MIS Quarterly

RESEARCH ARTICLE

IMPACT OF WIKIPEDIA ON MARKET INFORMATION ENVIRONMENT: EVIDENCE ON MANAGEMENT DISCLOSURE AND INVESTOR REACTION[1]

Sean Xin Xu
School of Economics and Management, Tsinghua University
Beijing, CHINA {xuxin@sem.tsinghua.edu.cn}

Xiaoquan (Michael) Zhang
School of Business and Management, Hong Kong University of Science and Technology,
Clear Water Bay, HONG KONG {zhang@ust.hk}

In this paper, we seek to determine whether a typical social media platform, Wikipedia, improves the information environment for investors in the financial market. Our theoretical lens leads us to expect that information aggregation about public companies on Wikipedia may influence how management's voluntary information disclosure reacts to market uncertainty with respect to investors' information about these companies. Our empirical analysis is based on a unique data set collected from financial records, management disclosure records, news article coverage, and a Wikipedia modification history of public companies. On the supply side of information, we find that information aggregation on Wikipedia can moderate the timing of managers' voluntary disclosure of companies' earnings disappointments, or bad news. On the demand side of information, we find that Wikipedia's information aggregation moderates investors' negative reaction to bad news. Taken together, these findings support the view that Wikipedia improves the information environment in the financial market and underscore the value of information aggregation through the use of information technology.

Keywords: Social media, Wikipedia, information environment, financial market, management disclosure, information aggregation

Introduction

One primary function of markets is to aggregate information. When market participants cannot communicate with each other freely, they collect their information piecemeal and it "never exists in concentrated or integrated form" (Hayek 1945, p. 519). In this context, market price makes important contributions to information transmission, but such markets may still be inefficient, with price dispersion often observed in practice (Ba et al. 2012; Brynjolfsson and Smith 2000). Researchers thus ask whether market participants can use any explicit mechanism to directly exchange and aggregate information (e.g., Chen et al. 2011; Dellarocas and Wood 2008; Fan et al. 2000; Granados et al. 2010; Sun 2011; Sun and Tyagi 2012; Zhu and Zhang 2010).

A burgeoning literature examines whether social media can serve as such a mechanism (e.g., Gu et al. 2007; Tetlock et al.

[1]Ravi Bapna was the accepting senior editor for this paper. Ming Fan served as the associate editor.

The appendix for this paper is located in the "Online Supplements" section of the *MIS Quarterly*'s website (http://www.misq.org).

2008). Social media are Internet-based applications "that allow the creation and exchange of user-generated content" (Kaplan and Haenlein 2010, p. 61). On social media, users are not only consumers, but also active contributors of content. Following this line of research, we focus on Wikipedia, a typical social media platform (Greenstein and Zhu 2012a, 2012b; Kane and Fichman 2009; Ransbotham and Kane 2011). Wikipedia is an online encyclopedia based on the "wiki" technology that allows individuals to make collaborative contributions to web pages through a web browser (Kane and Fichman 2009). This technology allows people who have access to the Internet to freely contribute and modify the content of its web pages. Volunteers contributed all of the entries in the encyclopedia. After its creation in January 2001, Wikipedia quickly developed to become one of the largest reference websites, with currently 400 million unique visitors monthly and more than 90,000 active contributors.[2] The English version of Wikipedia, the most popular of the 270 different language versions, attracts hundreds of thousands of visitors who make tens of thousands of modifications every day. As of September 2011, Wikipedia is the seventh-most-popular website worldwide, receiving 2.7 billion page views every month.

On Wikipedia, several thousand articles are about traded companies. For these companies, Wikipedia keeps a good record of their significant events. Appendix A depicts a section of the Wikipedia entry for Dell. All sentences in this section follow the format "On [Month Day, Year], Dell [did something]." As such, Wikipedia effectively maintains a virtual memory where qualitative information exists in a concentrated and integrated form (Hayek 1945, Kankanhalli et al. 2005; Stein and Zwass 1995). Visitors thereby can obtain *aggregated*, timely information about the firm's products, management structure, operations, marketing, sales, and other important firm events. With this rich information, Wikipedia has become an important source of knowledge about companies. In September 2011, for example, Wikipedia entries for Microsoft and Google received 400,000 and 2,000,000 visits, respectively.[3] In particular, there is some evidence that Wikipedia has substantial impact on active seekers of information. According to HitWise, an Internet traffic-monitoring service, more than 70 percent of Wikipedia's incoming traffic comes from search engines.[4] Because of Wikipedia's importance in search engines, Wikipedia pages for listed companies are often ranked among the highest in search results. As such, when investors conduct research online, Wikipedia entries of listed companies are highly visible and accessible to them. As noted by *Financial Times*, "the capital market is looking for information from companies in various social media channels and at a greater extent than we anticipated….Wikipedia is the most popular social media site for individuals looking for such information, used by more than three quarters of respondents."[5]

Our research examines the value of Wikipedia for aggregating firm information, and we choose the modern financial market as our research context. The market's information environment consists of *quantified* information provided by financial analysts (mainly their forecasts of firm earnings) and *qualitative* descriptions of firms' business environments, operations, and prospects published by media. These linguistic descriptions are a valuable complement to analyst forecasts, because summative earnings forecasts cannot entirely convey the limitless variety of firms' financial standings (Tetlock et al. 2008) and because analyst forecasts are often incomplete or even biased measures of firm performance (Jacob et al. 1999; Karamanou and Vafeas 2005). The accounting and finance literatures thus emphasize a mosaic notion of investors' information environment (Francis et al. 1997; Lang and Lundholm 1996; Loss and Seligman 1995). The central tenet of the ***mosaic*** perspective, that information is multifaceted and that both quantified and qualitative types of information are useful, implies that the advent of detailed, linguistic information on Wikipedia would profoundly change investors' information environment. Data on qualitative information aggregation have been hard to come by, however, and the value of linguistic descriptions of a firm's environments, operations, and prospects has been severely neglected in prior studies.

Fortunately, Wikipedia offers not only a practical means for information aggregation, but also a way for scholars to investigate the process of information aggregation. On Wikipedia, the history of user contributions is open to the public. This gives us a unique opportunity to develop a proxy for *information aggregation on Wikipedia*—that is, the number of user modifications of public firms' Wikipedia entries (called "Wikipedia modifications" hereafter). Through collaborative

[2]Statistics in this paragraph are from http://en.wikipedia.org/wiki/Wikipedia (accessed October 2011).

[3]http://stats.grok.se/en/201109/microsoft, and http://stats.grok.se/en/201109/google, respectively.

[4]http://weblogs.hitwise.com/leeann-prescott/2007/02/wikipedia_traffic_sources.html.

[5]http://www.ft.com/intl/cms/s/0/c9d0271a-bf49-11dd-ae63-0000779fd18c.html.

modifications, people contribute their information to the same common Wikipedia entries. The aggregation takes place not only in that there is a comprehensive memory of the firms' events, but also in that individuals' information can be pooled together (as illustrated in Appendix A).

With this proxy for information aggregation on Wikipedia, our approach to identify its impact is based on a rich literature on management disclosure (for literature reviews, see Dye 2001; Healy and Palepu 2001). We examine how information aggregation on Wikipedia may influence managers' and investors' behaviors. From the disclosure literature, the timing of management disclosure about firms' unfavorable news is related to investors' information set about the firms (Healy and Palepu 2001). The intuition is that, driven by self-interest, managers decide whether to admit unfavorable news based on what investors already know (Dye 1985, 2001). So far, the literature widely recognizes the quantified information provided by analysts as the foremost source of information for investors,[6] and, accordingly, relates management disclosure to characteristics of the quantified information. Our overall prediction, motivated by research on media impact, is that "linguistic media content captures otherwise hard-to-quantify aspects of firms' fundamentals" (Tetlock et al. 2008, p. 1437). If Wikipedia's qualitative information really informs investors, then it has a potential to alter how management discourse is conditioned on characteristics of the quantified information.

We address the following research questions: (1) Does Wikipedia's aggregation of *qualitative* information affect how management disclosure is conditioned on *quantified* information in the market about firm earnings? (2) If Wikipedia can really inform investors in the market, how would that change investors' reactions when managers disclose unfavorable news?

Theoretical Background

Investigative Context

Our investigative context is the information environment of the modern financial market. Investors need information about firm performance in order to make investment decisions. The Securities and Exchange Commission (SEC) sets official announcement dates when public firms must disclose their performance. Before an official announcement, managers have discretion to disclose performance information voluntarily. These voluntary disclosures are named *management disclosures* in the literature (Healy and Palepu 2001). The literature on management disclosure argues that managers make disclosure decisions to serve their self-interest, and this argument is based on two assumptions, as follows.

The first assumption is about information asymmetries in the financial market. That is, managers have an information advantage relative to investors; managers have more complete information about their firm's economic reality (Healy and Palepu 2001).

The second assumption is that managers observe (at least partially) what investors know.[7] There is evidence that managers announce news about cash flows "to meet investor demand for cash flow information" (Wasley and Wu 2006, p. 391). What is implied is that managers know investors' information demand and what investors already know, so that managers can leverage disclosure to meet investors' information demand. Prior studies also document evidence "consistent with managers being concerned with the risk for litigation and issuing preemptive earnings forecasts to adjust investor expectations" (Wasley and Wu 2006, p. 390). This, again, implies that managers are aware of investor expectations, so that they can use preemptive forecasts to influence investors. The literature on media coverage also assumes that consumers' prior beliefs are not proprietary information; rather, firms know it (Gentzkow and Shapiro 2006).

In particular, it is arguable that managers are aware that investors use Wikipedia to get information about their firms. An article in *Financial Times* reported how Klaus Kleinfeld, the former CEO of Siemens, described his reaction to the growing influence of Wikipedia and argued that "Wikipedia, the peer-produced online encyclopedia, is a popular way for people to gain information about companies and business people."[8] Steve Goodman, CEO of PacketTrap, believed that "if our potential customers want to learn about something, they either go to…Wikipedia, or to Google. And Google search results often lead right back to Wikipedia" (Zetlin 2010). A recent survey, by Webranking, found that 81 percent of surveyed companies cared about their information released to investors via Wikipedia; the survey suggested that

[6] A large body of literature supports this notion (e.g., Abarbanell et al. 1995; Francis et al. 1997; Healy and Palepu 2001; Kasznik and Lev 1995; Lang and Lundholm 1996; Roulstone 2003; Skinner 1994, 1997).

[7] We thank an anonymous reviewer for pointing out this key assumption.

[8] http://blogs.ft.com/businessblog/2008/01/wikipedia-is-pohtml/#axzz1mPeEZxlW.

"companies should definitely have some method for scanning different social media….You have to know what is being said about you."[9]

Management Disclosure

Based on the assumption that managers are aware of what investors know, the literature offers two perspectives to explain why managers choose to withhold or release information.

Withholding Perspective: When there is a high level of information asymmetry between managers and investors, managers tend to suppress or withhold unfavorable information (Dye 2001). A seminal model explains why managers' withholding of bad news hinges on information asymmetry between managers and investors (Dye 1985). The underlying rationale is that withholding bad news may give managers the necessary time to wait for the arrival of more favorable news or to make adjustments to accounting measures. Managers are better placed to do this when firm information is opaque to investors. By contrast, when there is lower information asymmetry between managers and investors, there is less room for managers to manipulate information.[10]

Preemption Perspective: When the market holds an overly optimistic view about firm performance, managers face several risks; to alleviate these risks, managers may choose to release information promptly to adjust market expectation (Healy and Palepu 2001). First, if the market has an overly optimistic expectation about firm performance, firm disclosure on the official announcement date would trigger a downward movement of stock price (Bartov et al. 2002). Managers may be held accountable for poor stock performance.[11] Because market reactions to bad news may be more negative on the official announcement date than on earlier dates, managers may use timely voluntary disclosure to reduce the likelihood of undervaluation and to explain away poor earnings performance (Brennan 1999; Healy and Palepu 2001). Second, managers may be subject to litigation risks. If delaying bad news until official announcement dates results in large stock-price declines, stockholders may sue, because they can allege that managers failed to disclose adverse news promptly. Given the time lag between the end of the fiscal quarter and the official announcement date, it is difficult for managers to argue credibly that they had no information about the bad news beforehand (Skinner 1994). As such, managers may resort to voluntary disclosures in order to reduce the likelihood of stockholder lawsuits. Third, the investment community (e.g., money managers, security analysts, trading institutions, etc.) dislikes negative earnings surprises and prefers firms whose managers are candid about potential earnings problems (Skinner 1994). Withholding bad news may render managers a bad reputation for failing to release adverse information in a timely manner. The investment community has long appeared to impose costs on firms when their managers have such an undesirable reputation (King 1988; Rose 1991). For example, their stocks are less likely to be followed by analysts and traded by money managers, resulting in reduced liquidity (Roulstone 2003). Candid and timely disclosure helps mitigate these negative consequences.

Information Aggregation Mechanisms

From each perspective above, management disclosure is conditioned on investors' information. As mentioned in the "Introduction," investors' information environment consists of two major sources: One is *quantified* information provided by financial analysts, who collect information from various sources, evaluate firms' current performance, and make quantitative forecasts about future firm profitability (Roulstone 2003). The other is *qualitative* information provided by media—that is, linguistic descriptions of firms' current and future profit-generating activities. Qualitative information sources include traditional media, such as individual journalists and the business press (Frankel and Li 2004; Mitchell and Mulherin 1994), and, more recently, social media, which plays an increasingly important role in disseminating firm information on the Internet (Gu et al. 2007; Tetlock et al. 2008).

[9]http://www.webranking.eu/Articles/Articles/2008/New-HH-Webranking-report/.

[10]Dye's (1985) model describes that, when investors can correctly infer management withholding of unfavorable information, investors will revise the stock price downward, and a downward price change signals to the market that the firm's value has been overestimated, which may trigger the stock price to cascade further downward. But if investors are unsure about whether a manager has received any news, the manager can withhold bad news, because investors cannot tell whether the manager has received news (but chosen not to release it). In another case, if investors know that the manager has received news but they do not know the rest of the manager's information set, then, the manager can still choose nondisclosure of bad news, because investors are unable to assess whether the stock price in the market is overestimated or underestimated (Dye 1985; Jung and Kwon 1988). In both cases of information asymmetry, withholding bad news would not induce the firm's stock price to plummet, and thus managers are tempted to withhold bad news.

[11]Prior studies find poor stock performance to be associated with management changes (Dahya et al. 2002; Warner et al. 1988; Weisbach 1988), as well as hostile takeovers, which in turn result in high CEO turnover (Franks and Mayer 1996; Morck et al. 1990).

With these various information sources, it is important to examine mechanisms that can help aggregate information for investors.[12] Prior studies suggest several mechanisms.

First, given the quantitative nature of analyst forecasts about firm earnings, it is convenient to combine analyst forecasts and create an aggregate measure (e.g., average analyst forecasts; Francis et al. 1997; Karamanou and Vafeas 2005; Lang and Lundholm 1996; Roulstone 2003). Such aggregation, however, may still result in inaccurate and incomplete information for investors. Analysts' forecasts are often biased in that they may overestimate a firm's earnings, resulting in an upward analyst bias (Jacob et al. 1999; Karamanou and Vafeas 2005). Analysts have incentives to make optimistic forecasts because they may be rewarded for providing information that generates trading volume and investment banking fees for their brokerage houses (Dechow et al. 2000; Lin and McNichols 1998). Also, analysts' summative earnings forecasts cannot entirely convey the limitless variety of firms' financial standings (Tetlock et al. 2008). Investors, however, need detailed, qualitative information about various aspects of firm operations in order to assess firm value, according to the "mosaic" perspective. The disclosure literature documents evidence in support of this mosaic view (e.g., Francis et al. 1997; Lang and Lundholm 1996).[13] We thus need to address mechanisms that can aggregate qualitative information for investors.

An approach to aggregate qualitative information is to store and present news articles and investor discussions in a unified depository. Mitchell and Mulherin (1994) find that the number of news articles reported by Dow Jones & Company is related to trading volume in the market, suggesting that investors become informed and thus buy and sell stocks. Antweiler and Frank (2004) find that stock-market discussions posted on Yahoo! Finance help predict market volatility. The huge number of articles, however, may give investors information overload. Also, a large proportion of online discussions may have repetitive content or even "noise"; as a result, investors bear high costs of information processing on online message boards (Gu et al. 2007).

One way to eliminate noise of online postings is through users' collective reporting. Gu et al. (2007, p. 74) examined virtual investing-related communities (VICs):

> One approach VIC providers use to improve posting quality is to actively monitor and filter low-quality postings… allowing users to report abusive postings, which are then investigated manually by VIC providers.

Another way to eliminate abusive postings is to have companies select and present relevant news articles in the "investor relations" section on corporate websites (Geerings et al. 2003). These approaches are controlled by VICs or companies, however, so information aggregation may be subject to their selection bias. For instance, companies behind some online communities may choose not to engage in noise filtering (Gu et al. 2007).

In Appendix B, we summarize these mechanisms as documented in previous research. In the next section, we explain how information aggregation on Wikipedia, the focus of our research, differs from these previous mechanisms.

Research Framework and Hypotheses

We develop a research framework for Wikipedia's information aggregation and managers' voluntary disclosure (Figure 1). The framework closely follows the premise of the disclosure literature that managers' decisions of voluntary disclosure are based on investors' information environment. This study moves one step forward by analyzing the additional value of Wikipedia's information aggregation in investors' information set.

Timing of Voluntary Disclosure

The dependent variable of our framework (Figure 1) is the timing of voluntary disclosure. Disclosure timing is determined by the tension between withholding information and releasing it in a timely manner. Following the literature, we

[12]We thank an anonymous reviewer for pointing this out and for suggesting a relevant literature review.

[13]For instance, the Financial Analysts Federation (FAF) Corporate Information Committee clearly states that FAF's rating of information transparency between a firm and investors involves evaluating aspects other than disclosing the summative earnings-per-share. Such additional factors include the clarity and candor of the financial highlights and president's letter, the amount of detail provided about corporate officers, the corporation's goals, and product and geographic segments, etc. (Lang and Lundholm 1996). Another study on corporate presentations to the New York Society of Securities Analysts (NYSSA) argues that market participants may view the presentations as opportunities to seek out "qualitative information about the firm's management, strategies and prospects. Such disclosures, when combined with existing information, could lead to significant market reactions even though the disclosures are not material if viewed in isolation" (Francis et al. 1997, p. 367).

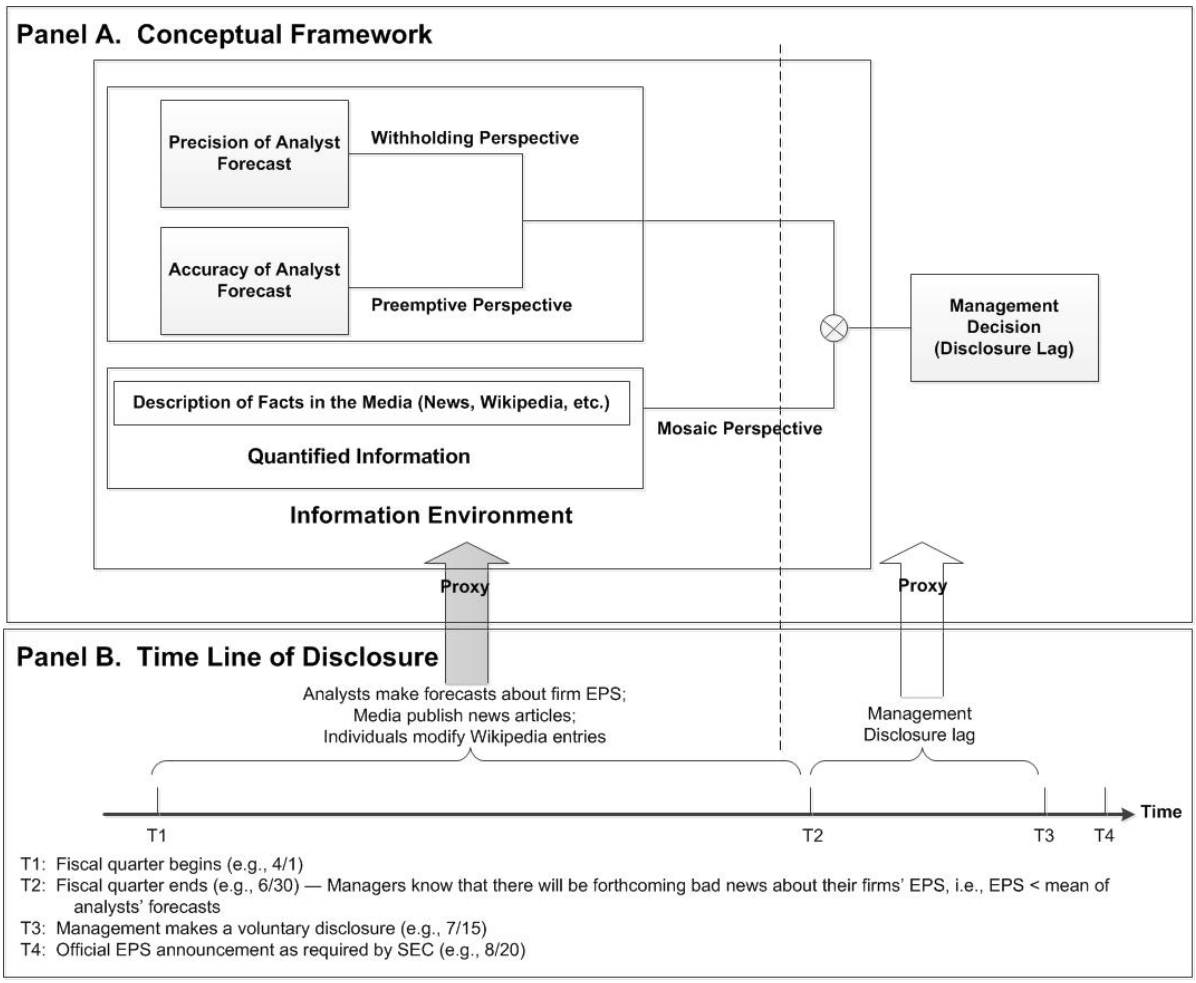


Figure 1. Research Design

restrict our attention to the disclosure of earnings performance (among others, Anilowski et al. 2007; Baginski et al. 1994; Cohen et al. 2007; Kasznki and Lev 1995; Miller 2002; Skinner 1994, 1997). We focus on voluntary disclosure and exclude mandatory disclosure, because the latter type is required by the SEC and thus may not be attributable to features of the information environment. Following the literature, we examine disclosure timing of bad news, that is, when forthcoming earnings are below market expectations (e.g., Baginski et al. 1994; Dye 2001; Healy and Palepu 2001; Skinner 1994, 1997; Wasley and Wu 2006). Disclosure timing of bad news can be leveraged because of the following considerations. First, managers generally have an incentive to withhold unfavorable information (Dye 1985, 2001). Second, according to the preemption perspective, it is a firm's poor rather than good earnings performance that may result in management turnover and damage the firm's relationship with the professional investment community (Kasznki and Lev 1995; Skinner 1994, 1997). In particular, litigation risks are attributable to bad rather than good news. Based on a review of prior studies on litigation cases, Skinner (1994) summarizes that more than 95 percent of firms that are targets of earnings-related shareholder lawsuits can be classified as having bad news on official earnings-announcement dates.[14] Therefore, managers face a crucial decision of whether to

[14]Skinner (1994, p. 42) explains that "The legal reasons for this asymmetry [i.e., the dominant fraction of lawsuits related to bad news] appear to relate to proof of damages and the need to show a sufficient causal connection between the plaintiff's injury and the wrongful conduct. If an investor purchases a stock whose price subsequently declines, it is relatively easy to show both an out-of-pocket loss and (if the decline is accompanied by the release of information) causation. Conversely, a plaintiff who sells before good news is revealed suffers an opportunity loss and must show that he or she would not have sold had the information been available."

release unfavorable news promptly prior to the mandatory earnings announcement.

Following this line of studies, the dependent variable in our framework (Figure 1) is *disclosure lag*, defined as the number of days between the end of a fiscal quarter and the date when managers voluntarily disclose bad news about earnings performance. The shorter the disclosure lag, the more timely is the voluntary disclosure of adverse information.

Quantified Information

Prior studies on investors' information environment use two variables to characterize analysts' quantified forecasts for firm earnings: dispersion and bias (Francis et al. 1997; Karamanou and Vafeas 2005; Lang and Lundholm 1996; Roulstone 2003).

The literature uses *analyst dispersion* to proxy for information asymmetry in the information environment (Abarbanell et al. 1995; Roulstone 2003). Analyst dispersion refers to the degree to which analysts' opinions are in disagreement. The withholding perspective suggests that when analysts cannot agree with each other, investors, who obtain information from analysts, are unsure about firm operations and performance. This gives managers an information advantage; that is, managers can keep silent (about their firm's bad earnings performance) for a longer time if the market is more uncertain about the information that they have. Prior studies in the disclosure literature suggest that when there is greater information asymmetry in the market, managers can delay their voluntary disclosure for a longer time, waiting for investors to acquire information and achieve a consensus about firm performance (Francis et al. 2008; Miller 2002; Wasley and Wu 2006). Greater information asymmetry generally gives managers more time to adjust accounting information to suppress bad news (Matsumoto 2002; Richardson et al. 2004). In sum, the withholding perspective suggests that *disclosure lag is positively related to analyst dispersion.*

The literature uses *analyst bias* to measure the difference between true firm performance and the quantified information accessible to investors (Jacob et al. 1999; Karamanou and Vafeas 2005). According to the preemption perspective, the greater the analyst bias, the larger is the negative earnings surprise on the official earnings announcement date (Kasznik and Lev 1995). The negative earnings surprise exposes managers to risks of job changes, reputation damage, and stockholder lawsuits, as reviewed above. Early disclosure is probably the best strategy to reduce these risks for three reasons (Skinner 1994): (1) only investors who bought or sold the firm's stock *before* management disclosure can sue, simply because investors who conduct transactions after the disclosure are already aware of the bad news; (2) early disclosure undercuts plaintiffs' argument that managers failed to disclose promptly, because managers made a relatively timely disclosure; and (3) early disclosure helps spread out (negative) market reactions over a longer time, which helps prevent the stock price from plummeting and thus helps mitigate the risk of management changes because of poor stock performance (Kasznik and Lev 1995). In sum, the preemption perspective suggests that *disclosure lag is negatively related to analyst bias*.

Qualitative Information

Prior literature characterizes media information using the *number of news articles* and the amount of *newsworthy content* (Antweiler and Frank 2004; Atiase 1985; Foster 1987; Mitchell and Mulherin 1994; Tetlock et al. 2008). We follow the literature and address these factors in our research. Since the focus of this study is the role of Wikipedia's information aggregation, they are posited as control variables in our model. We discuss their effects in detail later in the subsection "Empirical Model for Management Disclosure."

As defined earlier, *information aggregation on Wikipedia* refers to a process by which individuals synthesize their information by contributing to Wikipedia's company entries.

How Does Wikipedia Aggregate Information?

Earlier we summarized the various mechanisms of information aggregation as documented in prior research. Wikipedia differs from previous mechanisms of information aggregation in that it can better generate the "wisdom of crowds" (see Appendix B). Surowiecki (2004) suggests that a social media platform must satisfy four conditions for a crowd to be smart: diversity, independence, decentralization, and aggregation. Wikipedia meets these conditions.

Creating and revising any entry on Wikipedia is completely ***decentralized***. It allows anyone to freely edit any details of all entries (Kane and Fichman 2009; Te'eni 2009). Wikipedia thus supports "democratization of contributions (e.g., leveling the playing field so anyone can contribute an idea)" (Majchrzak 2009, p. 19; Zwass 2010). This guarantees ***diversified*** sources of contributors, in contrast to corporate websites or some communities that control information content (Geerings et al. 2003). Wikipedia ensures ***independent*** expression of

users' opinions and supports "deep profiling," allowing users to provide detailed information to define their own personal and social identities (Zhang and Zhu 2011). This encourages users to identify with Wikipedia and motivates them to contribute (Ma and Agarwal 2007). Contributors use the wiki technology to edit the same pages, thus effectively ***aggregating*** information (Kane and Fichman 2009) and reducing readers' information-processing costs (Gu et al. 2007). Wikipedia also provides functions to easily search and display other users' inputs. This leads to collaborative efforts to screen noise and eliminate abusive postings (Moon and Sproull 2008). The difference between a wiki and a traditional discussion board is that good content is retained on a wiki, and openness and transparency make a wiki naturally resistant to spam (Wagner and Majchrzak 2006).

A relevant literature examines the editing behavior of Wikipedia contributors. There is evidence that individuals' contributing patterns are often determined by their personalities and thus do not change much over time (Panciera et al. 2009). Some prominent Wikipedians argue that a group of prolific users is the driving force behind the success of Wikipedia (Wales 2005), because some contributors tend to express ownership of entries in the collaborative authoring (Thom-Santelli et al. 2009). Recent research, however, suggests a dramatic shift such that the collective contributions play an increasingly bigger role (Kittur et al. 2008). Not only quality evolves out of coordination (Kittur and Kraut 2008), but collaborative efforts (Swarts 2009), diversity (Chen et al. 2010), and conflicts (Arazy et al. 2011; Kittur and Kraut 2010) may also determine the success of Wikipedia. Collaborative authoring thus enables the aggregation of diverse, and even conflicting, opinions from individual contributors.

Nature of Company Information on Wikipedia

Wikipedia describes firms' basic profiles. Additionally, and importantly, Wikipedia aggregates comprehensive and timely information about firms' up-to-date activities.

Wikipedia aggregates a *comprehensive array of details* about a firm's activities, such as launching new products, changes in top management, outsourcing, etc. This type of information aggregation contrasts with the summative earnings forecasts issued by analysts (Carlson and Zmud 1999). Wikipedia also covers detailed firm operations more broadly than any single medium channel, because it aggregates information from various media. On March 23, 2009, for example, a modification on Dell's Wikipedia page suggested that Dell sold a call center to a French outsourcing firm, Teleperformance. Only a weekly magazine, *BusinessWorld*, reported this deal. In this case, traditional information channels may leave information holes between parties with asymmetric and incomplete information (Lin et al. 2005; Schultze and Leidner 2002), while Wikipedia helps reduce information asymmetry by keeping a comprehensive record of facts.

Wikipedia aggregates information in a *timely manner*. Wikipedia entries for Microsoft and Dell, for example, are edited multiple times a day, which is more frequent than how often analysts update their coverage of these companies. On June 1, 2009, almost immediately after the Dow Jones' announcement that Cisco was going to replace GM as a component of the Dow Jones Index, the Wikipedia entry for Cisco was modified to reflect it. The earliest newspaper coverage, by the *Washington Post*, came as late as June 2. These cases demonstrate that important information about firm fundamentals becomes public on Wikipedia very quickly.

Next we proceed to analyzing the impact of Wikipedia's information aggregation.

Wikipedia and Management Disclosure

Much prior research considers the impact of information aggregation on market price and returns. Classical asset-pricing models assume perfect information aggregation in the financial market (Banks 1985; Radner 1979). Diamond and Verrecchia's (1981) model shows that when the process of information aggregation via market pricing is not perfect, market outcomes can be different, and Axelson (2007) theoretically shows that managers should react (in the context of securities design) to different levels of investor private information. Forsythe and Lundholm (1990) and Sunder (1992) examine financial information aggregation with experiments and show evidence that various information aggregation mechanisms affect stock prices. DellaVigna and Pollet (2009) show that information aggregation eliminates the effect of limited attention, thus affecting stock returns. Collectively, these prior studies demonstrate the impact of information aggregation on investors. An important implication for our research is that information aggregation influences investors' information set.

We carry this implication to the setting of information aggregation on Wikipedia. We have offered some anecdotal evidence (in the "Introduction") suggesting that Wikipedia can benefit investors. In the same vein, Gu et al. (2007) suggest that investors value high-quality linguistic descriptions about firms. Butler (2001) argues that the key driver of participation in collaborative communities is that participants can gain

benefits from communicating and collaborating with others, and obtain information through information aggregation. Antweiler and Frank (2004) show that the aggregated sentiment of stock market discussions on Yahoo helps predict market volatility, even after controlling for news in the *Wall Street Journal*. This suggests that aggregation of linguistic descriptions informs investors. The aggregation in Antweiler and Frank, however, is based on computational methods for scholarly investigation. By contrast, Wikipedia offers a practical platform that individuals in the market use to aggregate and access qualitative information.

Our literature review motivates us to analyze how Wikipedia's information aggregation, by reshaping investors' information set, may alter management information-disclosure behaviors. Our overall logic is that, in the absence of Wikipedia, management disclosure is related to the information environment characterized by analyst dispersion and analyst bias (recall our discussions in a previous subsection, "Quantified Information"). If Wikipedia reshapes investors' information environment, then we expect different disclosure behaviors even though analyst characteristics (dispersion and bias) are the same. This expectation suggests a moderating effect of Wikipedia on the relationship between management disclosure and analyst characteristics.

According to the withholding perspective, when analyst dispersion is high, managers withhold information for a longer period (i.e., a positive relationship between disclosure lag and analyst dispersion). In this case, managers clearly have an information advantage, and their delay in disclosure indicates that they can benefit from information uncertainty in the market. Wikipedia can weaken the information advantage of managers, mainly because Wikipedia is a public platform to effectively aggregate private information possessed by individuals in the market. A high level of analyst dispersion indicates that, while managers generally have a complete set of information about their firm's economic activities, the set becomes incomplete and piecemeal when it goes to the separate information sources in the market (i.e., analysts) (Roulstone 2003). Individuals obtain their private information from analysts and, therefore, their information is a range of dispersed bits of incomplete and frequently contradictory messages. The democratization of the creation, distribution, and use of information on Wikipedia enables individuals to assemble their pieces of information into a unified whole. In doing so, information aggregation on Wikipedia would play a role in weakening managers' information advantage. If this is the case, managers would be less able to withhold unfavorable information even though there is high analyst dispersion. As such, the positive relationship between management disclosure lag and analyst dispersion would be weakened. We thus hypothesize:

> ***Hypothesis 1 (H1)****: Information aggregation on Wikipedia weakens the relationship between analyst dispersion and management disclosure lag.*

According to the preemption perspective, when analyst bias is high, managers' disclosure lag will be relatively shorter (i.e., a negative relationship between disclosure lag and analyst bias). This is because managers are concerned about the market's overly optimistic expectation of firm earnings. Wikipedia can help establish a more accurate expectation of firm earnings for investors in two ways. First, when investors are better informed through Wikipedia's information aggregation, they possess more comprehensive and timely information about firms' fundamentals. As a result, they can better evaluate the firm's true performance. Second, Wikipedia supports a neutral point of view (NPOV) stance (Majchrzak 2009). The NPOV policy explicitly forbids either avoiding or highlighting favorable or unfavorable facts.[15] Wikipedia encourages all users to comply with the NPOV policy and allows anyone to modify a firm's entry through rounds of addition, deletion, and reorganization. In such an information aggregation process, contributors are able to correct any details of an entry that violate the NPOV policy. Because this policy requires contributors to use neutral words and to only document facts, information aggregation on Wikipedia prevents one point of view from dominating. This is markedly different from analysts' recommendations, which are often biased (Jacob et al. 1999; Karamanou and Vafeas 2005). Traditional media also have a more emotional touch than Wikipedia entries, be it for ideological (Mullainathan and Shleifer 2005) or financial (Reuter and Zitzewitz 2006) reasons. When the market is replete with an overly optimistic sentiment, the NPOV policy can help investors obtain a more objective assessment of firm performance. Overall, information aggregation on Wikipedia should help alleviate managers' concerns about high analyst bias, thus weakening the negative relationship between management disclosure lag and analyst bias. Formally, we can write the following hypothesis:

> ***Hypothesis 2 (H2)****: Information aggregation on Wikipedia weakens the relationship between analyst bias and management disclosure lag.*

Investor Reaction to Disclosure

Now we use the lens of investor reaction to see how Wikipedia benefits investors. The disclosure literature has long

[15] http://en.wikipedia.org/wiki/NPOV.

observed investor reaction to management disclosure of firm earnings (Kasznik and Lev 1995; Miller 2002; Pownall and Waymire 1989; Waymire 1984). A literature review concludes that "disclosure is associated with stock price performance" (Healy and Palepu 2001, p. 431). In this line of research, Kasznik and Lev (1995) show a significant association between market reaction and unexpected earnings surprise to investors (because of analyst bias). The higher the bias, the more negative is the market reaction. Their study also suggests that if investors receive warnings before management disclosure of bad news, the disclosure is less of a surprise to investors. In the same vein, we propose that when investors obtain detailed and timely information about firms' economic reality, they adjust their firm valuation accordingly. Later, when managers release bad news about firm earnings, it will trigger a smaller market reaction because of the earlier market adjustment. We, therefore, hypothesize that information aggregation on Wikipedia, which occurs before management disclosure (see Figure 1), would weaken how analyst bias triggers investor reaction.[16]

> ***Hypothesis 3 (H3)****: Information aggregation on Wikipedia weakens the relationship between analyst bias and investor reaction to management disclosure.*

Method

Measures

To develop our measures for variables in the research framework (Panel A of Figure 1), we follow the disclosure literature and present a time line of disclosures in Panel B of Figure 1 (Skinner 1997). In a fiscal quarter (between T1 and T2), analysts make forecasts about firm earnings-per-share (EPS), media publish news articles, and contributors modify firm entries on Wikipedia. The disclosure literature assumes that managers know their own firms' EPS at fiscal quarter end T2, when they realize whether they are confronting forthcoming bad news about EPS (Skinner 1997). The SEC requires public companies to announce their EPS at T4 (the official announcement date). During the period between T2 and T4, managers may voluntarily make disclosures about their firms' EPS (e.g., at T3).

Consistent with the literature (Kasznik and Lev 1995;Skinner 1994), we measure disclosure lag (LAG) as the number of calendar days between fiscal quarter end (T2) and the date when management voluntarily discloses bad news about EPS (T3).

Analyst bias (BIAS) is the difference between the mean of analyst forecasts of a firm's EPS during a fiscal quarter (between T1 and T2) and the firm's actual EPS (Francis et al. 1997;Karamanou and Vafeas 2005). Analyst dispersion (DISPERSION) is the standard deviation of analyst forecasts of a firm's EPS during a fiscal quarter (between T1 and T2) (Abarbanell et al. 1995; Roulstone 2003).

We measure information aggregation on Wikipedia by Wikipedia modifications (WikiMOD)—specifically, the number of times modifications are made on Wikipedia about a firm in one quarter (between T1 and T2). WikiMOD is a proxy for information aggregation, because by modifying the same firm entry on Wikipedia, individuals add their information, combine different viewpoints about the firm's operations and performance, and filter out noise and biased statements, thus generating synthesized information about the firm. In order for WikiMOD to play a role in our analysis, a premise is that managers can be informed about the magnitude of WikiMOD, that is, the extent of information aggregation about their firms. This premise is likely to hold. For each company's Wikipedia entry, there is a "revision history" page on Wikipedia, listing information about each historical modification on that entry (including the time of each modification and the content modified). For example, if one wants to check the modification history of IBM's Wikipedia entry, s/he can easily see the entire history (http://en.wikipedia.org/w/index.php?title=IBM&action=history) with no need for any special technical skills. In addition, managers can use some publicly available tools, like "Wikipedia Page History Statistics," which provide the number of modifications for any single Wikipedia entry in the last day/week/month. As an illustration, Appendix C shows the summary statistics for Apple's Wikipedia page modifications. Using such tools, it may not be too difficult for managers to track and thus realize how many changes were made on their company's Wikipedia entry. Finally, although our estimation uses the number of modifications as a proxy for information arrivals on Wikipedia, managers do not really need to know the number of modifications of their Wikipedia entries in order to assess how information gets leaked through Wikipedia. A quick skimming of the whole page is likely to give a manager a very good idea about the transparency of the information environment.

It is possible that not all modifications are related to a company's financial performance (e.g., modifications related to word choice, grammar mistakes, etc.). We refrain from removing these unrelated modifications because that would unavoidably involve subjective human judgment. These modifications are likely to be orthogonal to the variables of

[16]We appreciate the Associate Editor's suggestion to formally putting forth this hypothesis.

interest. Hence, including these modifications would not affect the direction of our estimation results, while reducing the efficiency of the estimation. As such, the introduction of noise would bias against our finding evidence. Another potential concern is that WikiMOD may involve modifications in "edit wars" (i.e., back-and-forth changes due to disagreement between contributors). As reported later in our sensitive analysis in Appendix F, our results hold up after we remove possible back-and-forth changes.[17] We also note that if there are other social-media platforms that can achieve information aggregation, focusing on Wikipedia is only likely to bias our results downward. Therefore, our conclusion regarding the effect of information aggregation would be conservative.

As for investor reaction to bad news, we follow the disclosure literature (Kasznik and Lev 1995) and develop a proxy measure by computing the sum of market-adjusted returns in two windows: a five-day window around voluntary disclosure and a five-day window around the official EPS announcement. Market-adjusted returns (RET) in the combined two windows represent market reaction to bad news. In addition, we estimate cumulative abnormal returns (CAR) in the combined two windows.[18] These two methods yield highly consistent results, so the discussion below is based on one method (CAR).

Empirical Model for Management Disclosure

To test H1 and H2, we use a hazard regression model, developed in the statistics literature to assess the impact of explanatory variables on the timing of an event. In our research context, the event under investigation is management transition from withholding to disclosure, and the timing of the event refers to the disclosure lag (LAG). Because the distribution of event timing is often far from normal, a hazard model is usually superior to ordinary least squares regression (Kalbfleisch and Prentice 1980). The dependent variable of the hazard analysis is the hazard rate $h(t)$, which is the probability of management transition from withholding to disclosure at time t. Following Kauffman et al. (2000), we specify a Cox proportional-hazard model as follows:

$$h(t) = h_0(t)\exp(-\beta X) \quad (0)$$

where X is a vector of explanatory variables, β is a row vector of coefficients to be estimated, and $h_0(t)$ is a hazard function with $X = 0$. It is worth noting that we specify $\exp(-\beta X)$ rather than $\exp(\beta X)$, because with this specification, the coefficient β tells how the explanatory variables X affect the disclosure lag as follows: if an explanatory variable has a ***positive*** coefficient, it reduces the hazard rate $h(t)$ and thus ***increases*** LAG; conversely, a ***negative*** coefficient indicates an effect to ***decrease*** LAG. In the literature, Kauffman et al. use a similar approach for exposition reasons.

Because H1 and H2 are about whether WikiMOD moderates how LAG relates to analyst characteristics (DISPERSION and BIAS), the explanatory variables X in model (0) include DISPERSION, BIAS, WikiMOD, WikiMOD*DISPERSION, and WikiMOD*BIAS. We have the following hazard model for testing H1 and H2:

$$\begin{aligned} h(t) = \; & h_0(t)\exp[-(\theta_1 DISPERSION + \theta_2 BIAS + \delta_1 WikiMOD \\ & + \delta_2 WikiMOD*DISPERSION + \delta_3 WikiMOD*BIAS \\ & + \xi\,\text{controls})] \end{aligned} \quad (1)$$

We need two sets of control variables. First, we need to include controls identified by the disclosure literature as directly affecting management disclosure:

- We control for firm risk, measured as earnings variability (VAR)—that is, the standard deviation of earnings across the eight previous fiscal quarters (Kothari et al. 2002). To compensate for high risk resulting from high earnings variability, investors tend to demand high returns that increase the firm's costs of capital. Voluntary disclosure may help reduce costs of capital by providing more information to investors (Francis et al. 2008). As such, high earnings variability may motivate firms to disclose earlier.

- The literature suggests the need to control for firm size (MV), measured as firm market value at the beginning of a fiscal quarter, because larger firms are more exposed to litigation risks (Kasznik and Lev 1995).

- The literature suggests the need to control for two specific industry sectors: high tech (HIGHTECH) and regulated industry (REG). High-tech firms are likely to be exposed to a larger-than-average risk of shareholder lawsuits and thus may be motivated to disclose to deter investors' litigation (Kasznik and Lev 1995).[19] Yet,

[17]We thank an anonymous reviewer for suggesting this important check.

[18]We estimate an equation $R_t = a + b * R_{mt}$, where R_t is a firm's actual return on day t and R_{mt} is the market return. The difference between R_t and the predicted value based on the equation is the firm's abnormal returns on day t.

[19]Kasznik and Lev (1995, p. 124) argue "high technology ('high tech') firms appear to be exposed to a larger-than-average risk of shareholder lawsuits, particularly at the early stage of operations. Among the reasons for the

among the reasons for shareholder lawsuits are high-tech firms' aggressive accounting techniques (e.g., excessive capitalization of software development costs). Disclosing poor EPS may not help fend off such lawsuits (Kasznik and Lev 1995). In regulated industries, firms may be obliged to release information about their operations, in addition to quarterly financial reports. But firms in regulated industries may be perceived as having low risk (Dewan et al. 2007), which may reduce firms' incentive to disclose. We thus include the two industry dummies (HIGHTECH and REG) but do not predict their signs.

- We also control for fiscal quarters. Managers are more likely to withhold bad news in the fourth quarter than in earlier quarters (Baginski et al. 1994), as releasing bad news in the last fiscal quarter affects firm performance for the entire fiscal year.

Second, we need to control for news coverage, because modifications on Wikipedia and news coverage may be triggered by common events. If news coverage also informs investors, we need to address their moderation effects in model (1) as well. This helps rule out alternative explanations for the moderating role of Wikipedia modifications.

- The recent literature on media impact suggests that readers are sensitive to newsworthy content, which can be measured with the amount of positive words (POSITIVE) and negative words (NEGATIVE) in news articles (Tetlock et al. 2008). Given the same level of dispersion and uncertainty, an increased level of positive news affords managers the possibility to take advantage of it, causing further delay in disclosure.[20] We therefore expect a positive interaction effect between POSITIVE and DISPERSION on disclosure lag; conversely, we expect a negative interaction effect between NEGATIVE and DISPERSION. Regarding the moderation on BIAS, because a positive media bias means that EPS is already below expectation, negative news may further hasten managers to disclose bad news to minimize damage. We therefore expect a negative interaction effect between NEGATIVE and BIAS on disclosure lag; conversely, we expect a positive interaction effect between POSITIVE and BIAS.

- We count the number of news articles (NEWS) in the business press about a firm in one quarter (Antweiler and Frank 2004; Frankel and Li 2004). We expect the role of NEWS is similar to that of NEGATIVE, because prior research finds that readers are generally more affected by negative than positive words (Tetlock et al. 2008).

Empirical Model for Investor Reaction

To test H3, our analysis builds on the management disclosure literature (Kasznik and Lev 1995), which has established a regression model relating analyst bias (BIAS) to investor reaction (i.e., cumulative abnormal returns—CAR—as defined above). We extend the model by adding the moderating effect of WikiMOD, and have an ordinary least squares (OLS) model as follows:

$$CAR = \alpha + \eta\, BIAS + \delta_1\, WikiMOD + \delta_2\, WikiMOD*BIAS + \xi\, \text{controls} + \varepsilon \quad (2)$$

Regarding controls in the above model, we follow prior research and include firm size, which moderates the effect of BIAS, in that investors are more tolerant of analyst bias for large firms than for small firms (Kasznik and Lev 1995). The reason is that poor performance in one quarter may have less impact on large firms' long-term economic viability as compared with that of small firms (Kasznik and Lev 1995). We also include news articles and newsworthy content, for the abovementioned reasons. If negative news content can inform investors about forthcoming bad news, it can mitigate investor reaction to bad news, while positive news content may play an opposite role. Regarding the number of news articles, we expect that its role is similar to that of negative news (Tetlock et al. 2008).

Data and Sample

We collect data from several sources: (1) management disclosure data from the First Call Historical Database (FCHD), (2) editing history data from Wikipedia, (3) firm data from Compustat and Center for Research in Security Prices (CRSP), and (4) news-coverage data from Lexis-Nexis. Table 1 summarizes the measures and data sources of all variables, which are described in detail below.

---

prevalence of shareholder lawsuits against high tech firms is their relatively high risk, resulting in large price fluctuations and potential losses to investors. The aggressive accounting techniques sometimes used by such firms (e.g., front loading of gains from long-term contracts, excessive capitalization of software development costs) may also contribute to litigation exposure." Along the same line, Chen et al. (2002, p. 232) argue "high-tech firms also operate in rapidly changing environments that make their future operations, and hence future earnings, relatively more uncertain."

[20]We thank an anonymous reviewer for providing insights into how newsworthy content may interact with analyst dispersion and analyst bias in affecting management disclosure lag.

Table 1. Variable Definitions

Variable	Computation	Source
Theoretical Variables		
LAG (disclosure lag)	The number of calendar days between fiscal quarter end and the date when management voluntarily discloses bad news about EPS (i.e., between T2 and T3 in Figure 1)	First Call
DISPERSION (analyst dispersion)	The standard deviation of analyst forecasts about a firm's EPS in a fiscal quarter (i.e., between T1 and T2 in Figure 1), scaled by the mean EPS forecast	First Call
BIAS (analyst bias)	The absolute value of the difference between firm EPS in a quarter and the mean EPS forecast made by analysts in the quarter (i.e., between T1 and T2 in Figure 1), scaled by the mean EPS forecast	First Call
WikiMOD (Wikipedia modifications— a proxy for information aggregation on Wikipedia)	Natural logarithm of one plus the number of Wikipedia modifications for a company in a fiscal quarter (between T1 and T2 in Figure 1)	Wikipedia
CAR (cumulative abnormal returns)	Sum of abnormal returns in a combined window including 5 days around voluntary disclosure (i.e., T3 in Figure 1) and 5 days around the official earnings announcement (i.e., T4 in Figure 1)	CRSP
RET (market adjusted returns)	Sum of market adjusted returns (i.e., less market average returns) in a combined window including 5 days around voluntary disclosure (i.e., T3 in Figure 1) and 5 days around the official earnings announcement (i.e., T4 in Figure 1)	CRSP
Control Variables		
NEWS (number of news articles)	Natural logarithm of one plus the number of news articles about a company in a fiscal quarter (between T1 and T2 in Figure 1), as documented in Lexis-Nexis	Lexis-Nexis
POSITIVE (positive news)	Standardized proportion of positive words. Proportion of positive words (*POS*%) is calculated as total number of positive words for a company in a fiscal quarter divided by total number of words for that company in that quarter. We calculate the mean (*Mean_POS*) and the standard deviation (*Sd_POS*) of the proportion of positive words for the prior calendar year, *POSITIVE*=(*POS*%-*Mean_POS*)/*Sd_POS*.	Lexis-Nexis
NEGATIVE (negative news)	*NEGATIVE* is calculated similar to *POSITIVE*, with the proportion of negative words.	Lexis-Nexis
VAR (earnings variability)	The standard deviation of quarterly earnings across eight fiscal quarters before the quarter under examination (i.e., prior to T1 in Figure 1)	Compustat
MV (market value)	Natural logarithm of firm market value at the beginning of the fiscal quarter (i.e., at T1 in Figure 1)	Compustat
HIGHTECH (high-tech industries)	*HIGHTECH* equals 1 when the firm belongs to Drugs (SIC 2833-2836), R&D Services (8731-8734), Programming (7371-7379), Computers (3570-3577), or Electronics (3600-3674), and 0 otherwise	Compustat
REG (regulated industries)	*REG* equals 1 when the firm belongs to Telephone (SIC 4812-4813), TV (4833), Cable (4841), Communications (4811-4899), Gas (4922-4924), Electricity (4931), Water (4941), or Financial sectors (6021-6023, 6035-6036, 6141, 6311, 6321, 6331), and 0 otherwise	Compustat

We obtain access to FCHD from Wharton Research Data Services (WRDS). FCHD reports the history of analysts' estimates of companies' EPS, based on which we compute analyst bias (BIAS) and dispersion (DISPERSION). FCHD also records company-issued guidelines about forthcoming EPS and labels whether the guidelines are bad news.[21] We use this information to identify bad news and calculate disclosure lag (LAG).

FCHD contains quarterly earnings information for 8,500 U.S. securities; we limit our attention to common stocks. For each stock, we obtain firm identity (provided by FCHD) and then search on Wikipedia for the corresponding entry. A PERL program is used to search our list of companies on Wikipedia. After obtaining the URLs of these entries, we manually go over all of these Wikipedia pages to ensure that the entries are correctly matched to the companies. This practice yields 375 Wikipedia entries of public companies. To obtain Wikipedia revision information, we use a software robot program to collect the complete "revision history" of all these entries.[22] After comparing each revision with the previous revision, we calculate the number of words added or deleted and record the number of modifications (WikiMOD) in each fiscal quarter for each firm. Overall, for the period between March 21, 2001, and May 19, 2006, 8,789 registered users and 5,450 unregistered users contributed a total of 77,921 modifications on these Wikipedia firm entries.[23]

We manually search the Lexis-Nexis database for news articles about each of the 375 companies in our final sample. For Lexis-Nexis's HTML output, we develop a program to parse the result pages and record the newspaper name, date, and content of each piece of news. These are our raw data for quantifying news coverage (NEWS) and analyzing news content. We follow prior research to conduct a content analysis to quantify the language used in news stories. We compute the proportions of positive words (POSITIVE) and negative words (NEGATIVE) in all news stories related to a firm in one quarter (between T1 and T2).[24] The algorithm of our computation comes directly from the literature on stock-market responses to news content (Das and Chen 2007; Tetlock 2007; Tetlock et al. 2008).[25]

We obtain stock-return data from CRSP and use the data to calculate market-adjusted returns (RET) and cumulative abnormal returns (CAR), as defined above. We use firm data from Compustat for earnings variability (VAR), market value (MV), a high-tech industry dummy (HIGHTECH), and a regulated industry dummy (REG).

Our research context adds two necessary restrictions to our sample. First, a firm's EPS in a quarter should be below the mean EPS forecast made by analysts during the same quarter. This implies a situation in which managers faced forthcoming bad news. Second, the firm should voluntarily disclose bad news about EPS (between T2 and T4 in Figure 1). These two restrictions reduce our sample to a set of 161 warnings (i.e., voluntary disclosures of bad news about EPS) released by 96 public companies during the observation period when the Wikipedia modification history is available to us. These 161 warnings form our final sample used in the subsequent analysis.

Summary Statistics

Table 2 reports summary statistics. In the final sample, the mean disclosure lag is 14.11 days, and the maximum is 43

[21] When companies make voluntary earnings disclosures in press releases and interviews, the disclosures can be a number (e.g., $2.35), a range of numbers (e.g., from $2.35 to $2.68), or a simple text item (e.g., the earnings will be above/below analysts' expectations). FCHD reports the form of each specific disclosure and uses a field to describe the disclosure and indicate whether it is a positive or negative surprise. Please refer to FCHD Technical Guide for further details.

[22] A detailed explanation of Wikipedia's revision history can be found at http://en.wikipedia.org/wiki/Help:Page_history.

[23] Wikipedia attributes modifications to users' registered names if contributors log into the system. Anonymous contributors who do not log in are identified only by their IP addresses. We exclude all modifications contributed by anonymous users in our subsequent analysis because, given the low cost of anonymous editing, spammers and abusers often contaminate entries and because knowledge contribution on social media is found to be strongly associated with identity verification (Ma and Agarwal 2007; Wasko and Faraj 2005; Zhang and Zhu 2011).

[24] *NEGATIVE* is a standardized measure of the proportion of negative words in all news stories related to a firm (Tetlock et al. 2008). Specifically, we calculate the proportion of negative words (*NEG*%), the total number of negative words for a firm in a fiscal quarter divided by the total number of words for that firm in the same quarter. We then calculate the mean (μ_{NEG}) and the standard deviation (σ_{NEG}) of the proportion of negative words for the *prior* calendar year, and define $NEGATIVE = (NEG\%-\mu_{NEG})/\sigma_{NEG}$. Similarly, we compute $POSITIVE = (POS\%-\mu_{POS})/\sigma_{POS}$. To categorize the words as positive or negative, we follow Tetlock (2007) and Tetlock et al. (2008) and use the Harvard-IV-4 psychosocial dictionary.

[25] Das and Chen (2007) propose an algorithm to identify investor sentiments from stock message boards. Tetlock (2007) examines the relation between media content, specifically *The Wall Street Journal*'s "Abreast of the Market" column, and daily stock-market activity. He finds evidence to support that the media directly influence investors' sentiments toward securities. Tetlock et al. (2008) find that a quantitative measure of language used in *The Wall Street Journal* and Dow Jones News Service stories can be used to predict individual firms' accounting earnings and stock returns.

Table 2. Summary Statistics (Variable Definitions Are in Table 1)

	LAG	DISPER-SION	BIAS	NEWS	Wiki-MOD	POSITIVE	NEGA-TIVE	VAR	MV	HIGH TECH	REG	CAR	RET
Mean	2.2667	0.3246	1.0083	2.0316	1.5947	–1.0366	0.0491	0.4418	18.1280	0.3625	0.0188	–0.0484	–0.0684
Std. Dev.	0.8765	0.8691	3.0376	1.3455	1.0907	1.6965	1.3778	1.4266	1.2513	0.4822	0.1361	0.1112	0.1126
Bivariate Correlations (p-values shown in parentheses)													
LAG	1												
DISPERSION	0.151 (0.056)	1											
BIAS	0.068 (0.391)	0.499 (0.000)	1										
NEWS	–0.231 (0.003)	–0.155 (0.051)	–0.008 (0.924)	1									
WikiMOD	–0.158 (0.046)	–0.132 (0.096)	–0.051 (0.519)	0.332 (0.000)	1								
POSITIVE	–0.136 (0.087)	0.054 (0.497)	0.056 (0.482)	–0.027 (0.736)	–0.019 (0.816)	1							
NEGATIVE	–0.018 (0.825)	0.086 (0.282)	–0.002 (0.984)	–0.021 (0.796)	–0.061 (0.440)	–0.108 (0.175)	1						
VAR	–0.122 (0.124)	0.220 (0.005)	0.418 (0.000)	0.066 (0.406)	–0.140 (0.077)	0.091 (0.250)	0.030 (0.708)	1					
MV	–0.260 (0.001)	–0.246 (0.002)	–0.303 (0.000)	0.445 (0.000)	0.354 (0.000)	–0.056 (0.484)	–0.128 (0.105)	–0.227 (0.004)	1				
HIGHTECH	0.121 (0.129)	0.272 (0.000)	0.203 (0.010)	–0.267 (0.001)	–0.184 (0.020)	0.045 (0.570)	–0.082 (0.304)	0.224 (0.004)	–0.234 (0.003)	1			
REG	0.026 (0.748)	–0.050 (0.529)	–0.046 (0.563)	0.009 (0.910)	–0.144 (0.069)	0.038 (0.635)	0.007 (0.932)	–0.037 (0.639)	0.016 (0.842)	–0.104 (0.190)	1		
CAR	–0.080 (0.316)	–0.226 (0.004)	–0.250 (0.001)	0.074 (0.351)	0.007 (0.926)	–0.054 (0.501)	–0.056 (0.485)	–0.195 (0.013)	0.215 (0.006)	–0.254 (0.001)	0.004 (0.965)	1	
RET	–0.113 (0.156)	–0.221 (0.005)	–0.259 (0.001)	0.086 (0.278)	0.020 (0.798)	–0.071 (0.373)	–0.053 (0.506)	–0.198 (0.012)	0.225 (0.004)	–0.247 (0.002)	0.007 (0.934)	0.993 (0.000)	1

days. The mean number of modifications for a sample company on Wikipedia is 8.19 in one quarter, and the maximum number is 129. The mean number of news articles about a sample company in one quarter is 15.28, and the maximum is 75. We log-transform these variables because they are skewed.

Both Wikipedia modifications (WikiMOD) and news coverage (NEWS) are positively correlated with firm size (MV), suggesting that larger firms receive more media coverage. Firm size is negatively correlated with earnings variability (VAR) and analyst dispersion (DISPERSION), suggesting that larger firms have less uncertainty in earnings and may release more information to analysts (Kothari et al. 2002; Lang and Lundholm 1996). These correlations together may explain why WikiMOD is negatively correlated with VAR and DISPERSION.

Disclosure lag is negatively correlated with WikiMOD and NEWS, suggesting that information arrivals may result in earlier disclosure. Given the positive correlation between WikiMOD and NEWS, however, this univariate analysis cannot tell us whether LAG is attributable to information aggregation or news arrivals per se, or both. LAG is also negatively correlated with firm size, because large firms are more exposed to litigation risks than small firms (Kasznik and Lev 1995). Hence, large firms are more likely to disclose earlier to immunize themselves from litigation. LAG is positively correlated with analyst dispersion, suggesting that announcements occur later, in the presence of higher information uncertainty in the market.

Regarding analyst forecasts, both analyst bias and analyst dispersion are negatively correlated with firm size and positively correlated with earnings variability and the high-tech dummy. It could be the case that more information is available for larger firms, making analysts' forecasting relatively easier (Lang and Lundholm 1996). In addition, high-tech firms feature greater earnings variability, which may make it more difficult for analysts to predict future EPS (Dewan et al.

2007). The correlation between BIAS and DISPERSION is positive, suggesting that analysts' forecasts are more accurate when they agree more with each other. Overall, the correlations are consistent with observations in previous research.

Results

Results about Management Disclosure

To test H1 and H2, Table 3 reports regression results for model (1). The software package used is R and the function is coxph. From columns (1) through (5), Table 3 presents results when explanatory variables are added step by step. The purpose is to compare different model specifications to check the robustness of the results. We mean-center variables involved in interactions to ease the interpretation of results.

We discuss the results in turn below.

Analyst Forecasts (DISPERSION and BIAS)

We first examine the impact of analyst dispersion and bias. In column (1) of Table 3, the coefficients on DISPERSION and BIAS represent their ***average*** (or *constant*) effects across all values of moderators (Aiken and West 1991). In columns (2) through (5), as we add moderators, the coefficients on DISPERSION and BIAS change in magnitude and significance. This is not surprising because in columns (2) through (5), the coefficients on DISPERSION and BIAS represent their impact ***conditioned*** on the mean values of the moderators (Aiken and West 1991).

As show in column (1) of Table 3, the positive sign of DISPERSION indicates that its average effect is to increase disclosure lag. This confirms the notion that when the external information environment regarding firm EPS is highly dispersed, managers tend to delay disclosure. The negative sign of BIAS indicates that its average effect is to decrease disclosure lag. This supports the notion that managers are more likely to disclose unfavorable information when analysts are far off in estimating firm earnings.

Information Aggregation on Wikipedia (WikiMOD)

Regression results in Table 3 support our hypotheses about information aggregation on Wikipedia (H1 and H2). As seen in Table 3, the results for WikiMOD are consistent across columns (2) through (5). Below we use column (2) to present the results.

H1, supported: The average effect of DISPERSION in column (1) suggests that managers tend to withhold adverse information if they enjoy an information advantage. The interaction with Wikipedia modification (–0.1321, $p < 0.05$, column (2) of Table 3), however, moderates the effect. The interaction's negative sign suggests that, in the presence of more Wikipedia modifications, the extent to which managers leverage their information advantage is weakened. This supports our hypothesis that WikiMOD may mitigate the impact of DISPERSION on disclosure lag, such that the positive relationship between DISPERSION and disclosure lag is less conspicuous in the presence of higher WikiMOD.

H2, supported: The average effect of BIAS in column (1) suggests that when analysts are too optimistic in estimating earnings, managers typically shorten the lag to offer information more promptly. The significant interaction between WikiMOD and BIAS (0.0623, $p < 0.05$, column (2) of Table 3), however, moderates the effect. If the reason behind an earlier release of bad news is managers' concern about large stock-price declines on official announcement dates, then Wikipedia modifications may play the role of releasing information to the public earlier, thereby alleviating that concern. As such, when there are more Wikipedia modifications, managers are not as worried as before. The positive sign on WikiMOD*BIAS supports our hypothesis that WikiMOD may mitigate the impact of BIAS on disclosure lag, such that the negative relationship between BIAS and disclosure lag is less severe in the presence of higher WikiMOD.

Number of News Articles and News Content

As shown in Table 3, the coefficients on NEGATIVE are negative, suggesting that when there are more negative words about a company, managers tend to disclose earlier. This is consistent with our expectation and suggests a plausible "backfire" effect associated with NEGATIVE. Managers in our research context face forthcoming bad news given an upward market bias introduced by analysts. If investors follow analysts first and realize analyst bias later (on the official announcement date), then the inconsistency between market expectation and NEGATIVE would lead investors to suspect that managers had withheld unfavorable information. The inconsistency would backfire, exposing managers to risks resulting from negative earnings surprises.

The interaction between NEGATIVE and BIAS is negative as expected, although the significance level is weak. The interaction between NEGATIVE and DISPERSION is significant and positive, different than our expectation. The positive interaction suggests that NEGATIVE's effect of shortening disclosure lag would be salient when DISPERSION is low.

Table 3. Wikipedia and Management Disclosure (H1 and H2) (Regression specification is based on model (1) in text. Variable definitions are in Table 1.)

	Analysts	Aggregation via Wikipedia	Number of News Articles	Content of News (Positive & Negative Words)	
	(1)	(2)	(3)	(4)	(5)
Theoretical Variables					
Wikipedia Modifications (*WikiMOD*)		–0.1350** (0.0747)	–0.1048* (0.0744)	–0.1255** (0.0731)	–0.1206* (0.0737)
WikiMOD*DISPERSION		**–0.1321** (0.0621)**	**–0.1495*** (0.0580)**	**–0.3480*** (0.1005)**	**–0.3473*** (0.1308)**
WikiMOD*BIAS		**0.0623*** (0.0148)**	**0.0672*** (0.0128)**	**0.1036*** (0.0227)**	**0.0990*** (0.0408)**
Analyst Dispersion (*DISPERSION*)	0.1732** (0.1009)	0.1395*** (0.0464)	0.1711** (0.0827)	–0.0724 (0.1416)	–0.0508 (0.1890)
Analyst Bias (*BIAS*)	–0.0349* (0.0261)	–0.0378*** (0.0157)	–0.0322** (0.0168)	0.0121 (0.0285)	0.0091 (0.0460)
Control Variables					
Number of News Articles (*NEWS*)			–0.0995* (0.0675)	–0.0987* (0.0689)	–0.0947* (0.0702)
*NEWS*DISPERSION*			0.0656 (0.0620)	0.2613** (0.1513)	0.2476 (0.1960)
*NEWS*BIAS*			–0.0113 (0.0134)	–0.0562** (0.0302)	–0.0531 (0.0469)
Negative Words (*NEGATIVE*)				–0.0804** (0.0487)	–0.0929* (0.0650)
*NEGATIVE*DISPERSION*				0.2862*** (0.1109)	0.2661* (0.1748)
*NEGATIVE*BIAS*				–0.0667* (0.0480)	–0.0519 (0.0800)
Positive Words (*POSITIVE*)					–0.0571 (0.0545)
*POSITIVE*DISPERSION*					–0.0201 (0.1730)
*POSITIVE*BIAS*					–0.0052 (0.0779)
Earnings Variability (*VAR*)	–0.1413*** (0.0278)	–0.1240*** (0.0271)	–0.1099*** (0.0295)	–0.0920*** (0.0329)	–0.0922*** (0.0316)
Market Value (*MV*)	–0.2973*** (0.0828)	–0.2536*** (0.0860)	–0.2325*** (0.0878)	–0.2295*** (0.0843)	–0.2489*** (0.0913)
High-tech (*HIGHTECH*)	0.2423 (0.1735)	0.2129 (0.1654)	0.1387 (0.1780)	0.1785 (0.1746)	0.1664 (0.1935)
Regulation (*REG*)	0.0420 (0.2075)	–0.1332 (0.2308)	–0.0930 (0.2625)	–0.1137 (0.2699)	–0.0957 (0.2587)
Quarter 1 Dummy (*Q1*)	–0.4517*** (0.1814)	–0.5662*** (0.2111)	–0.5431*** (0.2101)	–0.4925** (0.2140)	–0.4338** (0.2211)
Quarter 2 Dummy (*Q2*)	–0.0826 (0.1958)	–0.1424 (0.2007)	–0.1405 (0.2071)	–0.0869 (0.2116)	–0.1051 (0.2069)
Quarter 3 Dummy (*Q3*)	–0.4644** (0.2132)	–0.5340*** (0.2064)	–0.5413*** (0.2054)	–0.5587*** (0.2048)	–0.4858** (0.2251)
Observations	161	161	161	161	161
Likelihood Ratio Test	p = 0.001	p < 0.001	p < 0.001	p = 0.001	p = 0.003
Pseudo R-Squared	0.158	0.203	0.215	0.230	0.235

Note: Robust estimates of standard errors are reported in parentheses. Significance levels are one-tailed for directional predictions and two-tailed otherwise. *p < 0.10; **p < 0.05; ***p < 0.01. The pseudo R^2 refers to Nagelkerke's (1991) R^2.

This implies that managers may perceive the risk associated with negative media content (i.e., the threat of backfire as mentioned above) to increase as DISPERSION decreases. When analysts are consistent with each other (i.e., low DISPERSION), their forecasts may become more convincing to investors. This would expose managers to a greater threat of backfire; consequently, managers may disclose earlier to minimize damage.

The impacts of NEWS are similar to those of NEGATIVE. The impacts of POSITIVE are not significant. These findings are consistent with prior research suggesting that readers are more affected by negative than positive news content (Tetlock et al. 2008).

Overall, our results seem to indicate that, unlike the salient role played by WikiMOD, some variables capturing media coverage turn out to be not significant. Similar to our result, Frankel and Li (2004) find that media coverage does not seem to improve investors' information environment. A possible explanation offered by the literature is media bias. The literature on media bias suggests that media are often biased for both demand- and supply-side reasons (Knight and Chiang 2008). On the demand side, consumers may prefer to consume information that confirms their prior beliefs; profit-maximizing publishers thus have incentives to bias their reports (Gentzkow and Shapiro 2006; Mullainathan and Shleifer 2005). On the supply side, individual journalists and financial analysts may also bias reports to reflect their own views or achieve their own financial objectives (Baron 2006; Reuter and Zitzewitz 2006). As a result, financial reports in the media often slant information through "selective omission, choice of words, and varying credibility ascribed to the primary source" (Gentzkow and Shapiro 2006, p. 281). Wikipedia, owing to its NPOV policy and because contributions on Wikipedia follow a democratic and collaborative model, is less susceptible to media bias. Prior studies establish that this kind of "wisdom-of-crowds" production model tends to generate facts (Kittur and Kraut 2008; Swarts 2009). If this is the case, one may expect that Wikipedia, instead of media, plays a salient role in improving the market's information environment. Our results offer supportive evidence, although we cannot completely rule out the possibility of the existence of bias on Wikipedia.

Results about Investor Reaction

To test H3, Table 4 presents the regression results for model (2). Columns (1) and (2) report regressions using CAR over a five-day time window to measure investors' reaction; columns (3) and (4) examine CAR over a three-day time window, for a robustness check; and columns (5) through (8) examine market-adjusted returns (RET) based on a five-day time window and a three-day time window. These columns show highly consistent results. Below, we use columns (1) and (2) to discuss our findings.

Column (1) of Table 4 establishes the baseline result. Since our sample includes only bad news, the market reaction should be negative. As expected, the more optimistic the analysts, the more disappointed are the investors (i.e., a negative coefficient on BIAS).

H3, supported: In column (2), WikiMOD*BIAS has a positive sign, suggesting that WikiMOD moderates investor reaction to bad news. Had there existed a medium that disclosed information to the public, this medium would have offset the surprise caused by BIAS. Our results suggest that WikiMOD seems to play such a role, in support of H3.

Regarding controls, firm size plays a positive moderating role as expected. NEWS and NEGATIVE play a negative moderating role. These results suggest that (1) more news coverage exacerbates the problem of overly optimistic analyst estimates, and (2) when there is an upward analyst bias, investors are more disappointed if traditional media used many negative words to describe that firm. The first effect seems to suggest that more news coverage may merely magnify the effect of analyst bias. Or, investors are generally affected by negative instead of positive words, so NEWS and NEGATIVE play similar roles. Judging from the sign, we can at least argue that more news coverage does not offer more information to reduce the impact of analyst bias or improve the information environment (Frankel and Li 2004). The second effect is consistent with the backfire story discussed above. Given the negative coefficient on BIAS, and the nonsignificant coefficient on NEGATIVE, it is plausible to argue that investors tend to "trust" analysts more than newspapers. That is, investors seem to follow analysts' upward-biased recommendation, even when newspapers use many negative words to describe a company. When analysts are wrong, the inconsistency between news and analysts' optimistic estimates would backfire. This may help explain why the interaction between NEGATIVE and BIAS is significantly negative.

Additional Analysis

Possible Endogeneity of WikiMOD: As with other empirical studies, we cannot control for *all* possible information sources to completely rule out the existence of alternative channels through which managers and investors get informed. Endogeneity concerns are legitimate in this context for two reasons.

Table 4. Investor Reaction to Disclosure (H3) (Regression specification is based on model (2) in text. Variable definitions are in Table 1.)

	Cumulative Abnormal Returns: [–2,2] Window		Cumulative Abnormal Returns: [–1,1] Window		Market-adjusted Returns: [–2,2] Window		Market-adjusted Returns: [–1,1] Window	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Theoretical Variables								
Analyst Bias (*BIAS*)	–0.0085*** (0.0019)	–0.0827*** (0.0321)	–0.0053*** (0.0022)	–0.0557** (0.0332)	–0.0088*** (0.0019)	–0.0808** (0.0367)	–0.0055*** (0.0022)	–0.0552* (0.0357)
WikiMOD		–0.0084 (0.0099)		–0.0087 (0.0102)		–0.0074 (0.0099)		–0.0075 (0.0103)
WikiMOD*BIAS		**0.0035*** (0.0015)**		**0.0033** (0.0018)**		**0.0035** (0.0015)**		**0.0033** (0.0019)**
Control Variables								
Market Value (*MV*)		0.0057 (0.0096)		0.0133 (0.0099)		0.0061 (0.0099)		0.0134 (0.0100)
*MV*BIAS*		0.0046*** (0.0017)		0.0034** (0.0019)		0.0045*** (0.0019)		0.0035** (0.0020)
NEWS		0.0075 (0.0088)		0.0034 (0.0086)		0.0087 (0.0088)		0.0045 (0.0087)
*NEWS*BIAS*		–0.0039*** (0.0013)		–0.0056*** (0.0015)		–0.0042*** (0.0015)		–0.0059*** (0.0015)
NEGATIVE		0.0025 (0.0045)		0.0022 (0.0035)		0.0031 (0.0044)		0.0025 (0.0034)
*NEGATIVE*BIAS*		–0.0100*** (0.0042)		–0.0104** (0.0036)		–0.0109** (0.0047)		–0.0109*** (0.0038)
POSITIVE		–0.0016 (0.0036)		–0.0022 (0.0019)		–0.0029 (0.0036)		–0.0033* (0.0020)
*POSITIVE*BIAS*		0.0000 (0.0033)		–0.0015 (0.0034)		0.0008 (0.0037)		–0.0007 (0.0037)
Observations	161	161	161	161	161	161	161	161
R–Squared	0.0752	0.1498	0.0307	0.1323	0.0789	0.1574	0.0327	0.1383

Note: Robust estimates of standard errors are reported in parentheses. Significance levels are one-tailed for directional predictions and two-tailed otherwise. *p < 0.10; **p < 0.05; ***p < 0.01.

First, it is possible that managers may release information through Wikipedia themselves. If a manager decides to release information through Wikipedia, and if the decision to edit the entry correlates with the decision about disclosure lag (e.g., because of an incentive to release information to the public), then the manager's self-editing may confound Wikipedia's effect.

Second, if there exists an alternative information channel that influences both Wikipedia modifications and disclosure lags, then the identified effect of WikiMOD cannot be attributed to Wikipedia. More broadly, if there are some omitted factors that affect disclosure lags and are correlated with Wikipedia modifications, endogeneity is a concern. We address these endogeneity issues by conducting an instrumental variable (IV) estimation with two possible IVs. Our results suggest that these concerns are not serious. Appendix D reports the details of our IV estimation.

An Alternative Explanation:[26] We need to control for the possible impact of firm visibility. It is an alternative explanation for the impact of WikiMOD, in that firm visibility is correlated with WikiMOD and may influence management disclosure. We reviewed prior studies on firm visibility and found that they used three sets of proxy variables for firm visibility. The first set refers to fundamental firm characteristics that are associated with firm visibility, including firm size (e.g., market value), firm profitability (e.g., ROA), firm age, and advertising expenditure. Firms with a larger size and greater profitability and older firms may be associated with higher levels of visibility (Bushee and Miller 2012, Grullon et al. 2004), and advertising may also help increase firm visibility (Grullon et al. 2004). The second set refers to media coverage in that the amount of news reports tends to covary

[26]We thank an anonymous reviewer for suggesting this important test.

with firm visibility (Bushee and Miller 2012). The third set concerns firm visibility in the capital market (Baker et al. 1999; Bushee and Miller 2012; Grullon et al. 2004) and suggests that firm visibility may be positively related to NYSE listing (a dummy variable indicating firms listed at NYSE), institutional ownership (the percentage of a firm's common stock held by institutions), and analyst following (the number of analysts following a firm).

In addition to the traditional measures of firm visibility, we seek to control for online firm visibility, given that the interest of our research is the effect of an online information channel. In a very recent paper, Da et al. (2011) use Google's Search Volume Index (SVI) of company names as a proxy for investor attention. We collect Google SVI data from Google Insights.[27] SVI for a search term is the number of searches for that term scaled by its time-series average. To make search volumes of different companies comparable, we obtain all firm SVIs relative to the generic word "cotton," which remains stable throughout the observation period. Figure 2 gives an example of how the relative search volume for "Dell" is obtained. The blue (lower) line is the search volume for "cotton." The red (upper) line is the search volume for "Dell." Google reports relative measures of the two search volume indices. Throughout the observation period, the search volume for "cotton" is normalized to 1. Some company names, such as Palm and Gap, may have alternative meanings that are not directly related to the respective companies. We manually go through our list and create an indicator for such companies. When reporting our results, we include all such companies to avoid a subjectivity bias. The results remain qualitatively unchanged after removing such companies from our sample.

Appendix E presents the results. For ease of comparison, Column (1) shows the base model (i.e., Column (5) in Table 3). Column (2) includes all of the traditional controls for firm visibility. Column (3) further includes the control for Google SVI. Appendix E conveys two key messages. First, firm visibility generally plays a role in leading firms to disclose earlier, as evident in the significant and negative coefficients on firm size, NYSE listing, analyst following, and firm profitability. Second, importantly, we see qualitatively unchanged effects of WikiMOD, after incorporating the various controls for firm visibility.

[27]http://www.google.com/insights/search/. For more information about how the indices were constructed, please refer to Google's help documents: http://support.google.com/insights/. Similar to Da et al. (2011), we search for both the company names and their stock tickers. The results are highly consistent.

In Appendix F, we examine the economic significance of Wikipedia modifications. We show that information aggregation on Wikipedia has an economically meaningful impact on market reaction. Appendix G shows that our results are robust to alternative samples (e.g., excluding firms with very low and very high values of WikiMOD, removing modifications that are involved in edit wars), and to alternative measures (e.g., using the number of words added to proxy for information aggregation on Wikipedia).

Concluding Remarks

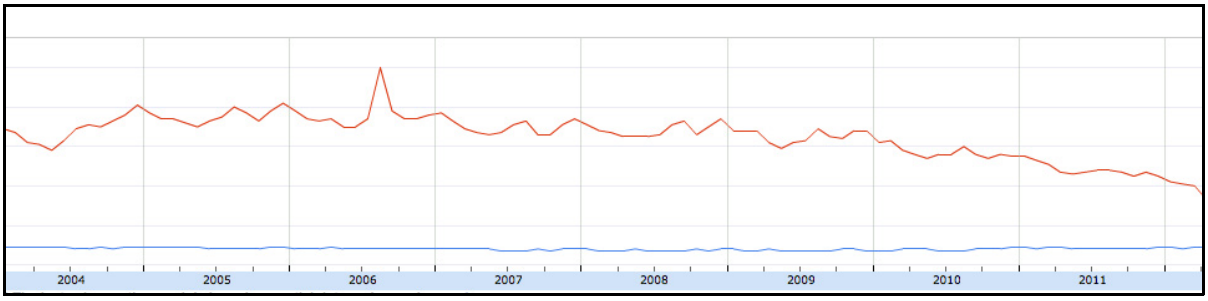
Major Findings

In this paper, we explore the question of whether Wikipedia improves the information environment for financial market investors, and if so, how public-firm managers change their voluntary disclosure behavior concerning bad news when there are variations in Wikipedia coverage. Our analysis is based on a unique data set collected from financial records of public firms, management-disclosure records, news articles related to these firms, and the modification history of firm entries on Wikipedia. The following new findings add to the existing literature.

First, we find that information aggregation on Wikipedia about public firms can moderate the timing of management disclosure of earnings disappointments. The timing of management disclosure is a function of the information environment, as described by the characteristics of analyst forecasts, a finding consistent with the disclosure literature. Although managers generally tend to delay disclosure of bad news when analyst forecasts are highly dispersed, the delay is shortened by Wikipedia's information aggregation, suggesting that Wikipedia weakens information asymmetry between managers and investors.

Second, although the literature points out that when facing high analyst bias, managers tend to release bad news early to avoid various risks, this study finds that information aggregation on Wikipedia helps alleviate managers' concerns, thus making management's disclosure timing less dependent on analyst bias.

Finally, by examining how investors react to management disclosure of bad news, we find that investors' negative reaction to bad news is moderated by Wikipedia modifications. Although Wikipedia may not be the only means by which investors can enjoy the benefits of social media, our results suggest a market adjustment associated with information aggregation on Wikipedia before a firm announces bad news.



Note: The blue (lower) line is the search volume for "cotton"; it remains relatively stable over the observation period. The red (upper) line is the search volume for "Dell." Google reports relative measures of the two search volume indices. Throughout the observation period, the search volume for cotton is normalized to 1.

Figure 2. Google Search Volume

Theoretical Implications

This paper makes several theoretical contributions. Most prior studies of Wikipedia focus on users' contribution behavior (e.g., Greenstein and Zhu 2012a; Kittur and Kraut 2008, 2010; Thom-Santelli et al. 2009; Zhang and Zhu 2011). In many ways, these studies have improved our understanding of how individuals may be motivated to contribute to Wikipedia. Without establishing the value of the outcomes of such collaborations, however, research on contribution incentives is limited. This study fills such a gap and examines how Wikipedia makes an impact to the real world. To this end, we document IT's capability of aggregating information (Benbasat and Zmud 2003) and analyze the impact of an IT artifact (Wikipedia) directly (Orlikowski and Iacono 2001; Venkatesh et al. 2007).

Our finding suggests that the informational impact of social media such as Wikipedia can be derived from social interactions and collaborative efforts. This study supports a shift to conceptualizing users as social actors rather than individualistic entities (Lamb and Kling 2003). With such a shift, our study goes beyond the view of Wikipedia contributors as disjointed individuals, with each following his or her own agenda. Rather, we promote a theoretical anchor for identifying the value of information aggregation arising from collaborative efforts.

In building the theory, this paper taps into the accounting literature of management disclosure. This not only introduces a new lens to the information systems literature but also demonstrates the complementarity of information systems research with that of other related fields. The increasing reliance of business on IT opens a door for information systems research to contribute to these other business disciplines. As this study shows, the analysis of the impact of Wikipedia sheds light on how managers and investors obtain and use information in the financial market.

Managerial Implications

This study also has implications for investors, managers, and policy makers in the financial market. Importantly, different from news stories, which fade away quickly, Wikipedia's aggregating and accumulating information keeps a good record of a company's important events. This suggests that Wikipedia's effects on investors, managers, and regulators can be long-lasting.

Our results suggest that investors can benefit greatly from the value of information aggregation offered by services such as Wikipedia. This type of information aggregation stands in stark contrast to traditional ways for investors to get informed. In the traditional model of information acquisition, individual investors suffer a significant information disadvantage. The channel to obtain information is very limited and both newspapers and analysts can introduce significant biases. Unlike these channels, the decentralized information aggregation on Wikipedia seems to be offering a new channel that is either free of these biases or resistant to them. Interestingly, some individual investors have already started to build an information aggregation portal that utilizes the wiki technology. This new service is called "Wikinvest"[28] On their "About Us" page, they write, "We're regular, everyday investors who are

[28]The website is http://www.wikinvest.com.

sick of the level of innovation at the major finance portals." One major difference between Wikinvest and traditional financial information websites is that Wikinvest relies heavily on user-generated content, and it extracts the value of information aggregation.

For managers, our study suggests that their information advantage is increasingly compromised with the advent of information aggregation channels like Wikipedia. In the pre-Wikipedia era, they could utilize their private information and manipulate what and when the market could learn about their firms. With Wikipedia and other types of social-media channels, it becomes increasingly difficult to maintain the same level of control over information. It may be necessary for managers to change their mindset and embrace the benefits brought about by new media. One of our findings suggests that Wikipedia can alleviate their litigation concerns when analyst bias is high. Our view is that the information asymmetry between managers and investors is not necessarily a zero-sum game. If managers can recognize the unavoidable impact of IT, and make good use of it, they will be able to better capture opportunities and resolve challenges.

Regulators such as the SEC should welcome the improved information transparency introduced by Wikipedia and other social-media channels. To restore investors' confidence in the market after a series of corporate scandals, the Sarbanes-Oxley Act (SOX) was enacted in 2002. SOX aims to reform public company accounting and protect investors, and one major objective is to enhance corporate transparency (Coates 2007). The information transparency achieved through information aggregation on Wikipedia offers a complementary mechanism to achieve the same goals as those of regulators. This paper examines information aggregation on Wikipedia for only public companies, and thus it is important to note that many private companies also have Wikipedia pages. If Wikipedia can improve the information transparency of these privately held companies, given that SOX does not apply to them, then our results suggest that technology-enabled information aggregation by individuals plays a broader role than the legal mechanism in enhancing companies' information environment.

Limitations and Future Research

This paper has several limitations that may need future work to address. First, as explained above, we focus on only Wikipedia and do not capture the full range of social-media channels (e.g., blogs, Twitter, Facebook, etc.). Although other social-media channels can also aggregate information, the mechanisms through which these other channels affect management information disclosure may be different. Blogs, for example, are typically managed in a centralized manner, such that one blog can have only one or a few contributing bloggers (Sun and Zhu 2012). Only some frequent bloggers can attract the public's attention, and their potential influence is limited to loyal followers of these blogs. Although it is easy to find some information about almost all public companies on Wikipedia, blogging information of these companies is scattered and may be difficult to find. Twitter is different from Wikipedia in that Twitter's social influence follows a directional path. In addition, similar to traditional news, information about specific companies on Twitter has short-lived popularity. On Facebook, network ties are based on friends, and the tie strength is much stronger than the poster–reader relation on Wikipedia. As a result, the information-exchange mechanism is different. As a result of the structural differences between Wikipedia and other social media platforms, whether other social media channels affect management disclosure in a similar way needs careful examination before our results can be generalized. Our prediction of such effects is rather optimistic. If the unmeasured information aggregations were to be correlated with what we have incorporated in our analysis, then our estimates would partially reflect this broader range of information over social media. This is likely to be true, because information can diffuse across social-media channels (Watson-Manheim and Bélanger 2007). A future study could deepen our understanding by studying other types of social-media channels.

Second, Wikipedia may play other roles in the financial market. This study merely touches the tip of the iceberg. We employ a narrow lens to examine managers' behavior of information disclosure. There are other significant issues, such as investor sentiments, insider trading, changes in liquidity, and stock-price volatility associated with social-media coverage. These would be fruitful avenues for future research.

Third, this study adopts a positivist perspective to infer the relationship between Wikipedia and the financial market's information environment. Our results could be complemented by case studies, surveys, or ethnographic methods. For instance, it remains an assumption that managers know that information aggregation over Wikipedia provides information to investors. Future research can conduct a survey to ask managers whether they are aware of information aggregation over Wikipedia and how they perceive the extent to which investors get informed from that. While our results present evidence of Wikipedia's role in the financial market, these alternative methods offer important insights about why and how such roles take place.

Acknowledgments

We thank the senior editor, the associate editor, and the two anonymous reviewers for their constructive feedbacks, which enabled us to improve the paper significantly. We appreciate the suggestions that we received from Kalok Chan, Eric Clemons, Sudipto Dasgupta, Alok Gupta, Chia-Chun Hsieh, Kirill Novoselov, Mike So, Feng (Alex) Tian, Bruce Weber, Hong Xu, Mingshan Zhang, Yong Zhang, Feng Zhu, and participants at the 2009 International Conference on Information Systems. Sean Xin Xu acknowledges financial support from the Program for New Century Excellent Talents in University (NCET-10-0508) and the National Natural Science Foundation of China (#70890080 and #71001056). Michael Zhang acknowledges generous financial support from Hong Kong Research Grant Council and HKUST (GRF 647311, GRF 645211, RPC07/08.BM16, and DAG_S09/10.BM03). An early version of this paper was presented and nominated for the Best Paper Award at the International Conference on Information Systems in 2009. Both authors contributed equally to this research.

References

Abarbanell, J. S., Lanen, W. N., and Verrecchia, R. E. 1995. "Analysts' Forecasts as Proxies for Investor Beliefs in Empirical Research," *Journal of Accounting and Economics* (20:1), pp. 31-60.

Aiken, L. S., and West, S. G. 1991. *Multiple Regression: Testing and Interpreting Interactions*, Newbury Park, CA: Sage Publications.

Anilowski, C., Feng, M., and Skinner, D. J. 2007. "Does Earnings Guidance Affect Market Returns? The Nature and Information Content of Aggregate Earnings Guidance," *Journal of Accounting and Economics* (44:1-2), pp. 36-63.

Antweiler, W., and Frank, M. Z. 2004. "Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards," *Journal of Finance* (59:3), pp. 1259-1294.

Arazy, O., Nov, O., Patterson, R., and Yeo, L. 2011. "Community-Based Collaboration in Wikipedia: The Effects of Group Composition and Task Conflict on Information Quality," *Journal of Management Information Systems* (21:4), pp. 71-98.

Atiase, R. K. 1985. "Predisclosure Information, Firm Capitalization, and Security Price Behavior Around Earnings Announcements," *Journal of Accounting Research* (23:1), pp. 21-36.

Axelson, U. 2007. "Security Design with Investor Private Information," *Journal of Finance* (62:6), pp. 2587-2632.

Ba, S., Stallaert, J., and Zhang, J. 2012. "Online Price Dispersion: A Game-Theoretic Perspective and Empirical Evidence," *Information Systems Research* (23:2), pp. 575-592.

Baginski, S., Hassell, J., and Waymire, G. 1994. "Some Evidence on the News Content of Preliminary Earnings Estimates," *The Accounting Review* (69:1), pp. 265-273.

Baker, H. K., Powell, G. E., and Weaver, D. G. 1999. "Does NYSE Listing Affect Firm Visibility? ," *Financial Management* (28:2), pp. 46-54.

Banks, J. S. 1985. "Price-Conveyed Information Versus Observed Insider Behavior: A Note on Rational Expectations Convergence," *Journal of Political Economy* (93:4), pp. 807-815.

Baron, D. P. 2006. "Persistent Media Bias," *Journal of Public Economics* (90:1-2), pp. 1-36.

Bartov, E., Givoly, D., and Hayn, C. 2002. "The Rewards to Meeting or Beating Earnings Expectations," *Journal of Accounting and Economics* (33:1), pp. 173-204.

Benbasat, I., and Zmud, R. W. 2003. "The Identity Crisis Within the Is Discipline: Defining and Communicating the Discipline's Core Properties," *MIS Quarterly* (27:2), pp. 183-194.

Brennan, N. 1999. "Voluntary Disclosure of Profit Forecasts by Target Companies in Takeover Bids," *Journal of Business Finance and Accounting* (26:7-8), pp. 883-918.

Brynjolfsson, E., and Smith, M. D. 2000. "Frictionless Commerce? A Comparison of Internet and Conventional Retailers," *Management Science* (46:4), pp. 563-585.

Bushee, B., and Miller, G. S. 2012. "Investor Relations, Firm Visibility, and Investor Following," *The Accounting Review* (87:3), pp. 867-897.

Butler, B. S. 2001. "Membership Size, Communication Activity, and Sustainability: A Resource-Based Model of Online Social Structures," *Information Systems Research* (12:4), pp. 346-362.

Carlson, J. R., and Zmud, R. W. 1999. "Channel Expansion Theory and the Experiential Nature of Media Richness Perceptions," *Academy of Management Journal* (42:2), pp. 153-170.

Chen, J., Ren, Y., and Riedl, J. 2010. "The Effects of Diversity on Group Productivity and Member Withdrawal in Online Volunteer Groups," in *Proceedings of the 28th International Conference on Human Factors in Computing Systems*, April 10-15, pp. 821-830.

Chen, J., Xu, H., and Whinston, A. B. 2011. "Moderated Online Communities and Quality of User-Generated Content," *Journal of Management Information Systems* (28:2), pp. 237-268.

Chen, S., Defond, M., and Park, C. 2002. "Voluntary Disclosure of Balance Sheet Information in Quarterly Earnings Announcements," *Journal of Accounting and Economics* (33:1), pp. 229-251.

Coates, J. C. 2007. "The Goals and Promise of the Sarbanes-Oxley Act," *Journal of Economic Perspectives* (21:1), pp. 91-116.

Cohen, D. A., Dey, A., Lys, T. Z., and Sunder, S. V. 2007. "Earnings Announcement Premia and the Limits to Arbitrage," *Journal of Accounting and Economics* (43:2-3), pp. 153-180.

Da, Z., Engelberg, J., and Gao, P. 2011. "In Search of Attention," *Journal of Finance* (66:5), pp. 1461-1499.

Dahya, J., McConnell, J. J., and Travlos, N. 2002. "The Cadbury Committee, Corporate Performance, and Top Management Turnover," *Journal of Finance* (57:1), pp. 461-483.

Das, S. R., and Chen, M. Y. 2007. "Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web," *Management Science* (53:9), pp. 1375-1388.

Dechow, P. M., Hutton, A. P., and Sloan, R. G. 2000. "The Relation Between Analysts' Forecasts of Long-Term Earnings Growth and Stock Price Performance Following Equity Offerings," *Contemporary Accounting Research* (17:1), pp. 1-32.

Dellarocas, C., and Wood, C. A. 2008. "The Sound of Silence in Online Feedback: Estimating Trading Risks in the Presence of Reporting Bias," *Management Science* (54:3), pp. 460-476.

DellaVigna, S., and Pollet, J. M. 2009. "Investor Inattention and Friday Earnings Announcements," *Journal of Finance* (64:2), pp. 709-749.

Dewan, S., Shi, C., and Gurbaxani, V. 2007. "Investigating the Risk–Return Relationship of Information Technology Investment: Firm-Level Empirical Analysis," *Management Science* (53:12), pp. 1829-1842.

Diamond, D. W., and Verrecchia, R. E. 1981. "Information Aggregation in a Noisy Rational Expectations Economy," *Journal of Financial Economics* (9:3), pp. 221-235.

Dye, R. A. 1985. "Disclosure of Nonproprietary Information," *Journal of Accounting Research* (23:1), pp. 123-145.

Dye, R. A. 2001. "Commentary on Essays on Disclosure," *Journal of Accounting and Economics* (32:1-3), pp. 181-235.

Fan, M., Stallaert, J., and Whinston, A. B. 2000. "The Internet and the Future of Financial Markets," *Communications of the ACM* (43:11), pp. 82-88.

Foster, G. 1987. "Rambo IX: Briloff and the Capital Market," *Journal of Accounting, Auditing and Finance* (2:4), pp. 409-430.

Forsythe, R., and Lundholm, R. 1990. "Information Aggregation in an Experimental Market," *Econometrica* (58:2), pp. 309-347.

Francis, J., Hanna, J. D., and Philbrick, D. R. 1997. "Management Communications with Securities Analysts," *Journal of Accounting and Economics* (24:3), pp. 363-394.

Francis, J., Nanda, D., and Olsson, P. 2008. "Voluntary Disclosure, Earnings Quality, and Cost of Capital," *Journal of Accounting Research* (46:1), pp. 53-99.

Frankel, R., and Li, X. 2004. "Characteristics of a Firm's Information Environment and the Information Asymmetry Between Insiders and Outsiders," *Journal of Accounting and Economics* (37:2), pp. 229-259.

Franks, J. R., and Mayer, C. 1996. "Hostile Takeovers and the Correction of Managerial Failure," *Journal of Financial Economics* (40), pp. 163-181.

Geerings, J., Bollen, L. H. H., and Hassink, H. F. D. 2003. "Investor Relations on the Internet: A Survey of the Euronext Zone," *European Accounting Review* (12:3), pp. 567-579.

Gentzkow, M., and Shapiro, J. M. 2006. "Media Bias and Reputation," *Journal of Political Economy* (114:2), pp. 280-316.

Granados, N., Gupta, A., and Kauffman, R. 2010. "Information Transparency in Business-to-Consumer Markets: Concepts, Framework, and Research Agenda," *Information Systems Research* (21:2), pp. 207-226.

Greenstein, S, and Zhu, F. 2012a. "Collective Intelligence and Neutral Point of View: The Case of Wikipedia," NBER Working Paper No. 18167, National Bureau of Economic Research.

Greenstein, S, and Zhu, F. 2012b. "Is Wikipedia Biased?," *American Economic Review Papers and Proceedings* (102:3), pp. 343-348.

Grullon, G., Kanatas, G., and Weston, J. P. 2004. "Advertising, Breadth of Ownership, and Liquidity," *Review of Financial Studies* (17:2), pp. 439-461.

Gu, B., Konana, P., Rajagopalan, B., and Chen, H-W. M. 2007. "Competition among Virtual Communities and User Valuation: The Case of Investing-Related Communities," *Information Systems Research* (18:1), pp. 68-85.

Hayek, F. A. 1945. "The Use of Knowledge in Society," *American Economic Review* (35:4), pp. 519-530.

Healy, P. M., and Palepu, K. G. 2001. "Information Asymmetry, Corporate Disclosure, and the Capital Markets: A Review of the Empirical Disclosure Literature," *Journal of Accounting and Economics* (31:1-3), pp. 405-440.

Jacob, J., Lys, T. Z., and Neale, M. A. 1999. "Expertise in Forecasting Performance of Security Analysts," *Journal of Accounting and Economics* (28:1), pp. 51-82.

Jung, W., and Kwon, Y. 1988. "Disclosure When the Market Is Unsure of Information Endowment of Managers," *Journal of Accounting Research* (26:1), pp. 146-153.

Kalbfleisch, J., and Prentice, R. 1980. *The Statistical Analysis of Failure Time Data*, New York: John Wiley & Sons.

Kane, G. C., and Fichman, R. G. 2009. "The Shoemaker's Children: Using Wikis for Information Systems Teaching, Research, and Publication," *MIS Quarterly* (33:1), pp. 1-22.

Kankanhalli, A., Tan, B. C-Y., and Wei, K-K. 2005. "Contributing Knowledge to Electronic Knowledge Repositories: An Empirical Investigation," *MIS Quarterly* (29:1), pp. 113-143.

Kaplan, A. M., and Haenlein, M. 2010. "Users of the World, Unite! The Challenges and Opportunities of Social Media," *Business Horizons* (53:1), pp. 59-68.

Karamanou, I., and Vafeas, N. 2005. "The Association Between Corporate Boards, Audit Committees, and Management Earnings Forecasts: an Empirical Analysis," *Journal of Accounting Research* (43:3), pp. 453-486.

Kasznik, R., and Lev, B. 1995. "To Warn or Not to Warn: Management Disclosures in the Face of an Earnings Surprise," *The Accounting Review* (70:1), pp. 113-134.

Kauffman, R. J., McAndrews, J., and Wang, Y-M. 2000. "Opening the 'Black Box' of Network Externalities in Network Adoption," *Information Systems Research* (11:1), pp. 61-82.

King, R. 1988. "The Liar's Discount," *Forbes,* May 30.

Kittur, A., Chi, E., Pendleton, B. A., Suh, B., and Mytkowicz, T. 2008. "Power of the Vew vs. Wisdom of the Crowd: Wikipedia and the Rise of the Bourgeiosie," in *Proceedings of the 25th ACM Conference on Human Factors in Computing Systems*, San Jose, CA, April 28-May 3.

Kittur, A., and Kraut, R. E. 2008. "Harnessing the Wisdom of Crowds in Wikipedia: Quality Through Coordination," in *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work*, November 8-12, pp. 37-46.

Kittur, A., and Kraut, R. E. 2010. "Beyond Wikipedia: Coordination and Conflict in Online Production Groups," in *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work*, February 6-10, pp. 214-244.

Knight, B. G., and Chiang, C. F. 2011. "Media Bias and Influence: Evidence from Newspaper Endorsements," *Review of Economic Studies* (78:3), pp. 795-820.

Kothari, S. P., Laguerre, T. E., and Leone, A. J. 2002. "Capitalization Versus Expensing: Evidence on the Uncertainty of Future Earnings from Current Investments in PP&E Versus R&D," *Review of Accounting Studies* (7:4), pp. 355-382.

Lamb, R., and Kling, R. 2003. "Reconceptualizing Users as Social Actors in Information Systems Research," *MIS Quarterly* (27:2), pp. 197-235.

Lang, M. H., and Lundholm, R. J. 1996. "Corporate Disclosure Policy and Analyst Behavior," *The Accounting Review* (71:4), pp. 467-492.

Lin, H. W., and McNichols, M. F. 1998. "Underwriting Relationships, Analysts' Earnings Forecasts and Investment Recommendations," *Journal of Accounting and Economics* (25:1), pp. 101-128.

Lin, L., Geng, X., and Whinston, A. B. 2005. "A Sender–Receiver Framework for Knowledge Transfer," *MIS Quarterly* (29:2), pp. 197-219.

Loss, L., and Seligman, J. 1995. *Fundamentals of Securities Regulation*, Boston: Little, Brown & Co.

Ma, M., and Agarwal, R. 2007. "Through a Glass Darkly: Information Technology Design, Identity Verification, and Knowledge Contribution in Online Communities," *Information Systems Research* (18:1), pp. 42-67.

Majchrzak, A. 2009. "Comment: Where Is the Theory in Wikis?," *MIS Quarterly* (33:1), pp. 18-20.

Matsumoto, D. A. 2002. "Management's Incentives to Avoid Negative Earnings Surprise," *The Accounting Review* (77:3), pp. 483-514.

Miller, G. S. 2002. "Earnings Performance and Discretionary Disclosure," *Journal of Accounting Research* (40:1), pp. 173-204.

Mitchell, M. L., and Mulherin, J. H. 1994. "The Impact of Public Information on the Stock Market," *Journal of Finance* (49:3), pp. 923-950.

Moon, J. Y., and Sproull, L. S. 2008. "The Role of Feedback in Managing the Internet-Based Volunteer Work Force," *Information Systems Research* (19:4), pp. 494-515.

Morck, R., Shleifer, A., and Vishny, R. 1990. "Do Managerial Objectives Drive Bad Acquisitions? ," *Journal of Finance* (45:1), pp. 31-50.

Mullainathan, S., and Shleifer, A. 2005. "Market for News," *American Economic Review* (95:4), pp. 1031-1053.

Nagelkerke, N. J. D. 1991. "A Note on a General Definition of the Coefficient of Determination," *Biometrika* (78:3), pp. 691-692.

Orlikowski, W. J., and Iacono, C. S. 2001. "Research Commentary: Desperately Seeking the 'IT' in IT Research – A Call to Theorizing the IT Artifact," *Information Systems Research* (12:2), pp. 121-134.

Panciera, K., Halfaker, A., and Terveen, L. 2009. "Wikipedians Are Born, Not Made: A Study of Power Editors on Wikipedia," in *Proceedings of the ACM 2009 International Conference on Supporting Group Work*, May 10-13, pp. 51-60.

Pownall, G., and Waymire, G. 1989. "Voluntary Disclosure Credibility and Securities Prices: Evidence from Management Earnings Forecasts," *Journal of Accounting Research* (27:2), pp. 227-245.

Radner, R. 1979. "Rational Expectations Equilibrium: Generic Existence and the Information Revealed by Prices," *Econometrica* (47:3), pp. 654-677.

Ransbotham, S., and Kane, G. C. 2011. "Membership Turnover and Collaboration Success in Online Communities: Explaining Rises and Falls from Grace in Wikipedia," *MIS Quarterly* (35:3), pp. 613-627.

Reuter, J., and Zitzewitz, E. 2006. "Do Ads Influence Editors? Advertising and Bias in the Financial Media," *Quarterly Journal of Economics* (121:1), pp. 197-227.

Richardson, S., Teoh, S. H., and Wysocki, P. D. 2004. "The Walk-Down to Beatable Analyst Forecasts: The Role of Equity Issuance and Insider Trading Incentives," *Contemporary Accounting Research* (21:4), pp. 885-924.

Rose, R. L. 1991. "SEC Investigating Harley Stock Plunge; Angry Investors See Leak of Results," *The Wall Street Journal*, October 25, p. A.6.

Roulstone, D. T. 2003. "Analyst Following and Market Liquidity," *Contemporary Accounting Research* (20:3), pp. 552-578.

Schultze, U., and Leidner, D. E. 2002. "Studying Knowledge Management in Information Systems Research: Discourses and Theoretical Assumptions," *MIS Quarterly* (26:3), pp. 213-242.

Skinner, D. J. 1994. "Why Firms Voluntarily Disclose Bad News," *Journal of Accounting Research* (32:7-8), pp. 38-61.

Skinner, D. J. 1997. "Earnings Disclosures and Stockholder Lawsuits," *Journal of Accounting and Economics* (23:3), pp. 249-282.

Stein, E. W., and Zwass, V. 1995. "Actualizing Organizational Memory with Information Systems," *Information Systems Research* (6:2), pp. 85-117.

Sun, M. 2011. "Disclosing Multiple Product Attributes," *Journal of Economics and Management Strategy* (20:1), pp. 195-224.

Sun, M., and Tyagi, R. 2012. "When Does a Manufacturer Disclose Product Match Information?" Working Paper, Stanford University.

Sun, M., and Zhu, F. 2012. "Ad Revenue and Content Commercialization: Evidence from Blogs," NET Institute Working Paper #11-32.

Sunder, S. 1992. "Markets for Information: Experimental Evidence," *Econometrica* (60:3), pp. 667-695.

Surowiecki, J. 2004. *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies, and Nations*, New York: Doubleday.

Swarts, J. 2009. "The Collaborative Construction of 'Fact' on Wikipedia," in *Proceedings of the 27th ACM International Conference on Design of Communication*, October 4-7, pp. 281-288.

Te'eni, D. 2009. "Comment: The Wiki Way in a Hurry—The ICIS Anecdote," *MIS Quarterly* (33:1), pp. 20-22.

Tetlock, P. C. 2007. "Giving Content to Investor Sentiment: The Role of Media in the Stock Market," *Journal of Finance* (62:3), pp. 1139-1168.

Tetlock, P. C., Saar-Tsechansky, M., and Macskassy, S. 2008. "More than Words: Quantifying Language to Measure Firms' Fundamentals," *Journal of Finance* (63:3), Pp. 1437-1467.

Thom-Santelli, J., Cosley, D., and Gay, G. 2009. "What's Mine Is Mine: Territoriality in Collaborative Authoring," in *Proceedings of the 27th International Conference on Human Factors*, April 4-9, pp. 1481-1484.

Venkatesh, V., Davis, F. D., and Morris, M. G. 2007. "Dead or Alive? The Development, Trajectory, and Future of Ttechnology Adoption Research," *Journal of the Association for Information System* (8:4), pp. 267-286.

Wagner, C., and Majchrzak, A. 2006. "Enabling Customer-Centricity Using Wiki and Wiki Way," *Journal of Management Information Systems* (23:3), pp. 17-43.

Wales, J. 2005. "Wikipedia, Emergence, and the Wisdom of Crowds," http://mail.wikipedia.org/pipermail/wikipedia-l/2005-May/039397.html

Warner, J., Watts, R., and Wruck, K. 1988. "Stock Prices and Top Management Changes," *Journal of Financial Economics* (20), pp. 461-493.

Wasko, M. M., and Faraj, S. 2005. "Why Should I Share? Examining Social Capital and Knowledge Contribution in Electronic Networks of Practice," *MIS Quarterly* (29:1), pp. 35-57.

Wasley, C. E., and Wu, J. S. 2006. "Why Do Managers Voluntarily Issue Cash Flow Forecasts?," *Journal of Accounting Research* (44:2), pp. 389-429.

Watson-Manheim, M. B., and Bélanger, F. 2007. "Communication Media Repertories: Dealing with the Multiplicity of Media Choices," *MIS Quarterly* (31:2), pp. 267-293.

Waymire, G. 1984. "Additional Evidence on the Information Content of Nanagement Earnings Forecasts," *Journal of Accounting Research* (22:2), pp. 703-719.

Weisbach, M. 1988. "Outside Directors and CEO Turnover," *Journal of Financial Economics* (20), pp. 431-461.

Zetlin, M. 2010. "Use Wikipedia as a Marketing Tool," Inc.com (http://www.inc.com/managing/articles/201001/wikipedia.html).

Zhang, X. M., and Zhu, F. 2011. "Group Size and Incentives to Contribute: A Natural Experiment at Chinese Wikipedia," *American Economic Review* (101:4), pp. 1601-1615.

Zwass, V. 2010. "Co-Creation: Toward a Taxonomy and an Integrated Research Perspective," *International Journal of Electronic Commerce* (15:1), pp. 11-48.

Zhu, F., and Zhang, X. M. 2010. "Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics," *Journal of Marketing* (74:2), pp. 133-148.

About the Authors

Sean Xin Xu is a professor at the School of Economics and Management, Tsinghua University. He received his Ph.D. from the University of California, Irvine. His current research interests include IT value under different corporate governance structures, post-adoptive behaviors of IT use, social media, enterprise systems and information risk, and IT use in supply chain contexts. His papers have been published in *Management Science, MIS Quarterly, Information Systems Research, Journal of MIS, Strategic Management Journal, European Journal of Information Systems,* and *International Journal of Technology Management*. He serves on the editorial board of *Information Systems Research* as an associate editor, and that of *MIS Quarterly* as a guest associate editor. He was nominated for the Best Paper Award at the International Conference on Information Systems (ICIS) in 2009. He won the Vernon Zimmerman Best Paper Award at the Asian-Pacific Conference on International Accounting Issues in 2007, two Best Paper Awards at ICIS in 2002 and 2003, and a Best Paper Award (International Track) at the 2004 Americas' Conference on Information Systems.

Xiaoquan (Michael) Zhang is an associate professor of Information Systems, Business Statistics, and Operations Management at the Hong Kong University of Science and Technology. He holds a Ph.D. in Management from MIT Sloan School of Management, an M.Sc. in Management, a BE in Computer Science, and a BA in English from Tsinghua University. Before joining academia, he worked as a consultant for an investment bank, and as an international marketing manager for a high-tech company. He holds a U.S. patent, and started a social-network company. Michael's research interests are on issues related to creation, dissemination, and processing of information in business and management contexts. His works study pricing of information goods, online word-of-mouth, online advertising, incentives of creation in open source and open content projects, and use of information in financial markets. His research has appeared in *American Economic Review*, *Management Science*, *Journal of Marketing*, *MIS Quarterly*, *Journal of Management Information Systems*, *Decision Support Systems*, and *Journal of Interactive Marketing*. He currently serves as an associate editor for *Information Systems Research* and as a guest associate editor for *MIS Quarterly*, and is also on the editorial boards of *Production and Operations Management* and *Electronic Commerce Research and Applications.*

MIS Quarterly

RESEARCH ARTICLE

IMPACT OF WIKIPEDIA ON MARKET INFORMATION ENVIRONMENT: EVIDENCE ON MANAGEMENT DISCLOSURE AND INVESTOR REACTION

Sean Xin Xu
School of Economics and Management, Tsinghua University
Beijing, CHINA {xuxin@sem.tsinghua.edu.cn}

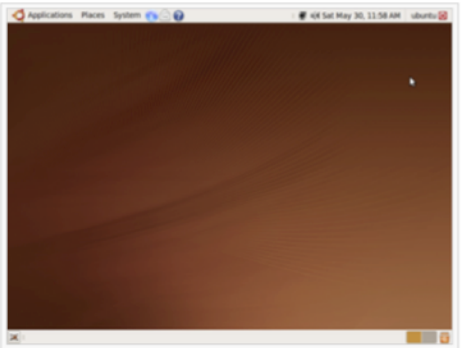
Xiaoquan (Michael) Zhang
School of Business and Management, Hong Kong University of Science and Technology,
Clear Water Bay, HONG KONG {zhang@ust.hk}

Appendix A

Example of Information Aggregation on Wikipedia

Ubuntu on Dell systems [edit]

On February 26, 2007 Dell announced that it had commenced a program to sell and distribute a range of computers with pre-installed Linux distributions as an alternative to Microsoft Windows. Dell indicated that Novell's SUSE Linux would appear first.[26] However, Dell on February 27, 2007 announced that its previous announcement related to certifying the hardware as ready to work with Novell SUSE Linux and that it (Dell) had no plans to sell systems pre-installed with Linux in the near future.[27] On March 28, 2007, Dell announced that it would begin shipping some desktops and laptops with Linux pre-installed, although it did not specify which distribution of Linux or which hardware would lead.[28] On April 18, 2007 a report appeared suggesting that Michael Dell used Ubuntu on one of his home systems.[29] On May 1, 2007, Dell announced it would ship the Ubuntu Linux distribution.[30] On May 24, 2007, Dell started selling models with Ubuntu Linux 7.04 pre-installed: a laptop, a budget computer, and a high-end PC.[31]



Dell has supplied Ubuntu on some of its computers since 2007

On June 27, 2007, Dell announced on its Direct2Dell blog that it planned to offer more pre-loaded systems (the new Dell Inspiron desktops and laptops). After the IdeaStorm site supported extending the bundles beyond the US market, Dell later announced more international marketing.[32] On August 7, 2007, Dell officially announced that it would offer one notebook and one desktop in the UK, France and Germany with Ubuntu "pre-installed". At LinuxWorld 2007 Dell announced plans to provide Novell's SUSE Linux Enterprise Desktop on selected models in China, "factory-installed".[33] On November 30, 2007 Dell reported shipping 40,000 Ubuntu PCs.[34] On January 24, 2008 Dell in Germany, Spain, UK and France launched a second laptop, a XPS M1330 with Ubuntu 7.10, for 849 euro or GBP 599 upwards.[35] On February 18, 2008, Dell announced that the Inspiron 1525 would have Ubuntu as an optional operating system.[*citation needed*] On February 22, 2008 Dell announced plans to sell Ubuntu in Canada and in Latin America[36] From September 16, 2008, Dell has shipped both Dell Ubuntu Netbook Remix and Windows XP Home versions of the Inspiron Mini 9 and the Inspiron Mini 12. As of 2009 Dell ships the Inspiron Mini 9 and the Inspiron Mini 12 laptops with Ubuntu version 8.04.[*citation needed*] Unsourced speculation[*who?*] opines that in the future[*weasel words*] Dell will offer Ubuntu 8.10 on the Mini 9 and Mini 12.

Source: http://en.wikipedia.org/wiki/Dell, accessed in January, 2010 (emphasis added).

Appendix B

Literature Review: Information Aggregation Mechanisms

Information aggregation mechanisms	Evidence and implications	How does information aggregation on Wikipedia differ from previous mechanisms?
Combining analyst forecasts	• Financial analysts as the foremost source of quantified information for investors (Abarbanell et al. 1995; Francis et al. 1997; Healy and Palepu 2001; Kasznik and Lev 1995; Lang and Lundholm 1996; Roulstone 2003; Skinner 1994, 1997). • Analyst forecasts, on average, are subject to upward bias (Francis et al. 1997; Karamanou and Vafeas 2005; Lang and Lundholm 1996; Roulstone 2003).	Wikipedia provides detailed qualitative information about firm operations, and information aggregation on Wikipedia retains information of neutral-point-of-view. This is different than analyst forecasts which are often subject to an upward bias.
A unified depository to store news and discussions (e.g., online message boards)	• Antweiler and Frank (2004) examined 1.5 million messages posted on online stock message boards. Collectively online messages affect trading activities in the market, suggesting that online messages provide information to investors. • Investors, however, may bear high costs of processing the large volume of information, i.e., the huge number of online messages (Gu et al. 2007).	Wikipedia contributors modify the same entry about a company, through which they aggregate information. Users, therefore, use a synthesized set of information, so they do not bear high costs of processing information.
Collective reporting by community participants	• Gu et al. (2007) examined 500,000 postings from three virtual investing-related communities (VICs). They showed that investors value high-quality postings. • "One approach VIC providers use to improve posting quality is to actively monitor and filter low-quality postings…allowing users to report abusive postings, which are then investigated manually by VIC providers" (Gu et al. 2007, p. 74).	Wikipedia removes noise in a decentralized way (i.e., by individual contributors). Also, information aggregation on Wikipedia involves, but not limited to, user screening to remove noise.
Corporate website	• Geerings et al. (2003) examined the "investor relations" webpage of 50 large companies. • On the corporate webpage of investor relations, "information available through other sources is combined to better inform investors" (Geerings et al. 2003 p. 567).	Wikipedia is a typical example of "wisdom of crowds" which is in sharp contrast to the "investor relations" section on corporate website, where companies select what information to present.

Appendix C

An Example of "Wikipedia Page History Statistics"

Page title:	Apple Inc.
Total revisions:	9,903
Number of minor edits:	2,770 (27.97%)
Number of IP edits:	3,141 (31.72%)
First edit:	03 November 2001, 13:19:06 (by 212.53.104.xxx)
Most recent edit:	16 March 2012, 19:10:01
Average time between edits:	0.38 days
Average number of edits per month:	79.56
Average number of edits per year:	954.73
Number of edits in the last day:	0
Number of edits in the last week:	2
Number of edits in the last month:	45
Number of edits in the last year:	507
Number of users:	4,530
Average edits per user:	2.19
Number of edits made by the top 10% of active users:	4,601 (46.46%)

Source: http://toolserver.org/~petrb/soxred/articleinfo/index.php?article=Apple_Inc.&lang=en&wiki=wikipedia (accessed March 22, 2012)

Appendix D

Instrumental Variable (IV) Estimation

We examine the possible endogeneity of Wikipedia modifications. First, managers may release information through Wikipedia themselves. If a manager decides to release information through Wikipedia, and if the decision to edit the entry correlates with the decision about disclosure lag (for example, because of an incentive to release information to the public), then the manager's self-editing may confound social media's effect. Second, if there exists an alternative information channel that influences both Wikipedia modifications and disclosure lags, then the identified effect cannot be attributed to Wikipedia. We employ two instrument variables for the possibly endogeneous variable *WikiMOD*.

We first consider *WikiMOD* from the previous quarter. It is correlated with *WikiMOD* of the current quarter because contributors' efforts spent on a firm's entry in a quarter are related to those in previous quarters. Since *LAG* is a variable contingent on a firm's quarterly performance, it is unlikely related to any modifications of the Wikipedia entry before the focal fiscal quarter.

We can also consider an alternative instrument as follows: For a firm's modifications in one quarter, we first identify the contributors. Then, we count the number of modifications that these contributors made on Wikipedia, but on different topics (i.e., entries other than companies). This instrument is correlated with *WikiMOD* because they both represent the degree to which the contributors add content to Wikipedia. This instrument is unlikely to be related to *LAG* because the instrument is about modifications on entries other than companies. Therefore, it is unlikely related to *LAG*.

In sum, the two instruments satisfy the two conditions for valid IV (i.e., correlated with the possibly endogenous variable while uncorrelated with regression residuals). In the below table, we follow Terza e al. (2008) and implement IV analysis using the software package R and the function coxph. We obtain results for *WikiMOD* that are highly consistent with results from our main analysis (Table 3). Hence, this IV approach gives us confidence that our main results are not due to alternative explanations.

Table D1. Results of IV Regressions (Regression specification is based on model (1) in text. Variable definitions are in Table 1.)

	Aggregation Via Wikipedia	Number of News Articles	Content of News (Positive & Negative Words)	
	(1)	(2)	(3)	(4)
Wikipedia Modifications (*WikiMOD*)	–0.1791** (0.1064)	–0.1170 (0.1095)	–0.1435* (0.1122)	–0.1153 (0.1131)
WikiMOD*DISPERSION	**–0.1395** (0.0623)**	**–0.1505*** (0.0584)**	**–0.3500*** (0.0990)**	**–0.3470*** (0.1309)**
WikiMOD*BIAS	**0.0617*** (0.0140)**	**0.0669*** (0.0128)**	**0.1033*** (0.0227)**	**0.0992*** (0.0411)**
Analyst Dispersion (*DISPERSION*)	0.1275*** (0.0498)	0.1653** (0.0921)	–0.0816 (0.1487)	–0.0484 (0.1918)
Analyst Bias (*BIAS*)	0.0392*** (0.0150)	–0.0325** (0.0169)	0.0116 (0.0285)	0.0093 (0.0464)
Number of News Articles (*NEWS*)		–0.0973* (0.0669)	–0.0957* (0.0678)	–0.0955* (0.0695)
*NEWS*DISPERSION*		0.0618 (0.0670)	0.2551** (0.1539)	0.2496 (0.2025)
*NEWS*BIAS*		–0.0109 (0.0137)	–0.0557** (0.0304)	–0.0533 (0.0474)
Negative Words (*NEGATIVE*)			–0.0821** (0.0481)	–0.0928* (0.0651)
*NEGATIVE*DISPERSION*			0.2862*** (0.1102)	0.2664* (0.1755)
*NEGATIVE*BIAS*			–0.0670* (0.0481)	–0.0518 (0.0797)
Positive Words (*POSITIVE*)				–0.0576 (0.0543)
*POSITIVE*DISPERSION*				–0.0207
				(0.1740)
*POSITIVE*BIAS*				–0.0050 (0.0784)
Earnings Variability (*VAR*)	–0.1232*** (0.0270)	–0.1098*** (0.0294)	–0.0918*** (0.0329)	–0.0923*** (0.0317)
Market Value (*MV*)	–0.2570*** (0.0872)	–0.2336*** (0.0889)	–0.2315*** (0.0855)	–0.2485*** (0.0921)
High-tech (*HIGHTECH*)	0.2446 (0.1712)	0.1491 (0.1870)	0.1934 (0.1823)	0.1622 (0.1986)
Regulation (*REG*)	–0.1929 (0.2487)	–0.1100 (0.2757)	–0.1385 (0.2835)	–0.0882 (0.2764)
Quarter 1 Dummy (*Q1*)	–0.5729*** (0.2108)	–0.5456*** (0.2127)	–0.4961** (0.2177)	–0.4320** (0.2251)
Quarter 2 Dummy (*Q2*)	–0.1254 (0.2000)	–0.1368 (0.2064)	–0.0816 (0.2111)	–0.1071 (0.2055)
Quarter 3 Dummy (*Q3*)	–0.5389*** (0.2083)	–0.5425*** (0.2069)	–0.5600*** (0.2061)	–0.4846** (0.2277)
Observations	161	161	161	161
Likelihood Ratio Test	p < 0.001	p = 0.001	p = 0.002	p = 0.004
Pseudo R-Squared	0.205	0.215	0.230	0.236

Note: Robust estimates of standard errors are reported in parentheses. Significance levels are one-tailed for directional predictions and two-tailed otherwise. *p < 0 10; **p < 0.05; ***p < 0.01. The pseudo R^2 refers to Nagelkerke's (1991) R^2.

Appendix E

Regressions Controlling for Firm Visibility Based on (1) Firm Characteristics and (2) Google Search Volume Index (SVI)

In this appendix we report how our results remain unchanged after controlling for firm visibility as we discuss in the "Additional Analysis" subsection of the "Results." Column (1) in the following table replicates the estimates we obtain in Column (5) of Table 3. In Column (2), we add common controls for firm visibility based on firm characteristics that we identify from the literature. Column (3) further considers Google search volume index (SVI). The regression specification is based on model (1) in text. Variable definitions are in Table 1.

	Base Model	Base Model + Controls for Visibility Based on Firm Characteristics	Base Model + Controls for Visibility Based on Firm Characteristics + Google SVI
	(1)	(2)	(3)
Wikipedia Modifications (*WikiMOD*)	–0.1206*	–0.2353***	0.0185
	(0.0737)	(0.1017)	(0.1744)
WikiMOD*DISPERSION	**–0.3473*** **	**–0.3797** **	**–1.1473*** **
	(0.1308)	**(0.1863)**	**(0.4592)**
WikiMOD*BIAS	**0.0990*** **	**0.0615** **	**0.4168***
	(0.0408)	**(0.0498)**	**(0.2571)**
Analyst Dispersion (*DISPERSION*)	–0.0508	–0.1836	0.2559
	(0.1890)	(0.2488)	(0.3764)
Analyst Bias (*BIAS*)	0.0091	–0.0343	0.0392
	(0.0460)	(0.0716)	(0.1394)
Number of News Articles (*NEWS*)	–0.0947*	0.1166	0.0836
	(0.0702)	(0.1034)	(0.2104)
*NEWS*DISPERSION*	0.2476	0.3783	1.7293***
	(0.1960)	(0.2409)	(0.6382)
*NEWS*BIAS*	–0.0531	0.0028	–0.3524*
	(0.0469)	(0.0499)	(0.2535)
Negative Words (*NEGATIVE*)	–0.0929*	–0.1436**	–0.1962*
	(0.0650)	(0.1008)	(0.1202)
*NEGATIVE*DISPERSION*	0.2661*	0.4015**	1.1486***
	(0.1748)	(0.2258)	(0.4829)
*NEGATIVE*BIAS*	–0.0519	–0.0730	–0.2692***
	(0.0800)	(0.1013)	(0.0963)
Positive Words (*POSITIVE*)	–0.0571	–0.0170	–0.0497
	(0.0545)	(0.0740)	(0.0762)
*POSITIVE*DISPERSION*	–0.0201	0.1279	0.5186
	(0.1730)	(0.2287)	(0.5118)
*POSITIVE*BIAS*	–0.0052	0.0032	–0.1257
	(0.0779)	(0.0695)	(0.1698)
Earnings Variability (*VAR*)	–0.0922***	–0.5048**	–0.3179
	(0.0316)	(0.2843)	(0.3472)
Market Value (*MV*)	–0.2489***	–0.1761*	–0.1872*
	(0.0913)	(0.1338)	(0.1252)
High-tech (*HIGHTECH*)	0.1664	–0.1728	–0.2020
	(0.1935)	(0.2916)	(0.3074)

	Base Model	Base Model + Controls for Visibility Based on Firm Characteristics	Base Model + Controls for Visibility Based on Firm Characteristics + Google SVI
	(1)	(2)	(3)
Regulation (*REG*)	–0.0957 (0.2587)	–0.0780 (0.6797)	–0.1151 (0.5917)
Quarter 1 Dummy (*Q1*)	–0.4338** (0.2211)	–0.5117** (0.2871)	–0.4272* (0.2809)
Quarter 2 Dummy (*Q2*)	–0.1051 (0.2069)	–0.2637 (0.2751)	–0.1654 (0.3294)
Quarter 3 Dummy (*Q3*)	–0.4858** (0.2251)	–0.6049*** (0.2881)	–0.9893*** (0.2653)
NYSE Listing (*NYSE*)		–0.5192** (0.3326)	–0.8105** (0.3600)
Analyst Following (*ANALYST*)		–0.0278** (0.0171)	–0.0292** (0.0173)
Institutional Ownership (*IO*)		0.2344 (0.5275)	–0.2389 (0.6926)
Firm Profitability (*ROA*)		–2.1351** (1.1372)	–1.7059* (1.1965)
Firm Age (*AGE*)		–0.0024 (0.0052)	–0.0024 (0.0047)
Advertising Expenditure (*AD*)		0.0001 (0.0004)	0.0002 (0.0003)
Advertising Dummy (*AD_DUM*)		0.2534 (0.2686)	0.0769 (0.3011)
Search Volume Index (*SVI*)			0.0664 (0.0617)
Observations	161	147	119
Likelihood Ratio Test	p = 0.003	p = 0.001	p = 0.000
Pseudo R-Squared	0.235	0.315	0.403

Note: Robust estimates of standard errors are reported in parentheses. Significance levels are one-tailed for directional predictions and two-tailed otherwise. *p < 0.10; **p < 0.05; ***p < 0.01. The pseudo R^2 refers to Nagelkerke's (1991) R^2.

Appendix F

Economic Significance of WikiMOD

We conduct an additional analysis to evaluate the economic significance of WikiMOD. The analysis is based on column (2) of Table 4. We estimate separate regressions by conditioning WikiMOD at its LOW, MEAN, and HIGH levels, defined as the sample mean minus one standard deviation, the sample mean, and the sample mean plus one standard deviation, respectively. For instance, when we subtract the LOW level from WikiMOD, the estimated coefficient on BIAS represents its effect expected at the LOW level of WikiMOD (Aiken and West 1991). As reported in column (1) of the table below, the coefficient on BIAS is –0.0062. We multiply the coefficient by one standard deviation of BIAS (3.0376), and the result (–1.88%) indicates the market reaction induced by an increasing BIAS (increase by one standard deviation). The market reaction becomes –0.73% in column (2) when WikiMOD is at the MEAN level, and it becomes statistically nonsignificant in column (3) when WikiMOD is at the HIGH level. A comparison of columns (1) and (2) shows that, when WikiMOD moves from the LOW to MEAN level, this increase in information aggregation is associated with a change in market reaction by 1.15% (=(–0.73%) – (–1.88%)). This indicates an economically meaningful impact of information aggregation.

Additional Analysis on Market Reaction			
	WikiMOD = LOW	**WikiMOD = MEAN**	**WikiMOD = HIGH**
	(1)	**(2)**	**(3)**
Coefficient on BIAS	–0.0062*** (p = 0.0013)	–0.0024* (p = 0.0818)	0.0014 (p = 0.2886)
Market reaction due to analyst bias (i.e., the abnormal return induced by change of BIAS by one standard deviation)			
	–1.88%	–0.73%	N.S.

Note: The LOW, MEAN, and HIGH levels of WikiMOD are equal to sample mean minus one standard deviation, sample mean, and sample mean plus one standard deviation, respectively. *p < 0.10; **p < 0.05; ***p < 0.01; N.S. = nonsignificant.

Appendix G

Robustness to Samples and Measures

Several additional tests show that our results are robust to alternative samples and measures. We present the regression results in the following two tables, one for management disclosure lag and the other for investor reaction.

In each of the two tables, for convenience, column (1) shows our base model as presented in text. In column (2), we sort all observations by the number of analysts' forecasts for EPS, and we keep only observations between the 5th and 95th percentiles. Because *BIAS* and *DISPERSION* are computed based on analyst forecasts, using this subsample helps remove extreme cases.

Columns (3) and (4) examine the extent to which edit wars may affect our result. Since Wikipedia is an open platform and anyone is allowed to post and modify entries, some entries can be very contentious. It is necessary to make sure that our results are not driven by the edit wars. Our first check (reported in Column 3) is based on the "revert" activities. For each modification j by user i, m_{ij}, on Wikipedia, we search for a future (within a month) edit, m_{kj}, by user k that reverts the contribution. Then we search again (within a month from the revert) for i's further revert of k's modification. When we find such patterns we remove the second and third modifications. If we find additional reverts in the same war (i.e., within one month k reverts i's work again), we also remove these additional modifications. The result of this check is a new data set that contains fewer modifications in each quarter with all "back and forth" edits removed. Our second check (reported in Column 4) adopts a more aggressive approach. In this check, we only consider the first contribution by each contributor in each financial period. With this approach, we not only eliminate all future edit wars between this user and others, but also remove additional edits by the same user. The first check yields highly significant effect of WikMOD, and the second also gives significant effect, though relatively weaker (because it is an aggressive approach). Overall, our results regarding information aggregation over Wikipedia remain qualitatively unchanged.

Column (5) uses the number of words added, instead of the number of times an entry is modified, to proxy for information aggregation on Wikipedia. Column (6) measures *NEWS* by the total number of words in all news stories, instead of the total number of news stories, about a firm. Column (7) deflates *BIAS* and *DISPERSION* by stock price (Francis e al. 1997). These two tables demonstrate qualitatively the same results for our hypothesis testing.

Robustness Checks: Wikipedia and Management Disclosure (Regression specification is based on model (1) in text. Variable definitions are in Table 1.)

	Base model	Number of forecasts: 5th–95th percentile	*WikiMOD*: removing "reverts"	*WikiMOD*: keeping the first edit only	*WikiMOD*: number of words added	*NEWS*: number of news words	*BIAS* & *DISPERSION*: deflated by stock price
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Wikipedia Modifications (*WikiMOD*)	–0.1206* (0.0737)	–0.1316* (0.0848)	–0.1240** (0.0749)	–0.1437* (0.1006)	–0.0869** (0.0534)	–0.0991* (0.0754)	–0.2185** (0.1027)
***WikiMOD*DISPERSION**	**–0.3473*** (0.1308)**	**–0.2700** (0.1466)**	**–0.3447*** (0.1313)**	**–0.4143*** (0.1710)**	**–0.2456*** (0.0750)**	**–0.3158*** (0.0955)**	**–52.6100* (33.1300)**
***WikiMOD*BIAS**	**0.0990*** (0.0408)**	**0.0801** (0.0430)**	**0.0991*** (0.0409)**	**0.1244* (0.0780)**	**0.0544** (0.0259)**	**0.0845*** (0.0283)**	**16.7800* (13.0000)**
Analyst Dispersion (*DISPERSION*)	–0.0508 (0.1890)	0.0115 (0.2211)	–0.0464 (0.1884)	–0.0694 (0.2204)	0.0093 (0.1728)	–0.0457 (0.2379)	21.3800 (51.7400)
Analyst Bias (*BIAS*)	0.0091 (0.0460)	–0.0105 (0.0530)	0.0080 (0.0457)	0.0017 (0.0534)	0.0094 (0.0492)	0.0001 (0.0524)	2.7360 (15.5500)
Number of News Articles (*NEWS*)	–0.0947* (0.*702)	–0.0933* (0.0697)	–0.0935* (0.0702)	–0.1033* (0.0712)	–0.0911* (0.0667)	–0.0230 (0.0350)	–0.0734 (0.0716)
*NEWS*DISPERSION*	0.2476 (0.1960)	0.1728 (0.1920)	0.2465 (0.1960)	0.1826 (0.1878)	0.3519** (0.1646)	0.2105*** (0.0836)	47.2500 (38.6900)
*NEWS*BIAS*	–0.0531 (0.0469)	–0.0318 (0.0502)	–0.0531 (0.0470)	–0.0483 (0.0585)	–0.0529 (0.0433)	–0.0222 (0.0264)	–7.6860 (6.1500)
Negative Words (*NEGATIVE*)	–0.0929* (0.0650)	–0.0582 (0.0672)	–0.0925* (0.0651)	–0.0944 (0.0818)	–0.0994* (0.0681)	–0.0788* (0.0553)	–0.0992* (0.0618)
*NEGATIVE*DISPERSION*	0.2661* (0.1748)	0.2050 (0.1891)	0.2649* (0.1752)	0.2345 (0.1964)	0.3971** (0.1889)	0.2647** (0.1193)	252.6000** (124.7000)
*NEGATIVE*BIAS*	–0.0519 (0.0800)	–0.0130 (0.0840)	–0.0520 (0.0801)	–0.0503 (0.1095)	–0.0706 (0.1008)	–0.0208 (0.0682)	–35.8200** (20.0900)
Positive Words (*POSITIVE*)	–0.0571 (0.0545)	–0.0693 (0.0561)	–0.0569 (0.0545)	–0.0665 (0.0592)	–0.0697 (0.0617)	–0.0709** (0.0420)	–0.0020 (0.1735)
*POSITIVE*DISPERSION*	–0.0201 (0.1730)	0.1156 (0.2101)	–0.0205 (0.1731)	0.0350 (0.1877)	0.1762 (0.2278)	0.0807 (0.1524)	–7.9730 (73.9100)
*POSITIVE*BIAS*	–0.0052 (0.0779)	–0.0481 (0.0891)	–0.0052 (0.0779)	–0.0290 (0.0858)	–0.0604 (0.1034)	–0.0357 (0.0553)	8.8550 (33.7100)
Earnings Variability (*VAR*)	–0.0922*** (0.0316)	–0.0938*** (0.0316)	–0.0924*** (0.0316)	–0.0769** (0.0402)	–0.0993*** (0.0357)	–0.1100*** (0.0340)	–0.1569*** (0.0597)
Market Value (*MV*)	–0.2489*** (0.0913)	–0.1432 (0.1171)	–0.2494*** (0.0909)	–0.2537*** (0.0939)	–0.2655*** (0.0957)	–0.2803*** (0.0966)	–0.2131** (0.1071)
High-tech (*HIGHTECH*)	0.1664 (0.1935)	0.2802 (0.2177)	0.1674 (0.1934)	0.1932 (0.2005)	0.1447 (0.1962)	0.2148 (0.1849)	0.1157 (0.2353)
Regulation (*REG*)	–0.0957 (0.2587)	–0.0866 (0.2132)	–0.1002 (0.2587)	–0.0569 (0.2591)	0.0698 (0.2912)	–0.1155 (0.3063)	–0.3109 (0.2278)
Quarter 1 Dummy (*Q1*)	–0.4338** (0.2211)	–0.4393** (0.2305)	–0.4318** (0.2206)	–0.4346** (0.2216)	–0.4177** (0.2075)	–0.4149** (0.2157)	–0.3727** (0.2220)
Quarter 2 Dummy (*Q2*)	–0.1051 (0.2069)	–0.1740 (0.2526)	–0.1047 (0.2069)	–0.0743 (0.2075)	–0.1053 (0.2085)	–0.0608 (0.2103)	–0.3114* (0.2210)
Quarter 3 Dummy (*Q3*)	–0.4858** (0.2251)	–0.4705** (0.2405)	–0.4855** (0.2245)	–0.4636** (0.2308)	–0.4221** (0.2275)	–0.4558** (0.2303)	–0.6235*** (0.2331)
Observations	161	145	161	161	161	161	161
Likelihood Ratio Test	p = 0.003	p = 0.030	p = 0.003	p = 0.005	p = 0.003	p = 0.001	p = 0.002
Pseudo R-Squared	0.235	0.213	0.236	0.228	0.238	0.251	0.241

Note: Robust estimates of standard errors are reported in parentheses. Significance levels are one-tailed for directional predictions and two-tailed otherwise. *p < 0.10; **p < 0.05; ***p < 0.01. The pseudo R^2 refers to Nagelkerke's (1991) R^2.

Robustness Checks: Investor Reaction to Disclosure (Regression specification is based on model (2) in text. Variable definitions are in Table 1.)

	Base model	Number of forecasts: 5th–95th percentile	*WikiMOD*: removing "reverts"	*WikiMOD*: keeping the first edit only	*WikiMOD*: number of words added	*NEWS*: number of news words	*BIAS* & *DISPERSION*: deflated by stock price
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Analyst Bias (*BIAS*)	–0.0827*** (0.0321)	–0.0768*** (0.0320)	–0.0829*** (0.0321)	–0.0837*** (0.0329)	–0.0812*** (0.0338)	–0.1021*** (0.0250)	–29.7532*** (7.0075)
WikiMOD	–0.0084 (0.0099)	–0.0102 (0.0102)	–0.0082 (0.0100)	–0.0104 (0.0121)	–0.0045 (0.0062)	–0.0083 (0.0098)	–0.0087 (0.0099)
***WikiMOD*BIAS**	**0.0035*** (0.0015)**	**0.0036*** (0.0015)**	**0.0035*** (0.0015)**	**0.0042* (0.0029)**	**0.0026** (0.0015)**	**0.0036*** (0.0014)**	**0.9843*** (0.2978)**
Market Value (*MV*)	0.0057 (0.0096)	0.0092 (0.0094)	0.0056 (0.0096)	0.0051 (0.0095)	0.0062 (0.0103)	0.0056 (0.0089)	0.0069 (0.0095)
*MV*BIAS*	0.0046*** (0.0017)	0.0043*** (0.0017)	0.0046*** (0.0017)	0.0046*** (0.0018)	0.0046*** (0.0018)	0.0065*** (0.0015)	1.5831*** (0.3737)
NEWS	0.0075 (0.0088)	0.0073 (0.0088)	0.0075 (0.0088)	0.0074 (0.0087)	0.0059 (0.0084)	0.0033 (0.0036)	0.0038 (0.0086)
*NEWS*BIAS*	–0.0039*** (0.0013)	–0.0039*** (0.0013)	–0.0039*** (0.0013)	–0.0038*** (0.0016)	–0.0041*** (0.0013)	–0.0030*** (0.0007)	0.0571 (0.2644)
NEGATIVE	0.0025 (0.0045)	0.0027 (0.0046)	0.0026 (0.0045)	0.0027 (0.0048)	0.0025 (0.0040)	0.0028 (0.0046)	–0.0019 (0.0027)
*NEGATIVE*BIAS*	–0.0100*** (0.0042)	–0.0099*** (0.0042)	–0.0100*** (0.0042)	–0.0103** (0.0049)	–0.0096*** (0.0035)	–0.0100*** (0.0034)	–0.9694*** (0.2206)
POSITIVE	–0.0016 (0.0036)	–0.0015 (0.0036)	–0.0015 (0.0036)	–0.0015 (0.0037)	–0.0017 (0.0036)	–0.0009 (0.0039)	–0.0034 (0.0027)
*POSITIVE*BIAS*	0.0000 (0.0033)	0.0000 (0.0033)	0.0000 (0.0033)	–0.0003 (0.0032)	–0.0006 (0.0030)	–0.0009 (0.0025)	1.2430*** (0.5228)
Observations	161	154	161	161	161	161	161
R–Squared	0.1498	0.1657	0.1494	0.1464	0.1505	0.1613	0.1674

Note: Robust estimates of standard errors are reported in parentheses. Significance levels are one-tailed for directional predictions and two-tailed otherwise. *p < 0.10; **p < 0.05; ***p < 0.01.

References

Abarbanell, J. S., Lanen, W. N., and Verrecchia, R. E. 1995. "Analysts' Forecasts as Proxies for Investor Beliefs in Empirical Research," *Journal of Accounting and Economics* (20:1), pp. 31-60.

Aiken, L. S., and West, S. G. 1991. *Multiple Regression: Testing and Interpreting Interactions*, Newbury Park, CA: Sage Publications.

Antweiler, W., and Frank, M. Z. 2004. "Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards," *Journal of Finance* (59:3), pp. 1259-1294.

Francis, J., Hanna, J. D., and Philbrick, D. R. 1997. "Management Communications with Securities Analysts," *Journal of Accounting and Economics* (24:3), pp. 363-394.

Geerings, J., Bollen, L. H. H., and Hassink, H. F. D. 2003. "Investor Relations on the Internet: A Survey of the Euronext Zone," *European Accounting Review* (12:3), pp. 567-579.

Gu, B., Konana, P., Rajagopalan, B., and Chen, H-W. M. 2007. "Competition among Virtual Communities and User Valuation: The Case of Investing-Related Communities," *Information Systems Research* (18:1), pp. 68-85.

Healy, P. M., and Palepu, K. G. 2001. "Information Asymmetry, Corporate Disclosure, and the Capital Markets: A Review of the Empirical Disclosure Literature," *Journal of Accounting and Economics* (31:1-3), pp. 405-440.

Karamanou, I., and Vafeas, N. 2005. "The Association Between Corporate Boards, Audit Committees, and Management Earnings Forecasts: an Empirical Analysis," *Journal of Accounting Research* (43:3), pp. 453-486.

Kasznik, R., and Lev, B. 1995. "To Warn or Not to Warn: Management Disclosures in the Face of an Earnings Surprise," *The Accounting Review* (70:1), pp. 113-134.
Lang, M. H., and Lundholm, R. J. 1996. "Corporate Disclosure Policy and Analyst Behavior," *The Accounting Review* (71:4), pp. 467-492.
Nagelkerke, N. J. D. 1991. "A Note on a General Definition of the Coefficient of Determination," *Biometrika* (78:3), pp. 691-692.
Roulstone, D. T. 2003. "Analyst Following and Market Liquidity," *Contemporary Accounting Research* (20:3), pp. 552-578.
Skinner, D. J. 1994. "Why Firms Voluntarily Disclose Bad News," *Journal of Accounting Research* (32:7-8), pp. 38-61.
Skinner, D. J. 1997. "Earnings Disclosures and Stockholder Lawsuits," *Journal of Accounting and Economics* (23:3), pp. 249-282.
Terza, J. V., Basu, A., and Rathouz, P. J. 2008. "Two-Stage Residual Inclusion Estimation: Addressing Endogeneity in Health Econometric Modeling," *Journal of Health Economics* (27:3), pp. 531-543.