Investor Demand for Contextual Information: Evidence from Wikipedia*

Chenqi Zhu
czhu@stern.nyu.edu
New York University

January 2019

Abstract:

I use the web traffic (wiki-visits) to the Wikipedia pages of S&P 500 companies to study the determinants and consequences of investors' contextual information acquisition. Employing the launch of Google Knowledge Graph as a natural experiment, I first establish that wiki-visits reflect information acquisition by investors, especially retail investors. I then document that wiki-visits spike at key informational events such as earnings announcements. To better understand investors' demand for contextual information, I focus on earnings announcements and find that the abnormal wiki-visits at these announcements increase with the announcing firms' abnormal tone in earnings press releases, textual complexity of previous financial reports, and off-balance sheet intangible intensity. Consistent with contextual information improving investors' interpretation of financial information, I find that the abnormal wiki-visits at earnings announcements are positively associated with the earnings response coefficient and the speed of price discovery measured by intra-period timeliness. Collectively, these results suggest that investors *supplement* financial information with highly readable contextual information, thereby improving their understanding of the financial information.

Key Words: Information Acquisition, Information Quality, Wikipedia, Price Discovery

JEL Classification: D83, M41, G11, G14

*I sincerely thank my dissertation committee for their invaluable guidance and continuous support: Mary Billings, Yiwei Dou, Deepak Hegde, Baruch Lev (Chair), and Stephen Ryan. I am also thankful for valuable comments and suggestions from Sophie Calder-Wang, Justin Deng, Svenja Dube, Disen Huang, Kai-Wen Hsueh, Paul Zarowin, and seminar participants at the 2018 Deloitte Doctoral Consortium, NYU, and Fordham. I am grateful to Nuria Ruiz and Leon Ziemba at Wikimedia Foundation for their institutional insights. Finally, I acknowledge NYU and the Deloitte Foundation for their generous financial support. Any errors are my own.

"The capital market is looking for information from companies in various social media channels [Wikipedia, Facebook, LinkedIn, and Twitter] and at a greater extent than we anticipated."

--Marcus Eriksson, Hallvarsson & Halvarsson

I. Introduction

The emergence of social media has brought significant changes to the provision, dissemination, and processing of information in the capital market (SEC 2012; Miller and Skinner 2015). A burgeoning literature documents that investors frequently create and share information at online sites (e.g., message boards, Twitter, Seeking Alpha, and Estimize).[1] Most research in this strand examines whether such peer-provided information generates or disseminates *new* signals beyond that provided by traditional information intermediaries. Meanwhile, *background* information and *historical* facts about public firms are also shared online. The most active platform for such contextual information is Wikipedia. Survey evidence shows that investors frequently visit Wikipedia for firm-specific information (Bradshaw 2008; Comprend 2015). Using web traffic (wiki-visits) to corporate Wikipedia pages (wiki-pages), I investigate the usefulness of contextual information by examining the determinants of wiki-visits and the consequences of these visits for stock price discovery.

Wikipedia is a website that everyone can use and edit. Corporate wiki-pages, the focus of this paper, provide qualitative information on firms' business, background, top executives, major investments (e.g., M&A, R&D, and brands), scandals and controversies, etc. (see Appendix A for the example of Microchip Technology's wiki-page). In principle, all information in wiki-pages is publicly available, so Wikipedia must mitigate certain costs that impede investors from directly accessing the original sources. Investors incur awareness cost to identify the existence of information, acquisition cost to collect it, and integration cost to process it (Blankespoor, deHaan, Wertz, and Zhu 2018). By compiling relevant content from numerous sources into one web page, Wikipedia reduces investors' awareness and acquisition costs. By leveraging the collective wisdom of the crowd to simplify and neutralize the original content, Wikipedia reduces investors'

[1] Papers that study the usefulness of user-provided information in the capital market include Blankespoor, Miller, and White (2013); Chen, De, Hu, and Hwang (2014); Lee, Hutton, and Shu (2015); Jame, Johnston, Markov, and Wolfe (2016); Bartov, Faurel, and Mohanram (2017); Lawrence, Ryans, Sun, and Soni (2017); Lerman (2017); Cade (2018); and Campbell, DeAngelis, and Moon (2018).

integration cost. I perform textual analyses on the wiki-pages of S&P 500 companies and confirm that they aggregate a broad range of information sources into a highly readable record.

It is *ex ante* unclear whether wiki-pages are useful for investors, particularly to interpret newly released financial information. On the one hand, the information in wiki-pages is general and not necessarily compiled by professionals, whereas abundant information produced by specialists (e.g., financial analysts) is available at a very low cost. On the other hand, investors, especially retail investors, may find the information in wiki-pages useful for three reasons. First, it quickly brings investors up to speed with the firm's background and latest developments. Second and relatedly, it enables investors to interpret newly released information in proper context. Third, wiki-pages' neutral narratives, thanks to the edits by thousands of volunteers who have little conflict of interest, allow investors to unravel tone-hyping or spin in corporate disclosures.

Consistent with the above-conjectured usage of wiki-pages, a consulting company conducted a recent survey and found that 78% of buy-side analysts use Wikipedia to locate firm-specific information (Comprend 2015). Although I cannot observe the identity of wiki-visitors, I validate that wiki-visits reflect information acquisition by investors, especially retail investors.[2] I find the average wiki-page of S&P 500 companies is visited 996 times per day. Compared to the mere 28 requests of the average 10-K filing at EDGAR on the day of and the day following the filing date (Loughran and McDonald 2017),[3] the frequent wiki-visits are prima facie evidence of the perceived usefulness of contextual information.

The first part of my main analysis is to examine the circumstances under which investors acquire contextual information from Wikipedia. Consistent with investors using contextual information to better interpret new events, I find large spikes in daily wiki-visits on the days of earnings announcements, management forecasts, M&A announcements, and extreme stock price movements. Specifically, wiki-visits jump by 15.0% at earnings announcements, which are comparable to the spikes at management forecasts, but smaller than those on the days of M&A

[2] These validation analyses are discussed later in the introduction briefly and in Section V at length.

[3] Drake, Roulstone, and Thornock (2016) report that an average 10-K filing is only requested 4.55 times from EDGAR on a typical day. This number understates the total utilization of 10-K filings as they are also accessed in other venues such as Yahoo! Finance and companies' official websites. However, the web traffic to Yahoo! Finance's SEC filings and key financial statistics only account for 6% of the total traffic (Lawrence, Ryans, Sun, and Soni 2017). In an interview with the Wall Street Journal, General Electric's CFO stated that the company's 2013 financial report was only downloaded 800 times throughout the year, i.e, 2.19 times per day on average (see the article at https://www.wsj.com/articles/the-109-894-word-annual-report-1433203762).

announcements and extreme stock movements. Moreover, wiki-visits roughly stay at the elevated levels one day after these events but do not increase considerably on the day before, with the exception of extreme stock movements whose significant pre-event abnormal wiki-visits are probably due to return momentums. This overall pattern indicates that investors visit Wikipedia to acquire contextual information to help them interpret, rather than predict, new events.

To better understand how wiki-pages help investors interpret new events, I focus on earnings announcements and investigate how the abnormal wiki-visits during the announcements vary with three firm characteristics.[4] The first two characteristics under investigation (abnormal tone in earnings press releases and textual quality of prior financial reports) capture the quality of firm-issued qualitative information. The third characteristic is off-balance sheet intangible intensity, which intends to capture the deficiencies in balance sheets. I expect that these characteristics affect the marginal benefits of wiki-pages to investors, thus influencing wiki-visits at earnings announcements.

Following Huang, Teoh, and Zhang (2013), I measure abnormal tone as the tone residual from annual regressions of tone on the financial information in earnings press releases and firm characteristics. I find that the abnormal wiki-visits at earnings announcements are significantly larger when the earnings press releases carry an abnormally positive tone, although the wiki-visits do not differ significantly for negative versus positive earnings. This evidence suggests that investors value the contextual information prepared by neutral third parties when they are concerned with opportunism in firm-issued disclosures.

Considering Wikipedia's simple and concise writing style, I expect that wiki-pages are more useful when prior 10-K filings' Item 1 "Business" and Item 2 "Properties", whose content has the most overlap with wiki-pages (SEC 2011), are more difficult to process.[5] Consistent with this expectation, I find that the abnormal wiki-visits at earnings announcements are more pronounced when the contextual information in the most recent 10-K filing is less readable or more boilerplate.

[4] I focus on earnings announcements because they are the most prominent type of accounting events and their news content can be quantified (e.g., Bushee, Core, Guay, and Hamm 2010; Drake et al. 2012).

[5] I do not assume investors read historical financial statements immediately before visiting wiki-pages. Instead, I assume that investors' perceived textual quality of 10-Ks correlates with 10-Ks' readability and boilerplate language. Investors can develop their perception through, for example, prior personal experience, media, or word of mouth.

Intangible assets are often unrecognized in the balance sheet, rendering it largely irrelevant to the valuation of securities (e.g., Amir and Lev 1996). Given that wiki-pages often discuss intangible assets such as brands and R&D based on the information gathered from the patent office, customer surveys, etc., I expect that wiki-pages are more useful to evaluate intangible-intensive firms. In line with this expectation, I find that the abnormal wiki-visits at earnings announcements are significantly larger for firms with intensive off-balance sheet intangibles and that they do not vary significantly with on-balance sheet (acquired) intangible intensity.

The second part of the main analysis investigates the consequences of contextual information acquisition for the efficiency of price discovery. If wiki-pages enable investors to better interpret new financial information, then stock prices should be more responsive to this new information immediately after its release, all else being equal. Consistent with this conjecture, I find that the earnings response coefficient (ERC) increases with the abnormal wiki-visits at earnings announcements.[6] A one-standard-deviation increase in the abnormal wiki-visits over the two-day window of earnings announcements is associated with a 15.6% increase in the ERC.

To distinguish wiki-pages from other online information and to highlight the difference between information acquisition for the purpose of interpretation versus prediction (Kim and Verrecchia 1991, 1994), I investigate how pre-announcement wiki-visits affect price discovery around earnings announcements. Some online information, for example, crowd-sourced earnings forecasts, may provide investors with insights about future earnings. Thus, the pre-announcement acquisition of such information allows stock prices to incorporate upcoming earnings news before the announcements (Drake et al. 2012; Jame et al. 2016). Considering that contextual information is useful to interpret rather than predict earnings, I do not expect that pre-announcement wiki-visits affect the incorporation of upcoming earnings news into stock prices in the week prior to earnings announcements (measured by future ERC). Consistent with this expectation, I do not find that pre-announcement abnormal wiki-visits significantly affect future ERC. I also use abnormal trading volumes to measure market reactions and find consistent results. Both the return and volume results underscore the usefulness of wiki-pages in improving investors' interpretation, rather than prediction, of new financial information.

[6] It is plausible that investors visit wiki-pages because their prior information about the firm is inadequate. This plausibility biases against finding a positive impact of wiki-visits on the ERC and the speed of price discovery.

Lastly, I study the association between the abnormal wiki-visits at earnings announcements and the speed of price discovery, as measured by intra-period timeliness (IPT). The key advantage of this IPT measure is that it captures the speed with which news is incorporated into stock prices while holding the magnitude of news constant (Butler, Kraft, and Weiss 2007; Twedt 2016; Blankespoor, deHaan, and Zhu 2018).[7] Therefore, this analysis is less susceptible than the analysis of ERC and future ERC to endogeneity issues such as omitted variables (e.g., the "newsworthiness" of earnings announcements) and reverse causality (e.g., larger market reactions trigger more wiki-visits). I find that the abnormal wiki-visits at earnings announcements are positively associated with IPT, consistent with investors visiting wiki-pages to improve their interpretation, thereby expediting the speed of price discovery.

An important caveat is that Wikipedia visitors are not necessarily investors. To mitigate this concern, in all analyses I adjust raw wiki-visits by the median wiki-visits for the same day of the week over the past 10 weeks to partially remove non-investment-related wiki-visits. I also examine the example of the clickstreams to Microchip Technology's wiki-page (websites visited immediately before the focal wiki-page) and find that these clickstreams evidence investors navigating from one web page to another to gather information.

To further verify that wiki-visits reflect information acquisition by investors, I examine the relationship between abnormal wiki-visits and subsequent trades, classified as either retail or institutional following Boehmer, Jones, and Zhang (2017) and Bushee, Cedergren, and Michels (2018). Consistent with the intuition that retail investors are more likely than institutional investors to acquire information from wiki-pages, I find that abnormal wiki-visits are positively associated with subsequent net buying by retail investors but not by institutional investors. One interpretation of this result is that retail investors visit wiki-pages as part of their due diligence before making investment decisions. An alternative interpretation is that retail investors trade on attention-grabbing stocks. For example, shocks to firms' visibility could simultaneously trigger wiki-visits by the general public and trades by retail investors.[8] Thus, wiki-visits will positively correlate with retail trades even if no retail investors visit wiki-pages.

[7] The news content is measured as the unexpected earnings in ERC and as the cumulative abnormal returns over the entire window (i.e., from day 0 to day 5 starting with the earnings announcement date) in IPT.

[8] The non-results of institutional trades mitigate the concern that the positive association between wiki-visits and retail trades is driven by shocks to firms' fundamentals.

I employ two approaches to distinguish between the above two interpretations. The first approach examines the association between abnormal wiki-visits and subsequent returns. The alternative interpretation predicts zero subsequent returns or positive returns followed by reversals due to temporary price pressure caused by net retail purchases. However, if the investment decision is justified by the information acquired from Wikipedia, as implied by the information acquisition interpretation, we should not observe any return reversals. I find that abnormal wiki-visits positively predict the next week's abnormal returns without reversals, which is inconsistent with the alternative interpretation. More interestingly, the predictive power of abnormal wiki-visits for subsequent returns is driven by wiki-pages that contain richer or more up-to-date information, suggesting that higher-quality wiki-pages have larger impacts.

The second approach exploits the launch of Google Knowledge Graph (GKG) in May 2012 as a natural experiment. GKG lists basic facts about the firm (e.g., a one-sentence summary of the firm's business, headquarters, CEO, subsidiaries) in an information box shown alongside the Google search results. I expect the wiki-visits post GKG to involve more meaningful information acquisition (beyond what is in the GKG information box). Consistent with this expectation, I find that the association between abnormal wiki-visits and subsequent retail trades or abnormal returns is only significant in the post-GKG period, providing further support for the information acquisition interpretation.[9]

This study makes the following contributions. First, it adds to our understanding of financial reporting quality and information demand from the perspective of investors, especially retail investors who are of great importance to regulators. For example, former SEC chairperson Mary Jo White stated: "The retail investor must be a constant focus of the SEC."[10] The frequent visits to wiki-pages, relative to the infrequent usage of SEC filings, reveal investors' preferences and imply their dissatisfaction with financial reports. Consistent with the prior literature (e.g., You and Zhang 2009; Miller 2010; Lawrence 2013; Asay, Elliott, and Rennekamp 2016), my analyses indicate that part of the dissatisfaction stems from the poor readability and boilerplate language of financial reports. Such revealed preferences provide micro-underpinnings for the "plain English"

[9] The association between wiki-visits and institutional trades is insignificant both before and after the launch of GKG, undercutting the possibility that confounding events simultaneously affect wiki-visits, trades, and stock returns.

[10] See Mary Jo White's speech at https://www.sec.gov/news/speech/mjw-speech-032114-protecting-retail-investor. Current SEC chairperson Jay Clayton stated, "Our analysis starts and ends with the long-term interests of the Main Street investor [retail investor]." (see: https://www.sec.gov/news/speech/remarks-economic-club-new-york).

initiative promoted by the SEC in 1998 and the AICPA in 2013 and for the designation of "understandability" as one of the four enhancing qualitative characteristics in the *Statements of Financial Accounting Concepts No. 8*.

This study also fits into the broader literature examining the impact of technological and institutional changes on the provision and dissemination of information in the capital market (see Miller and Skinner (2015) for a review). Recent research documents that social media harnesses the wisdom of the crowd to produce or disseminate information incremental to traditional information sources (e.g., Chen et al. 2014; Jame et al. 2016; Bartov et al. 2017). My results show that the crowd's compiling, simplification, and neutralization of publicly available information also facilitate investors' decision making.

Lastly, this study adds to the literature on information complementarity (see Antle, Demski, and Ryan (1994) and Amir and Lev (1996) for complementarity between non-accounting and accounting information; see Lev, Ryan, and Wu (2008) and Drake et al. (2016) for complementarity between historical and current financial information). The innovation of this paper is that it provides evidence on the complementarity between new financial information and qualitative information on firms' background and history. This evidence complements the prior research which demonstrates that managers supplement quantitative information with qualitative information to affect investors' assessment of disclosure credibility without observing investors' actual usage of the qualitative information (e.g., Hutton, Miller, and Skinner 2003; Baginski, Demers, Wang, and Yu 2016). I find that investors indeed have a high demand for qualitative information for interpretation purposes as they visit wiki-pages frequently and more so at the time of important events such as earnings announcements. Moreover, I find that wiki-visits are associated with faster price discovery of earnings news, highlighting the importance of contextual information in interpreting financial information.

An overview of the rest of the paper is as follows. Section II describes the background, Section III develops the hypotheses, and Section IV summarizes the sample and data. Section V validates that wiki-visits reflect information acquisition by investors. Sections VI analyzes the causes and consequences of contextual information acquisition from Wikipedia. Section VII concludes.

II. Background

Wikipedia is a non-profit website that provides free and editable content. By allowing everyone to create new or edit existing wiki-pages, Wikipedia leverages millions of volunteers to keep its content up-to-date and correct misleading edits.[11] To further ensure accuracy and minimize bias, Wikipedia has developed and enforced editing policies such as "neutral point of view" and "verifiability."[12] Comparing 42 science entries in Wikipedia and Encyclopedia Britannica, Giles (2005) concludes that Wikipedia is comparable to Britannica in terms of accuracy. Focusing on wiki-pages about U.S. political topics, which are most prone to biases, Greenstein and Zhu (2012) find that wiki-pages are neutral on average. In addition, Wikipedia has guidelines such as to be "plain, direct, unambiguous, and specific," and "as concise as possible" to improve readability. For example, Microchip Technology's (MT) wiki-page, listed in Appendix A, generally describes events following the format of "on XXX date, MT did something" in plain English. As a result of reliable and neutral information content as well as a simple and concise writing style, Wikipedia receives 16 to 20 billion visits (pageviews) per month, making it the 5th (6th) most frequently visited website globally (in the U.S.) according to Alexa Internet, Inc.[13]

Compared to other information sources, Wikipedia has the following unique features which make it a good setting to study investors' demand for contextual information. First, the information contained in wiki-pages is primarily contextual and thus wiki-visits reveal demand for contextual information. Other sources either seldom provide contextual information or integrate contextual information with quantitative information (e.g., traditional information sources such as financial reports or social media such as Seeking Alpha), making it impossible to separate the role of contextual information. Second, wiki-pages in principle are created and maintained by volunteer editors, who have less conflict of interest than preparers of traditional information (e.g., managers or financial analysts). Unlike social media such as Twitter and Facebook that allow interested parties to communicate freely and directly (Blankespoor et al. 2013; Lee et al. 2015), Wikipedia

[11] For example, Microchip Technology's wiki-page was quickly updated after it announced the acquisition of Atmel on January 19, 2016. Its revision history is available at: https://en.wikipedia.org/w/index.php?title=Microchip_Technology&action=history .

[12] See https://en.wikipedia.org/wiki/Wikipedia:Policies_and_guidelines for Wikipedia's editing policies.

[13] Alexa Internet Inc. provides web traffic statistics at https://www.alexa.com/siteinfo/wikipedia.org (accessed on May 2, 2018).

strives to set the record straight by detecting and eliminating false or self-serving edits.[14] Third, corporate wiki-pages are widely used by investors (Bradshaw 2008; Comprend 2015). The large readership of Wikipedia ensures a relatively representative sample of investors, thereby increasing the external validity of my results regarding investors' demand for contextual information in general. Lastly, Wikipedia keeps its entire database open to the public. The data outline what information is accessed and when, enabling me to examine the circumstances under which investors search for firm-specific contextual information.

A few papers examine Wikipedia's role in the capital market. Moat et al. (2013) show that a trading strategy of holding a short (long) position of the Dow Jones Industrial Average Index for a week when the total wiki-visits of the companies in the index increase (decrease) generates a significant return of 0.5%.[15] Prior research also shows that wiki-pages mitigate the information asymmetry between informed and less informed parties. Specifically, Rubin and Rubin (2010) find that firms with more frequent revisions to their wiki-pages (hence more efficient information aggregation) are associated with smaller analyst forecast dispersions and forecast errors. Xu and Zhang (2013) find that wiki-pages moderate the opportunistic timing of managers' voluntary disclosures of disappointing earnings. These two papers study the association between the characteristics of wiki-pages and firms' information environment or managerial responses (voluntary disclosures) without identifying the underlying channels. My study complements these two papers by documenting investors' demand for contextual information to interpret financial information as one channel through which wiki-pages influence the firm's information environment and disclosure practices. To the best of my knowledge, my paper is the first to use

[14] Wikipedia reveals the IP addresses of editors, and volunteer technologists have developed algorithms to trace the sources of edits. With these data, the media have exposed big companies which attempted to revise their pages in their favor (e.g., Anheuser Busch, ExxonMobil, Microsoft, Wal-Mart, and PepsiCo). See some anecdotal stories at: https://www.nytimes.com/2007/08/19/technology/19wikipedia.html; http://www.washingtonpost.com/wp-dyn/content/article/2007/01/23/AR2007012301025.html.

[15] I find wiki-visits are positively associated with subsequent abnormal returns in Section V. My results differ from those of Moat et al. (2013) for the following reasons. First, Moat et al. (2013) trade the Dow Jones Industrial Average Index, not the individual stocks, on a weekly basis based on the changes in total wiki-visits of the companies in the index. Second, their sample period starts on December 10th, 2007, and ends on April 30th, 2012, before the launch of GKG, a period my results show that wiki-visits have little impact on trades and returns. The wiki-visit data during their sample period do not differentiate human visits from robot visits. Lastly, the negative association documented in Moat et al. (2013) is driven by the year 2008, when the market crashed, and the year 2011, when the market went on a roller coaster (see their Figure 3).

web traffic to wiki-pages to investigate the causes and consequences of investors' acquisition of contextual information.

To illustrate the linguistic features of wiki-pages, I perform textual analyses on the wiki-pages for S&P 500 firms as of the beginning of each month from July 2015 to December 2017 in Panel A of Table 1. The unit of observation is a firm-month. The average wiki-page (excluding references and footnotes) has 1779 words, 66 sentences, 10 sections, and 42 references to information sources outside Wikipedia. The median page length only increases slightly from 1760 words (65 sentences) to 1790 (66 sentences) throughout the sample period, as shown in Panel A of Figure 1. However, this word count (or sentence count) understates the content changes in wiki-pages as volunteer editors can add or cut content as they see fit. On average, a typical wiki-page in my sample is edited by 3.2 unique volunteer editors in a four-week window.

To quantify the information content of wiki-pages, I first classify the sections of wiki-pages using keyword searches in section titles. An average wiki-page has 0.33 sections on intangibles with section titles containing keywords such as "R&D" and "brands" and 0.67 sections on segments with keywords such as "operating units", "product lines", and "world presence".[16] Panel B of Figure 1 shows that the numbers of these two types of sections remain stable during the sample period. The vast majority of sections are unclassified, most likely discussing the corporate history and major events. The "verifiability" principle requires volunteer editors to provide references for their edits, allowing me to classify these references into four categories, based on the domain of the referenced web addresses, to decipher the original information sources. If the reference links to the firm's official website, EDGAR filings, or press releases, I classify it as a reference to firm-prepared documents. If the reference comes from a major news outlet, I classify it as a news reference. If the reference stems from books.google.com, educational institutions, or government agencies, I classify it as a third-party reference. An average wiki-page has 6.8 references to firm-prepared documents, 10.6 to news articles, and 2.5 to third-party sources. The remaining unclassified references (52.3% of total references) are from blogs or other online information sources. Panel C of Figure 1 shows that unclassified references and news references rise steadily while the other two types of references remain stable over time. Overall, these

[16] To the extent that my keyword list is incomplete, I under-identify sections on intangibles and segments.

descriptive statistics indicate that wiki-pages are a rich information repository that aggregates information from many sources.

Lastly, I analyze wiki-pages in terms of readability, boilerplate language, and tone. For comparison, I extract the texts of Item 1. "Business" and Item 2. "Properties" from 10-K filings, whose content largely overlaps with wiki-pages (SEC 2011), and perform the same textual analyses on them.[17] Table 1 shows that wiki-pages are more readable than 10-Ks across five different readability indexes. For example, the average Gunning-Fog index of wiki-pages is 15, which is equivalent to the reading level of a college junior. Similar to the statistics reported in Li (2008) and Dyer, Lang, and Stice-Lawrence (2017), the average Gunning-Fog index of 10-Ks is 19.7, which suggests that 10-Ks require almost 20 years of formal education to understand. I also compare the boilerplate language (Lang and Stice-Lawrence 2015) and tone (Loughran and McDonald 2011) and find that wiki-pages are less boilerplate and more neutral than 10-Ks.[18]

III. Hypotheses Development

Due to information acquisition cost and limited attention, investors do not acquire information randomly. They are motivated to do so when new events arrive which potentially affect their trading decisions (e.g., see Kim and Verrecchia (1991, 1994) for theoretical evidence and Drake et al. (2012) for empirical evidence). This holds for information acquisition in general and is true for contextual information acquisition from Wikipedia in particular, whose content alone is not informative about firms' future prospects and should be used in conjunction with other information to derive tradable inferences (Kim and Verrecchia 1994). Moreover, investors using Wikipedia are less sophisticated on average and tend to trade on attention-grabbing stocks that are going through major events (Barber and Odean 2008). These events create demand for the information in wiki-pages by raising firms' visibility thereby attracting new investors or triggering

[17] I use Python's regular expression function to identify the relevant text from 10-K filings. The text starts from "Item 1" and ends with "Item 3". Any mentioning of these items in the table of contents and by reference (e.g., "refer to Item 3") is ignored. If the identified text is shorter than 10,000 characters or ends with an incomplete sentence, I manually collect the section from EDGAR filings. I exclude sentences with more than 50% numbers, more than 40% capitalized letters, or less than 30 characters, following the prior literature such as Li (2008) and Lang and Stice-Lawrence (2015).

[18] The average proportion of boilerplate sentences in my sample (14%) is higher than that reported in Lang and Stice-Lawrence (2015) for an international sample of financial reports but smaller than that in Dyer et al. (2017) for the entire 10-Ks of U.S. firms. The differences stem from the texts under examination as well as the decision to exclude or include innocuous common phrases such as "as a result of" when defining boilerplate sentences.

existing investors' information acquisition to re-evaluate their positions. Accordingly, I state the first hypothesis below, in an alternative form.

H1. *Wiki-visits are more frequent around major informational events.*

In the next set of hypotheses, I discuss the circumstances that change the marginal benefits of acquiring contextual information from Wikipedia. I focus on earnings announcements as they are one of the most common and prominent accounting events. First, I expect that wiki-pages enable investors to better interpret the new financial information released by firms with opportunistic financial disclosures. Many studies document that managers embed linguistic cues in earnings press releases to inform or misinform investors (e.g., Huang et al. 2013). However, few extant studies examine how investors assess the credibility of linguistic cues (Baginski et al. 2016). In order to distinguish opportunistic cue manipulation from genuine private signals, investors can place the disclosures in context or contrast the disclosures with the narratives prepared by third parties. Wikipedia has gained public trust for its immunity to advertising influence and its effort to restrict self-serving edits in order to maintain a neutral point of view (Greenstein and Zhu 2012; Xu and Zhang 2013). Therefore, investors are more likely to visit wiki-pages when they have credibility concerns about firms' earnings press releases.

H2a. *The abnormal wiki-visits at earnings announcements are more pronounced when managers are suspected of manipulating financial disclosures.*

Relatedly, I posit that investors find the contextual information in wiki-pages more useful when the contextual information in prior financial reports is more difficult to process. Financial reports are arguably the most comprehensive source of information about the firm, and Part I of Form 10-K is designed to help investors understand the firm's business (SEC 2011). However, financial reports are often criticized for their poor readability and excessive complexity, deterring investors, especially less sophisticated ones, from fully incorporating the information of financial reports into their investment decisions (You and Zhang 2009; Miller 2010; Lawrence 2013).[19] When firm-issued information is less readable, investors are more likely to seek out and rely on

[19] Drake et al. (2016) document that investors download historical financial reports from EDGAR more frequently at earnings announcements than normal days, although the increased frequency is still economically low. As quantitative and qualitative information is integrated in financial reports, this download frequency cannot differentiate investors' acquisition of quantitative versus qualitative information.

external information sources (Asay et al. 2016). This heightened propensity to acquire external information holds in particular for Wikipedia because of its simple and concise writing style, as confirmed in Section II. Therefore, I predict that investors are more likely to visit wiki-pages to interpret newly released financial numbers when existing financial reports are hard to read.

H2b. *The abnormal wiki-visits at earnings announcements are more pronounced when the contextual information in prior financial reports is more difficult to understand.*

Next, I expect that investors are more likely to visit wiki-pages to interpret the newly released financial information when the announcing firm has a higher intangible intensity. Prior research shows that, for new economy firms, financial information alone is basically value-irrelevant and only becomes value-relevant when combined with non-financial information (e.g., Amir and Lev 1996). One reason is that many intangible assets are not recognized in the balance sheet, leaving a hole in the balance sheet. To make it worse, firms normally provide only limited voluntary disclosures about intangible assets, probably due to proprietary concerns (e.g., Glaeser 2018). However, volunteer editors in Wikipedia are largely free from such concerns and they gather information from sources outside the firm to draw a complete picture of the firm. In fact, they often discuss intangible assets based on information gathered from patent offices, customer surveys, etc., as shown in Section II. Considering that intangible-intensive firms offer investors opportunities to make greater investment returns by acquiring more information to gain an informational advantage, I predict that investors are more likely to visit these firms' wiki-pages to supplement financial information.

H2c. *The abnormal wiki-visits at earnings announcements are more pronounced for firms with higher intangible intensities.*

If investors visit wiki-pages at earnings announcements to assist their interpretation of earnings news, stock prices should be more responsive to earnings news immediately after its release, thereby accelerating the speed of price discovery. This prediction is not without tension. In a perfect market with immediate and efficient information processing by rational investors equipped with all necessary contextual information, the market should react to news instantaneously and completely, regardless of wiki-visits. In reality, price discovery is an ongoing process (Lee 2001), hence the impact of wiki-visits on the price discovery depends on the marginal

contribution of wiki-visits to the efficient processing of earnings news. Wiki-pages summarize firms' background and key historical transactions, whose usefulness to the interpretation of current earnings may turn out to be limited because alternative information such as analyst reports and news articles is sufficient for investors. Furthermore, retail investors, as opposed to institutional investors who have access to professional data sources, are more likely to acquire information from Wikipedia. These retail investors may have limited impacts on price discovery as they are not the dominant players in the capital market or even negative impacts as their overall information set and information processing ability are inferior to those of institutional investors (Han, Tang, and Yang 2016). Taken together, I state the third set of hypotheses below in an alternative form:

H3a. *The immediate market reaction towards earnings news is more pronounced when more investors acquire contextual information from Wikipedia at earnings announcements.*

H3b. *The speed of price discovery at earnings announcements is faster when more investors acquire contextual information from Wikipedia at the announcements.*

IV. Sample and Data

I obtain daily wiki-visits to the wiki-pages of S&P 500 companies from July 1st, 2015 to December 31st, 2017 through Wikimedia Foundation's RESTBase Application Programming Interface (API). The sample starts on July 1st, 2015, when the API for wiki-visits became available. I only include human visits and exclude robot visits (pageviews through API, spiders, or web crawlers) as classified by Wikimedia Foundation based on the header information sent by the visitor's web browser or mobile device. [20] To capture the information demand of investors rather than customers, I focus on the wiki-pages of the company (e.g., Google LLC) rather than its products (e.g., Google Chrome or Gmail). [21] Occasionally, companies have multiple wiki-pages, for example, Google LLC has one titled "Google" and another titled "Alphabet Inc.". I use total wiki-visits in reported analysis and get almost identical results when using average wiki-visits.

[20] The data include visits by editors, although their visits should account for a small proportion of total visits as an average wiki-page only has 3.2 unique editors in a four-week window.

[21] I include wiki-visits redirected from other pages (e.g., a search in Wikipedia for Google's ticker symbol "Googl" will be automatically redirected to the target wiki-page titled "Google"). If a user browses to a redirect, Wikipedia records a page view for the redirect but not the target wiki-page.

To quantify the content and revision frequency of wiki-pages, I use API to obtain the revision history of the sample wiki-pages, including revision size (measured in bytes), revision date, and volunteer editors (online username or IP address if the editor is not registered), as well as the content of wiki-pages (including section titles and references) at a particular point in time.

I obtain firm characteristics, stock performance, analyst forecasts, management forecasts, and institutional ownership from standard sources including Compustat, CRSP, IBES, and Thomas Reuters. To control for the influence of media coverage on wiki-visits, I obtain news coverage from Capital IQ Key Development. To control for the overall information demand for the firm, I obtain daily Google search volume (GSV) of the firm's ticker from Google Trends, which is an index ranging from 0 to 100 to quantify the search intensity.[22] Trading activities by retail investors or institutional investors are identified from the TAQ millisecond datasets. Data on SEC filings are obtained from the SEC EDGAR system. Please refer to Appendix B for detailed variable definitions and data sources.

Panel A of Table 2 presents the summary statistics. To minimize the influence of outliers, all variables are winsorized at 1% and 99% except dummy variables, count variables, and variables that have been taken logarithm. In total, the sample includes 564 unique firms that are in the S&P 500 index at any time from July 2015 to December 2017. These firms yield 510,592 firm-calendar-day observations and 337,732 firm-trading-day observations.[23] The average visits to a sample wiki-page (*Raw_Wiki*) are 996 per day. In contrast, only 112 requests are made for any SEC filings at EDGAR on a typical day. The frequent visits to wiki-pages provide initial evidence that wiki-pages contain information useful to investors.

With respect to the primary independent variables, the sample firms are large, complex, and fairly visible. These firms on average have significant intangible assets (on-[off-]balance sheet

[22] GSVs are available on a daily (weekly) if the search period is less than 90 days (5 years). Since GSVs are scaled by the highest search volume within the search period, daily GSVs over two 90-day windows are inconsistent. Following Chi and Shanthikumar (2018), I download daily GSVs every 90 days and weekly GSVs for the entire sample period, and multiply daily GSVs by the corresponding weekly GSVs, dividing by 100, to accumulate a consistent time-series of daily GSVs. Note that these adjusted daily GSVs are still not comparable across firms (a 100 GSV could mean different search intensities for firm A and B). I construct abnormal daily GSVs (*Ab_GSV)* as the log difference between the adjusted daily GSV and its same day-of-week median over the past 10 weeks to further ensure comparability.

[23] I include all calendar days in the analyses of the determinants of contextual information acquisition. I only include trading days in the validation tests and the analyses of the consequences of contextual information acquisition as daily data on trades or returns are only available on trading days.

intangible assets are worth 26.2% [23.9%] of the book value of total assets), 2.57 business segments, 75.8% institutional ownership, and 4% short interest. They are followed by 13.61 analysts and covered by media on 13.6% of the sample calendar days.

Panel A also reports the frequency of the six information events of interest. 1.1% firm-days have earnings announcements, 0.8% have management forecasts, and 1.0% have 10-K or 10-Q filings. M&A announcements, stock price sudden crashes, and jumps are much less frequent. The summary statistics of market variables show that retail investors are net buyers and institutions are net sellers on average. Not surprisingly, both the average wiki-visits (1,092 versus 996) and news coverage (18.6% versus 13.6%) are higher on trading days than on calendar days.

Panel B of Table 2 tabulates the sample observations based on the wiki-visit quintiles over two consecutive days. It shows that wiki-visits are highly persistent. For example, 85.0% of firms with bottom-quintile current raw wiki-visits stay in the bottom quintile the next day. Only 14.5% of bottom-quintile firms switch to the second quintile, and almost none of them go beyond that. The persistence is weaker for intermediate firms but stronger for top-quintile firms. When I adjust the raw wiki-visits by its rolling median to partially remove firm heterogeneities and construct abnormal wiki-visits (*Ab_Wiki* [*Ab_Wiki%*], the level [percentage] difference between *Raw_Wiki* and its median for the same day of the week over the past 10 weeks), the persistence becomes smaller.

Panel C of Table 2 reports the Pearson's pair-wise correlation matrix between the selected key variables. The three measures of wiki-visits are significantly correlated, although *Raw_Wiki* is less correlated with the two measures of abnormal wiki-visits (*Ab_Wiki* and *Ab_Wiki%*). Since both GSV and requests of SEC filings at EDGAR capture information acquisition, they positively correlate with wiki-visits. Below I discuss how they differ from wiki-visits, conceptually and empirically.

Wiki-visits differ from GSV in the following respects. First, given the widely-held perception of the type of information contained in wiki-pages, it is reasonable to assume people visiting a firm's wiki-page are looking for contextual information about the firm. However, people googling the same term may seek different information (e.g., quantitative information such as a firm's stock price, qualitative information such as CEOs' shareholder letters, or even information unrelated to the firm), hence it is difficult, if not impossible, to infer from GSV the type of

information investors intend to acquire *ex ante*. Second and relatedly, researchers can *ex post* quantify the information accessed by investors from Wikipedia by reverse-engineering the historical wiki-page, but we do not know the exact information obtained by investors through Google search. Lastly, Drake et al. (2012) document a substitution between accounting information and information acquired through Google search while my evidence shows that the information acquired from wiki-pages is complementary to accounting information.

Wiki-visits also differ from requests of SEC filings through EDGAR (EDGAR-visits), conceptually and empirically. Wiki-pages contain primarily contextual information about the firm. The contents of SEC filings vary a lot depending on the filing type. In general, they are a complicated mix of qualitative and quantitative information, which requires a certain level of sophistication to process. Therefore, the user base of EDGAR is tilted toward sophisticated investors (Crane, Crotty, and Umar 2018), although some of them are retail investors (Chi and Shanthikumar 2018). In Appendix C, I examine the time-series properties of wiki-visits and EDGAR-visits and find that both wiki-visits and EDGAR-visits capture investor attention although wiki-visits lead EDGAR-visits more often than the other way around. In addition, I run a horse race between wiki-visits and EDGAR-visits in explaining the subsequent retail trades. I find that wiki-visits have a slightly stronger predictive power for retail trades than EDGAR-visits, suggesting that wiki-visits better capture information acquisition by retail investors.

The use of wiki-visits as a proxy for contextual information acquisition comes with several limitations. First, Wikipedia is just one of many sources of contextual information, although it is a prominent one. Second, accessing wiki-pages does not guarantee that investors use the information contained in these pages in their investment decision process, although accessing the information is a necessary step toward its use.[24] Third, I do not observe the identity of the readers of wiki-pages, who may access wiki-pages for non-investment purposes. I will address this issue at length in Section V. Lastly, information acquisition from Wikipedia could correlate with other online information acquisition in timing. In all analyses, I include daily abnormal GSV in an attempt to control for overall online information acquisition.

[24] Many studies of online information acquisition face similar issues (e.g., Drake et al. 2012, 2015, 2016; Lawrence, Ryans, and Sun 2017; Lawrence, Ryans, Sun, and Soni 2017).

V. Wiki-visits, Retail Trades, and Subsequent Returns

In this section, I empirically examine whether wiki-visits capture information acquisition by investors, and if so, by what types of investors (retail versus institutional investors). Before turning to regression analysis, I tabulate the clickstreams that lead to MT's wiki-page in January 2017 in Panel D of Appendix A to shed light on the purpose of wiki-visits. External search engines such as Google or Bing account for 58.3% of the incoming traffic, suggesting that people visiting wiki-pages are actively seeking information. Internal search through Wikipedia's home page ("Microchip (disambiguation)" or "Main Page") or lists (e.g., "List of semiconductor fabrication plants") also account for a large proportion of the incoming traffic. In addition, visits to MT's wiki-page are often preceded by visits to MT's product and technology (e.g., "PIC microcontroller", "Microcontroller"), major customers (e.g., "Apple Inc.", "Microsoft"), or industry peers (e.g., "Qualcomm", "Cypress Semiconductor"). This clickstream pattern evidences investors navigating from one web page to another to gather information.

To further validate that wiki-visits capture information acquisition by investors, I conduct regression analysis between wiki-visits and subsequent trades. Intuitively, retail investors are more likely to turn to Wikipedia for contextual information, as institutional investors are usually equipped with necessary background knowledge about public firms and have access to professional data sources. I, therefore, predict wiki-visits to be positively associated with retail trades but not institutional trades. I then employ two approaches to rule out alternative interpretations for the regression results.

Wiki-visits and Subsequent Trades

I follow Boehmer et al. (2017) and Bushee et al. (2018) to identify retail trades and institutional trades from TAQ millisecond consolidated files, respectively. Retail trades are often executed off-exchange, thus recorded with exchange code "D" in TAQ, and wholesalers often offer a small price improvement (typically a fraction of a penny per share) to attract retail orders (Boehmer et al. 2017). I hence classify retail sale (buy) trades as those with TAQ exchange code "D" and prices just above (below) a round penny.[25] One caveat is that this approach does not

[25] To be conservative, trades with prices at a round penny or near the half-penny (0.4-0.6 cents, inclusive) are not classified as retail trades. I determine sale or buy transactions using the Lee and Ready (1991) algorithm. The

identify retail trades if they are limit orders or market orders that are executed on the exchange. As most retail trades arise from market orders and receive price improvements (e.g., 90% for TD Ameritrade in the second quarter of 2016), this approach should pick up the majority of retail trades (Boehmer et al., 2017) and should not bias my analysis systematically. Non-retail trades larger than $50,000 are classified as institutional trades (Bushee et al. 2018). Unclassified trades are excluded from this analysis.

I report the results in Panel A of Table 3. The purpose is to see whether wiki-visits predict the cross-section of subsequent trades, hence I use panel regressions with day fixed effects.[26] To further control for time-invariant firm heterogeneity, I also add firm fixed effects as a robustness check. The analysis is restricted to trading days and all independent variables are lagged by one trading day. The dependent variables are the net trading activities by retail investors and institutional investors, respectively, where the net trading activities are the number of buy transactions minus the number of sale transactions (*NetTrade*) or net shares traded (*NetVolume*). The key variable of interest is abnormal wiki-visits (*Ab_Wiki*), measured as the raw wiki-visits adjusted by the median wiki-visits for the same day of the week over the past 10 weeks. Following Baber and Odean (2008) and Da et al. (2011), I include five proxies for investor attention, defined at the daily level: the absolute value of abnormal return (|*AR*|--individual stock returns adjusted by 5 x 5 value-weighted portfolio returns matched by size and book-to-market ratio), the absolute value of cumulative abnormal stock returns in the past four weeks (|*CAR_Month*|), abnormal turnover (*AVol*--total shares traded scaled by shares outstanding, minus the same day-of-the-week median over the past 10 weeks), media coverage (*News_Dummy*--a dummy variable indicating whether the firm is in the news), and abnormal GSV (*Ab_GSV,* which is also a proxy for overall online information acquisition).[27] I also control for firm-specific time-varying factors as of the most recent data date available to investors. These control variables include common firm characteristics such as firm size, firm age, ROA, loss indicator, financial leverage, book-to-market ratio, investor base measured by institutional ownership, firm visibility proxied by analyst following and short interest. To ease the comparison of the coefficients, I standardize all

algorithm requires the data on the National Best Bid and Offer (NBBO) which are constructed by combining the NBBO files and Quotes files in TAQ following Holden and Jacobsen (2014) who show that the NBBO files alone miss the best quotes when a single exchange provides the best bid and offer quotes at the same time.

[26] The untabulated results using Fama-MacBeth regressions yield similar inferences. All untabulated results mentioned throughout the paper are available upon request.

[27] The results are similar if I use the natural logarithm of one plus the number of news articles.

independent variables to have a mean of zero and standard deviation of one. See Appendix B for detailed variable definitions.

Consistent with the intuition that retail investors are more likely than institutional investors to acquire contextual information from Wikipedia, I find that abnormal wiki-visits are positively associated with net purchases by investors but not by institutional investors. Economically, a one-standard-deviation increase in *Ab_Wiki* is associated with a 21.5% (=5.325/24.779) increase in subsequent retail *NetTrade* or a 27.0% (=0.625/2.316) increase in retail *NetVolume*.[28] I also scale the dependent variable *NetTrades* and *NetVolume* by the total number of transactions or shares traded, repeat the regressions, and obtain consistent results.

Comparing the coefficients on control variables, I find interesting differences between retail trades and institutional trades. First, retail investors are net buyers of attention-grabbing stocks, i.e., stocks with larger prior price movements, higher turnovers, or more media coverage, consistent with Baber and Odean (2008). Meanwhile, institutional trades exhibit the opposite or insignificant association with these factors. Relatedly, Google search volume is positively associated with retail net buying, consistent with Da et al. (2011), but the evidence on its association with institutional net buying is mixed. Lastly, retail investors tend to buy relatively large-cap stocks and growth stocks, while institutional investors tend to buy relatively small-cap stocks and value stocks.

Wiki-visits and Subsequent Returns

The positive association between wiki-visits and subsequent retail trades is consistent with retail investors visiting wiki-pages as part of their due diligence before making investment decisions. It is also consistent with retail investors trading attention-grabbing stocks (Barber and Odean 2008). For example, some shocks to a firm's visibility induce the general public to visit its wiki-page. These wiki-visits could be spuriously correlated with the attention-driven trades made by retail investors even if none of them visits the firm's wiki-page. This alternative interpretation implies high wiki-visits are associated with zero subsequent returns if these attention-driven trades do not move the market, or positive subsequent returns followed by return reversals due to

[28] In unreported results, I regress retail buys and sales separately on wiki-visits. I find the coefficients on wiki-visits are significantly positive in both regressions, but the magnitude is larger for retail buys than retail sales.

temporary price pressures. However, the information acquisition interpretation suggests that net purchases by retail investors are justified by the information collected from wiki-pages, which should generate positive returns without reversals.

To distinguish the above-discussed two interpretations, I examine the relationship between wiki-visits and subsequent abnormal returns in Panel B of Table 3. I accumulate the abnormal returns over the next week (week 1, *CAR[1W]*), the week after the next (week 2, *CAR[2W]*), weeks 3 and 4 (*CAR[3W-4W]*), and weeks 5 to 12 (*CAR[5W-12W]*). I find a significant positive correlation between abnormal wiki-visits (*Ab_Wiki*) and abnormal returns cumulated over the next week (*CAR[1W]*). The coefficient on *Ab_Wiki* remains positive but becomes insignificant for longer windows. A one-standard-deviation increase in *Ab_Wiki* is associated with positive abnormal returns of 2.5 basis points (bps) in week 1, 1.3 bps in week 2, 1.1 bps in weeks 3 and 4 (0.55 bps per week), and 2.9 bps from week 5 to week 12 (0.36 bps per week). The evidence that the coefficients on *Ab_Wiki* taper off gradually but do not change signs undercuts the plausibility of the alternative interpretation.

In Columns 5-8, I split the sample based on the quality of wiki-pages and repeat the same regression procedures in the two subsamples. If contextual information improves investors' decision making, its impact should increase with its quality. I use the number of references or the unique number of editors who made edits in the past four weeks to proxy for the information quality, assuming that the information in wiki-pages with more references (recent edits) is richer and more credible (more up-to-date).[29] The regression results show that the positive association between *Ab_Wiki* and *CAR[1W]* found in the whole sample is mainly driven by firms with high-quality wiki-pages, providing further support for the information acquisition interpretation.

The Launch of Google Knowledge Graph

To further differentiate the two interpretations, in this section I exploit the launch of Google Knowledge Graph (GKG) on May 16, 2012, as a natural experiment. GKG presents an information box on the top right of the first page of Google search results. The information box usually includes a brief summary description of the firm, its official website, wiki-page link, and other basic facts

[29] Results hold if I use alternative quality measures, such as the number of words or sections in the wiki-page. I also use the quality scores assigned by Wikipedia based on machine learning techniques and find similar results. The page score data are described at length by Halfaker (2016), who further updated the data through July 2017.

such as its ticker symbol, headquarters, and key executives (see Panel C of Appendix A for an example). The purpose of GKG is to make search results more intelligent by providing direct answers. Firms other than Google LLC have no control as to whether or when to launch GKG. [30]

GKG resolves trivial inquiries by providing basic facts about firms, thus, I assume that wiki-visits post GKG are more likely to involve meaningful information acquisition, at least information beyond what is shown in the information box.[31] To verify this assumption, I plot the median *Raw-Wiki* around the launch of GKG in Figure 2. I classify firms into four groups based on the *Raw-Wiki* in the 10 weeks before March 2012. Wiki-visits of relatively lesser-known firms (Panel A) slightly decrease after the launch of GKG, relative to the pre-trend depicted in the dashed line. For relatively better-known firms (Panel D), wiki-visits drop significantly in the post-GKG period, consistent with post-GKG wiki-visits involving more meaningful information acquisition. As a result, wiki-visits will have larger consequences for subsequent trading activities and returns after the launch of GKG than before, as the information acquisition interpretation implies. On the contrary, if the alternative interpretation holds, there is no reason to expect the spurious relation between wiki-visits and subsequent trades or returns to change systematically around the launch of GKG.

I extract the number of visits to sample firms' wiki-pages in 2012 from the entire page-view database maintained by the Wikipedia Media Foundation, as Wikipedia RESTBase API is not available for wiki-visits before July 2015.[32] The 2012 wiki-visits differ from the API wiki-visits mainly in two respects. First, the 2012 data do not include visits via mobile phones. Second, the data do not allow me to exclude visits by web crawlers. These two limitations add noise to the data, biasing the coefficient estimates towards zero. Nevertheless, these limitations should not

[30] Google unveiled GKG on May 16, 2012, but I do not observe the exact date when the information box became available for each individual firm. I verified that all sample firms were included in Freebase before the launch of GKG. As GKG was powered in part by Freebase, I assume that these firms' information boxes were immediately available after the launch. Please see https://googleblog.blogspot.com/2012/05/introducing-knowledge-graph-things-not.html for more details about GKG and https://developers.google.com/freebase/ for Freebase.

[31] Anecdotal evidence supports this point. See discussion at:
https://www.theregister.co.uk/2014/01/13/google_stabs_wikipedia_in_the_front/ ;
https://www.forbes.com/sites/jaysondemers/2015/09/03/is-the-google-knowledge-graph-killing-wikipedia/#925f8c0138cd

[32] This analysis includes 526 out of the 564 firms in the main analyses because some wiki-pages had not been created yet in 2012. The entire wiki-visit database can be accessed at https://dumps.wikimedia.org/other/pagecounts-raw/2012/. The data record page-views for each wiki-page by hour (UTC time, further adjusted to EST) and I aggregate these hourly visits by day.

affect the comparison around the launch of GKG as both the pre- and post-subperiods exhibit the same issue.

Panel A (B) of Table 4 reports the impact of abnormal wiki-visits on trades (returns) before or after the introduction of GKG, estimated by panel (Fama-MacBeth) regressions. Consistent with the idea that more meaningful information acquisition has bigger consequences for subsequent investment decisions, I find a significant positive association between wiki-visits and the next-day retail net purchases after the launch of GKG but not before. Moreover, the association between wiki-visits and institutional trades is insignificant both before and after the launch of GKG. This non-result mitigates the possibility that the change in the association between wiki-visits and retail trades is caused by market condition changes that coincide with the introduction of GKG. Lastly, I find that wiki-visits positively predict the next-week abnormal returns only in the post-GKG period and that the predictive power is driven by firms with higher quality wiki-pages of higher quality. To summarize, these results further undercut the plausibility of the alternative interpretation and provide more support for the information acquisition interpretation.

VI. Wiki-visits around Earnings Announcements

In this section, I investigate the circumstances under which investors acquire contextual information from Wikipedia to understand how investors assemble information to assist their decision making. I first investigate the changes in wiki-visits around six key events. I then focus on earnings announcements and examine how wiki-visits vary with the characteristics of earnings announcements to uncover when contextual information is more useful to investors. Lastly, I study how wiki-visits affect the price discovery of earnings news to shed light on the consequences of contextual information acquisition.

When do investors acquire contextual information?

I test H1 by examining the changes in wiki-visits around the following events: earnings announcements, management forecasts, annual and quarterly financial statement filings, M&A announcements, and extreme stock movements that do not coincide with the other four events. The first three events are common financial reporting events. M&A deals are important corporate events. I also include extreme stock movements because they induce investors to gather more

information to adjust their portfolios (Drake et al. 2016). Please refer to Appendix B for the definitions of these events.

To illustrate the changes in wiki-visits around the above-mentioned events, Figure 3 plots the average abnormal wiki-visits in an 11-day window centered around the events. On average, wiki-visits rise by 101.5 (15.0%) on the day of earnings announcements and remain at this elevated level the following day before gradually returning to the normal level. There is also a small increasing trend leading up to the announcements. The spike of wiki-visits at earnings announcements is close to the spike at the issuance of management forecasts (16.4%), much larger than that at the filings of 10-Ks or 10-Qs (8.3%), but much smaller than that at M&A announcements (41.4%). The spike at earnings announcements is comparable to the 13.2% increase of Google search volume documented by Drake et al. (2012).

Extreme stock movements also trigger more visits to wiki-pages. Interestingly, stock price jumps trigger more wiki-visits than crashes (42.4% versus 31.6%). The difference is statistically significant at the 1% level. This pattern is the opposite of Drake et al. (2016), who document that investors seek out more historical accounting information from EDGAR in response to crashes than jumps. This discrepancy is likely caused by the difference in the user base and the information content between Wikipedia and EDGAR. Compared to those using EDGAR, investors using Wikipedia tend to be less sophisticated (Crane et al. 2018). These investors rarely short stocks, hence irrational herding and limited attention predict that they pay more attention to stock price jumps than to crashes unless they already own the stock (Barber and Odean 2008). Moreover, Wikipedia provides investors with contextual information about firms' background and historical key transactions, hence investors are more likely to turn to Wikipedia to get a sense of the company before buying the stock than before selling it as shown in Section V. On the contrary, relatively more sophisticated investors are likely to use SEC filings to understand the reasons for the crashes and to identify under-valued stocks after crashes (Drake et al. 2016).

In addition to the visual examination, I estimate the following regression model to study investors' contextual information acquisition around key information events:

$$Ab\_Wiki_{it} = \alpha_0 + \alpha_1 EA_{it} + \alpha_2 MF_{it} + \alpha_3 10K\&10Q_{it} + \alpha_4 M\&A_{it} + \alpha_5 Crash_{it} + \alpha_6 Jump_{it} + \alpha W_{it} + Firm_i + Date_t + \epsilon_{it} \quad (1)$$

where i represents the firm and t the date. In the baseline regression, the dependent variable is abnormal wiki-visits (*Ab_Wiki*). As robustness checks, I also use abnormal percentage wiki-visits (*Ab_Wiki%*) and the natural logarithm of one plus raw wiki-visits (*Ln(Raw_Wiki)*). The variables of interest are dummy variables that indicate the occurrence of earnings announcements (*EA*), management forecasts (*MF*), 10-K and 10-Q filings (*10K&10Q*), merger & acquisition announcements (*M&A*), and stock price crashes and jumps (*Crash* and *Jump*). As firm fixed effects remove time-invariant firm heterogeneities, the coefficients on the event dummies can be interpreted as the responses of wiki-visits to these events. I include date fixed effects to control for changes in the popularity of Wikipedia over time. I include all control variables used in Table 3 plus two variables (intangible intensity and the number of business segments) to control for the impact of business complexity on investors' general propensity to acquire more information.

Consistent with the graphic evidence, the regression results reported in Table 5 demonstrate significant spikes of wiki-visits on the event day across all events except 10-K or 10-Q filings.[33] Overall, the economic magnitude is smaller than the univariate analysis shown in Figure 3, probably for two reasons. One is the inclusion of firm and date fixed effects as well as other control variables. Particularly, I control for media coverage (*News_Dummy*), prior stock movements (|*CAR_Month*|), and overall online information acquisition (*Ab_GSV*) whose coefficients are significantly positive.[34] The second reason is the correlation across events, especially the bundling of earnings announcements and management forecasts. Consistent with this point, comparing Columns 2 and 3, I find that the spikes at earnings announcements decrease considerably after controlling for management forecasts. In the untabulated analysis, I find that earnings announcements bundled with management forecasts on average are associated with fewer wiki-visits than standalone earnings announcements. My interpretation is that bundled earnings announcements are likely to be supplemented with contextual information (Hutton et al. 2003), reducing investors' demand for the contextual information in wiki-pages.

In terms of the timing, again consistent with the graphic evidence, Columns 2-4 of Table 5 show that wiki-visits increase slightly one day before earnings announcements and markedly on the announcement day and the calendar day after. A plausible reason for the continuing high

[33] The potential reasons for the weak response are that few investors pay attention to 10-K and 10-Q filings and that those who read these filings, most likely sophisticated investors, have little demand for the information in wiki-pages.
[34] I also use the natural logarithm of one plus the number of news articles and obtain very similar results.

abnormal wiki-visits one day after the announcements is that the predominant users of wiki-pages are less sophisticated investors, who tend to respond to events slowly (Ben-Rephael et al. 2017). The increase in wiki-visits one day before earnings announcements is economically small across all regression specifications and becomes statistically insignificant when I use the raw wiki-visits as the dependent variable. The pre-announcement increase is even smaller for management forecasts and M&A, as these two events are less anticipated than earnings announcements. The pre-event increase is statistically significant for extreme stock movements, although economically smaller than the corresponding event-day spike. This pre-event increase is probably due to return momentum. Overall, the evidence indicates that investors acquire information from Wikipedia in response to informational events with a lag but not proactively, consistent with information acquisition for interpretation purposes modeled by Kim and Verrecchia (1994).

I obtain consistent results using abnormal percentage wiki-visits (Columns 5-6) and raw wiki-visits (Columns 7-8). The R-squared goes up to 96% for raw wiki-visits, largely because firm fixed effects explain a significant proportion of the total variation in raw wiki-visits. In unreported results, I use the reverse Fama-MacBeth regressions (i.e., estimate the regression by firm and take the average of coefficients across firms) to isolate the estimation from cross-sectional correlations (Peterson 2009) and find similar results.

To test H2a-H2c, I examine how the abnormal wiki-visits at earnings announcements vary with three characteristics: the opportunism in earnings press releases, the textual complexity of prior financial reports, and the deficiency of balance sheets as measured by off-balance sheet intangible intensity. The regression model is similar to Model 1 with two modifications. First, I focus on the day of and the day following earnings announcements (*EA[0,1]*) as the previous results show that wiki-visits stay at the elevated levels one day after the announcements. Second, I split earnings announcements into two categories based on the median of the characteristics under examination. I report the results in Table 6 where the dependent variables are abnormal wiki-visits (*Ab_Wiki*). In unreported results, I use abnormal percentage wiki-visits and find qualitatively similar results. The results using raw wiki-visits are weaker but consistent.

Before testing H2a, I first distinguish announcements of positive and negative earnings measured relative to analyst consensus forecasts. The results, reported in Column 1 of Table 6, show that the abnormal wiki-visits do not differ for positive and negative earnings

announcements.[35] This insignificant difference is probably due to two offsetting effects. On the one hand, negative earnings raise red flags to investors who subsequently seek out more information. On the other hand, positive earnings surprises draw the attention of investors with herding behaviors, who visit wiki-pages to better understand the firm's business. These results show that positive or negative earnings news per se does not trigger different responses in wiki-visits.

To test H2a, I identify earnings announcements that are delivered in a tone misaligned with the earnings news. Following Huang et al. (2013), I measure abnormal tone as the residual from annual regressions of unadjusted tone (the proportion of positive words minus that of negative words) on earnings news and firm characteristics.[36] I then compare wiki-visits at opportunistic announcements (above-median abnormal tone) to other announcements. The results, reported in Column 2, show significantly larger abnormal wiki-visits at opportunistic earnings announcements than other announcements (61.70 versus 48.60). I obtain similar results if I identify opportunistic earnings announcements as those with below-median earnings surprises and above-median unadjusted tone. These results support H2a that investors are more likely to seek out contextual information from Wikipedia when they suspect that earnings press releases are opportunistic.

To test H2b, I extract the text from Item 1. "Business" and Item 2. "Properties" from the most recent 10-K filings, an alternative and perhaps more authoritative source of contextual information, which largely overlaps with the information in wiki-pages. I then perform textual analysis to measure its quality in terms of readability and boilerplate language as discussed in Section II. The results in Columns 3-4 show that investors are more likely to visit wiki-pages at earnings announcements when the announcing firms have poorer readability and more boilerplate language in their 10-Ks. These results support H2b that wiki-pages are more useful for investors when financial reports are more difficult to understand. These results also imply that investors vote

[35] I do not find significant differences in the abnormal wiki-visits at the announcements of above- versus below-median earnings surprises. I also construct earnings surprises relative to the earnings of the same quarter last year or zero, and do not find any significant differences in the abnormal wiki-visits, either.

[36] I download earnings press releases from filings of Form 8-K with Item 2.02 within [-1,4] where day 0 is the earnings announcement date. I use Python to extract the attached exhibits and count the positive and negative words using the dictionary developed by Loughran and McDonald (2011). If no exhibits are filed, the body of the 8-K filing (starting with various forms of "Item 2.02" and ends before the signature) is used instead. Less than 5% of earnings announcements cannot be matched with earnings press releases. The estimation of abnormal tone further reduces the sample size due to missing value of the variables used in Huang et al. (2013)'s tone model.

with their feet and that the textual complexity of 10-K filings prevents investors from fully appreciating the information contained.

To test H2c, I obtain data of both on- and off-balance sheet intangible assets from Peters and Taylor (2017).[37] I measure intangible intensity as intangible assets scaled by the book value of total assets. Column 5 (6) presents the results with earnings announcements split by firms' on- (off-) balance sheet intangible intensity. I find that earnings announcements by firms with a higher off-balance sheet intangible intensity trigger significantly more wiki-visits, but do not find significant differences in wiki-visits for on-balance sheet intangible intensity. These results support H2c that the information in wiki-pages is more useful to interpret earnings announced by new economy firms with a higher off-balance sheet intangible intensity.

Taken together, the analyses in this section document that wiki-visits spike at key informational events and that the abnormal wiki-visits at earnings announcements increase with the marginal usefulness of wiki-pages. These results are consistent with investors acquiring contextual information in order to better interpret the informational events.

Consequences of Contextual Information Acquisition

I test H3a by examining whether the abnormal wiki-visits at earnings announcements affect the market reactions towards earnings news around earnings announcements. Specifically, I estimate the following cross-sectional regressions:

$$CAR[0,1] = \beta_0 + \beta_1 SUE + \beta_2 Ab\_Wiki[0,1] + \beta_3 (SUE * Ab\_Wiki[0,1])$$
$$+\beta_x Controls + \beta_y (Controls * SUE) + \epsilon \quad (2a)$$

$$CAR[-5,-1] = \gamma_0 + \gamma_1 SUE + \gamma_2 Ab\_Wiki[-5,-1] + \gamma_3 (SUE * Ab\_Wiki[-5,-1])$$
$$+\gamma_x Controls + \gamma_y (Controls * SUE) + \epsilon \quad (2b)$$

$$CAR[0,1] = \delta_0 + \delta_1 SUE + \delta_2 Ab\_Wiki[-5,-1] + \delta_3 (SUE * Ab\_Wiki[-5,-1])$$
$$+\delta_x Controls + \delta_y (Controls * SUE) + \epsilon \quad (2c)$$

where *CAR* is cumulative abnormal returns and *SUE* is unexpected earnings relative to analyst forecast consensus, scaled by stock prices. Model 2a is an expansion of the earnings response

[37] The data are downloaded from https://wrds-web.wharton.upenn.edu/wrds/query_forms/navigation.cfm?navId=421

coefficient (ERC) model with additions of *Ab_Wiki[0,1]* (the average of *Ab_Wiki* during the two-trading-day window of earnings announcements, divided by one thousand) and its interaction with *SUE*. The coefficient of interest is β_3. A positive β_3 is consistent with H3a, which suggests that contextual information improves investors' interpretation of earnings news, thereby enhancing the association between the immediate market reaction and earnings news.

To explore the timing, following Drake et al. (2012), I define the window to be one week prior to earnings announcements, i.e. trading days [-5,-1], as specified in Model 2b. Essentially, this model is a variation of the future ERC model, testing whether the pre-announcement stock price incorporates more of the upcoming earnings news when more investors visit the firm's wiki-page in the week prior to the announcement. Relatedly, I also estimate Model 2c to see whether wiki-visits before earnings announcements affect ERC. The variable of interest in both Model 2b and Model 2c is the interaction of unexpected earnings and pre-announcement wiki-visits (*SUE*Ab_Wiki[-5,-1]*). Considering that the information in wiki-pages by itself (e.g., firms' background and historical key transactions) is unlikely to predict future earnings, I do not expect that the pre-announcement wiki-visits accelerate the price discovery of future earnings (γ_3 is 0). Considering that it is difficult to seek relevant contextual information for the purpose of interpretation before the information to be interpreted is announced, I also do not expect the pre-announcement wiki-visits to increase ERC (δ_3 is 0). These predictions are different from Drake et al. (2012)'s findings that online information search speeds up the price discovery of upcoming earnings news before the announcements, thereby attenuating the market reactions at the announcements. Taken together, the purpose of Models 2a-2c is to differentiate Wikipedia from other online information sources and to highlight the role of contextual information in interpreting new financial information.

I include in Models 2a-2c the same set of control variables used in Model 1 with three modifications. First, I exclude *ROA* as it highly correlates with *SUE*. Second, I exclude *CAR_Month* (cumulative abnormal returns over the last four weeks) as it overlaps with the dependent variable *CAR[-5,-1]*. Third, I modify the dummy variable indicating whether the firm is in the news (*News_Dummy*) to be the number of news articles in the corresponding window to capture information discovery and dissemination by the media (*News[-5,-1]* and *News[0,1]*). Lastly, I compute the average abnormal GSV over the corresponding window (*Ab_GSV[-5,-1]* and *Ab_GSV[0,1]*) to capture overall information acquisition. Following the prior literature (see

Kothari (2001) for a review), I also add earnings persistence (*Earn_Rho*), earnings volatility (*Earn_Std*), return volatility over the fiscal quarter (*Ret_Std*), and a fourth fiscal quarter indicator (*Fqtr4*). The regressions also include the interactions between all control variables and earnings news (*SUE*). To simplify the computation of economic magnitude, I standardize these control variables (all independent variables except *SUE* and wiki-visits) to have a mean of zero and a standard deviation of one. To ease the presentation, the coefficients on these control variables and their interactions with *SUE* are not tabulated. The sample includes earnings announcements made by the sample firms from July 2015 to December 2017. Panel A of Table 7 reports the summary statistics for this sample.

I report the estimates of Models 2a-2c in Panel B of Table 7. Consistent with my expectation, the coefficient on *SUE*Ab_Wiki[-5,-1]* is insignificant in Column 1 where the dependent variable is pre-announcement cumulative abnormal returns (*CAR[-5,-1]*). This result means that pre-announcement wiki-visits do not affect the price discovery of upcoming earnings news, highlighting the difference between contextual information acquisition from Wikipedia and online information acquisition in general as studied in Drake et al. (2012). More interestingly, the coefficient on *SUE*Ab_Wiki[-5,-1]* is insignificant in Column 2 while that on *SUE*Ab_Wiki[0,1]* is significantly positive in Column 3, indicating that ERC significantly increases when investors pay more visits to wiki-pages at earnings announcements but not before. Economically, a one-standard-deviation increase in *Ab_Wiki[0,1]* is associated with a 15.6% increase in the ERC, holding all other variables at sample mean.[38] Moreover, this significant increase is driven by firms with higher analyst forecast dispersions, as shown in Columns 4-5, consistent with wiki-pages being more useful to investors of opaque firms. Overall, the results provide support for H3a that the immediate market reactions towards earnings news are stronger when more investors acquire contextual information from Wikipedia.

Next, I use the cumulative abnormal trading volume to measure market reactions. Stock returns reflect the change in investors' expectations as a whole while volumes capture changes in expectations of each investor, hence volume reactions are more sensitive than return reactions

[38] The economic magnitude is computed as 4.012*0.331/(8.067+4.012*0.115), as other control variables are standardized to have a mean of zero.

(Beaver 1968). To test volume reactions around earnings announcements, I estimate the following model, which is analogous to that of return reactions.

$$CAV[\cdot] = \eta_0 + \eta_1|SUE| + \eta_2 Ab\_Wiki[\cdot] + \eta_3(Ab\_Wiki[\cdot] * |SUE|)$$
$$+\eta_x Controls + \eta_y(Controls * |SUE|) + \epsilon \quad (3)$$

Model 3 differs from Models 2a-2c in two respects. First, the dependent variable is *CAV*, cumulative average abnormal volume where abnormal volume is measured as turnover ratio minus its average over the past four weeks. Second, I take the absolute value of *SUE*. As before, I use two windows to measure *CAV[.]* and *Ab_Wiki[.]*: [-5,-1] represents five trading days prior to earnings announcements and [0,1] represents two trading days after the announcements.

I report the results in Panel C. The coefficient on *|SUE|*Ab_Wiki[-5,-1]* in Column 1 is positive, significant at 10% level, different from the return analyses reported in Panel B. Given that volume tests are more sensitive than return tests (Beaver 1968), this evidence provides only weak support for the notion that pre-announcement wiki-visits accelerate the price discovery of upcoming earnings news. Consistent with the return analyses, Columns 2-3 show that only wiki-visits after earnings announcements, but not before, are associated with a stronger market reaction towards earnings news immediately after its release. Moreover, Columns 4-5 show that this strengthened market reaction is concentrated in firms with more opaque information environment. Collectively, both the return and volume analyses indicate that the market is more responsive to earnings news when the abnormal wiki-visits immediately after the announcements are more pronounced, consistent with contextual information enabling investors to better interpret earnings news.

I test H3b by examining the association between the abnormal wiki-visits over the two-trading-day window starting with the earnings announcement date (*[0,1]*) and intra-period timeliness (IPT) that captures the speed with which information is incorporated into stock prices over a six-trading-day window (*[0,5]*).[39] This test complements the analyses of ERC and future ERC in two respects. First, IPT quantifies the speed of price discovery while holding the magnitude of news (the cumulative abnormal returns over the entire event window--*CAR[0,5]*) constant, while ERC relies on analyst forecasts to measure the "newsworthiness" of earnings

[39] I also compute the IPT over the window [0, 10] and find consistent results.

announcements. Second, as I construct abnormal wiki-visits in a two-day window immediately after earnings announcements and construct IPT over a longer window, this test is less subject to reverse causality concerns.

The standard IPT is the area under the IPT curve, which is a plot of the percentage of *CAR[0,5]* earned as of the end of each event day (Butler et al. 2007). Figure 4 plots the average IPT curves for earnings announcements with above- or below- median abnormal wiki-visits over the two-day window after the announcements (*Ab_Wiki[0,1]*). It shows that price discovery is faster when more investors visit wiki-pages over the window [0, 1]. Specifically, 82.5% of the *CAR[0,5]* is earned over the first two days after earnings announcements with above-median *Ab_Wiki[0,1]* versus 71.2% when *Ab_Wiki[0,1]* is below the median (a difference of 11.3%).

The standard IPT adjusts for under-reactions (intermediate CAR has the opposite sign as the final CAR) but does not adjust for over-reactions (intermediate CAR is larger than the final CAR in magnitude). In other words, the standard IPT is larger when the price overshoots and reverses than when the price quickly reaches the final return and stays at that level. I follow Blankespoor, deHaan, and Zhu (2018) to adjust for over-reactions (see Appendix B for the variable construction of the adjusted IPT—*IPT[0,5]*). I then regress this adjusted IPT on the abnormal wiki-visits at earnings announcements (*Ab_Wiki[0,1]*) using the following specification:

$$IPT[0,5] = \theta_0 + \theta_1 Ab\_Wiki[0,1] + \theta_2 SUE + \theta_3 |SUE| + \theta_x Controls + \epsilon \qquad (4)$$

H3b predicts θ_1 to be positive. Consistently, as shown in Columns 1 of Table 8, θ_1 is 1.125, significant at 5%, suggesting that more contextual information acquisition is associated with faster price discovery at earnings announcements. Economically, one standard deviation increase in *Ab_Wiki[0,1]* is associated with a 10-percentage-point increase in the sample median (i.e., increase the median firm to the 60th percentile).[40] Moreover, Columns 2-3 show that this significant association is driven by firms with more opaque information environment. To mitigate the noise in the IPT measure due to a small denominator (*CAR[0,5]*), I exclude observations with the absolute value of *CAR[0,5]* below 1% and find consistent results, as reported in Columns 4-6. To alleviate the influence of outliers, in untabulated results, I run median regressions and find consistent results. Taken together, these results provide further support for the hypothesis that

[40] I gauge the economic magnitude using percentiles because the average IPT is negative after adjusting for overreactions.

investors acquire contextual information to improve their interpretation of new financial information, thereby expediting the speed of price discovery.

VII. Conclusion

This paper studies the causes and consequences of contextual information acquisition using the web traffic to wiki-pages. Focusing on S&P 500 companies, I document that wiki-visits positively correlate with subsequent retail trading and returns, in the main sample period from July 2015 to December 2017 as well as a one-year sample around the launch of Google Knowledge Graph in May 2012, which imposes exogenous changes to wiki-visits. This evidence validates that wiki-visits reflect information acquisition by investors, especially retail investors.

The first part of the main analysis examines the determinants of contextual information acquisition. Consistent with investors acquiring contextual information to assist their interpretation of new events, I find that wiki-visits rise significantly upon major informational events such as earnings announcements. Additional analysis reveals that the abnormal wiki-visits at earnings announcements are amplified when earnings press releases are opportunistic, existing financial reports are difficult to read, or the announcing firms have higher intangible intensities. These results strengthen the inference that investors acquire contextual information to better interpret financial information.

The second part of the main analysis focuses on the consequences of contextual information acquisition. I find the abnormal wiki-visits at earnings announcements are associated with a larger ERC while the pre-announcement wiki-visits do not affect the price discovery of upcoming earnings news. These results suggest that contextual information improves the interpretation of earnings news after its announcement rather than the prediction of earnings news before the announcement. Lastly, using intra-period timeliness to measure the speed with which stock prices incorporate earnings news, I find price discovery is faster when the abnormal wiki-visits at earnings announcements are larger. Overall, these results are consistent with contextual information improving investors' interpretation of new financial information, thereby increasing the speed of price discovery.

I study Wikipedia as one particular source of contextual information because its information boundary is well-defined, it is widely used by investors, and the data on its web traffic

and historical information content are publicly available. My evidence should be of interest to regulators and managers. The web traffic to wiki-pages reveals investors' preferences for highly readable contextual information, suggesting directions that we could take to improve the current financial reporting system. Managers should note that thousands of investors turn to Wikipedia as well as other information sources when corporate disclosures are vague, arcane, or suspicious. At the very least, they can learn from Wikipedia's simple and concise style to facilitate effective communication with investors.

References

Amir, E., & Lev, B. (1996). Value-relevance of nonfinancial information: The wireless communications industry. *Journal of Accounting and Economics*, 22(1-3), 3-30.

Antle, R., Demski, J. S., & Ryan, S. G. (1994). Multiple sources of information, valuation, and accounting earnings. *Journal of Accounting, Auditing & Finance*, 9(4), 675-696.

Asay, H. S., Elliott, W. B., & Rennekamp, K. (2016). Disclosure readability and the sensitivity of investors' valuation judgments to outside information. *The Accounting Review*, 92(4), 1-25.

Baginski, S., Demers, E., Wang, C., & Yu, J. (2016). Contemporaneous verification of language: evidence from management earnings forecasts. *Review of Accounting Studies*, 21(1), 165-197.

Barber, B. M., & Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *The Review of Financial Studies*, 21(2), 785-818.

Bartov, E., Faurel, L., & Mohanram, P. S. (2017). Can Twitter help predict firm-level earnings and stock returns?. *The Accounting Review*, 93(3), 25-57.

Beaver, W. H. (1968). The information content of annual earnings announcements. *Journal of Accounting Research*, 67-92.

Ben-Rephael, A., Da, Z., & Israelsen, R. D. (2017). It depends on where you search: Institutional investor attention and underreaction to news. *The Review of Financial Studies*, 30(9), 3009-3047.

Blankespoor, E., Miller, G. S., & White, H. D. (2013). The role of dissemination in market liquidity: Evidence from firms' use of Twitter™. *The Accounting Review*, 89(1), 79-112.

Blankespoor, E., deHaan, E., & Zhu, C. (2018). Capital market effects of media synthesis and dissemination: Evidence from robo-journalism. *Review of Accounting Studies*, 23(1), 1-36.

Blankespoor, E., deHaan, E., Wertz, J., & Zhu, C. (2018). Why do individual investors disregard accounting information? The roles of information awareness and acquisition costs. Available at: https://ssrn.com/abstract=3059073

Boehmer, E., Jones, C., & Zhang, X. (2017). Tracking retail investor activity. Available at: https://ssrn.com/abstract=2822105

Bradshaw, T. 2008. Companies woo investors via social websites. Financial Times. Available at: http://www.ft.com/cms/s/0/c9d0271a-bf49-11dd-ae63-0000779fd18c.html

Bushee, B., Cedergren, M., & Michels, J. (2018). Does the Media Help or Hurt Retail Investors during the IPO Quiet Period?. Available at https://ssrn.com/abstract=3160949

Bushee, B. J., Core, J. E., Guay, W., & Hamm, S. J. (2010). The role of the business press as an information intermediary. *Journal of Accounting Research*, 48(1), 1-19.

Butler, M., Kraft, A., & Weiss, I. S. (2007). The effect of reporting frequency on the timeliness of earnings: The cases of voluntary and mandatory interim reports. *Journal of Accounting and Economics*, 43(2-3), 181-217.

Cade, N. L. (2018). Corporate social media: How two-way disclosure channels influence investors. *Accounting, Organizations and Society*.

Campbell, J., DeAngelis, M., & Moon, J. (2018). Skin in the game: Personal stock holdings and investors' response to stock analysis on social media. Working paper.

Chen, H., De, P., Hu, Y. J., & Hwang, B. H. (2014). Wisdom of crowds: The value of stock opinions transmitted through social media. *The Review of Financial Studies*, 27(5), 1367-1403.

Chi, S., & Shanthikumar, D. M. (2018). Do Retail Investors Use SEC Filings? Evidence from EDGAR Search. Evidence from EDGAR Search. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3281234

Comprend. (2015). Comprend's Capital Market Report 2015. Available at: https://comprend.com/spotlight/2015/capital-market/capital-market-report-2015.pdf

Crane, A. D., Crotty, K., & Umar, T. (2018). Do Hedge Funds Profit from Public Information?. Available at: https://ssrn.com/abstract=3127825

Da, Z., Engelberg, J., & Gao, P. (2011). In search of attention. *The Journal of Finance*, 66(5), 1461-1499.

Drake, M. S., Roulstone, D. T., & Thornock, J. R. (2012). Investor information demand: Evidence from Google searches around earnings announcements. *Journal of Accounting Research*, 50(4), 1001-1040.

Drake, M. S., Roulstone, D. T., & Thornock, J. R. (2015). The determinants and consequences of information acquisition via EDGAR. *Contemporary Accounting Research*, 32(3), 1128-1161.

Drake, M. S., Roulstone, D. T., & Thornock, J. R. (2016). The usefulness of historical accounting reports. *Journal of Accounting and Economics*, 61(2-3), 448-464.

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), 221-245.

Giles, J. (2005.) Internet encyclopedias go head to head. *Nature*, 438, 900-901.

Glaeser, S. (2018). The effects of proprietary information on corporate disclosure and transparency: Evidence from trade secrets. *Journal of Accounting and Economics*.

Greenstein, S., & Zhu, F. (2012). Is Wikipedia Biased?. *American Economic Review*, 102(3), 343-48.

Halfaker, Aaron (2016): Monthly Wikipedia article quality predictions. Available at: figshare. doi.org/10.6084/m9.figshare.3859800

Han, B., Tang, Y., & Yang, L. (2016). Public information and uninformed trading: Implications for market liquidity and price efficiency. *Journal of Economic Theory*, 163, 604-643.

Holden, C. W., & Jacobsen, S. (2014). Liquidity Measurement Problems in Fast, Competitive Markets: Expensive and Cheap Solutions. *The Journal of Finance* 69(4), 1747-1785.

Huang, X., Teoh, S. H., & Zhang, Y. (2013). Tone management. *The Accounting Review*, 89(3), 1083-1113.

Hutton, A. P., Marcus, A. J., & Tehranian, H. (2009). Opaque financial reports, R2, and crash risk. *Journal of Financial Economics*, 94(1), 67-86.

Hutton, A. P., Miller, G. S., & Skinner, D. J. (2003). The role of supplementary statements with management earnings forecasts. *Journal of Accounting Research*, 41(5), 867-890.

Jame, R., Johnston, R., Markov, S., & Wolfe, M. C. (2016). The value of crowdsourced earnings forecasts. *Journal of Accounting Research*, 54(4), 1077-1110.

Kelley, E. K., & Tetlock, P. C. (2013). How wise are crowds? Insights from retail orders and stock returns. *The Journal of Finance*, 68(3), 1229-1265.

Kim, O., & Verrecchia, R. E. (1991). Market reaction to anticipated announcements. *Journal of Financial Economics*, 30(2), 273-309.

Kim, O., & Verrecchia, R. E. (1994). Market liquidity and volume around earnings announcements. *Journal of Accounting and Economics*, 17(1-2), 41-67.

Kothari, S. P. (2001). Capital markets research in accounting. *Journal of Accounting and Economics*, 31(1-3), 105-231.

Lang, M., & Stice-Lawrence, L. (2015). Textual analysis and international financial reporting: Large sample evidence. *Journal of Accounting and Economics*, 60(2-3), 110-135.

Lawrence, A. (2013). Individual investors and financial disclosure. *Journal of Accounting and Economics*, 56(1), 130-147.

Lawrence, A., Ryans, J., and Sun, E. (2017). Investor Demand for Sell-Side Research. *The Accounting Review*, Vol. 92(2).

Lawrence, A., Ryans, J., and Sun, E., and Soni, A. (2017). Investor Reliance on the Crowd. Available at SSRN: https://ssrn.com/abstract=2945411.

Lee, C. M., and Mark J. R. (1991). "Inferring Trade Direction from Intraday Data." *The Journal of Finance* 46(2): 733–46.

Lee, C. M. (2001). Market efficiency and accounting research: a discussion of 'capital market research in accounting' by SP Kothari. *Journal of Accounting and Economics*, 31(1-3), 233-253.

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.

Lerman, A. (2017). Individual investors' attention to accounting information: Message board discussions. Available at SSRN: https://ssrn.com/abstract=1540689

Lev, B., Ryan, S. G., & Wu, M. (2008). Rewriting earnings history. *Review of Accounting Studies*, 13(4), 419-451.

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

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, 66(1), 35-65.

Loughran, T., & McDonald, B. (2017). The use of EDGAR filings by investors. *Journal of Behavioral Finance*, 18(2), 231-248.

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

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.

Moat, H. S., Curme, C., Avakian, A., Kenett, D. Y., Stanley, H. E., & Preis, T. (2013). Quantifying Wikipedia usage patterns before stock market moves. *Nature Scientific Reports*, 3, 1801.

Peters, R. H., & Taylor, L. A. (2017). Intangible capital and the investment-q relation. *Journal of Financial Economics*, *123*(2), 251-272.

Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *The Review of Financial Studies*, 22(1), 435-480.

Rubin, A., & Rubin, E. (2010). Informed investors and the internet. *Journal of Business Finance & Accounting*, 37(7-8), 841-865.

SEC. (2011). How to Read a 10-K. Available at: https://www.sec.gov/fast-answers/answersreada10khtm.html

SEC. (2012). Investment Adviser Use of Social Media. Available at: https://www.sec.gov/about/offices/ocie/riskalert-socialmedia.pdf

Twedt, B. (2016). Spreading the word: Price discovery and newswire dissemination of management earnings guidance. *The Accounting Review*, 91(1), 317-346.

Wulczyn, E., & Taraborelli, D. (2015). Wikipedia Clickstream. Available at: figshare.doi:10.6084/m9.figshare.1305770

Xu, S. X., & Zahng, X. M. (2013). Impact of Wikipedia on Market Information Environment: Evidence on Management Disclosure and Investor Reaction. *MIS Quarterly*, 37(4), 1043-1068.

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

Appendix A. Microchip Technology (MT)

Panel A is a screenshot of MT's wiki-page (https://en.wikipedia.org/wiki/ Microchip_Technology) and Panel B is a screenshot of its daily wiki-visits. Panel C presents the Google Knowledge Graph information box when googling "Microchip Technology". Panel D lists the clickstreams leading to MT's wiki-page in January 2017. The clickstream data are shared and updated by Wulczyn and Taraborelli (2015). "External-search" means search engines such as Google and Bing, "unidentified-external" or "unclassified-external" means unidentified or unclassified websites outside Wikipedia, "unidentified-internal" includes all wiki-pages that are clicked before MT's, for less than 10 times, "Main Page" means searching from the home page of Wikipedia, and the others are the titles of individual wiki-pages.

Panel A. MT's wiki-page

Microchip Technology Inc. is an American publicly-listed corporation that is a manufacturer of microcontroller, mixed-signal, analog and Flash-IP integrated circuits. Its products include microcontrollers (PIC, dsPIC, AVR and SAM), Serial EEPROM devices, Serial SRAM devices, embedded security devices, radio frequency (RF) devices, thermal, power and battery management analog devices, as well as linear, interface and wireless solutions. Examples of these solutions include USB, zigbee, MiWi, LoRa, SIGFOX and Ethernet.

Corporate headquarters are located in Chandler, Arizona, with wafer fabs in Tempe, Arizona, Gresham, Oregon, and Colorado Springs, Colorado, assembly/test facilities in Chachoengsao, Thailand and Calamba, Philippines. Sales for the fiscal year ending on March 31, 2018 were $3.981 billion.[2]

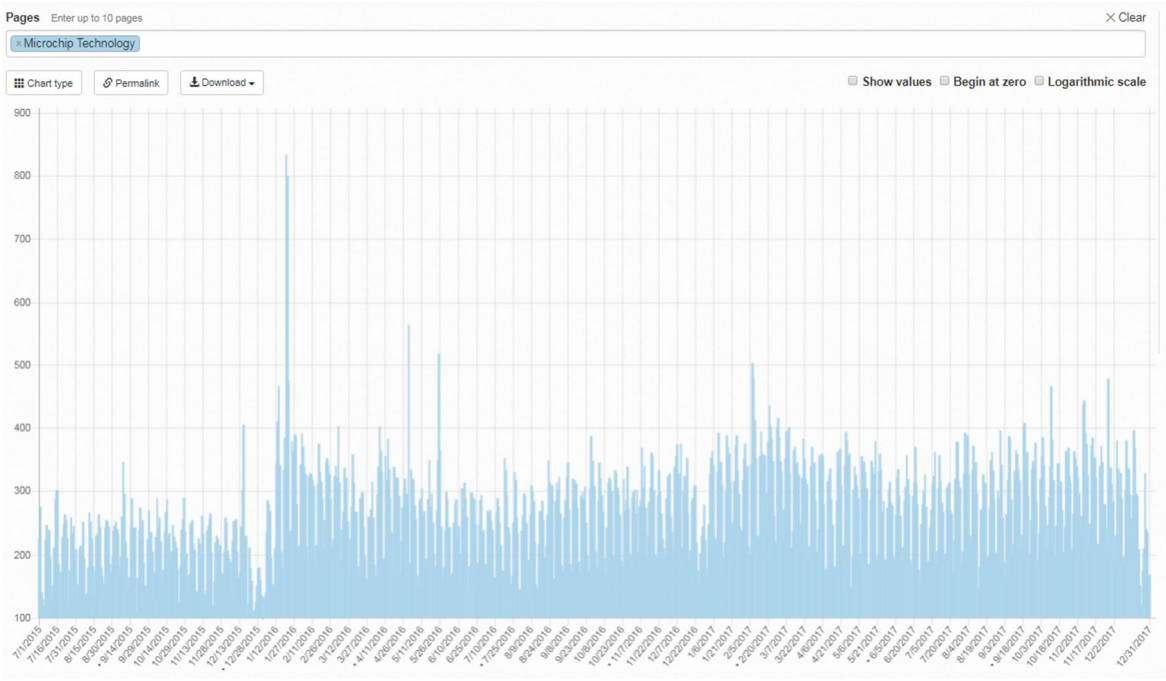
Contents [hide]

1 History
2 Products
  2.1 Microcontrollers
  2.2 Integrated circuits
3 Leadership[19]
  3.1 Steve Sanghi
  3.2 Ganesh Moorthy
  3.3 J. Eric Bjornholt
  3.4 Stephen V. Drehobl
  3.5 Mitchell R. Little
  3.6 Richard J. Simoncic
  3.7 Appointed Officers
4 Acquisitions
  4.1 HI-TECH Software

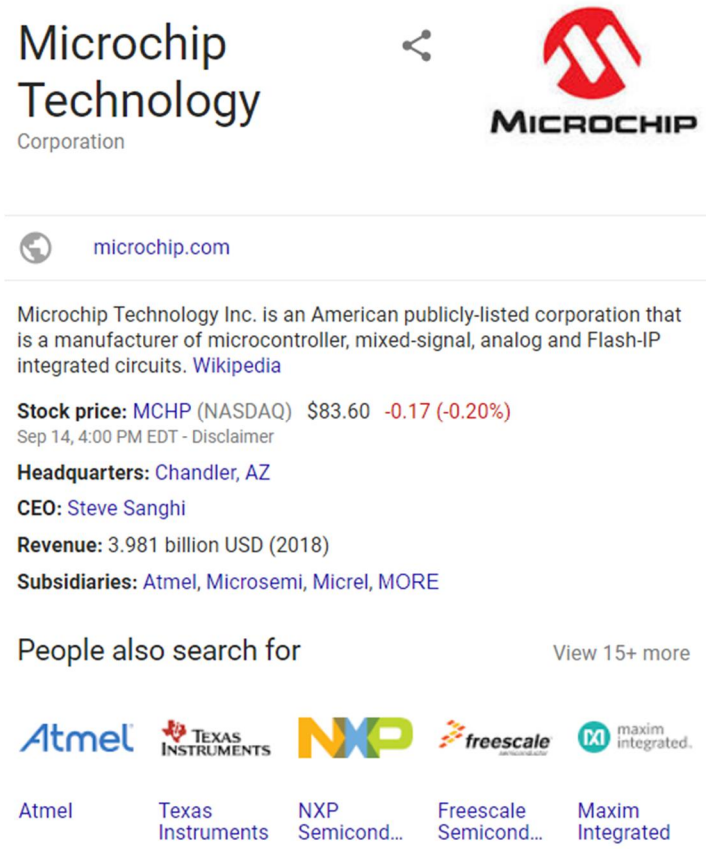
Microchip Technology Inc.

Type	Public corporation
Traded as	NASDAQ: MCHP NASDAQ-100 Component S&P 500 Component
Industry	Semiconductors
Founded	1989; 29 years ago
Headquarters	2355 W Chandler Blvd Chandler, AZ 85224, USA
Key people	Steve Sanghi, Chairman & CEO J. Eric Bjornholt, CFO Ganesh Moorthy, President & COO
Products	Microcontrollers Serial EEPROMs Serial SRAM Analog ICs
Revenue	▲$3.981 billion (2018)[1]
Number of employees	18,000+ (2018)
Website	www.microchip.com

Panel B. MT's wiki-visits



Panel C. Knowledge graph in the google search results



Panel D. Clickstreams leading to MT's wiki-page

Previous web page	# clickstreams	% clickstreams
External-search	5290	58.3
Unidentified-external	1587	17.5
Atmel	398	4.4
Unidentified-internal	288	3.2
PIC microcontroller	183	2.0
List of semiconductor fabrication plants	154	1.7
Microchip (disambiguation)	125	1.4
Apple Inc.	121	1.3
Qualcomm	120	1.3
Unclassified-external	100	1.1
Microsoft	81	0.9
Microcontroller	68	0.7
List of common microcontrollers	50	0.6
Cypress Semiconductor	44	0.5
NASDAQ-100	42	0.5
SMSC	38	0.4
ELM327	37	0.4
MPLAB	27	0.3
Steve Sanghi	26	0.3
Freescale Semiconductor	26	0.3
List of integrated circuit manufacturers	24	0.3
In-system programming	23	0.3
EEPROM	23	0.3
MOST Bus	22	0.2
Integrated circuit	19	0.2
List of corporations in Phoenix	19	0.2
List of Arizona companies	19	0.2
Main Page	16	0.2
List of CPU architectures	16	0.2
MiWi	15	0.2
List of companies of the United States	14	0.2
Harvard architecture	13	0.1
General Instrument	12	0.1
Chandler, Arizona	11	0.1
Dialog Semiconductor	11	0.1
SST	11	0.1

Appendix B. Variable definitions

Variable	Definition	Source
Information Acquisition		
Raw_Wiki	The daily number of human visits to a firm's Wikipedia page. A prefix of "Ln" means to take the natural logarithm of one plus the variable value, henceforth for all other variables.	Wikipedia
Ab_Wiki	Abnormal wiki-visits, *Raw_Wiki* minus the median wiki-visits for the same day of the week during the past 10 weeks.	Wikipedia
Ab_Wiki%	Abnormal percentage wiki-visits, *Raw_Wiki* divided by the median wiki-visits for the same day of the week during the past 10 weeks, minus one and multiplied by 100.	Wikipedia
GSV	Daily GSV of the firm's ticker symbol. The data are downloaded every 90 days and adjusted by a consistent time series of weekly GSVs, following Chi and Shanthikumar (2018).	Google Trends
Ab_GSV	Abnormal GSV, the log of *GSV* (defined above) and the median *GSV* for the same day of the week during the past 10 weeks.	Google Trends
EDGAR	The daily number of unique IP addresses that request a firm's SEC filings at EDGAR, excluding automated requests following Drake et al. (2016).	SEC
Ab_EDGAR	Abnormal EDGAR requests, *EDGAR* (define above) minus the median *EDGAR* for the same day of the week over the past 10 weeks.	SEC
Firm Characteristics		
Size	The natural logarithm of market capitalization (prccq*cshoq).	CRSP
Age	The natural logarithm of one plus number of years since entering CRSP.	CRSP
ROA	Return on assets as of the latest quarter (ibq/atq).	Compustat
Loss	One if earnings (ibq) is negative.	Compustat
Leverage	Book leverage as of the latest quarter, total debt (dlttq+dlcq) scaled by total asset (atq).	Compustat
BTM	Book-to-market ratio as of the latest quarter, book value of equity (ceqq) scaled by market capitalization	Compustat & CRSP
IO	Total institutional ownership (13F holdings) scaled by shares outstanding, as of the most reporting date.	Thomson Reuters
Analyst	Number of analysts following the company in the most recent quarter.	IBES
Short_Interest	Short sales of the firm's stock as of the latest short interest publication date (the 15th and the last business date of each month) scaled by shares outstanding.	Compustat
News_Dummy	Dummy variable, one if the firm is covered in at least one news article.	Capital IQ
Intangibles	Total intangible assets (both on- and off- balance sheet) scaled by total assets. Data on intangible assets are obtained from Peters and Taylor (2017). Prefixes of *OnBS* and *OffBS* indicates on- and off-balance sheet intangibles, respectively.	WRDS
Bus_Segments	The number of business segments.	Compustat
Information_Events		
EA	One when the firm announces its earnings, which is determined by the earlier date of *rdq* in Compustat and *anndats* in IBES.	Compustat & IBES
MF	One when the firm issues a management guidance for EPS or revenue of any horizon.	IBES
10K&10Q	One when the firm files From 10-K or 10-Q with the SEC.	SEC

M&A	One when the firm announces a merger or acquisition deal.	SDC
Crash or Jump	*Crash* (*Jump*) equals one when the natural logarithm of one plus the residual from the firm-specific regression of daily stock returns is 3.09 standard deviations below (above) its mean. The independent variables in the firm-specific regression include value-weighted market returns and Fama-French 48 industry returns for the current day, one trading day before and after, following Hutton et al. (2009). *Crash* and *Jump* confounded with *EA*, *MF*, *10-K&10-Q*, or *M&A* are excluded.	CRSP
Market Variables		
NetTrade-(Retail/Institution)	The number of buy transactions minus sell transactions. The suffix indicates whether the transaction is initiated by retail or institutional investors. Retail trades are identified based on TAQ exchange code (D) and a small price improvement (0-0.4 cents, exclusive, above (below) a round cent for sale (buy) transactions), following Boehmer et al. (2017). Institutional trades are non-retail trades with trade size above $50,000, following Bushee et al. (2018).	TAQ
NetVolume-(Retail/Institution)	The shares of buy transactions minus sale transactions initiated by retail or institutional investors, indicated by the suffix (unit: 1000).	TAQ
AR	Daily abnormal return, computed as individual stock returns (ret) minus the corresponding 5 x 5 value-weighted portfolio returns matched by size and book-to-market ratio. $\|\cdot\|$ means taking the absolute value, henceforth.	CRSP & Kenneth French's website
CAR_Month	Cumulative abnormal return (*AR*, defined above) of the past 4 weeks.	Same as above
AVol	Abnormal trading volume, the current turnover (vol/shrout) minus the median turnover for the same day of the week over the past 10 weeks.	CRSP
Earnings Announcement Characteristics		
CAR[X]	Cumulative abnormal return (*AR*, defined above) of the period X, where X takes the value of -1W, 2W, 3W-4W, 5W-12W meaning the next week 1, week 2, week 3 to week 4, week 5 to week 12, respectively. CAR[-5,-1], CAR[0,1], CAR[0,5] are abnormal returns accumulated over the corresponding windows around earnings announcements.	CRSP & Kenneth French's website
CAV[X]	Average abnormal volume (*AVol*, defined above) over the window X.	CRSP
IPT[0,5]	Intra-period timeliness, adjusted for overreaction and reversals following Blankespoor et al. (2018). The formula is specified below where *CAR[0,t]* is the cumulative abnormal return over trading days [0,t] with day 0 as the earnings announcements date. $$IPT = \sum_{t=0}^{5} 1 - \frac{\|CAR[0,5] - CAR[0,t]\|}{\|CAR[0,5]\|}$$	CRSP & Kenneth French's website
IPT[0,5]Truncated	*IPT* defined above, and set it to missing when *CAR[0,5]* <+/-1%.	CRSP
Ab_Wiki[X]	Average *Ab_Wiki* over the window X, scaled by 1000.	Wikipedia
Ab_Wiki%[X]	Average *Ab_Wiki%* over the window X.	Wikipedia
Ab_GSV[X]	Average *Ab_GSV* over the window X.	Google Trends
News[X]	The natural logarithm of one plus the daily average number of news articles over the window X.	Capital IQ
Fqtr4	One if the earnings announcement is the fourth quarter of the fiscal year.	Compustat

Earn_Rho	Earnings persistence, measured as the coefficient on lagged earnings (the same quarter last year) in firm-specific regressions of quarterly earnings (ibq/atq) estimated over the past five years.	Compustat
Earn_Std	Earnings volatility, measured as the standard deviation of quarterly earnings (ibq/atq) over the past five years.	Compustat
Ret_Std	Annualized return volatility, standard deviation of stock returns during the fiscal quarter, multiplied by the square root of 252.	
SUE	Actual earnings minus analyst forecast consensus scaled by stock prices where the consensus is computed as the median forecasts issued 90 days before earnings announcements excluding stale forecasts.	Compustat & IBES
Dispersion	Analyst forecast dispersion, the standard deviation of forecasts issued 90 days before earnings announcements, excluding stale forecasts.	Compustat & IBES
Textual Analysis		
Words	The number of words in the corresponding text under examination.	Wikipedia &EDGAR
Sentences	The number of sentences in the corresponding text under examination.	Wikipedia &EDGAR
Sections	The number of sections in the historical wiki-pages.	Wikipedia
References	The number of references to information sources outside Wikipedia in the historical wiki-pages.	Wikipedia
Gunning-Fog	A readability index which measures the years of formal education a person needs to understand the text. A higher value means lower readability. Flesch-Kincaid Grade Level, the SMOG Index, LIX, and RIX also measure readability, although the formulas are different.	Wikipedia &EDGAR
Unreadable	The average of the standardized value of five readability indexes (Gunning-Fog, Flesch-Kincaid Grade Level, the SMOG Index, LIX, and RIX). A higher value means the text is more difficult to read.	Wikipedia &EDGAR
Boilerplate	The proportion of words in sentences with at least one tetragram (any consecutive four words in the sentence) that appear in more than 60% of disclosures in all sample documents.	Wikipedia &EDGAR
Tone	The proportion of positive words minus that of negative words. The dictionary of positive and negative words are downloaded from Loughran and McDonald's website.	Wikipedia &EDGAR
Atone	Abnormal tone, the residual of annual regressions of *Tone* on firm characteristics, following Huang et al. (2013).	Compustat
Uncertainty	The proportion of uncertainty words whose dictionary is downloaded from Loughran and McDonald's website.	Wikipedia &EDGAR
Editors[4W]	The number of unique users who made at least one edit to the Wikipedia page during the past 4 weeks.	Wikipedia
Revisions[4W]	The number of revisions in the Wikipedia page made in the past 4 weeks, weighted by revision size in bytes.	Wikipedia

Appendix C. The relationship between wiki-visits and EDGAR-visits

This section examines the relationship between wiki-visits and EDGAR-visits. Wiki-visits are the daily visits to a firm's wiki-page and the EDGAR-visits are daily requests for the firm's filings at the SEC EDGAR system. Both variables only include human visits. The sample starts on July 1st, 2015, the first day when the API of wiki-visits becomes available and ends on June 30, 2017, the last day when EDGAR-visits are available. To capture the visits that are more likely initiated by investors, I focus on abnormal wiki-visits (*Ab_Wiki*) and abnormal EDGAR-visits (*Ab_EDGAR*).[1] Conceptually, both variables reflect the effort to acquire firm-specific information, and they are highly correlated as shown in Panel C of Table 2. To shed more light on the similarity and difference between the two, I conduct the following analyses.

I first examine the lead-lag relation between abnormal wiki-visits and EDGAR-visits using vector autoregressions (VAR) models. The vector comprises *Ab_Wiki*, *Ab_EDGAR*, abnormal GSV (*Ab_GSV*), abnormal turnover (*AVol*), the absolute value of abnormal return (|*AR*|), news coverage (*News_Dummy*), and a constant (unreported for brevity), following Da et al. (2011). The VAR analysis is restricted to sample firms with at least 100 observations on trading days (539 unique firms, 242,279 firm-trading days). I run the VAR model by firm and then take the average coefficients across firms. The standard errors are computed using the block-bootstrap method to deal with the time-series and cross-sectional correlation.[2] Panel A of Table C1 presents the average coefficients. I find that, on average, *Ab_Wiki* and *Ab_EDGAR* significantly predict each other. Next, I tabulate the percentage of firms with significant coefficients (5% level) on *Ab_Wiki* (*Ab_EDGAR*) when predicting *Ab_EDGAR* (*Ab_Wiki*) in Panel B. For comparison, I also tabulate the theoretical distribution given the 5% significance level. The largest deviation from the benchmark is observed when lagged *Ab_Wiki* positively predicts future *Ab_EDGAR* while lagged *Ab_EDGAR* has insignificant predictive power for future *Ab_Wiki*. Overall, in 20.59% (12.43%) of firms, wiki-visits (EDGAR-visits) significantly predict subsequent EDGAR-visits (wiki-visits), suggesting that wiki-visits lead EDGAR-visits more often.

Second, I run a horse race between *Ab_Wiki* and *Ab_EDGAR* to predict subsequent trades, using the same regression specification used in Panel A of Table 3. The results are reported in Table C2. Similarly to the results reported in Table 3, *Ab_Wiki* significantly correlates with net purchases by retail investors. The association between *Ab_EDGAR* and net purchases by retail investors is also significant when the trades are unweighted (*NetTrade*), but becomes insignificant when the trades are weighted by trading volume (*NetVolume*). In terms of net purchases by institutional investors, neither *Ab_Wiki* nor *Ab_EDGAR* has a significant explanatory power. This analysis suggests wiki-visits are slightly better than EDGAR-visits at explaining subsequent retail trades.

[1] The results using raw wiki-visits and EDGAR-visits yield similar inferences.

[2] To conserve the autocorrelation structure in the time-series, for a given firm in the sample, I block-bootstrap with replacement using a block-length of seven trading days (the block length is determined by the one-third root of average number of observations in each VAR model). Firms are bootstrapped with replacement. The procedure is repeated 1000 times. Each time the same VAR model is run in the bootstrapped panel. After 1000 repetitions, an empirical distribution of VAR coefficients is constructed. The VAR coefficients estimated from the actual sample are then compared against this empirical distribution to see whether they are significant at the 10%, 5%, and 1% level, respectively.

Table C1. Vector autoregression (VAR) results

The table examines the lead-lag relation between wiki-visits and EDGAR-visits using VARs. I run the VAR model for each firm and report the average coefficients across firms in Panel A. To determine the statistical significance, I use a block bootstrap procedure to adjust for the cross-sectional correlation and autocorrelations in the error terms and report the standard errors underneath the coefficients. Specifically, for each firm in the sample, I block-bootstrap with replacement using a block-length of 7 trading days (the block length is determined by the one-third root of the average number of observations in each VAR model) to conserve the autocorrelation structure. The procedure is repeated 1000 times, hence 1000 bootstrapped panels are built. In each panel, I estimate the same VAR model to build up an empirical distribution of VAR coefficients. Panel B tabulates the percentage of firms with significantly positive (5% significance level), significantly negative, or insignificant coefficients on lagged *Ab_Wiki* (*Ab_EDGAR*) in predicting *Ab_EDGAR* (*Ab_Wiki*). ***, **, and * stand for statistical significance at the 1%, 5%, and 10% level, respectively.

Panel A. Average VAR Coefficients

	Lagged by one trading day						
	Ab_ Wiki	Ab_ EDGAR	Ab_ GSV	AVol	|AR|	News_ Dummy	R-Squared
Ab_Wiki	0.501***	11.61***	0.252***	1192***	-136.7**	5.461	0.327
	(0.007)	(1.722)	(0.108)	(269.360)	(88.905)	(1.320)	
Ab_EDGAR	0.001***	0.414***	0.002***	4.57***	-0.494***	0.064***	0.250
	(0.000)	(0.005)	(0.000)	(0.536)	(0.208)	(0.005)	
Ab_GSV	0.003***	0.268***	0.655***	43.76***	0.420	0.126**	0.497
	(0.001)	(0.055)	(0.008)	(9.802)	(3.537)	(0.072)	
AVol	0.000***	0.000	0.000***	0.463***	-0.012***	0.001***	0.056
	(0.000)	(0.000)	(0.000)	(0.005)	(0.002)	(0.000)	
|AR|	0.000***	0.000***	0.000***	0.214***	0.06***	0.001***	0.250
	(0.000)	(0.000)	(0.000)	(0.010)	(0.004)	(0.000)	
News_Dummy	0.000**	0.002	0.000***	5.215***	-0.663***	0.018***	0.022
	(0.000)	(0.003)	(0.000)	(0.490)	(0.176)	(0.003)	

Panel B. Distribution of the lead-lag relation between *Ab_Wiki* and *Ab_EDGAR*

	Coef. on lagged *Ab_EDGAR*							
Coef. On lagged *Ab_Wiki*	**Empirical Distribution**				**Benchmark Distribution**			
	Sig. Negative	Insig.	Sig. Positive	Total	Sig. Negative	Insig.	Sig. Positive	Total
Sig. Negative	0.00%	0.74%	0.00%	0.74%	0.25%	4.50%	0.25%	5.00%
Insig.	0.56%	71.61%	7.24%	79.41%	4.50%	81.00%	4.50%	90.00%
Sig. Positive	0.00%	15.21%	4.64%	19.85%	0.25%	4.50%	0.25%	5.00%
Total	0.56%	87.57%	11.87%	100.00%	5.00%	90.00%	5.00%	100.00%

Table C2. Wiki-visits, EDGAR-visits, and subsequent trading

This table reports the panel regressions of subsequent net buying of retail and institutional investors on wiki-visits and EDGAR-visits. Retail trades and institutional trades are identified from TAQ millisecond consolidated files following Boehmer et al. (2017) and Bushee et al. (2018). Control variables, whose coefficients are not reported for brevity, include *Size, Age, Leverage, BTM, ROA, Loss, IO, Ln(Analyst),* and *Short_Interest*. To facilitate comparison of the coefficients, I standardize all the independent variables. Standard errors reported in parentheses are clustered by firm. ***, **, and * stand for statistical significance at the 1%, 5%, and 10% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Retail Investors		**Institutional Investors**		**Retail Investors**		**Institutional Investors**	
VARIABLES	$NetTrade_{t+1}$	$NetTrade_{t+1}$	$NetTrade_{t+1}$	$NetTrade_{t+1}$	$NetVolume_{t+1}$	$NetVolume_{t+1}$	$NetVolume_{t+1}$	$NetVolume_{t+1}$
Ab_Wiki	5.406***	5.224***	0.096	0.115	0.819***	0.783**	-0.051	0.004
	(1.531)	(1.476)	(0.251)	(0.258)	(0.308)	(0.308)	(1.332)	(1.349)
Ab_EDGAR	3.533***	3.260***	-0.182	-0.171	-0.115	-0.136	-0.148	-0.169
	(0.956)	(0.915)	(0.191)	(0.186)	(0.217)	(0.216)	(1.009)	(1.023)
|AR|	-1.139*	-1.124*	0.433***	0.381***	0.096	0.047	1.313*	1.073
	(0.598)	(0.629)	(0.113)	(0.118)	(0.176)	(0.186)	(0.696)	(0.764)
|CAR_Month|	2.773***	2.548***	-0.491***	-0.525***	0.037	-0.086	-2.470***	-2.721***
	(0.916)	(0.899)	(0.123)	(0.133)	(0.205)	(0.201)	(0.624)	(0.713)
AVol	3.681***	3.602***	-0.463**	-0.423**	1.321***	1.318***	-1.561	-1.394
	(1.208)	(1.177)	(0.179)	(0.176)	(0.367)	(0.360)	(1.201)	(1.208)
News_Dummy	0.174	-0.422	0.175**	0.222***	0.147	0.077	0.632	0.705
	(0.419)	(0.429)	(0.082)	(0.084)	(0.131)	(0.137)	(0.543)	(0.545)
Ab_GSV	2.551***	2.575***	-0.071	-0.074	0.358**	0.374**	-1.440**	-1.430**
	(0.965)	(0.937)	(0.113)	(0.112)	(0.176)	(0.173)	(0.596)	(0.598)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	242,279	242,279	242,279	242,279	242,279	242,279	242,279	242,279
R-squared	0.063	0.069	0.008	0.010	0.016	0.018	0.018	0.020

Figure 1. Characteristics of wiki-pages

Panel A plots the monthly average wiki-page length (word count and sentence count). Panel B plots the monthly average number of wiki-pages' sections which are classified into three types based on their titles: intangibles (e.g., R&D, brand development or recognition, promotion campaigns), business or geography segments (e.g., operation segments, world presences, product lines), and others. Panel C plots the monthly average number of wiki-pages' external references which are classified into four types based on the domain of the referenced web addresses: firm references (firm's website, SEC filings, or press releases), news references (major news outlets), third-party references (books, education institutions, and government agencies), and others.

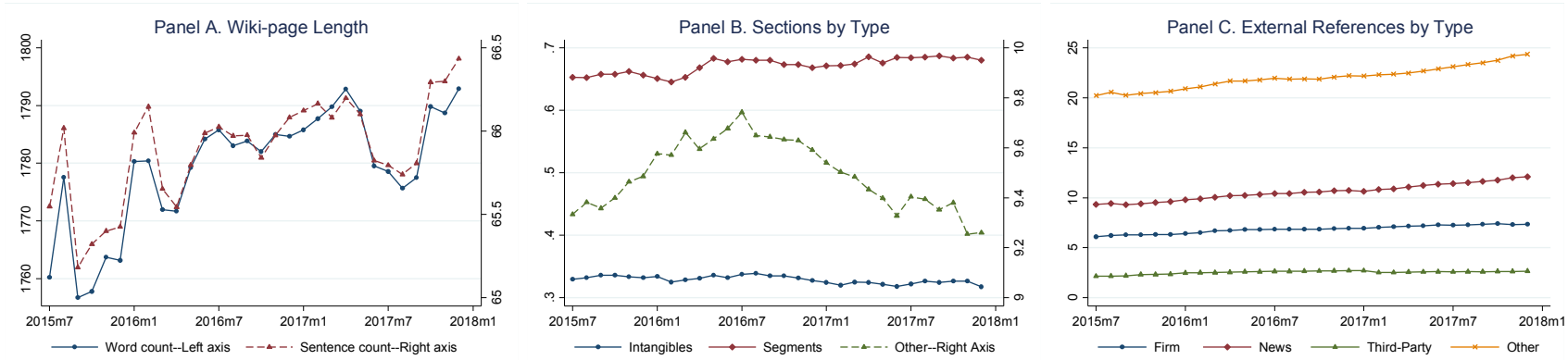


Figure 2. Wiki-visits around the launch of GKG

This figure plots the daily median wiki-visits of S&P 500 firms around the launch of GKG (May 16, 2012, indicated by the vertical dashed line). Sample firms are grouped into four subsamples based on the wiki-visits in the 10 weeks before 2012 March. The red solid line in each graph represents an OLS regression line fit estimated using the pre-GKG wiki-visits. The green dashed line depicts the predicted wiki-visits based on the estimates from the pre-GKG period.

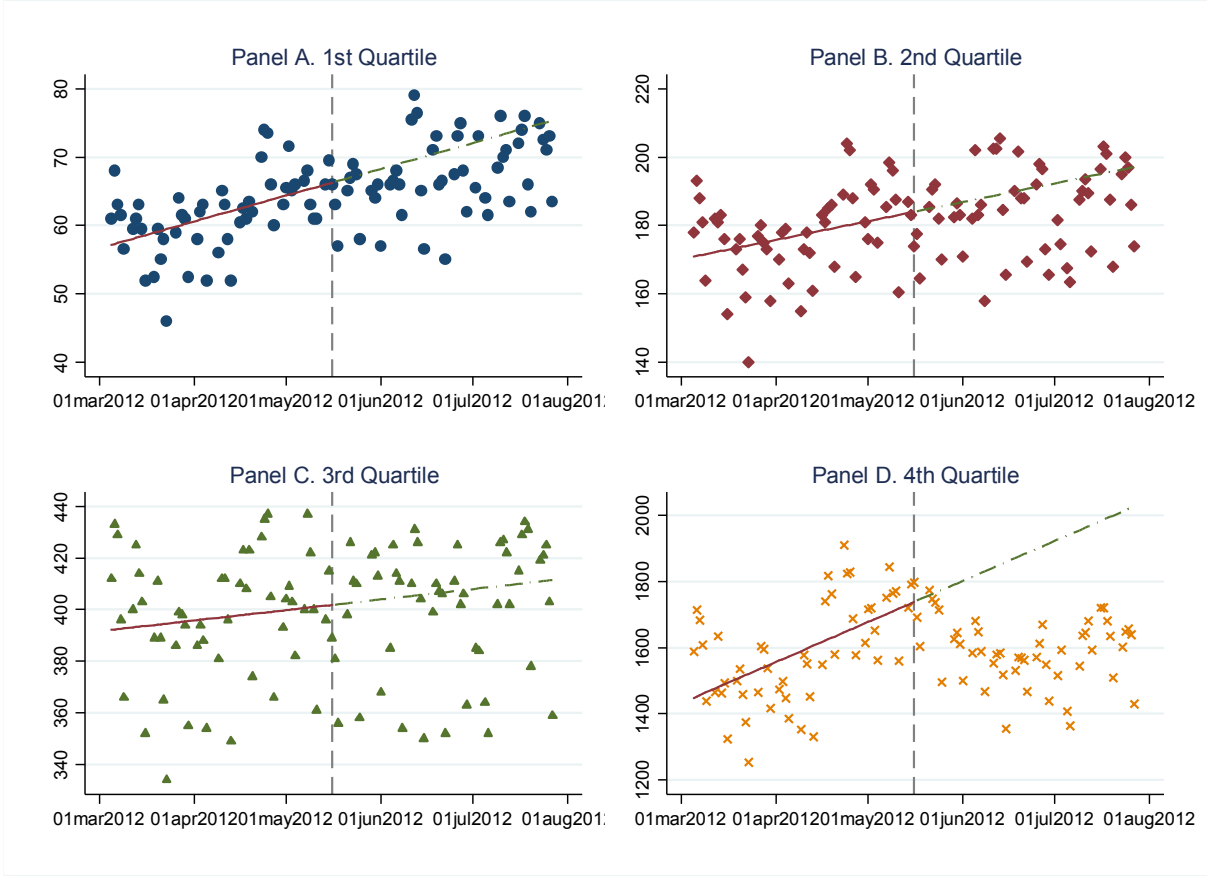


Figure 3. Abnormal wiki-visits around key information events

I plot the abnormal wiki-visits in an eleven-day window around earnings announcements, issuances of management guidance, and filings of 10-K and 10-Q in the first two graphs and around stock price crashes, jumps, and M&A announcements in the last two graphs. Please refer to Appendix B for the definitions of these events.

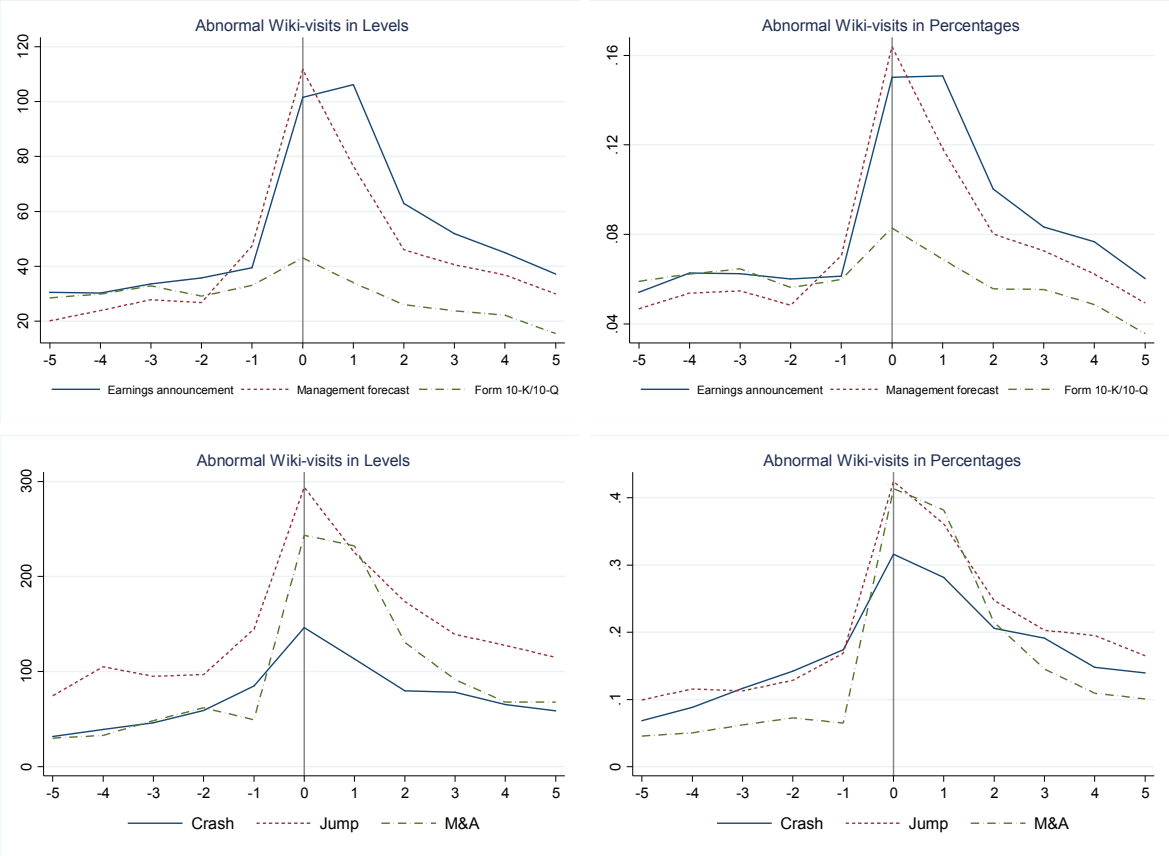


Figure 4. The speed of price discovery at earnings announcements

This figure plots the intra-period timeliness curve over days [0,5] with day 0 as earnings announcement date. The y-axis is the percentage of *CAR[0,5]* earned as of the end of each event day. *CAR[0,5]* is the individual stock's cumulative abnormal returns minus the corresponding portfolio returns matched by size and book-to-market ratio, accumulated from day 0 to day 5. The solid (dashed) line represents earnings announcements with above-median (below-median) abnormal wiki-visits over days [0,1] (*Ab_Wiki[0,1]*).

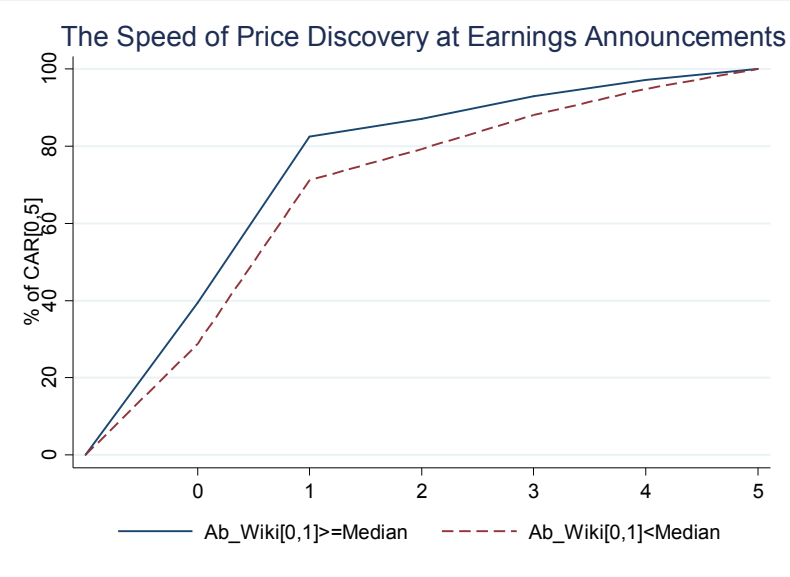


Table 1. Comparison of wiki-pages and Part I of 10-Ks

Panel A of this table reports the characteristics of historical wiki-pages as of the beginning of each sample month. The unit of observation is a wiki-page per month. Panel B reports the characteristics of Item 1. "Business" and Item 2. "Properties" of 10-Ks filed during the sample period. The unit of observation is one 10-K filing.

Panel A. Characteristics of Wiki-pages						
	# Obs.	**Mean**	**Std Dev**	**P25**	**P50**	**P75**
Words	16734	1779.376	1906.355	569.000	1111.000	2195.000
Sentences	16734	65.893	68.280	21.000	43.000	80.000
Gunning-Fog	16734	15.438	2.564	13.972	15.174	16.484
Fleschkincaid	16734	12.543	2.395	11.213	12.283	13.464
Smog	16734	13.645	1.753	12.545	13.526	14.498
LIX	16734	59.859	7.493	55.459	59.167	63.247
RIX	16734	8.671	2.410	7.321	8.333	9.483
Boilerplate	16734	0.000	0.000	0.000	0.000	0.000
Tone	16734	0.002	0.005	-0.001	0.001	0.004
Uncertainty	16734	0.012	0.009	0.004	0.010	0.018
Sections	16734	10.492	8.385	6.000	8.000	13.000
Sections_Intangibles	16734	0.330	0.570	0.000	0.000	1.000
Sections_Segments	16734	0.672	0.802	0.000	1.000	1.000
Sections_Other	16734	9.491	7.993	4.000	8.000	12.000
References	16734	41.972	52.631	12.000	24.000	49.000
Ref_Firm_Prepared	16734	6.847	9.076	1.000	4.000	9.000
Ref_News	16734	10.574	16.637	2.000	5.000	11.000
Ref_3rd_Party	16734	2.534	5.413	0.000	1.000	3.000
Ref_Others	16734	22.017	29.036	5.000	12.000	25.000
Editors[-4W]	16734	3.179	6.181	0.000	1.000	4.000
Revisions[-4W]	16734	3015.735	28683.900	0.000	35.000	422.000
Panel B. Characteristics of Item 1 and Item 2 of 10-Ks						
Words	1151	16161.26	7931.41	10383.00	14920.00	20204.00
Sentences	1151	529.50	247.97	349.00	497.00	649.00
Gunning-Fog	1151	19.70	1.31	18.87	19.78	20.52
Fleschkincaid	1151	16.09	1.22	15.33	16.16	16.82
Smog	1151	16.79	0.83	16.24	16.86	17.33
LIX	1151	67.86	3.45	65.55	67.87	70.20
RIX	1151	11.37	1.21	10.55	11.36	12.15
Boilerplate	1151	0.14	0.05	0.11	0.14	0.18
Tone	1151	-0.02	0.01	-0.02	-0.02	-0.01
Uncertainty	1151	0.02	0.00	0.02	0.02	0.03

Table 2. Summary statistics

Panel A of this table reports the summary statistics of the key variables used in this study. Each observation is a firm-calendar day and firm characteristics are as of the most recent data date available to investors. Variables under "*Market Variables*" are only defined on trading days. Panel B tabulates the distribution of the quintiles of the current and the next day's wiki-visits. Panel C presents the pairwise correlation matrix between the selected proxies for information acquisition. Correlation coefficients significant at 1% are indicated in bold. See Appendix B for detailed variable definitions.

Panel A. Summary Statistics

	# Obs.	Mean	Std Dev	P25	P50	P75
Information Acquisition						
Raw_Wiki	510592	995.869	3902.267	135.000	321.000	757.000
Ab_Wiki	470961	13.463	205.502	-27.500	-0.500	29.000
Ab_Wiki%	470611	2.941	23.511	-9.653	-0.191	10.769
GSV	510592	35.221	23.740	14.700	33.300	53.200
Ab_GSV	510592	-0.003	0.445	-0.157	0.000	0.153
EDGAR	406232	111.913	795.750	33.000	56.000	95.000
Ab_EDGAR	404308	6.267	38.467	-11.500	1.000	17.000
Firm Characteristics						
Size	448400	9.911	1.299	8.951	9.768	10.671
Age	448400	3.326	0.820	2.890	3.401	3.892
ROA	448400	0.014	0.019	0.004	0.012	0.022
Loss	448400	0.104	0.305	0.000	0.000	0.000
Leverage	448400	0.311	0.184	0.183	0.297	0.421
BTM	448400	0.395	0.340	0.171	0.314	0.520
OnBS_Intangibles	448400	0.262	0.255	0.030	0.202	0.427
OffBS_Intangibles	448400	0.239	0.271	0.021	0.145	0.378
Bus_Segments	448400	2.566	2.333	1.000	2.000	4.000
IO	448400	0.758	0.248	0.707	0.823	0.915
Analyst	448400	13.607	8.042	7.000	13.000	18.000
Short_Interest	448400	0.040	0.040	0.015	0.025	0.048
News_Dummy	448400	0.136	0.343	0.000	0.000	0.000
|CAR_Month|	448400	0.047	0.044	0.015	0.034	0.063
Informational Events						
EA	510592	0.011	0.102	0.000	0.000	0.000
MF	510592	0.008	0.091	0.000	0.000	0.000
10K&10Q	510592	0.010	0.100	0.000	0.000	0.000
M&A	510592	0.002	0.046	0.000	0.000	0.000
Crash	510592	0.003	0.051	0.000	0.000	0.000
Jump	510592	0.002	0.049	0.000	0.000	0.000
Market Variables						
NetTrade-Retail	337163	24.779	177.232	-40.000	1.000	53.000
NetTrade -Institution	337163	-4.595	40.465	-12.000	-1.000	8.000
NetVolume-Retail	337163	2.316	48.643	-9.038	-0.139	9.313

NetVolume- Institution	337163	-8.343	232.738	-49.024	-1.125	41.510
Raw_Wiki	337732	1092.197	4080.850	165.000	379.000	821.000
Ab_Wiki	310809	17.373	213.048	-29.000	0.500	33.500
AR	337725	0.000	0.013	-0.006	0.000	0.006
|AR|	337725	0.009	0.010	0.003	0.006	0.012
CAR[1W]	337692	0.000	0.030	-0.015	0.000	0.015
CAR[2W]	337505	0.000	0.030	-0.015	0.000	0.015
CAR[3W-4W]	337315	-0.001	0.054	-0.029	0.000	0.029
CAR[5W-12W]	336219	-0.003	0.089	-0.053	-0.001	0.050
AVol	337732	0.001	0.005	-0.002	-0.000	0.002
News_Dummy	337732	0.186	0.389	0.000	0.000	0.000

Panel B. Variation of wiki-visits

		Percentage of Raw_Wiki_{t+1} in Each Quintile				
Quintile of Raw_Wiki_t	# Obs	Quintile1	Quintile2	Quintile3	Quintile4	Quintile5
Quintile1	102,751	84.94	14.52	0.51	0.03	0.01
Quintile2	101,774	15.04	68.61	15.45	0.85	0.05
Quintile3	101,501	0.24	16.4	68.85	14.28	0.23
Quintile4	101,977	0.01	0.38	14.93	78.02	6.66
Quintile5	102,019	0.00	0.01	0.11	6.86	93.02

		Percentage of Ab_Wiki_{t+1} in Each Quintile				
Quintile of Ab_Wiki_t	# Obs	Quintile1	Quintile2	Quintile3	Quintile4	Quintile5
Quintile1	94,222	62.76	16.12	6	6.68	8.44
Quintile2	94,583	16.43	34.26	24.22	18.11	6.99
Quintile3	93,615	6.22	24.82	38.03	24.62	6.32
Quintile4	94,082	6.88	18.49	24.94	33.59	16.1
Quintile5	93,891	7.98	6.84	6.37	16.94	61.88

		Percentage of $Ab\_Wiki\%_{t+1}$ in Each Quintile				
Quintile of $Ab\_Wiki\%_t$	# Obs	Quintile1	Quintile2	Quintile3	Quintile4	Quintile5
Quintile1	93,728	49.68	21.85	12.27	8.84	7.36
Quintile2	94,048	22.47	31.18	23.29	14.73	8.33
Quintile3	94,070	12.55	24.04	28.44	22.74	12.23
Quintile4	94,102	8.69	14.97	24.08	30.92	21.34
Quintile5	94,059	6.78	8.03	11.88	22.69	50.62

Panel C. Correlation Matrix

	Raw_Wiki	Ab_Wiki	Ab_Wiki%	GSV	Ab_GSV	EDGAR
Ab_Wiki	**0.135**					
Ab_Wiki%	**0.075**	**0.578**				
GSV	-0.002	**0.017**	**0.027**			
Ab_GSV	**0.015**	**0.079**	**0.099**	**0.242**		
EDGAR	**0.328**	**-0.043**	0.000	**0.005**	**0.007**	
Ab_EDGAR	**0.089**	**0.211**	**0.181**	**0.035**	**0.093**	**0.140**

Table 3. Wiki-visits, trading, and returns

Panel A reports the panel regressions of the next day's trading activities by retail or institutional investors on abnormal wiki-visits. Retail and institutional trades are identified from TAQ, following Boehmer et al. (2017) and Bushee et al. (2018). The dependent variables are indicated in the table header where *NetTrade* is the number of buy transactions minus that of sale transactions and *NetVolume* is the net shares traded. Panel B reports the daily Fama-MacBeth regressions of subsequent size and book-to-market adjusted cumulative abnormal returns on abnormal wiki-visits. For ease of presentation, cumulative abnormal returns are converted into percentages. In Columns 5-6 (7-8), I split the sample based on the median number of references (editors who made revisions over the past 4 weeks) and repeat the same regressions. To facilitate comparison of the coefficients, I standardize all the independent variables. See Appendix B for detailed variable definitions. Robust standard errors clustered by firm (Newey and West (1987) standard errors with lags of four) are reported in the parentheses of Panel A (B). ***, **, and * stand for statistical significance at the 1%, 5%, and 10% level, respectively.

Panel A. Subsequent Trades

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Retail Investors		**Institutional Investors**		**Retail Investors**		**Institutional Investors**	
VARIABLES	$NetTrade_{t+1}$	$NetTrade_{t+1}$	$NetTrade_{t+1}$	$NetTrade_{t+1}$	$NetVolume_{t+1}$	$NetVolume_{t+1}$	$NetVolume_{t+1}$	$NetVolume_{t+1}$
Ab_Wiki	6.390***	5.325***	-0.081	0.010	0.647**	0.625**	-0.394	-0.292
	(1.586)	(1.175)	(0.207)	(0.171)	(0.288)	(0.280)	(0.799)	(0.806)
|AR|	3.097***	0.545	0.134	0.106	1.222***	0.281	0.001	0.569
	(1.144)	(0.685)	(0.208)	(0.116)	(0.381)	(0.214)	(0.832)	(0.656)
|CAR_Month|	5.879***	2.925***	-0.821***	-0.613***	1.244***	0.203	-3.870***	-2.633***
	(1.625)	(0.898)	(0.232)	(0.135)	(0.416)	(0.187)	(0.899)	(0.648)
AVol	3.283***	4.393***	-0.398**	-0.363***	0.575	0.955***	-1.153	-1.298
	(1.187)	(1.044)	(0.176)	(0.140)	(0.350)	(0.303)	(0.948)	(0.917)
News_Dummy	3.441***	0.121	-0.690**	0.234***	0.592***	0.149	-1.388*	0.938*
	(0.985)	(0.397)	(0.277)	(0.090)	(0.197)	(0.106)	(0.836)	(0.503)
Ab_GSV	2.791***	2.645***	-0.060	-0.023	0.361**	0.372***	-1.533***	-1.537***
	(0.887)	(0.710)	(0.115)	(0.108)	(0.168)	(0.139)	(0.515)	(0.505)
Size	38.558***	51.323***	-7.009***	-3.018	2.517**	6.042*	-11.837***	-10.683
	(5.928)	(15.307)	(1.030)	(2.038)	(1.068)	(3.524)	(2.787)	(9.901)
Age	-1.035	5.225	-0.053	-0.676	-0.847	-1.074	-0.096	3.251
	(3.843)	(13.746)	(0.424)	(0.993)	(0.770)	(2.434)	(1.023)	(3.678)
ROA	13.881***	7.785***	-2.457***	-0.366	1.264	0.433	-3.422*	-2.057
	(4.801)	(2.987)	(0.661)	(0.372)	(1.130)	(0.796)	(1.887)	(1.615)
Loss	5.108*	3.065*	-0.982**	0.015	1.869**	0.947**	-1.594	-0.498
	(2.968)	(1.743)	(0.406)	(0.224)	(0.765)	(0.406)	(1.409)	(1.101)
Leverage	-5.087	-17.130**	0.763	2.917***	-0.011	-4.211*	-1.559	-2.631

	(3.697)	(7.944)	(0.514)	(1.054)	(0.681)	(2.173)	(1.183)	(3.724)
BTM	-25.336***	-30.472***	2.426**	2.097**	-0.793	-3.371**	-0.492	0.695
	(5.321)	(6.847)	(0.959)	(0.828)	(1.294)	(1.519)	(2.430)	(3.292)
IO	5.471	5.493	-0.527	0.224	-0.156	0.926	-2.154	1.421
	(3.863)	(5.227)	(0.436)	(0.568)	(0.676)	(0.909)	(1.403)	(2.966)
Ln(Analyst)	-4.636	-5.659	0.201	2.953	-1.012	0.711	0.727	2.561
	(3.455)	(15.246)	(0.478)	(2.552)	(0.663)	(3.775)	(1.430)	(10.713)
Short_Interest	-6.408***	7.677*	0.145	-0.812	-0.130	2.075***	-1.385	0.690
	(2.435)	(4.273)	(0.353)	(0.527)	(0.457)	(0.798)	(1.189)	(3.304)
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	308,643	308,643	308,643	308,643	308,643	308,643	308,643	308,643
R-squared	0.089	0.266	0.031	0.101	0.023	0.106	0.020	0.036

Panel B. Subsequent Returns

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
					References		**Page Revisions**	
VARIABLES	CAR[1W]	CAR[2W]	CAR [3W-4W]	CAR [5W-12W]	CAR[1W] Many	CAR[1W] Few	CAR[1W] Many	CAR[1W] Few
Ab_Wiki	0.025**	0.013	0.011	0.029	0.023**	-0.002	0.026**	-0.045
	(0.011)	(0.010)	(0.019)	(0.032)	(0.011)	(0.039)	(0.010)	(0.049)
AR	-0.019	-0.003	-0.005	0.111*	-0.011	-0.025	-0.016	-0.020
	(0.020)	(0.019)	(0.030)	(0.057)	(0.021)	(0.023)	(0.019)	(0.026)
CAR_Month	-0.071**	-0.029	0.029	0.236**	-0.072**	-0.071*	-0.068**	-0.083*
	(0.033)	(0.030)	(0.061)	(0.110)	(0.033)	(0.039)	(0.030)	(0.045)
AVol	-0.040**	-0.006	-0.022	0.084	-0.056***	-0.030	-0.059***	-0.005
	(0.017)	(0.017)	(0.032)	(0.053)	(0.021)	(0.022)	(0.018)	(0.028)
News_Dummy	0.016**	0.008	0.013	0.014	0.022**	0.003	0.011	0.024*
	(0.007)	(0.007)	(0.015)	(0.026)	(0.009)	(0.011)	(0.008)	(0.013)
Ab_GSV	0.005	-0.008	-0.002	0.053*	0.010	0.002	0.022*	-0.019
	(0.009)	(0.009)	(0.015)	(0.027)	(0.014)	(0.012)	(0.011)	(0.013)
Observations	309,169	309,000	308,825	307,797	156,629	152,540	196,808	112,361
R-squared	0.057	0.053	0.050	0.049	0.067	0.081	0.059	0.100
Number of days	582	582	582	582	582	582	582	582

Table 4. Subsequent trades and returns around the launch of GKG

This table examines the impact of wiki-visits on subsequent trades and abnormal returns in 2012, split into two subperiods by the launch of GKG on May 16, 2012. Panel A reports the panel regressions of subsequent retail or institutional trades on abnormal wiki-visits in the pre- and post-GKG period, respectively. Retail and institutional trades are identified from TAQ, following Boehmer et al. (2017) and Bushee et al. (2018). Control variables, whose coefficients are not reported for brevity, include *Size, Age, ROA, Loss, Leverage, BTM, IO, Ln(Analyst),* and *Short_Interest*. Panel B reports the daily Fama and MacBeth (1973) regressions of subsequent returns on abnormal wiki-visits. The dependent variable is size and book-to-market adjusted abnormal returns accumulated over the next week (*CAR[1W]*). For ease of presentation, *CAR[1W]* is converted into percentages. To facilitate comparison of the coefficients, I standardize all the independent variables. Robust standard errors clustered by firm (Newey and West (1987) standard errors with lags of three) are reported in the parentheses of Panel A (B). ***, **, and * stand for statistical significance at the 1%, 5%, and 10% level, respectively.

Panel A. Subsequent Trades

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	NetTrade$_{t+1}$				**NetVolume$_{t+1}$**			
	Retail Investors		**Institutional Investors**		**Retail Investors**		**Institutional Investors**	
VARIABLES	Pre-GKG	Post-GKG	Pre-GKG	Post-GKG	Pre-GKG	Post-GKG	Pre-GKG	Post-GKG
Ab_Wiki	1.199	4.117**	-0.218	0.173	0.271	1.211**	-0.259	0.745
	(2.277)	(1.780)	(0.461)	(0.410)	(0.619)	(0.661)	(2.670)	(2.181)
|AR|	1.121	2.565**	0.052	-0.023	0.961	1.086***	-0.090	-0.184
	(1.904)	(1.069)	(0.176)	(0.142)	(0.605)	(0.417)	(1.582)	(1.194)
|CAR_Month|	3.083	2.226	-0.182	-0.140	0.780	0.295	-1.504	-2.298
	(2.394)	(1.761)	(0.226)	(0.175)	(0.706)	(0.645)	(1.557)	(1.653)
AVol	8.784***	6.286***	0.213	0.297**	1.715***	1.553***	-0.230	0.550
	(1.972)	(1.383)	(0.181)	(0.148)	(0.553)	(0.498)	(1.629)	(1.289)
News_Dummy	-1.870	-2.572*	-0.232	-0.006	-0.930	-1.469***	-1.625	-0.718
	(1.953)	(1.391)	(0.173)	(0.185)	(0.675)	(0.524)	(1.540)	(1.319)
Ab_GSV	2.381*	0.639	-0.092	-0.122	-0.062	0.122	-0.919	-0.306
	(1.364)	(0.900)	(0.134)	(0.113)	(0.434)	(0.299)	(1.083)	(0.879)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	45,709	76,155	45,709	76,155	45,709	76,155	45,709	76,155
R-squared	0.024	0.044	0.031	0.034	0.017	0.044	0.010	0.038

Panel B. Subsequent Returns (*CAR[1W]*)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Full Sample		**More References**		**Few References**		**More Revisions**		**Few Revisions**	
VARIABLES	Pre-KG	Post-KG	Pre-KG	Post-KG	Pre-KG	Post-KG	Pre-KG	Post-KG	Pre-KG	Post-KG
Ab_Wiki	0.000	0.036**	-0.003	0.040**	-0.085	-0.031	-0.012	0.037*	0.125	-0.021
	(0.022)	(0.017)	(0.024)	(0.020)	(0.104)	(0.084)	(0.023)	(0.019)	(0.101)	(0.089)
AR	0.006	-0.014	-0.003	0.002	0.011	-0.031	0.002	0.001	0.023	-0.033
	(0.035)	(0.024)	(0.035)	(0.025)	(0.043)	(0.031)	(0.038)	(0.025)	(0.042)	(0.033)
CAR_Month	0.106*	-0.064	0.096*	-0.048	0.106	-0.082*	0.123**	-0.044	0.067	-0.093*
	(0.055)	(0.039)	(0.056)	(0.039)	(0.075)	(0.050)	(0.050)	(0.042)	(0.080)	(0.048)
AVol	0.024	-0.005	0.018	-0.016	0.050*	-0.012	0.029	0.013	0.016	-0.023
	(0.026)	(0.025)	(0.036)	(0.031)	(0.029)	(0.035)	(0.028)	(0.028)	(0.047)	(0.040)
News_Dummy	-0.032	0.007	-0.027	0.008	-0.031	0.013	-0.029	0.003	-0.043*	0.029
	(0.022)	(0.011)	(0.026)	(0.014)	(0.024)	(0.018)	(0.026)	(0.014)	(0.024)	(0.020)
Ab_GSV	-0.027*	0.010	-0.003	0.047***	-0.045**	-0.022	-0.020	0.015	-0.049**	0.008
	(0.014)	(0.012)	(0.024)	(0.017)	(0.018)	(0.016)	(0.017)	(0.017)	(0.022)	(0.017)
Observations	46,808	78,302	23,976	39,917	22,832	38,385	30,039	49,936	16,769	28,366
R-squared	0.041	0.040	0.052	0.052	0.062	0.062	0.044	0.049	0.078	0.069
Number of days	94	156	94	156	94	156	94	156	94	156

Table 5. Wiki-visits around informational events

This table examines wiki-visits around six important informational events (earnings announcements (*EA*), management forecasts (*MF*), 10-K or 10-Q filings (*10-K&10-Q*), merger or acquisition announcements (*M&A*), stock price crashes (*Crash*), and jumps (*Jump*)). -1, 0, +1 in the brackets indicate the day before, on, and after the event. The dependent variables are abnormal wiki-visits (*Ab_Wiki*) in Columns 1-4, abnormal percentage wiki-visits (*Ab_Wiki%*) in Columns 5 and 6, the natural logarithm of one plus the raw daily wiki-visits (*Ln(Raw_Wiki)*) in Columns 7 and 8. Control variables, whose coefficients are not reported for brevity, include *Size, Age, ROA, Loss, Leverage, BTM, IO, Ln(Analyst), Short_Interest, News_Dummy, |CAR_Month|, Ab_GSV, Intangibles,* and *Ln(Bus_Segments)*. See Appendix B for detailed variable definitions. Standard errors are clustered by date (Petersen 2009). ***, **, and * stand for statistical significance at the 1%, 5%, and 10% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Ab_Wiki	Ab_Wiki	Ab_Wiki	Ab_Wiki	Ab_Wiki%	Ab_Wiki%	Ln(Raw_Wiki)	Ln(Raw_Wiki)
EA[-1]		9.477**	7.973*	6.627	1.102***	0.781**	0.008**	0.004
		(3.731)	(4.213)	(4.366)	(0.332)	(0.379)	(0.004)	(0.004)
EA[0]	41.778***	69.089***	50.493***	33.683***	9.937***	5.139***	0.090***	0.041***
	(5.031)	(4.495)	(5.118)	(5.352)	(0.512)	(0.565)	(0.005)	(0.006)
EA[+1]		73.471***	52.142***	43.287***	10.164***	5.509***	0.092***	0.040***
		(4.333)	(4.709)	(4.839)	(0.498)	(0.535)	(0.005)	(0.005)
MF[-1]			4.548	2.107		0.048		0.010**
			(3.854)	(3.865)		(0.429)		(0.005)
MF[0]	63.238***		48.613***	45.545***		6.260***		0.075***
	(5.223)		(5.252)	(5.179)		(0.527)		(0.006)
MF[+1]			29.901***	27.577***		3.894***		0.045***
			(4.718)	(4.679)		(0.491)		(0.005)
10K&10Q[-1]				-0.001		0.312		0.003
				(3.554)		(0.333)		(0.004)
10K&10Q[0]	-1.294			-10.536***		-0.100		-0.001
	(3.330)			(3.354)		(0.355)		(0.004)
10K&10Q[+1]				-4.746		0.120		-0.007
				(3.514)		(0.350)		(0.004)
M&A[-1]				6.769		1.216*		0.008
				(7.972)		(0.707)		(0.008)
M&A[0]	121.855***			112.586***		19.010***		0.207***
	(12.266)			(12.175)		(1.419)		(0.017)
M&A[+1]				116.674***		18.703***		0.192***

				(11.843)		(1.357)		(0.015)
Crash[-1]				27.167***		4.851***		0.051***
				(6.804)		(0.847)		(0.011)
Crash[0]	92.693***			75.734***		14.626***		0.147***
	(8.775)			(8.535)		(1.048)		(0.013)
Crash[+1]				57.013***		12.400***		0.123***
				(7.030)		(0.960)		(0.012)
Jump[-1]				25.498***		3.323***		0.049***
				(7.836)		(0.867)		(0.010)
Jump[0]	125.062***			103.260***		16.637***		0.199***
	(11.124)			(10.650)		(1.321)		(0.016)
Jump[+1]				95.158***		14.435***		0.161***
				(9.939)		(1.216)		(0.015)
Controls	No	No	No	Yes	No	Yes	No	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Day FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	470,961	470,961	470,961	448,400	470,611	448,075	510,592	487,067
R-squared	0.072	0.070	0.071	0.078	0.141	0.156	0.961	0.964

Table 6. Earnings announcement characteristics

This table examines wiki-visits around earnings announcements. The dependent variable is abnormal wiki-visits (*Ab_Wiki*). Column 1 compares *Ab_Wiki* at earnings announcements with positive to negative surprises relative to the median analyst consensus. In Columns 2-6, earnings announcements are split into two categories based on the median of the variable indicated in the table header. *ATone* is abnormal tone in earnings press releases (Huang et al. 2013); *Unreadable* and *Boilerplate* quantify the difficulty to read or the proportion of boilerplate language in Item 1 and 2 of the most recent 10-K filing; *OnBS_Intangibles* and *OffBS_Intangibles* refer to the intensity of on- and off-balance-sheet intangible assets (Peters and Taylor 2017). Control variables, whose coefficients are not reported for brevity, include *Size, Age, ROA, Loss, Leverage, BTM, IO, Ln(Analyst), Short_Interest, News_Dummy, |CAR_Month|, Ab_GSV, Intangibles,* and *Ln(Bus_Segments)*. See Appendix B for detailed variable definitions. Standard errors are clustered by date (Petersen 2009). Two-tailed t-tests are conducted to test the difference in the coefficients on the two earnings announcement indicators. ***, **, and * stand for statistical significance at the 1%, 5%, and 10% level, respectively.

Dependent Variable: *Ab_Wiki*						
	(1)	(2)	(3)	(4)	(5)	(6)
Partition Variable:	SUE	ATone	Unreadable	Boilerplate	OnBS_Intangibles	OffBS_Intangibles
EA[0,1]&SUE>=0	57.747***					
	(3.846)					
EA[0,1]&SUE<0	52.753***					
	(5.293)					
EA[0,1]&Partition Variable>=Median		61.701***	61.697***	64.810***	54.450***	83.055***
		(4.633)	(4.246)	(4.360)	(4.082)	(4.510)
EA[0,1]&Partition Variable<Median		48.595***	53.350***	50.128***	61.881***	32.513***
		(3.631)	(4.253)	(3.960)	(4.724)	(3.936)
Difference	4.994	13.110***	8.347*	14.682***	-7.431	50.542***
S.E.	(5.961)	(4.930)	(4.977)	(4.689)	(5.545)	(5.214)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Day FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	434,306	385,577	440,745	440,745	448,400	448,400
R-squared	0.073	0.074	0.076	0.076	0.075	0.075

Table 7. Abnormal returns and abnormal turnovers around earnings announcement

This table reports the association between abnormal returns, abnormal trading volume, earnings surprises, and abnormal wiki-visits around earnings announcements. The unit of observation is an earnings announcement. Panel A reports the sample summary statistics. The dependent variable in Panel B is the size and book-to-market adjusted abnormal returns accumulated over the five trading days prior to earnings announcements (*CAR[-5,-1]*) or over the two-trading-day window during the announcements (*CAR[0,1]*). The dependent variable in Panel C is average daily abnormal turnover over the same windows. The sample is spilt into two subsamples based on the median analyst forecast dispersion in Columns 4-5 of Panel B and C. The variable of interest is the average daily abnormal wiki-visits over the five-day or two-day windows (*Ab_Wiki[-5,-1]* or *Ab_Wiki[0,1]*, unit: 1000). Control variables include *Size, Age, Loss, Leverage, BTM, IO, Ln(Analyst), Short_Interest, News[-5,-1]* (*News[0,1]* when the dependent variable is *CAR[0,1]* or *CAV[0,1]*), *Ab_GSV[-5,-1]* (*Ab_GSV[0,1]* when the dependent variable is *CAR[0,1]* or *CAV[0,1]*), *Fqtr4, Earn_Std, Earn_Rho,* and *Ret_Std*. See Appendix B for detailed variable definitions. For ease of interpretation, these control variables are standardized to have a mean of zero and a standard deviation of one, and their coefficients are not reported for brevity. Standard errors are clustered by FF 17 industry x calendar quarter. ***, **, and * stand for statistical significance at the 1%, 5%, and 10% level, respectively.

Panel A. Summary Statistics

VARIABLES	# Obs.	Mean	Std Dev	P25	P50	P75
CAR[-5,-1]	4735	-0.001	0.028	-0.015	0.000	0.016
CAR[0,1]	4735	0.000	0.054	-0.030	0.001	0.030
CAV[-5,-1]	4735	0.001	0.003	-0.001	0.000	0.002
CAV[0,1]	4735	0.011	0.015	0.003	0.006	0.012
IPT[0,5]	4735	-0.234	16.701	1.850	3.765	4.619
IPT[0,5]Truncated	3966	3.517	2.023	2.975	4.053	4.740
SUE	4689	0.001	0.002	0.000	0.000	0.001
|SUE|	4689	0.002	0.002	0.000	0.001	0.002
Ab_Wiki[-5,-1]	4735	0.035	0.193	-0.011	0.006	0.038
Ab_Wiki[0,1]	4735	0.115	0.331	-0.001	0.024	0.097
Ab_Wiki%[-5,-1]	4733	0.059	0.171	-0.039	0.029	0.107
Ab_Wiki%[0,1]	4733	0.155	0.289	-0.001	0.089	0.219
News[-5,-1]	4735	0.184	0.229	0.000	0.182	0.336
News[0,1]	4735	0.822	0.308	0.693	0.693	0.916
Ab_GSV[-5,-1]	4735	0.057	0.325	-0.081	0.024	0.167
Ab_GSV[0,1]	4735	0.307	0.621	-0.034	0.095	0.438
Size	4727	9.909	1.033	9.196	9.778	10.479
Age	4735	3.316	0.819	2.890	3.401	3.871
Loss	4735	0.105	0.307	0.000	0.000	0.000
Leverage	4729	0.310	0.183	0.182	0.296	0.420
BTM	4727	0.394	0.339	0.171	0.311	0.518
IO	4729	0.761	0.244	0.708	0.822	0.915
Analyst	4735	13.617	8.048	7.000	13.000	18.000
Short_interest	4735	0.040	0.040	0.015	0.025	0.049
Fqtr4	4735	0.231	0.422	0.000	0.000	0.000
Earn_Rho	4686	0.325	0.395	0.019	0.279	0.636
Earn_Std	4686	0.009	0.007	0.004	0.007	0.013
Ret_Std	4735	0.256	0.109	0.182	0.230	0.298
Dispersion	4620	0.056	0.076	0.017	0.031	0.061

Panel B. Abnormal Returns

	(1)	(2)	(3)	(4)	(5)
				Analyst Dispersion	
				High	**Low**
VARIABLES	CAR[-5,-1]	CAR[0,1]	CAR[0,1]	CAR[0,1]	CAR[0,1]
SUE	0.157	8.616***	8.067***	6.926***	12.144***
	(0.273)	(0.678)	(0.687)	(0.747)	(1.156)
Ab_Wiki[-5,-1]	0.005*	0.004			
	(0.002)	(0.004)			
SUE*Ab_Wiki[-5,-1]	-0.733	-2.270			
	(1.157)	(2.473)			
Ab_Wiki[0,1]			-0.004	-0.003	-0.002
			(0.003)	(0.004)	(0.005)
SUE*Ab_Wiki[0,1]			4.012**	3.826**	2.047
			(1.810)	(1.936)	(3.609)
Other Controls	Yes	Yes	Yes	Yes	Yes
SUE*Other Controls	Yes	Yes	Yes	Yes	Yes
Observations	4,633	4,633	4,633	2,300	2,264
R-squared	0.020	0.116	0.117	0.124	0.143

Panel B. Abnormal Volume

	(1)	(2)	(3)	(4)	(5)
				Analyst Dispersion	
				High	**Low**
VARIABLES	CAV[-5,-1]	CAV[0,1]	CAV[0,1]	CAV[0,1]	CAV[0,1]
|SUE|	-0.018	0.870***	0.661***	0.539***	0.798***
	(0.029)	(0.160)	(0.143)	(0.181)	(0.217)
Ab_Wiki[-5,-1]	0.001***	0.002			
	(0.000)	(0.001)			
|SUE|*Ab_Wiki[-5,-1]	0.381*	0.475			
	(0.193)	(0.693)			
Ab_Wiki[0,1]			0.008***	0.007***	0.009***
			(0.001)	(0.001)	(0.002)
|SUE|*Ab_Wiki[0,1]			1.835***	2.142***	0.738
			(0.576)	(0.646)	(1.036)
Other Controls	Yes	Yes	Yes	Yes	Yes
|SUE|*Other Controls	Yes	Yes	Yes	Yes	Yes
Observations	4,633	4,633	4,633	2,300	2,264
R-squared	0.133	0.403	0.441	0.485	0.422

Table 8. The speed of price discovery at earnings announcement

This table reports the association between the speed of price discovery and wiki-visits at earnings announcements. The dependent variable is intra-period timeliness (*IPT[0,5]*) adjusted for overreactions and reversals (Blankespoor, deHaan, and Zhu 2018). To reduce the noise in *IPT[0,5]*, I set the value of *IPT[0,5]* to be missing when the absolute value of *CAR[0,5]* is less than 1% in Columns 4-6. The sample is split into two subsamples based on the median analyst forecast dispersion in Columns 2-3 and Columns 5-6, respectively. The key variable of interest is the average abnormal wiki-visits during the two-day window at earnings announcements (*Ab_Wiki[0,1]*). See Appendix B for detailed variable definitions. Standard errors are clustered by FF 17 industry x calendar quarter. ***, **, and * stand for statistical significance at the 1%, 5%, and 10% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
		Analyst Dispersion			**Analyst Dispersion**	
		High	**Low**		**High**	**Low**
VARIABLES	IPT[0,5]	IPT[0,5]	IPT[0,5]	IPT[0,5] Truncated	IPT[0,5] Truncated	IPT[0,5] Truncated
Ab_Wiki[0,1]	1.125**	1.579***	0.337	1.060*	1.377***	0.382
	(0.564)	(0.540)	(1.007)	(0.568)	(0.502)	(1.120)
News[0,1]	-1.018	-1.755	-0.264	-1.588*	-1.852	-1.497
	(0.823)	(1.202)	(1.206)	(0.864)	(1.184)	(1.317)
Ab_GSV[0,1]	1.031***	1.423***	0.631	1.198***	1.552***	0.842
	(0.330)	(0.413)	(0.540)	(0.373)	(0.436)	(0.636)
|SUE|	93.676	88.745	70.356	55.336	50.011	-33.078
	(85.972)	(88.826)	(269.718)	(89.283)	(93.112)	(278.671)
Size	-0.519	-0.362	-0.536	-0.609	-0.462	-0.659
	(0.325)	(0.426)	(0.475)	(0.394)	(0.494)	(0.568)
Age	-0.030	0.091	-0.171	0.290	0.311	0.246
	(0.318)	(0.407)	(0.413)	(0.357)	(0.467)	(0.478)
Loss	0.634	1.050	0.094	-0.010	0.273	-0.406
	(0.810)	(1.047)	(1.606)	(0.860)	(1.056)	(1.815)
Leverage	-0.856	-1.424	-0.209	-1.126	-1.256	-1.065
	(1.703)	(2.258)	(1.864)	(1.744)	(2.469)	(1.939)
BTM	-1.637*	-1.634	-1.829	-1.489	-1.074	-2.276
	(0.988)	(1.063)	(1.870)	(1.045)	(1.166)	(2.174)
IO	-0.688	-0.991	-0.235	-1.371	-1.356	-1.254
	(0.951)	(1.186)	(1.424)	(1.018)	(1.249)	(1.533)
Ln(Analyst)	0.950**	0.437	1.432**	1.031**	0.676	1.390**
	(0.415)	(0.603)	(0.593)	(0.466)	(0.708)	(0.639)
Short_Interest	-0.774	3.655	-4.008	-0.891	5.755	-10.227
	(6.439)	(7.781)	(13.305)	(7.411)	(8.573)	(15.565)
Fqtr4	0.011	-0.388	0.235	-0.108	0.121	-0.441
	(0.648)	(0.866)	(0.838)	(0.626)	(0.899)	(0.913)
Earn_Rho	0.456	0.047	1.027	0.103	-0.066	0.383
	(0.605)	(0.877)	(0.845)	(0.645)	(0.924)	(0.953)
Earn_Std	3.693	0.798	12.255	0.065	-17.461	23.249
	(32.767)	(44.061)	(70.111)	(33.296)	(44.868)	(75.558)
Ret_Std	-3.677	-5.482	0.734	-3.320	-5.590	1.985
	(2.994)	(3.592)	(5.177)	(2.971)	(3.683)	(4.695)
Excluded Obs.	None	None	None	<+/-1%	<+/-1%	<+/-1%
Observations	4,633	2,300	2,264	3,811	1,893	1,864
R-squared	0.004	0.006	0.006	0.004	0.006	0.006