.12115

Miller, G. S., & Skinner, D. J. (2015). The evolving disclosure landscape: How changes in technology, the media, and capital markets are affecting disclosure. *Journal of Accounting Research*, *53*(2), 221–239. https://doi.org/10.1111/1475-679X.12075

Figure 1. Example Observations of Corporate Websites

Panel A. Example of a Primarily Rich Content Website (Tesla, 08/13/2019)



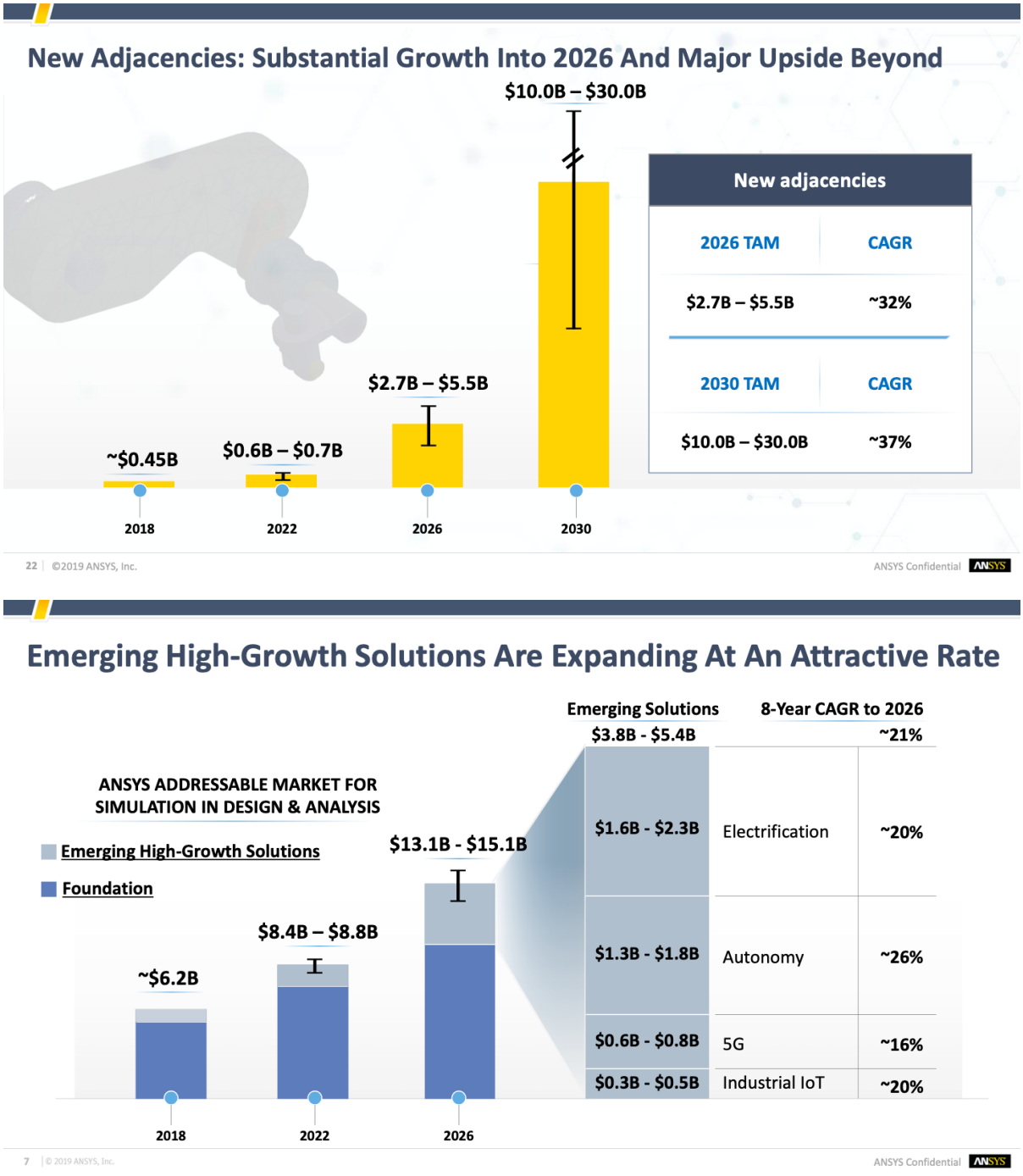
Panel B. Example of a Primarily Text Content Website (Berkshire Hathaway, 08/13/2019)

BERKSHIRE HATHAWAY INC.
3555 Farnam Street
Omaha, NE 68131
Official Home Page

- A Message From Warren E. Buffett
- Annual & Interim Reports
  Updated August 3, 2019
- Special Letters From Warren & Charlie RE:Past, Present and Future
- Link to SEC Filings
- Links to Berkshire Subsidiary Companies
- Corporate Governance
- Sustainability
- Owner's Manual
- Letters from Warren E. Buffett Regarding Pledges to Make Gifts of Berkshire Stock
- News Releases
  Updated August 3, 2019
- Warren Buffett's Letters to Berkshire Shareholders
  Updated February 23, 2019
- Charlie Munger's Letters to Wesco Shareholders
- Annual Meeting Information
- Celebrating 50 Years of a Profitable Partnership
  (A commemorative book first sold at the 2015 Annual Meeting and now for sale on eBay.)
- Comparative Rights and Relative Prices of Class A and B Stock
- Berkshire Activewear

This figure contains two observations from our daily sample of corporate websites. Panel A presents a website which utilizes a large amount of *Rich* and *Dynamic* elements. Panel B presents a website that utilizes a large amount of text-based content, measured by *Length* and *Formatting*.

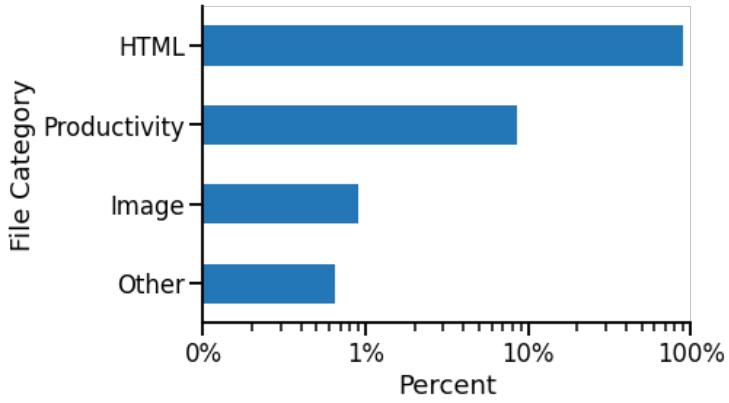
Figure 2. Example of Novel Website Disclosure



This figure contains two slides selected from the hundreds that were published to https://investors.ansys.com/home/default.aspx on September 6, 2019 in preparation for an investor conference the following week. The changes to this portion of their website included more than ten separate presentations given by the CEO and numerous Vice Presidents. This content was not posted to EDGAR, nor did ANSYS post anything to EDGAR about the conference in the time period around the conference (August or September 2019).

Figure 3. Content of Large Website Changes

Panel A. Distribution of Content Types



Panel B. Distribution of Topics in Website Changes and EDGAR Filings

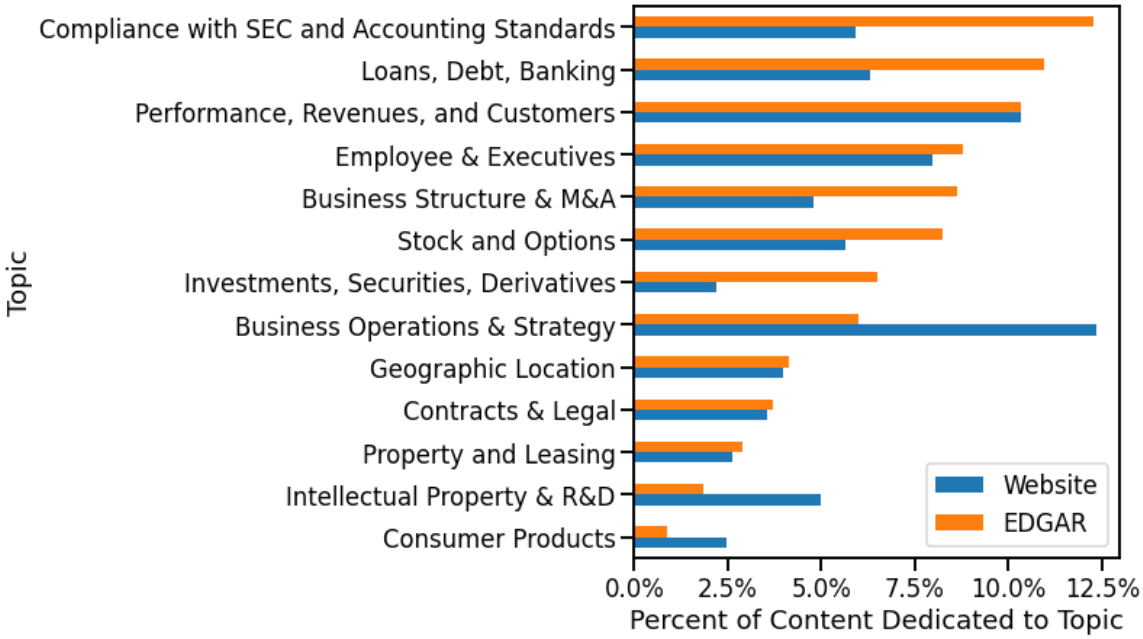
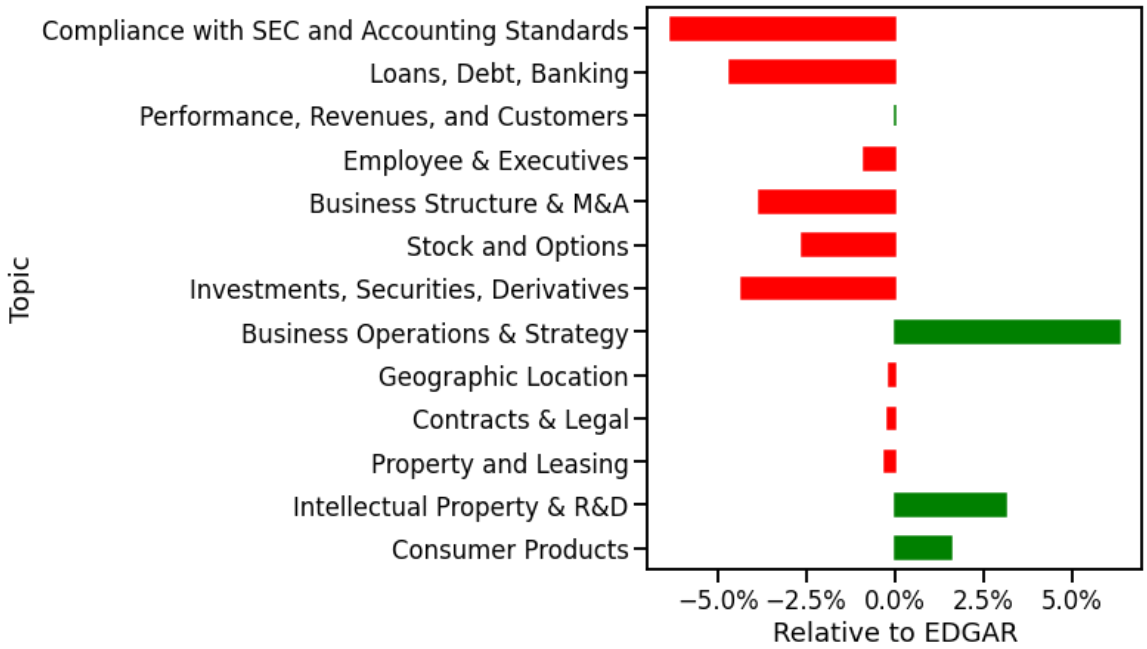


Figure 3. Content of Large Website Changes (cont'd)

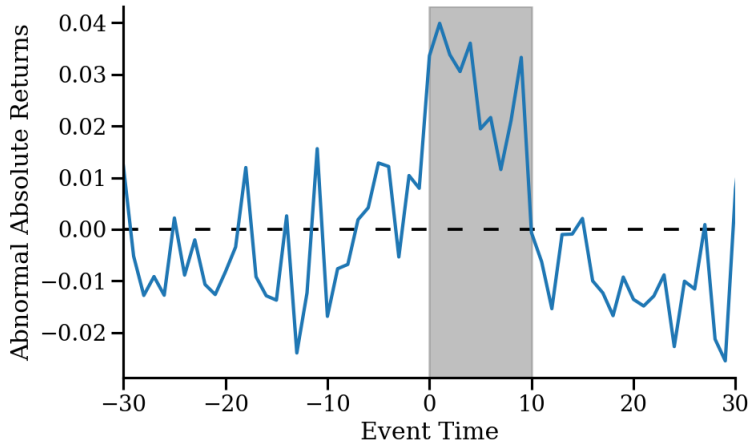
Panel C. Distribution of Topics Relative to EDGAR



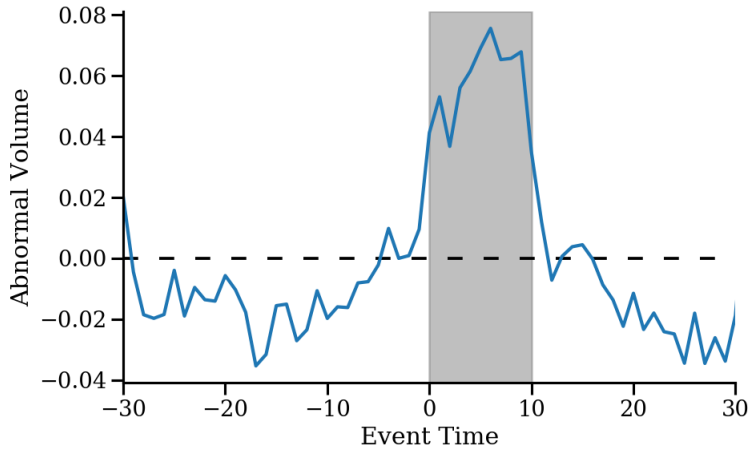
This figure presents descriptive statistics on the content contained in large website changes. Panel A presents the distribution of file types in the website changes; *HTML* includes file extensions that are associated with human readable webpages that are either static or dynamic (ashx, asp, aspx, axd, cfm, cgi, cn, com, do, fcgi, htm, html, jsp, jspx, page, php, shtml, zhtml), *Productivity* includes file extensions that are related to Office documents, PDFs, engineering drawings, etc. (csv, cw, doc, docx, dot, dwg, dxf, ics, lp, pdf, ppt, pptx, txt, xls, xlsx, zip), *Image* includes file extensions related to static visual content (eps, gif, jpg, png, svg), *Other* includes file extensions which are observed but not contained in the previous categories (action, dll, exe, faces, json, lpc, mi, mp3, mp4, xhtml, xml, xsd). Panel B presents the distribution of topics, per Dyer et al. (2017), in website changes and 10-K, 10-Q, and 8-K filings posted to EDGAR during our sample period. Panel C presents the difference in the distribution of topics in website changes relative to 10-K, 10-Q, and 8-K filings posted to EDGAR during our sample period.

Figure 4. Event Study: Information Content

Panel A. Abnormal *AbsoluteReturn* around an Increase in Website Content



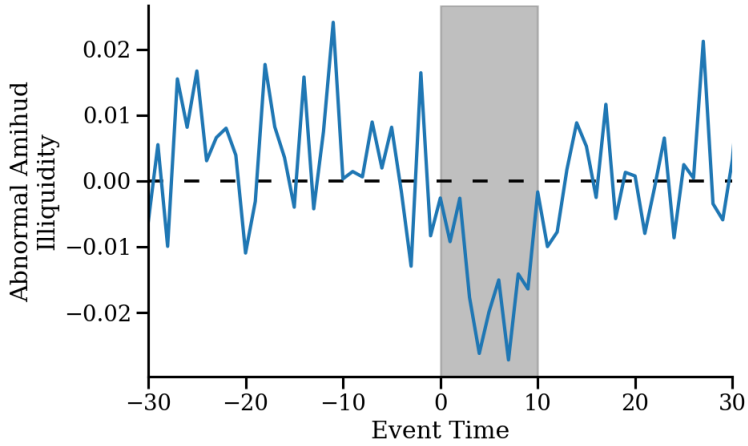
Panel B. Abnormal *Volume* around an Increase in Website Content



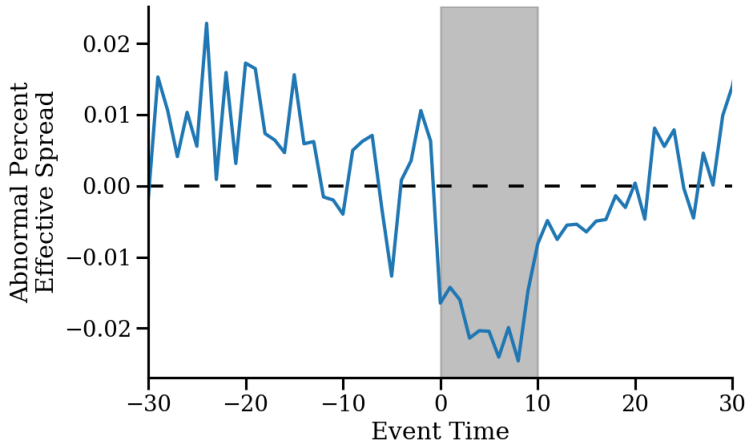
This figure plots the standardized residuals from regressions of absolute value of daily returns and daily trading volume on the control variables in Table 3 Column 1. Day 0 represents the day of the increase in website content, and the shaded area represents the [0,+10] event window. Panel A presents results when the dependent variable is *AbsoluteReturn*. Panel B presents results when the dependent variable is *Volume*. Sample of 699,060 firm-days in the [–30, +30] day window around the increase in website content.

Figure 5. Event Study: Information Asymmetry

Panel A. Abnormal *Illiquidity* around an Increase in Website Content



Panel B. Abnormal *Spread* around an Increase in Website Content



This figure plots the standardized residuals from regressions of illiquidity and bid-ask spread on the control variables in Table 3 Column 1. Day 0 represents the day of the increase in website content, and the shaded area represents the [0,+10] event window. Panel A presents results when the dependent variable is *Illiquidity*. Panel B presents results when the dependent variable is *Spread*. All variables are defined in Table 2. Sample of 699,060 firm-days in the [–30, +30] day window around the increase in website content.

Table 1. Website Content

Panel A. Descriptive Statistics

Variable	Mean	Std	Median
Length	3716.49	7847.95	1899.00
Formatting	396.53	413.86	284.00
Links	126.50	140.39	90.00
Rich	38.16	94.51	17.00
Dynamic	22.80	31.71	15.00

Panel B. Factor Analysis Output

Factor	Eigenvalue	Proportion of the variation explained	Cumulative proportion of the variation explained	First Factor Loading	
				Variable	Weight
1st	1.77	93.0%	93.0%	*Length*	0.398
2nd	0.12	6.5%	99.5%	*Formatting*	0.876
3rd	0.01	0.5%	100.0%	*Links*	0.782
				Rich	0.471
				Dynamic	0.236

Panel C. Descriptive Statistics for Latent Variable Content

Variable	Mean	Std	Median
WebContent	0.00	1.00	-0.29

This table describes our measure of website content (*WebContent*). We measure website content using the first factor from a factor analysis of the number of characters on the website (*Length*), the count of formatting tags (*Formatting*), the count of links (*Links*), the count of tags related to displaying images, audio, and video (*Rich*), and the number of dynamic assets (*Dynamic*). All variables are standardized prior to the factor analysis. Panel A presents descriptive statistics for the observed measures of content prior to standardization. Panel B presents the factor analysis output. Panel C presents descriptive statistics for *WebContent* (*WebContent* = 0.398 * *Length* + 0.876 * *Formatting* + 0.782 * *Links* + 0.471 * *Rich* + 0.236 * *Dynamic*). Daily sample of 198,756 unique website-days.

Table 2. Descriptive Statistics

Panel A. Distribution of Variables used in the Event Study

Variable	Mean	Std	P25	Median	P75
AbsoluteReturn	1.48	1.58	0.44	0.97	1.90
Volume	0.88	1.05	0.30	0.56	1.03
Illiquidity	4.31	9.05	0.14	0.75	3.62
Spread	0.33	0.65	0.05	0.09	0.26
DailyEDGAR	0.12	0.33	0.00	0.00	0.00
DailyPR	0.18	0.38	0.00	0.00	0.00
DailyReturn	0.00	0.07	-0.02	0.00	0.03
DailyVolat	0.00	0.00	0.00	0.00	0.00
Size	7.56	2.13	6.02	7.57	8.98
Leverage	0.63	0.27	0.45	0.63	0.81
MTB	1.95	1.94	1.03	1.32	2.19
AdExp	3.84	106.91	0.01	0.04	0.15
CorpAcq	0.29	0.45	0.00	0.00	1.00
CapIntens	0.38	1.98	0.05	0.16	0.42
CapEx	0.02	0.39	0.00	0.01	0.02
R&D	1.77	50.95	0.00	0.01	0.09
Financing	0.10	1.26	0.00	0.01	0.08
Loss	0.23	0.42	0.00	0.00	0.00
Returns	0.00	0.19	-0.09	0.01	0.09
IdioVol	0.02	0.01	0.01	0.02	0.02
Surprise	0.00	0.21	-0.01	0.00	0.00
SpecItems	-0.01	0.05	0.00	0.00	0.00
SmallBeat	0.03	0.16	0.00	0.00	0.00

Table 2. Descriptive Statistics (cont'd)

Panel B. Correlations Between Website Content and Variables Used in the Event Study

	Composite Measure	*Individual Components*				
Variable	*WebContent*	*Length*	*Formatting*	*Links*	*Rich*	*Dynamic*
AbsoluteReturn	0.02	0.03	0.01	0.02	0.01	0.04
Volume	0.15	0.04	0.09	0.14	0.11	0.08
Illiquidity	-0.03	0.00	-0.02	-0.02	-0.03	0.00
Spread	-0.06	0.05	-0.05	-0.07	-0.06	0.00
DailyEDGAR	-0.01	0.00	-0.01	-0.01	0.00	-0.01
DailyPR	0.02	-0.02	0.00	0.03	0.01	-0.02
DailyReturn	0.01	-0.01	0.00	0.01	0.02	0.00
DailyVolat	0.02	0.00	0.02	0.03	0.00	0.04
Size	0.05	-0.01	0.05	0.04	0.00	-0.03
Leverage	0.01	0.01	0.00	-0.01	0.00	-0.02
MTB	-0.01	-0.01	0.00	-0.03	-0.04	0.04
AdExp	-0.02	-0.01	-0.01	-0.01	-0.01	0.00
CorpAcq	-0.01	0.01	-0.02	-0.02	-0.02	0.00
CapIntens	-0.05	-0.01	-0.04	-0.02	-0.02	-0.01
CapEx	-0.01	0.00	-0.01	-0.01	0.00	0.00
R&D	-0.02	-0.01	-0.01	-0.01	-0.01	-0.01
Financing	-0.01	0.00	-0.01	-0.01	0.00	0.00
Loss	0.03	0.07	0.00	0.00	0.01	0.04
Returns	-0.07	0.01	-0.05	-0.11	-0.01	-0.07
IdioVol	0.03	0.03	0.02	0.02	0.03	0.06
Surprise	-0.02	0.00	-0.02	-0.01	-0.01	-0.01
SpecItems	-0.03	-0.02	-0.01	-0.01	-0.03	-0.03
SmallBeat	-0.04	-0.03	-0.03	-0.05	0.01	-0.03

This table presents descriptive statistics for the variables used in our daily event study tests of website content. Panel A presents the distribution of variables used in our tests. Panel B presents pearson correlations between our measures of website content and the variables used in our tests. Sample of 699,060 firm-days. *AbsoluteReturn* is the absolute value of daily returns, expressed as a percent. *Volume* is the ratio of daily volume of shares traded to shares outstanding, expressed as a percent. *Illiquidity* is the daily absolute return divided by the daily dollar trading volume multiplied by one million. *Spread* is twice the absolute difference between the trade execution price and the midpoint, scaled by the midpoint and averaged over all trades during date *t*, expressed as a percent. *DailyEDGAR* is an indicator variable equal to one if there is an EDGAR filing within the window [*t*-1, *t*+1] for each date *t*. *DailyPR* is an indicator variable equal to one if the firm issues a press release within the window [*t*-1, *t*+1] for each date *t*. *DailyReturn* is the buy and hold return during the window [*t*-5, *t*-1] for each date *t*. *DailyVolat* is the daily stock return volatility during the window [*t*-5, *t*-1] for each date *t*. *Size* is the natural logarithm of market value. *Leverage* is the ratio of total liabilities to total assets. *MTB* is the ratio of market value to book value of

equity. *AdExp* is the ratio of advertising expense to sales. *CorpAcq* is an indicator variable equal to one if an acquisition accounts for at least 20% of sales. *CapIntens* is the ratio of property, plant, and equipment to total assets. *CapEx* is the ratio of capital expenditures to total assets. *R&D* is the ratio of research and development expense to sales. *Financing* is of the sum of equity and debt issuances over the fiscal quarter scaled by total assets. *Loss* is an indicator variable equal to one if the firm had negative net income. *Returns* is the buy and hold return over the fiscal quarter. *IdioVol* is the unexplained variance in a Fama-French three factor expected returns model using the daily returns over the fiscal quarter. *Surprise* is the change in earnings from the same quarter one year prior scaled by market value of equity at the end of the prior quarter. *SpecItems* is special items scaled market value of equity. *SmallBeat* is an indicator variable equal to one if the year over year change in earnings per share is greater than zero and less than or equal to one cent.

Table 3. Daily Event Study of Information Content

	Dependent Variable: *AbsoluteReturn*				Dependent Variable: *Volume*			
	Coeff.	*t*-stat	Coeff.	*t*-stat	Coeff.	*t*-stat	Coeff.	*t*-stat
EventDay[0,10]	0.48***	(6.89)	0.48***	(6.91)	0.45***	(7.95)	0.45***	(8.01)
Controls								
DailyEDGAR	4.45***	(13.16)	4.47***	(13.18)	3.23***	(14.28)	3.23***	(14.23)
DailyPR	1.16***	(5.25)	1.16***	(5.23)	0.43***	(2.78)	0.43***	(2.78)
DailyReturn	3.93**	(2.37)	4.10**	(2.51)	1.94	(1.00)	2.04	(1.06)
DailyVolat	159.56***	(5.86)	154.00***	(5.70)	229.13***	(6.06)	224.53***	(5.89)
Size	-0.44	(-0.43)			-1.08*	(-1.88)		
Leverage	0.07	(0.02)			-0.55	(-0.30)		
MTB	-0.43	(-1.57)			-0.44	(-1.60)		
AdExp	0.00	(-0.51)			0.00	(1.03)		
CorpAcq	1.26**	(2.43)			0.29	(0.83)		
CapIntens	5.50***	(3.69)			-0.94	(-0.65)		
CapEx	-2.03	(-0.35)			-2.89	(-0.60)		
R&D	0.00*	(-1.74)			0.00	(-0.82)		
Financing	0.11	(0.82)			-0.12	(-1.26)		
Loss	-0.33	(-0.94)			-0.20	(-1.02)		
Returns	-0.53	(-0.74)			0.67	(1.60)		
IdioVol	-21.88	(-1.51)			-20.16**	(-2.02)		
Surprise	0.07	(0.26)			0.51***	(3.91)		
SpecItems	-2.37	(-0.70)			-2.12	(-1.59)		
SmallBeat	-0.63	(-1.62)			-0.09	(-0.28)		
Firm FE	Yes		No		Yes		No	
Date FE	Yes		Yes		Yes		Yes	
Firm-Year-Qtr FE	No		Yes		No		Yes	
N-obs	699,060		699,060		699,060		699,060	
N-events	11,294		11,294		11,294		11,294	

This table presents results from estimating Eq. (1). Events are defined as days where the change in website content (Δ*WebContent*) is in the top 33rd percentile, and where there are no contemporaneous press releases or EDGAR filings within a [-1,+1] window of the change in website content. *Firm-Year-Qtr* fixed effects subsume all variables calculated at the quarterly level. *t*-statistics appear in parentheses and are based on standard errors clustered by firm and date. ***, **, and * denote statistical significance at the 0.01, 0.05, and 0.10 levels (two–tail), respectively. Sample of 699,060 firm-days in the [–30,+30] day window around the change in website content.

Table 4. Daily Event Study of Information Content: Windows of No Disclosure

	Dep. Var. *AbsoluteReturn*		Dep. Var. *Volume*	
Non-disclosure Window [-*s*,+*s*]	Coeff. on *EventDay[0,10]*	*t*-stat on *EventDay[0,10]*	Coeff. on *EventDay[0,10]*	*t*-stat on *EventDay[0,10]*
Excluding observations with disclosures [-5,+5] *N-obs* = 444,446, *N-events* = 7,214	0.60***	(5.82)	0.55***	(6.51)
Excluding observations with disclosures [-10,+10] *N-obs* = 307,112, *N-events* = 4,987	0.63**	(5.00)	0.43***	(4.69)
Excluding observations with disclosures [-15,+15] *N-obs* = 228,548, *N-events* = 3,712	0.53***	(3.44)	0.30***	(2.97)
Controls	Yes		Yes	
Date FE	Yes		Yes	
Firm-Year-Qtr FE	Yes		Yes	

This table presents results from re-estimating the specifications in Table 3 after restricting the sample to events where the change in website content (Δ*WebContent*) is in the top 33rd percentile, and there are no contemporaneous press releases or EDGAR filings within a [-*s*,+*s*] window of the change in website content. All specifications include *Firm-Year-Qtr* and *Date* fixed effects. *t*-statistics appear in parentheses and are based on standard errors clustered by firm and date. ***, **, and * denote statistical significance at the 0.01, 0.05, and 0.10 levels (two–tail), respectively.

Table 5. Daily Event Study of Information Asymmetry

	Dependent Variable: *Illiquidity*				Dependent Variable: *Spread*			
	Coeff.	*t*-stat	Coeff.	*t*-stat	Coeff.	*t*-stat	Coeff.	*t*-stat
EventDay[0,10]	-0.11***	(-4.94)	-0.12***	(-5.04)	-0.06***	(-5.03)	-0.07***	(-5.18)
Controls								
DailyEDGAR	0.38***	(4.15)	0.39***	(4.21)	0.07**	(1.98)	0.08**	(2.08)
DailyPR	0.41***	(3.15)	0.41***	(3.14)	0.03	(0.71)	0.03	(0.73)
DailyReturn	0.03	(0.04)	0.13	(0.16)	-0.42	(-1.33)	-0.40	(-1.26)
DailyVolat	-5.25	(-0.35)	-6.23	(-0.43)	0.68	(0.20)	0.14	(0.04)
Size	0.36	(0.61)			-0.03	(-0.09)		
Leverage	-1.24	(-0.60)			-0.30	(-0.27)		
MTB	-0.18	(-1.22)			0.03	(0.35)		
AdExp	0.00***	(-3.20)			0.00	(-1.14)		
CorpAcq	0.77	(1.08)			0.32	(1.58)		
CapIntens	2.02**	(2.42)			-1.28	(-1.12)		
CapEx	1.36	(0.34)			-3.96	(-0.82)		
R&D	0.00	(0.34)			0.00	(0.56)		
Financing	0.07	(1.33)			0.09	(1.46)		
Loss	-0.35	(-1.62)			-0.09	(-0.64)		
Returns	-0.81*	(-1.84)			-0.43	(-1.59)		
IdioVol	3.06	(0.35)			-6.56	(-0.89)		
Surprise	0.10	(0.47)			0.33***	(3.05)		
SpecItems	-1.84	(-0.75)			-0.95	(-1.07)		
SmallBeat	-0.13	(-0.47)			-0.13	(-1.01)		
Firm FE	Yes		No		Yes		No	
Date FE	Yes		Yes		Yes		Yes	
Firm-Year-Qtr FE	No		Yes		No		Yes	
N-obs	699,060		699,060		699,060		699,060	
N-events	11,294		11,294		11,294		11,294	

This table presents results from estimating Eq. (1) using market-based measures of information asymmetry as the dependent variable. Events are defined as days where the change in website content (Δ*WebContent*) is in the top 33rd percentile, and where there are no contemporaneous press releases or EDGAR filings within a [-1,+1] window of the change in website content. *Firm-Year-Qtr* fixed effects subsume all variables calculated at the quarterly level. *t*-statistics appear in parentheses and are based on standard errors clustered by firm and date. ***, **, and * denote statistical significance at the 0.01, 0.05, and 0.10 levels (two–tail), respectively. Sample of 699,060 firm-days in the [–30,+30] day window around the change in website content.

Table 6. Daily Event Study of Information Asymmetry: Windows of No Disclosure

	Dep. Var. *Illiquidity*		Dep. Var. *Spread*	
Non-disclosure Window [-*s*,+*s*]	Coeff. on *EventDay[0,10]*	*t*-stat on *EventDay[0,10]*	Coeff. on *EventDay[0,10]*	*t*-stat on *EventDay[0,10]*
Excluding observations with disclosures [-5,+5] *N-obs* = 444,446, *N-events* = 6,013	-0.20***	(-5.65)	-0.09***	(-4.88)
Excluding observations with disclosures [-10,+10] *N-obs* = 307,112, *N-events* = 4,987	-0.20***	(-3.02)	-0.11***	(-4.07)
Excluding observations with disclosures [-15,+15] *N-obs* = 228,548, *N-events* = 3,712	-0.16***	(-3.97)	-0.11***	(-3.51)
Controls	Yes		Yes	
Date FE	Yes		Yes	
Firm-Year-Quarter FE	Yes		Yes	

This table presents results from re-estimating the specifications in Table 5 after restricting the sample to events where the change in website content (Δ*WebContent*) is in the top 33rd percentile, and there are no contemporaneous press releases or EDGAR filings within a [-*s*,+*s*] window of the change in website content. All specifications include *Firm-Year-Qtr* and *Date* fixed effects. *t*-statistics appear in parentheses and are based on standard errors clustered by firm and date. ***, **, and * denote statistical significance at the 0.01, 0.05, and 0.10 levels (two–tail), respectively.

Table 7. Website Content and Information Environment: Pre/Post-SEC Guidance

Panel A. Dependent Variable: *AmihudIlliquidityQtrly*$_{q+1}$

	Pooled sample *2004 to 2018*		Sample pre-*Aug 2008*		Sample post-*Aug 2008*	
	Coeff.	*t*-stat	Coeff.	*t*-stat	Coeff.	*t*-stat
WebContent	-0.45***	(-2.92)	-0.92**	(-2.28)	-1.95***	(-6.17)
Controls						
8Ks	-0.79***	(-4.18)	-3.00***	(-6.01)	-2.89***	(-6.02)
PRs	0.02	(0.30)	0.41*	(1.85)	0.74***	(4.33)
Size	-15.97***	(-15.50)	-16.50***	(-22.49)	-18.52***	(-21.68)
Leverage	3.83***	(9.33)	5.30***	(8.16)	4.09***	(8.83)
MTB	-1.46***	(-2.81)	3.86***	(5.78)	2.06***	(4.67)
AdExp	1.38***	(3.01)	0.71**	(1.99)	0.61*	(1.94)
CorpAcq	-0.37***	(-3.17)	-0.78***	(-3.71)	-1.34***	(-8.31)
CapIntens	0.51	(0.79)	-0.21	(-0.34)	1.37**	(2.52)
CapEx	-0.89***	(-4.53)	-0.46	(-0.92)	-2.07***	(-4.75)
R&D	0.79	(1.11)	-0.59	(-1.24)	-0.52	(-1.23)
Financing	-0.66***	(-5.27)	-1.14***	(-2.90)	-1.27***	(-5.26)
Loss	0.49***	(4.90)	2.98***	(8.77)	2.07***	(8.66)
Returns	-0.47***	(-2.62)	-1.41***	(-5.09)	-1.43***	(-5.53)
IdioVol	3.32***	(7.19)	1.21	(1.49)	1.93***	(2.96)
Surprise	0.42***	(4.38)	1.42***	(5.02)	0.90***	(6.69)
Turnover	-4.84***	(-8.14)	-5.83***	(-9.53)	-6.56***	(-8.98)
SpecItems	-0.09	(-0.34)	2.22***	(3.70)	1.54***	(3.47)
SmallBeat	-0.08	(-0.66)	1.87***	(3.01)	2.31***	(5.52)
InvPrice	1.32***	(2.61)	8.45***	(10.93)	4.57***	(8.78)
Firm FE	Yes		Yes		Yes	
Year-Qtr FE	Yes		Yes		Yes	
N-obs	178,689		47,065		131,624	
two-tailed *p*-value for the test of difference in coefficients pre/post Aug. 2008				[0.018]		

Table 7. Website Content and Information Environment: Pre/Post-SEC Guidance (cont'd)

Panel B. Dependent Variable: *EffectiveSpreadQtrly*$_{q+1}$

	Pooled sample *2004 to 2018*		Sample pre-*Aug 2008*		Sample post-*Aug 2008*	
	Coeff.	*t*-stat	Coeff.	*t*-stat	Coeff.	*t*-stat
WebContent	-0.20***	(-3.21)	-0.88***	(-2.94)	-1.38***	(-4.97)
Controls						
8Ks	-0.23***	(-3.92)	-0.99***	(-4.05)	-0.91***	(-3.08)
PRs	-0.01	(-0.33)	0.29**	(2.33)	0.16*	(1.86)
Size	-1.02***	(-5.46)	-2.77***	(-6.01)	-2.20***	(-5.97)
Leverage	0.62***	(5.76)	0.82**	(2.52)	1.25***	(4.20)
MTB	-0.25**	(-2.33)	-0.31	(-0.78)	-0.72*	(-1.76)
AdExp	0.08	(0.56)	0.31	(1.42)	0.18	(1.15)
CorpAcq	-0.02	(-0.74)	-0.41***	(-3.71)	-0.45***	(-4.78)
CapIntens	0.14	(0.68)	-0.12	(-0.29)	-0.09	(-0.24)
CapEx	-0.12*	(-1.69)	-0.52	(-1.47)	-0.58***	(-2.99)
R&D	-0.09	(-0.47)	0.03	(0.09)	0.24	(0.91)
Financing	-0.12***	(-3.92)	-0.55***	(-2.62)	-0.34**	(-2.40)
Loss	0.04	(1.55)	0.38***	(2.65)	0.18**	(2.27)
Returns	-0.04*	(-1.66)	0.06	(0.60)	0.02	(0.34)
IdioVol	0.28***	(2.93)	0.77**	(2.50)	0.31	(1.33)
Surprise	0.06***	(3.02)	0.22**	(2.24)	0.08**	(2.04)
Turnover	-0.58***	(-5.74)	-1.81***	(-5.32)	-1.18***	(-4.63)
SpecItems	-0.07	(-1.39)	0.27	(0.83)	0.61**	(2.52)
SmallBeat	0.00	(-0.17)	0.09	(0.64)	0.28***	(2.95)
InvPrice	0.09	(0.79)	-1.69**	(-2.57)	-0.88*	(-1.92)
Firm FE	Yes		Yes		Yes	
Year-Qtr FE	Yes		Yes		Yes	
N-obs	178,689		47,065		131,624	
two-tailed *p*-value for the test of difference in coefficients pre/post Aug. 2008				[0.033]		

This table presents results from estimating Eq. (2) for separate samples before and after the SEC issued guidance allowing corporate websites to substitute for EDGAR disclosure in August 2008. Panel A presents results from using *Illiquidity*$_{q+1}$ as the measure of information asymmetry. Panel B presents results from using *Spread*$_{q+1}$ as the measure of information asymmetry. All specifications include *Firm* and *Year-Qtr* fixed effects. *t*-statistics appear in parentheses and are based on standard errors clustered by firm and quarter. Two-tailed *p*-values appear in brackets and test for differences in coefficient across samples. ***, **, and * denote statistical significance at the 0.01, 0.05, and 0.10 levels (two–tail), respectively. Sample of 178,689 firm-quarters from 2004 to 2018.

Table 8. Daily Event Study of Information Content by Topic

Panel A. Dependent Variable: *AbsoluteReturns*

	Coeff.	*t*-stat	Coeff.	*t*-stat	Coeff.	*t*-stat	Coeff.	*t*-stat	Coeff.	*t*-stat
EventDay[0,10]	0.41***	(6.16)	0.32***	(4.12)	0.49***	(4.81)	0.50***	(5.48)	0.27**	(2.15)
Productivity	-0.18***	(-3.83)							-0.19***	(-3.84)
EventDay[0,10]·Productivity	0.83***	(4.27)							0.84***	(4.28)
BSO			-0.02	(-0.62)					-0.02	(-0.64)
EventDay[0,10]·BSO			0.33***	(3.18)					0.33***	(3.20)
CP					0.00	(0.08)			0.00	(0.08)
EventDay[0,10]·CP					-0.01	(-0.08)			-0.01	(-0.10)
IP							0.00	(0.11)	0.00	(0.16)
EventDay[0,10]·IP							-0.03	(-0.27)	-0.04	(-0.33)
Controls	Yes		Yes		Yes		Yes		Yes	
Date FE	Yes		Yes		Yes		Yes		Yes	
Firm-Year-Qtr FE	Yes		Yes		Yes		Yes		Yes	
N-obs	699,060		699,060		699,060		699,060		699,060	
N-events	11,294		11,294		11,294		11,294		11,294	

Table 8. Daily Event Study of Information Content by Topic (cont'd)

Panel B. Dependent Variable: *Volume*

	Coeff.	*t*-stat	Coeff.	*t*-stat	Coeff.	*t*-stat	Coeff.	*t*-stat	Coeff.	*t*-stat
EventDay[0,10]	0.36***	(6.85)	0.37***	(6.12)	0.46***	(6.20)	0.44***	(6.58)	0.31***	(3.53)
Productivity	-0.12***	(-2.95)							-0.12***	(-2.97)
EventDay[0,10]·Productivity	0.93***	(6.80)							0.93***	(6.80)
BSO			-0.01	(-0.40)					-0.01	(-0.42)
EventDay[0,10]·BSO			0.15**	(2.04)					0.15**	(2.07)
CP					0.00	(-0.22)			0.00	(-0.20)
EventDay[0,10]·CP					-0.04	(-0.54)			-0.04	(-0.58)
IP							0.01	(0.60)	0.01	(0.64)
EventDay[0,10]·IP							0.01	(0.10)	0.00	(0.00)
Controls	Yes		Yes		Yes		Yes		Yes	
Date FE	Yes		Yes		Yes		Yes		Yes	
Firm-Year-Qtr FE	Yes		Yes		Yes		Yes		Yes	
N-obs	699,060		699,060		699,060		699,060		699,060	
N-events	11,294		11,294		11,294		11,294		11,294	

This table presents results from re-estimating the specifications in Table 3 with the addition of variables for the file type and topic of the website change. *Productivity* is an indicator variable equal to one if the change contains files in the productivity category—PDFs and Microsoft Office documents are two of the top files in this category. *BOS*, *CP*, and *IP* are indicator variables equal to one if the amount of content in the website change is above the median amount of Business Operations and Strategy, Consumer Products, and Intellectual Property related disclosure, respectively. All specifications include *Firm-Year-Qtr* and *Date* fixed effects. *t*-statistics appear in parentheses and are based on standard errors clustered by firm and date. ***, **, and * denote statistical significance at the 0.01, 0.05, and 0.10 levels (two–tail), respectively.