# Informing the Market: The Effect of Modern Information Technologies on Information Production*

Meng Gao† Jiekun Huang‡

This Draft: August 2017

*We thank Heitor Almeida, Ferhat Akbas, Mara Faccio, Itay Goldstein, Jack He, Jianfeng Hu, Paul Irvine, Paul Koch, Tim Johnson, Mathias Kronlund, Laura Li, Lei Li, Wei Li, Vikram Nanda, Neil Pearson, Joshua Pollet, Jeff Pontiff, Matt Spiegel, Avanidhar Subrahmanyam, Alexei Tchistyi, Chengwei Wang, Rong Wang, Stijn Van Nieuwerburgh, Scott Weisbenner, Hao Zhang, Mengxin Zhao, Wei Zhu, and seminar participants at the 2017 China International Conference in Finance, the Securities and Exchange Commission, Kansas University, Nanjing University, University of Illinois at Urbana-Champaign, and Xiamen University for comments and helpful discussions. Zhongnan Xiang provided excellent research assistance. We retain responsibility for any remaining errors.

†Department of Finance, University of Illinois at Urbana-Champaign; e-mail: `menggao4@illinois.edu`.

‡Department of Finance, University of Illinois at Urbana-Champaign; e-mail: `huangjk@illinois.edu`.

# Informing the Market: The Effect of Modern Information Technologies on Information Production

**Abstract**

Modern information technologies have fundamentally changed how information is disseminated in financial markets. Using the staggered implementation of the EDGAR system in 1993–1996 as a shock to information dissemination technologies, we find evidence that internet dissemination of corporate information increases information production by corporate outsiders. Specifically, trades by individual investors in a stock become more informative about future stock returns after the stock becomes subject to mandatory filing on EDGAR. This effect is driven primarily by investors who have access to the internet. The amount and accuracy of information produced by sell-side analysts increase following the EDGAR implementation. Market responses to analyst revisions also become stronger after a firm becomes an EDGAR filer. Furthermore, stock pricing efficiency improves after the EDGAR implementation. Overall, these results suggest that greater and broader information dissemination facilitated by modern information technologies improves information production and stock pricing efficiency.

JEL Classification: G12, G14

Keywords: Information production, internet, informational efficiency, individual investors, financial analysts

# 1 Introduction

A well-functioning securities market requires that a broad base of investors have access to corporate information and process such information to promote price efficiency and facilitate capital formation. The advent of modern information technologies has dramatically changed how information is disseminated in financial markets by making a large amount of information available to a broad base of financial market participants in real time at low costs. Investors nowadays can get immediate access to corporate disclosures as well as other market participants' opinions disseminated through the internet to gain insights into firms' fundamental value. In the past few decades, a series of regulatory changes have been made to make use of modern information technologies to improve the accessibility of information to the public. For example, the SEC launched the EDGAR system in 1993 to move corporate disclosure from the print era to the digital age, and in 2013 the SEC allowed public companies to use social media sites to announce key information to investors. Yet, despite the dramatic changes brought about by modern information technologies in the dissemination of information, the effects of modern information technologies on information production by market participants remain underexplored.

Modern information dissemination technologies can have two opposite effects on information production by corporate outsiders. On the one hand, more timely and extensive dissemination of information facilitated by modern information technologies may crowd out information production by market participants. This may arise because of at least three reasons. First, when information is widely disseminated (i.e., more investors become informed about the information), prices may reveal more information (Grossman and Stiglitz, 1980). Since information processing takes time, the advantage of becoming an information processor decreases, resulting in reduced intensity of information processing activities (e.g., Dugast and Foucalt, 2017). Second, since widely disseminated public information can serve as a coordinating device for investors' beliefs, greater dissemina-

tion of information may cause investors to overweight public information and underweight private information. This may reduce stock price efficiency when the precision of private information is high (Morris and Shin, 2002; Amador and Weill, 2010). For example, Shiller (2006) argues that mass dissemination of information by the media may negatively impact the efficiency of asset prices by creating similar thinking among large groups of people, causing "an avoidance of individual assessment of quantitative data". Third, the availability of large amounts of information may create an information overload problem (e.g., Barber and Odean, 2001; Shapiro and Varian, 1999), reducing the attention allocated to information processing. Relatedly, when information is widely disseminated because of technological progresses, the marginal recipient of information might be less able to process it correctly (D'Avolio, Gildo, and Shleifer, 2002). These considerations suggest that the advent of modern information technologies may dampen the incentive to produce information and therefore reduce pricing efficiency.

On the other hand, there could be a crowding-in effect in that greater dissemination of information and the ensuing decline in information acquisition costs may induce greater intensity of information production by market participants. This may arise because, other things equal, the net profit information producers derive from producing information increases as the cost of information production declines (see, e.g., Verrecchia, 1982; Kim and Verrecchia, 1994). As Verrecchia (1982) argues, "[a]s technological improvements permit more information to be obtained at the same cost, traders' increased information acquisition results in prices revealing more information." Thus, greater dissemination of information facilitated by modern information technologies may increase the incentives of market participants to produce information and, as a result, improve pricing efficiency.

Therefore, the net effect of modern information technologies on information production is ultimately an empirical question. In this paper, we investigate this question by exploiting the staggered implementation of the EDGAR system in 1993–1996 as a shock to informa-

tion dissemination technologies. Before the implementation of EDGAR in 1993, publicly traded corporations had to transmit multiple paper copies of filings to the SEC, and the three public reference rooms of the SEC (in Washington DC, New York, and Chicago) were the ultimate sources of these filings. The SEC introduced the EDGAR system in February 1993 to enable companies to file electronically to facilitate the dissemination of information to the public in a timely manner. Importantly, the SEC required that all public companies began filing to EDGAR in 10 discrete groups, with companies in the first group starting to file on EDGAR in April 1993 and companies in the last group starting in May 1996. Thus, the staggered nature of the implementation of the EDGAR system provides a set of counterfactuals for how information production would have changed in the absence of a change in information dissemination technologies and so allows us to disentangle the effect of information technologies on information production from other confounding factors.

In this paper, we focus on information production by two groups of market participants, namely individual investors and sell-side financial analysts, for two reasons. First, both individual investors and sell-side analysts play the role of information producers in the financial markets. Specifically, there is growing evidence suggesting that individual investors produce information about stocks (e.g., Kaniel, Liu, Saar, and Titman, 2012; Kelley and Tetlock, 2013, 2017).[1] There is also a large literature on the role of sell-side financial analysts as information intermediaries in the stock market (see, e.g., Healy and Palepu, 2001, for a comprehensive review of this literature). Second, for both groups, we can directly observe their behavior at a relatively high frequency, which enables us to construct proxies of information production around specific points in time. In particular, we use the trading data from a large discount brokerage database (the LDB dataset) used

[1] As Kaniel, Liu, Saar, and Titman (2012) and Kelley and Tetlock (2013) argue, there are at least two reasons why individuals' trades may contain information. First, while each individual investor may have only noisy information, aggregating the information through the trades of a large number of individuals may result in signals that are relatively precise. Second, individuals might be especially well positioned to exploit private information through their trades, because they tend to trade in small quantities and are not subject to the agency problems, career concerns, or liquidity constraints that institutional managers typically face.

by Barber and Odean (2000) and analyst forecasts data from I/B/E/S database.[2] More important for our purposes, the LDB dataset allows us to identify investors with access to the internet who likely benefit directly from the EDGAR shock.

Using a comprehensive set of firms covered in the phase-in schedule of the EDGAR system, we find evidence suggesting that the crowding-in effect dominates the crowding-out effect for both individual investors and sell-side analysts. Specifically, we find that individual investors' net buying following an earnings announcement of a stock becomes more informative about future stock returns after the stock becomes subject to mandatory filing on EDGAR. The economic magnitude is non-trivial. For example, a one-standard-deviation increase in net buying by individual investors during the 20 trading days post-announcement is associated with 1.649 percentage points higher subsequent 12-month cumulative abnormal returns after the stock becomes an EDGAR filer than before, which is economically nontrivial considering that the 12-month CAR has a mean of 2.897% and a standard deviation of 49.574 percentage points. Importantly, we are able to identify which investors have access to the internet based on whether they placed a trade through the internet in the past. While internet users account for only 12% of the investors in our sample, the increase in stock return predictability after the EDGAR implementation is driven primarily by trades placed by these investors. These results suggest that the crowding-in effect dominates the crowding-out effect, thereby resulting in more information production by individual investors, especially those with ready access to information on the internet.

Turning to sell-side analysts, we find evidence suggesting that both the amount and the accuracy of information produced by sell-side analysts increase following the EDGAR implementation. Specifically, the number of analysts covering a firm increases and the forecast accuracy of analysts improves after the firm becomes subject to mandatory filing

[2] We do not examine information production by institutional investors, because the 13F institutional holdings data, used in institutional investor studies, provide quarterly snapshots of institutions' holdings and hence do not allow us to infer institutions' trades at a relatively high frequency in a specific window.

on EDGAR. In terms of economic magnitudes, the average firm experiences an increase of 0.234 analysts post-EDGAR, which is large considering that the mean and standard deviation of the number of analysts covering a firm are 2.488 and 3.917, respectively. Similarly, the average firm experiences an increase of 0.00119 in analysts' forecast accuracy, representing 13.2% (1.5%) of the mean (standard deviation) of the variable. Perhaps more important, stock market responses to analysts' revisions become significantly stronger after the firm becomes an EDGAR filer, suggesting that the market perceives analyst research as more informative. These results are consistent with the crowding-in effect dominating the crowding-out effect for sell-side analysts.

We conduct several additional tests to assess the robustness of our results. First, we include cohort-specific time trends as additional controls in the regressions. In this case, the identification of the effects of EDGAR implementation comes from whether the implementation leads to deviations from preexisting cohort-specific trends. We find that the observed effects continue to hold with the inclusion of these time trends. Second, our results are robust to redefining the post-EDGAR period for the first four groups of firms to start from January 1994 when the EDGAR system became publicly available to internet users without additional charges. Third, we conduct a placebo test using a period preceding the actual EDGAR implementation. We find insignificant changes in information production around these pseudo-events, alleviating the concern that the observed effects may be driven by unobserved characteristics that are generally correlated with both the relative timing of EDGAR implementation and changes in information production. Last, to address the concern that assignment to groups is not random, we construct a control sample using a propensity-score matching approach. Specifically, for each firm that switches from being a non-filer to an EDGAR filer in a given month, we identify a non-switching firm that has statistically the same size, book-to-market, profitability, leverage, R&D, and etc. We find that the above results continue to hold, suggesting that the observed effects are not driven by firm characteristics that are associated with assignment to groups.

Last but not least, we examine the effect of EDGAR implementation on stock price efficiency. Using various measures of pricing efficiency, namely stock price synchronicity (Morck, Yeung, and Yu, 2000), the absolute value of stock return autocorrelation, and the standard deviation of the pricing error of Hasbrouck (1993), we find evidence that EDGAR implementation improves stock price efficiency. This result suggests that internet dissemination of corporate information not only increases information production by market participants, but also improves pricing efficiency.

Our paper is the first in the literature to provide causal evidence on the effect of the EDGAR system on information production. As the first paper to exploit the staggered timing of the implementation of the EDGAR system, our study highlights the impacts of technological advances on information dissemination and production in financial markets. Our findings have important policy implications. Government regulations that aim to promote the availability of fundamental information, such as earnings reports and other corporate releases, to a broad base of investors in real time are likely to enhance the resource allocation role of financial markets by increasing the supply of information by corporate outsiders.

The rest of the paper is organized as follows. Section 2 provides a review of related research as well as background information on the implementation of EDGAR. Section 3 describes the data and summary statistics. Section 4 presents the empirical results, and Section 5 concludes.

# 2 Literature Review and Institutional Background

## 2.1 Related literature

Our paper contributes to three strands of literature. The first is the theoretical literature on costly information acquisition in financial markets. Existing theories provide ambiguous predictions regarding the effects of greater information dissemination brought about by modern information technologies on information production and pricing efficiencies (see, e.g., Grossman and Stiglitz, 1980; Verrecchia, 1982; Kim and Verrecchia, 1994; Barlevy and Veronesi, 2002; Dugast and Foucalt, 2017).[3] On the one hand, greater dissemination of information and lower information acquisition costs brought about by modern information technologies may crowd out information production by market participants. For example, in the model of Dugast and Foucalt (2017), investors can acquire and trade on two types of costly information, namely raw information, which is noisy but can be immediately traded upon, and processed information, which is more precise but takes time to process. They argue that when raw information is precise enough, lowering the cost of raw information leads to more trades on raw information and reduces the value of processed information. In this case, the decline in the cost of raw information due to technological advances can reduce the incentive to produce information and hence lower the informativeness of prices in the long run. On the other hand, Verrecchia (1982) contends that as information acquisition becomes less costly, the amount of costly diverse information investors acquire increases, which leads to more informative prices. Thus, whether modern information technologies facilitate or dampen information production is an empirical question. By exploiting the staggered implementation of the EDGAR system as plausibly exogenous shocks to information technologies, our paper provides evidence suggesting that greater and broader dissemination of fundamental information facilitated by modern

[3] See Goldstein and Yang (2017) for a thorough review of the theoretical literature on the effects of information disclosure on market quality and information production in financial markets.

information dissemination technologies positively impacts information production by market participants. Our findings are consistent with the crowding-in effect dominating the crowding-out effect, highlighting technological advances in information dissemination as a contributing factor to the informational efficiency of stock prices.

The second literature our paper is related to is the empirical literature on the role of corporate outsiders as information producers in financial markets. Recent studies find evidence suggesting that individual investors produce information about stocks (e.g., Kaniel, Liu, Saar, and Titman, 2012; Kelley and Tetlock, 2013, 2017). For instance, Kaniel, Liu, Saar, and Titman (2012) show that intense buying (selling) by individual investors in the 10 days prior to an earnings announcement predicts large positive (negative) abnormal returns following the earnings announcement. Coval, Hirshleifer, and Shumway (2005), Ivković and Weisbenner (2005), and Ivković, Sialm, and Weisbenner (2008) use the LDB dataset and find evidence suggesting that some individual investors possess an informational advantage about stocks. Also, it has been well established that sell-side financial analysts are among the most important information intermediaries in the stock market (see, e.g., Bhushan, 1989; O'Brien and Bhushan, 1990; Lang and Lundholm, 1996; Healy and Palepu, 2001). Our paper contributes to this literature by focusing on the effect of a plausibly exogenous shock to information dissemination technologies on information production by corporate outsiders. Our findings highlight the importance of timely and broad dissemination of information in influencing the extent of information production by individual investors and financial analysts.

Last, our paper is related to an emerging literature examining the effects of the information dissemination process on financial market outcomes (see, e.g., Engelberg and Parsons, 2011; Dougal, Engelberg, Garcia, and Parsons, 2012; Peress, 2014; Blankespoor, Miller, and White, 2014; Dong, Li, Lin, and Ni, 2016). For example, Engelberg and Parsons (2011) use extreme weather events as exogenous shocks that disrupt the delivery of

daily newspapers to identify the causal impact of media coverage on investor trading. Using newspaper strikes as shocks to information dissemination by the media, Peress (2014) finds evidence that the media improve stock pricing efficiency. Blankespoor, Miller, and White (2014) and Dong, Li, Lin, and Ni (2016) show that the adoption of the eXtensible Business Reporting Language (XBRL) increases the bid-ask spread and reduces stock return synchronicity, respectively. Our paper adds to this literature by focusing on the effects of a technological/regulatory shock, namely the implementation of the EDGAR system, that significantly increases the accessibility of corporate information to a broad base of investors.

The implementation of the EDGAR system is relatively underexplored. The only two papers that examine financial market outcomes before and after EDGAR implementation are Asthana and Balsam (2001) and Asthana, Balsam, and Sankaraguruswamy (2004). Asthana and Balsam (2001) show that market reactions to 10-K reports become stronger after a firm becomes an EDGAR filer. Asthana, Balsam, and Sankaraguruswamy (2004) use the TAQ data and show that the volume of small trades around 10-K filings increases after a company switches to electronic filings on EDGAR. They also show in a univariate setting that net buying of small trades around 10-K filings becomes more positively correlated with subsequent short-term (i.e., five-day) stock returns when the 10-K reports are filed electronically through EDGAR for the first time.[4] Neither of these papers, however, exploits the staggered nature of the EDGAR implementation. Moreover, because both papers focus on short-term stock returns, the results can be interpreted as consistent with an attention effect (in which greater dissemination of information through EDGAR causes investors to respond in a naïve fashion). In contrast, our paper provides more direct tests of an information effect (in which electronic dissemination of information facilitates information acquisition by market participants) by examining relatively long-term stock

[4] The multivariate regressions in Asthana, Balsam, and Sankaraguruswamy (2004) include interaction terms combining an indicator for initial EDGAR filers and changes in market capitalization, which do *not* allow for testing the unconditional effect of the EDGAR implementation.

returns and analyzing actual trades (as opposed to inferred trades) of investors, especially those with access to the internet. Our paper also conducts tests on financial analysts' information production around the EDGAR implementation.

## 2.2 The implementation of the EDGAR system

Prior to the implementation of EDGAR in 1993, public firms had to transmit multiple paper copies of filings to the SEC by mail, by courier, or by personal delivery. These paper copies of filings would then be filed in the SEC public reference rooms for public viewing after being reviewed by the SEC examiners. Thus, the three locations of the public reference rooms (in Washington DC, New York, and Chicago) are the ultimate source of corporate disclosures for the investing public. Since the paper filings can be inspected by one reader at a time, the limited availability of paper copies for each filing (typically one or two copies at each location) makes it hard for the information to reach a large audience. Moreover, the large volume of filings being filed with the SEC makes it difficult for the investing public to find and analyze specific data. For example, a *New York Times* (1982) article quotes reference room users as saying that,"[i]t's just incredible the number of problems you can run into trying to find something you need. [...] The place can be a zoo."

To meet the objective of providing information to the public in a timely and efficient manner, the SEC developed an automated system, the Electronic Data Gathering, Analysis and Retrieval (EDGAR) system, for electronic submission of company filings. The main goal of EDGAR was to enable companies to file electronically to facilitate the dissemination of information to the public in real time. By disseminating information through the internet, the EDGAR system increases the accessibility of corporate filings and thus significantly reduces corporate outsiders' information acquisition costs. Moreover, corporate outsiders can more readily process information in electronic filings than in paper filings,

e.g., by using the search function to locate specific information in an electronic document.

On February 23, 1993, the SEC issued rules requiring corporate filings be transmitted electronically to EDGAR. These rules specified a phase-in schedule for all public firms to begin filing to EDGAR. Specifically, the rules categorized public firms into 10 groups and each group was phased in at different times.[5] Companies in the first group, i.e., Group CF–01, had to commence mandated electronic filing to EDGAR in April 1993, and those in the last group, i.e., Group CF–10, became EDGAR filers in May 1996. The time-lapse between the starting date of one group and that of the next group ranges from three to six months. Figure 1 plots the number of firms that are subject to mandatory filing through EDGAR in each point in time from January 1993 through December 1996. Appendix A provides a timetable for the implementation of the EDGAR system.

[Insert Figure 1 about here]

# 3 Data and Summary Statistics

We retrieve the list of firms on the phase-in schedule for the implementation of the EDGAR system from Appendix B of SEC Release No. 33-6977 (released on February 23, 1993).[6] The list provides the firm name, CIK, and group number (from 1 through 10). We match the companies on the list to Compustat by CIK and company name. We are able to match $5,212$ firms that are on the phase-in schedule and have financial information available in Compustat as of January 31, 1993, i.e., the month-end immediately before the release

[5] We filed a Freedom of Information Act request to the SEC for information on how companies are assigned to different groups. The SEC responded that their staff "conducted a thorough search of the SEC's various systems of records, but did not locate or identify any information responsive to [the] request."

[6] We code a firm as being subject to mandatory filing to EDGAR based on the phase-in schedule in SEC Release No. 33-6977. According to the Release, the SEC may, in its discretion, grant or deny a request by a firm to participate in a phase-in group other than the group assigned in the phase-in schedule. It is worth noting that if the actual implementation date of a firm is different from that specified in the phase-in schedule, it will result in misclassifications in our coding and bias against finding significant results.

of the rules regarding EDGAR implementation. For most of our analysis, we focus on quarterly earnings announcements since they are accompanied by mandatory disclosure of quarterly financial results. Our sample period starts in April 1991 (i.e., two years before the starting date of the first batch of EDGAR filers) and ends in May 1998 (i.e., two years after the starting date of the last batch).

We obtain trading data from the large discount brokerage database used by Barber and Odean (2000), which cover the trades by 77,795 households between 1991 and 1996. The dataset is particularly appropriate for assessing the impact of internet dissemination of information on individual investors' trading decisions, because about a quarter of the investors in the dataset reside in California, which was one of the states with the highest rates of internet penetration in the early years of the internet (e.g., Greenstein, 1998). Therefore, individual investors in our sample may be more tech savvy and better positioned to take advantage of the internet technology than the average individual investor in a general sample.

We use the informativeness of individual investors' trades about subsequent stock returns to capture their information production activities. If investors produce information about a stock that is not yet incorporated into stock prices and trade on such information, their trades in the stock should be positively correlated with the subsequent stock returns. We focus on individuals' trades immediately following earnings announcements. Since earnings announcements are accompanied by the release of financial information that is critical for assessing the fundamental value of the firms (Kim and Verrecchia, 1994), we expect that investors should be especially active in processing such information when it is released. We calculate net buying by individual investors during the first 20 trading days following an earnings announcement (i.e., from day +1 to +20, with day 0 being the earnings announcement date) as the total number of shares bought by individual investors during the period minus the total number of shares sold by individual investors during

the same period normalized by the total number of shares outstanding. While individual investors' trades may be driven by non-informational factors such as liquidity shocks, aggregating the trades of a large number of investors can result in relatively precise signals about the information investors possess insofar as the non-informational factors are not systematically correlated.

We compute cumulative abnormal returns (CARs) following the trading window (i.e., starting from day +21) as the sum of daily DGTW characteristics-adjusted returns. We consider two holdings horizons, i.e., 6 months (i.e., 126 trading days from day +21 to +146) and 12 months (i.e., 252 trading days from day +21 to +272). We focus on relatively long holding horizons to reduce the noise in stock prices due to non-informational reasons such as temporary price pressure and liquidity effects. If information disseminated through EDGAR attracts investor attention and increases uninformed trading by these investors (e.g., Barber and Odean, 2001, 2008), one may expect short-run, but not long-run return predictabilities of investors' trades. Thus, focusing on relatively long windows to measure stock returns provides a cleaner test of the information story. Panel A of Table 1 shows that individual net buying has a mean of 0.034% and a standard deviation of 3.6 percentage points. The 6-month (12-month) cumulative abnormal returns starting from the 21st day post-announcement have a mean of 1.194% (2.897%) and a standard deviation of 32.053% (49.574%).

Since EDGAR makes information publicly accessible through the internet, it may have a direct impact on information production by investors who have access to the internet. We make use of the information on the channel through which investors place trades (i.e., by phone or internet) to classify investors into two categories. Internet users are those that placed a trade through the internet in the past and non-users are otherwise. About 12.049% of the investor-month observations are classified as internet users.[7] We then calculate net

[7] According to the Current Population Survey conducted in 1994 (the earliest year in which internet access is being surveyed), about 11.4% of the U.S. households owned a personal computer with a modem.

buying by internet users and non-users separately. The mean post-announcement net buying by internet users is 0.008% and that by non-users is 0.025%.

We retrieve quarterly earnings forecasts made within 90 days of the quarterly earnings report date from I/B/E/S. We construct three measures to capture information production by sell-side analysts. The first is the number of analysts following a firm, calculated as the number of quarterly earnings forecasts made by distinct analysts. The second is the forecast accuracy of analysts, calculated as the negative of the absolute value of the difference between the actual earnings per share and the median analyst forecast normalized by stock price (following Lang, Lins, and Miller, 2003). The third is market responses to analyst revisions. The idea is that if analyst revisions contain information that is not yet reflected in stock prices, the market should react positively (negatively) to upward (downward) revisions. We calculate analyst revision as the difference between two consecutive quarterly earnings forecasts of an analyst for the same stock-quarter scaled by stock price (following Clement and Tse, 2003). We calculate cumulative abnormal returns during a three-day window around the revision (i.e., from $-1$ to $+1$, with day 0 being the earnings revision date) as the sum of daily DGTW characteristics-adjusted returns. Panel B of Table 1 shows the summary statistics of the analyst sample. The mean and standard deviation of the number of analysts following a firm are 2.488 and 3.917, respectively. The mean and standard deviation of forecast accuracy are $-0.009$ and 0.079, respectively. The mean revision is $-0.184\%$ and the mean revision CAR is $-0.232\%$.

[Insert Table 1 about here]

# 4 Empirical Results

## 4.1 Informativeness of individual investors' trades

Since the implementation of the EDGAR system changes how corporate information is disseminated in the financial markets, we focus on the informativeness of individual investors' trades *following* the release of corporate information. As mentioned above, earnings announcements are accompanied by the release of financial information, which is of crucial importance to investors in evaluating the fundamental value of the firms (Kim and Verrecchia, 1994). Therefore, trades during the period following earnings announcements are likely to be motivated by informational reasons rather than other considerations.[8] If greater and broader information dissemination enables individual investors to produce information that is not yet incorporated into prices (i.e., when the crowding-in effect dominates), their trades in a firm's stock following earnings announcements should become more informative about future stock price movements after the firm becomes an EDGAR filer. On the other hand, if the crowding-out effect dominates the crowding-in effect, we should expect that individual investors' trades become less informative following the EDGAR implementation.

We construct a firm-quarter panel and run the following regression:

$$CAR_{i,q} = c_i + c_q + \beta_1 \times Netbuy_{i,q} \times Post\text{-}EDGAR_{i,q} + \beta_2 \times Netbuy_{i,q} + \beta_3 \times Post\text{-}EDGAR_{i,q} + \gamma \mathbf{X}_{i,q} + \varepsilon_{i,q}, \tag{1}$$

where $CAR_{i,q}$ is the cumulative DGTW-adjusted abnormal returns of stock $i$ during a 6- or 12-month window starting from the 21st trading day after quarter $q$'s earnings an-

[8] Individuals' trades may be motivated by non-informational reasons such as liquidity shocks, hedging, taxes, and behavioral biases, which may explain the observation that the overall performance of individual investors' trades is insignificant or even negative (e.g., Barber and Odean, 2000). Restricting the analysis to trades placed following earnings releases, therefore, allows us to focus on a period during which individuals are likely to process public information and trade on such information.

nouncement; $Netbuy_{i,q}$ is the net buying by individual investors in stock $i$ during the 20-day period immediately following the earnings announcement, $Post\text{-}EDGAR_{i,q}$ is an indicator that equals one if the firm-quarter is subject to mandatory filing on EDGAR; $c_i$ and $c_q$ are firm and quarter fixed effects, respectively; and $\mathbf{X}_{i,q}$ is a vector of lagged firm characteristics that are commonly used to predict stock returns, including firm size, book-to-market ratio, past stock return, ROA, leverage, and so on. The firm fixed effects and quarter fixed effects control for time-invariant differences across treatment and control firms and aggregate fluctuations in stock returns over time, respectively. Since the time-varying firm characteristics are likely affected by the EDGAR implementation, controlling for these variables might attenuate the total impact of the implementation on information production by corporate outsiders. We therefore run all of our regressions with and without these time-varying firm characteristics. We cluster standard errors by firm and by quarter (Petersen, 2009). The coefficient on the interaction term combining $Netbuy_{i,q}$ and $Post\text{-}EDGAR_{i,q}$ captures the incremental effect of filings to EDGAR on the informativeness of individuals' trades. If the crowding-in effect dominates the crowding-out effect, we should expect the coefficient to be positive and significant. On the other hand, if the crowding-out effect dominates the crowding-in effect, we should expect a negative and significant coefficient on the interaction term.

It is useful to note that because of the staggering of the different groups over time, firms in the sample are both treatment and control firms. For example, firms in Groups CF–02 through CF–10 serve as the control firms when firms in Group CF–01 switch from being non-EDGAR filers to EDGAR filer in April 1993, and firms in Group CF–01 as well as those in Groups CF–03 through CF–10 serve as the control firms when firms in Group CF–02 become subject to mandatory filings to EDGAR in July 1993. Thus, the staggered implementation of the EDGAR system mitigates the concern that the phase-in schedule may coincide with other firm-level shocks that may affect information production by corporate outsiders. In other words, for an omitted variable to explain our findings, it

would have to affect different groups of companies at discrete points in time as specified in the phase-in schedule. Also, it is unlikely that the phase-in schedule is designed in such a way that it anticipates changes in information production up to three years into the future, which casts doubt on reverse-causality stories.

Panel A of Table 2 reports the regression results for all trades by our sample of individual investors. The coefficient on the interaction term, $Netbuy \times Post\text{-}EDGAR$, is positive and significant in all specifications. Notably, the coefficient estimates are similar in magnitude regardless of whether firm-level control variables are included in the regressions, suggesting that the effect of the shock is largely independent from that of firm characteristics.[9] In terms of economic magnitudes, model 4 shows that a one-standard-deviation increase in net buying by individual investors during the 20 trading days post-announcement is associated with 1.649 percentage points higher subsequent 12-month cumulative abnormal returns after the stock becomes an EDGAR filer than before, which is economically non-trivial considering that the 12-month CAR has a mean of 2.897% and a standard deviation of 49.574 percentage points. This result provides evidence that EDGAR implementation increases the likelihood that individuals' trades are based on information not yet incorporated into prices, suggesting that the crowding-in effect dominates the crowding-out effect.[10]

We exploit heterogeneity across investors in terms of internet access to shed light on the sources of the increase in the informativeness of individual investors' trades after the implementation of the EDGAR system. Panel B of Table 2 replaces net buying by all individual investors with that by internet users and that by non-users separately. The coefficient on the interaction term combining the post-EDGAR indicator and net buying by internet users is positive and significant in all four specifications, whereas that combining

[9] Our results are little unchanged if we include earnings surprises or abnormal stock returns during the first 20 trading days following earnings announcements as additional controls.

[10] The sum of the coefficients on the interaction term, i.e., $Netbuy \times Post\text{-}EDGAR$, and $Netbuy$ is positive and significant in all specifications, indicating that individual investors' trades during the post-period are based on information not yet incorporated into stock prices.

the post-EDGAR indicator and net buying by non-users is insignificant. The difference in the two coefficients is significant at conventional levels when we use 12-month abnormal returns. Thus, although internet users account for a relatively small fraction (i.e., about 12%) of the sample of investors, they account for the bulk of the observed increase in the informativeness of individual investors' trades. This finding strengthens the interpretation that the EDGAR implementation enables individual investors, especially those with ready access to information on the internet, to acquire and process information.[11]

We then examine how the effect of EDGAR implementation on the informativeness of individuals' trades varies across stocks facing different levels of information asymmetry. If a firm faces a low level of information asymmetry (e.g., it is heavily covered by financial analysts and the news media), the implementation is likely to have a relatively muted effect on the informativeness of individual investors' trades because information about such firms is available from other sources. On the other hand, the implementation of EDGAR is likely to significantly improve the information environment of firms that face a high level of information asymmetry in the equity market, i.e., those whose information is otherwise costly to obtain, by increasing the amount of information that investors can access at low costs. We thus expect the effect of EDGAR implementation on the informativeness of individuals' trades to be driven mainly by firms with a high level of information asymmetry. We use analyst coverage and market capitalization to proxy for the level of information asymmetry. We measure analyst coverage and market cap as of January 31, 1993. We classify a firm as opaque if the firm has no analyst coverage and the market capitalization of the firm is in the bottom quartile. We interact the opaque indicator with the main variables, i.e., $Netbuy_{i,q}$, $Post\text{-}EDGAR_{i,q}$, and their interaction term and repeat

[11] It might be tempting to speculate that since markets must clear, individual investors as a whole gain an informational advantage over other investors such as institutions post-EDGAR. This reasoning, however, is invalid, because our data only cover a subset of individual investors and hence do not allow us to draw conclusions regarding individual investors as a whole. Instead, our evidence suggests that *some* individual investors, especially those with access to the internet, benefit from the implementation of the EDGAR system and are able to trade more profitably at the expense of other investors that presumably do not have access to the internet.

the regressions.

The results, reported in Panel C of Table 2, show that the triple interaction term combining $Netbuy_{i,q}$, $Post\text{-}EDGAR_{i,q}$, and the opaque indicator is positive and significant at the 1% level in all specifications. These findings suggest that EDGAR implementation lowers information acquisition costs for investors, especially in stocks facing a high level of information asymmetry.[12]

[Insert Table 2 about here]

## 4.2 Sell-side analyst research

To examine the effect of EDGAR implementation on information production by sell-side financial analysts, we conduct two sets of tests. The first examines analyst coverage and analyst forecast accuracy at the firm-quarter level, and the second examines market responses to analyst forecast revisions using analyst-level revision events. Specifically, for the first test, we construct a firm-quarter panel and run the following regression:

$$Analyst\ research_{i,q} = c_i + c_q + \theta_1 \times Post\text{-}EDGAR_{i,q} + \gamma \mathbf{X}_{i,q-1} + \varepsilon_{i,q}, \tag{2}$$

where $Analyst\ research_{i,q}$ is either the number of analysts making quarterly forecasts for stock $i$'s quarter $q$ earnings per share or the forecast accuracy of analysts; $Post\text{-}EDGAR_{i,q}$ is an indicator that equals one if the firm-quarter is subject to mandatory filing on EDGAR; $c_i$ and $c_q$ are firm and quarter fixed effects, respectively; and $\mathbf{X}_{i,q-1}$ is the same set of firm characteristics used in Eq. (1). We again cluster standard errors by firm and by quarter

[12] If individual investors simply trade on *raw* publicly released information by corporations (rather than processed information), one might expect a weaker predictability of their trades for subsequent long-run stock returns after the EDGAR implementation. This arises because as more investors become informed about the same public information due to EDGAR implementation, competition among these homogeneously informed investors leads to faster incorporation of the information into the prices (Holden and Subrahmanyam, 1992; Back, Cao, and Willard, 2000).

(Petersen, 2009). If the crowding-in effect dominates the crowding-out effect, the coefficient on the *Post-EDGAR* indicator should be positive and significant. On the other hand, if the crowding-out effect dominates the crowding-in effect, we should expect a negative and significant coefficient.

The results, reported in Table 3, show that both the number of analysts covering a firm and the forecast accuracy of analysts increase significantly after the firm becomes subject to mandatory filing on EDGAR. These results hold regardless of whether we control for firm size, market-to-book, prior stock return, ROA, and other variables that could be correlated with analysts' research. In terms of economic magnitudes, model 2 shows that the average firm experiences an increase of 0.234 analysts post-EDGAR, which is large considering that the mean and standard deviation of the number of analysts covering a firm are 2.488 and 3.917, respectively. Similarly, model 4 shows that the average firm experiences an increase of 0.00119 in analysts' forecast accuracy, representing 13.2% (1.5%) of the mean (standard deviation) of the variable. It is worth noting that the coefficient estimates are little unchanged when we include firm-level controls in the regressions, suggesting that the EDGAR implementation schedule is largely independent of time-varying firm characteristics.

[Insert Table 3 about here]

If financial analysts are able to produce more accurate information after a firm becomes an EDGAR filer, the market should respond more strongly to analysts' forecasts. Thus, our second set of tests investigates the impact of EDGAR implementation on market responses to analysts' forecast revisions. We estimate the following regression using each revision event as a unit of observation:

$$CAR_{i,a,d} = c_{i,q} + c_{a,q} + \kappa_1 \times Revision_{i,a,d} \times Post\text{-}EDGAR_{i,q} + \kappa_2 \times Revision_{i,a,d} + \varepsilon_{i,a,d}, \quad (3)$$

where $CAR_{i,a,d}$ is the three-day cumulative DGTW-adjusted abnormal returns of stock $i$ around analyst $a$'s forecast revision on day $d$; $Revision_{i,a,d}$ is the price-scaled changes in analyst $a$'s earnings forecasts for stock $i$ on day $d$; $Post\text{-}EDGAR_{i,q}$ is an indicator that equals one if the firm-quarter is subject to mandatory filing on EDGAR; $c_{i,q}$ and $c_{a,q}$ are firm $\times$ quarter and analyst $\times$ quarter fixed effects, respectively. In some specifications, we include firm fixed effects and the same set of firm characteristics as used in Eqs. (1) and (2) instead of firm $\times$ quarter fixed effects. In the most stringent specification, we include both firm $\times$ quarter and analyst $\times$ quarter fixed effects, which completely absorb time-varying firm attributes (e.g., prior performance, information asymmetry, and ownership structure) and time-varying analyst attributes (e.g., experience of the analyst, areas of expertise, and broker resources). The inclusion of firm $\times$ quarter fixed effects forces identification of the coefficient on the interaction term from variations across analysts covering a given firm-quarter, and that of analyst $\times$ quarter fixed effects forces identification from variations across firms covered by a given analyst in a given quarter. Standard errors are three-way clustered to allow for arbitrary correlation within firm, analyst, and quarter.

Table 4 reports the results. In all specifications, the coefficients on the interaction terms are positive and highly significant, suggesting that the market perceives analysts research as more informative after the firm becomes an EDGAR filer. The economic magnitudes are large: for example, model 4 shows that for a one-standard-deviation increase in the magnitude of the revisions, the three-day CAR is 0.414 percentage points ($= 0.00723 \times 0.573$) higher after EDGAR implementation than before. This result provides evidence that the market views analysts' research as more informative after the EDGAR implementation.

Overall, the two sets of regressions in Tables 3 and 4 show consistent patterns in the effect of EDGAR implementation on information production by sell-side analysts. These results suggest that greater dissemination of information facilitated by modern information technologies increases both the quantity and quality of sell-side analyst research.

[Insert Table 4 about here]

## 4.3 Additional tests

In this subsection, we perform a number of additional tests to assess the robustness of the main results.

**Controlling for group-specific time trends.** It is possible that time trends in our outcome variables may be different across groups that become subject to filings to EDGAR at discrete points in time. To account for this possibility, we include group-specific time trends as additional controls in the regressions (e.g., Angrist and Pischke, 2008). The identification of the effects of EDGAR implementation thus comes from whether the implementation leads to deviations from preexisting group-specific trends. We report the regression results using the most stringent specification for each test, i.e., Eqs. (1) through (3), in Table 5. The results show that the effects of EDGAR implementation on various outcomes continue to be positive and significant and the magnitude of the effects is little changed by the inclusion of these trends. These results suggest that the observed effects are not driven by differential time trends across groups.

[Insert Table 5 about here]

**Ease of access to EDGAR filings.** When the EDGAR system first got started, corporate filings on EDGAR were available electronically through Mead Data Central, a commercial data vendor, which provided access to the information for a fee (*New York Times*, 1993). The Internet Multicasting Service, a nonprofit organization, secured a National Science Foundation grant to New York University, which made EDGAR filings publicly accessible to internet users without additional charges starting from January 17, 1994. Therefore, for the first four groups of companies, there is an interim period when the filings are electronically filed but are available at a cost, which may limit the accessibility

of these filings. We thus redefine the *Post-EDGAR* indicator for the first four groups to take the value of one if the firm-quarter is after January 17, 1994 and zero otherwise, and create a new variable, *Interim*, which takes the value of one if the firm-quarter falls in the interim period for the first four groups of companies and zero otherwise. About 1% of the firm-quarters in the sample are classified as being in the interim period.

Table 6 reports the results when we replace the original *Post-EDGAR* indicator with the redefined *Post-EDGAR* indicator and the *Interim* indicator. The results show that the effects of the redefined *Post-EDGAR* indicator on various outcomes continue to be positive and significant and the magnitude of the effects is slightly larger than that obtained using the baseline specifications. Interestingly, we find positive, although statistically insignificant, effects of the interim period on our information production proxies. For example, model 2 shows that the number of analysts covering a firm in the first four groups increases by 0.205 when the firm moves from the pre-EDGAR period to the interim period. The insignificant results may be due to the low statistical power of the test given that only about 1% of the observations are in the interim period.

[Insert Table 6 about here]

**Placebo tests.** To strengthen the interpretation of the results, we repeat the tests using a period preceding the actual EDGAR implementation. We define pseudo-events as occurring two years prior to the actual implementation. We restrict the sample for this test to firm-quarters during a four-year window before the actual implementation, thus none of the firm-quarters switches during the four-year period. The *Post-EDGAR* indicator takes the value of one if the firm-quarter is in the two-year period after the pseudo-event dates and zero if it is in the two-year period before. This falsification test helps rule out alternative explanations for our results. For example, there could be unobservable characteristics that are generally correlated with both the relative timing of the implementation and an

increase in information production. In this case, we should expect significant increase in information production around these pseudo-events.

Table 7 reports the results from the placebo tests. The coefficients on our variables of interest are generally close to zero and statistically insignificant. For example, the coefficient estimates on the *Post-EDGAR* indicator are 0.055 and 0.00002, respectively, in the regressions of the number of analysts and forecast accuracy, as compared to 0.234 and 0.00119 in the baseline specifications in Table 3. These results show that there is little change in information production in the absence of shocks to information dissemination, suggesting that the observed effects are not driven by omitted variables that are generally correlated with both the relative timing of the implementation and an increase in information production.

[Insert Table 7 about here]

**Propensity-score matching.** To address the concern about nonrandom assignment of groups, we use a propensity-score matching approach. We first construct a sample of control firms that are statistically identical to firms that switch from being a non-filer to an EDGAR filer. Specifically, for each month in which a group of firms start to become subject to mandatory filings to EDGAR, we create a cohort consisting of firms that switch from being a non-filer to an EDGAR filer in that month (i.e., the treatment firms) and firms that do not switch in that month or in the 18 months before or 18 months after that month (i.e., the control firms). Note that a control firm can be an EDGAR filer or a non-filer as long as the firm retains that status during the 37-month period around the month under consideration. We then stack the 10 cohorts into a panel and run a logistic regression to predict whether a firm becomes treated. We use a comprehensive list of firm characteristics, including the full set of control variables in our main regression as well as industry fixed effects and cohort fixed effects, as the explanatory variables. We then use the predicted probabilities, or propensity scores, from this logit estimation and perform a one-to-one

nearest-neighbor matching with replacement. Panel A of Table 8 reports the pre- and post-matching firm characteristics for treatment and control firms. We cluster standard errors by firm and by cohort. Treatment stocks on average have lower capital expenditures than control firms pre-matching, but the two groups of stocks do not differ significantly in other characteristics.[13] After matching, none of the matching variables are significantly different between the treatment and matched control stocks. Hence, the matching process seems effective in removing any meaningful observable differences between the two groups of stocks.

We compare the change in various information production proxies between treatment firms and matched control firms. We use the four quarters immediately before the switching event (i.e., quarters $-4$ through $-1$, with quarter 0 being the switching quarter) as the pre-period and a four-quarter period after the switching event (i.e., quarters $+3$ through $+6$) as the post-period. We skip the first two quarters immediately following the event to allow time for market participants to start processing information.

To test whether individual investors' trades in treatment stocks, relative to those in matched control stocks, become more informative about subsequent stock returns after the implementation than before, we pool the treatment and matched control stocks and regress the subsequent 12-month cumulative abnormal returns on net buying by individual investors, an indicator for treatment stocks, an indicator for whether the observation is from the post-event period, and interaction terms for each of these variables, as well as firm fixed effects and quarter fixed effects. The coefficient on the triple interaction term is the difference-in-differences estimator comparing the change in trade informativeness between treatment and matched control firms. Panel B of Table 7 shows that the coefficient on the triple interaction term is 0.673 and significant at the 5% level, which is comparable to the magnitude obtained in our baseline specification in Table 2.

[13] Since a given firm can be both a treated and control observation (at different points in time), it is expected that the differences in characteristics between treatment and control groups are largely insignificant.

We conduct similar tests for analyst research. Panel B of Table 8 shows that, compared to matched control stocks, treatment stocks experience an increase in the number of analysts, forecast accuracy, and market responses to forecast revisions after the firms become EDGAR filers. The difference-in-differences for these outcome variables are again significant at conventional levels with magnitudes similar to those obtained in the baseline specifications in Tables 3 and 4. These results lend further support to our findings that greater information dissemination facilitated by modern information technologies increases information production by corporate outsiders.

[Insert Table 8 about here]

**Transitional filers.** Prior to the mandatory phase-in of the EDGAR system starting in April 1993, the SEC tested the system by allowing volunteers to file electronically. These voluntary filers are assigned to Group CF–01 in the phase-in schedule and are referred to as "transitional" filers in the SEC release adopting the rules for EDGAR implementation. Since transitional filers elect to switch to electronic filings on a voluntary basis, they are not required to submit *all* filings electronically before the mandatory phase-in (see SEC Release No. 33-6977). Also, transitional filers can choose not to file electronically at any time and submit all filings in paper format until mandated to file electronically. Once phased in, however, firms are required to submit all documents electronically and will not be permitted to file in paper absent a hardship exemption. Since the mandated phase-in to electronic filings limits the discretion of transitional filers in their filing decisions (i.e., whether to file electronically and, if so, what documents to file electronically), it still represents a shock to the dissemination of these firms' disclosures. Therefore, we include these transitional filers in the main tests. Nevertheless, to mitigate the concern that these firms drive the observed effects, we conduct a robustness check by excluding firms assigned to Group CF–01.

Table 9 reports the regression results using the most stringent specification for each

test. The results show that the effects of EDGAR implementation on various outcomes continue to hold and the magnitude of the effects is little changed when Group CF–01 firms are excluded. For example, models 2 and 3 show that the coefficient estimates for the post-EDGAR indicator in the regression of the number of analysts and forecast accuracy are 0.248 and 0.00127, respectively, as compared to 0.234 and 0.00119 obtained using the full sample of firms reported in Table 3.

[Insert Table 9 about here]

**Stock pricing efficiency.** Since greater information dissemination facilitated by modern information technologies increases information production by corporate outsiders, it may lead to more efficient stock prices. To test this, we use three measures, namely stock price synchronicity (Morck, Yeung, and Yu, 2000), the absolute value of stock return autocorrelation, and Hasbrouck's (1993) pricing error, as inverse measures of market efficiency. Price synchronicity is the $R$-square from the regression of a stock's daily return on the contemporaneous market return and industry return (following Chen, Goldstein, and Jiang, 2007). Durnev et al. (2003) show that firms with low stock price synchronicity are associated with a stronger predictive ability of current stock returns for future earnings, suggesting that the current stock price reflects more information about future earnings. We compute stock return autocorrelation for a stock-month as the first-order autocorrelation coefficient for the daily stock return series. A lower absolute value of return autocorrelation implies more efficient stock pricing (e.g., Lo and MacKinlay, 1988).

To construct the pricing error measure, we first decompose the log transaction price $p_t$ as $p_t = m_t + s_t$, where $m_t$ is a random walk process representing the market efficient price conditional on all public information available at $t$; $s_t$ is a zero-mean covariance-stationary process capturing the transient deviation of the transaction price from the efficient price due to factors such as inventory control by market makers, price discreteness, and temporary liquidity effects. The standard deviation of the pricing error, denoted as

$\sigma(s_t)$, captures the extent to which the transaction price deviates from the efficient price and thus can be interpreted as an inverse measure of market efficiency. We follow Boehmer and Kelly (2009) to use a vector autoregressive (VAR) system to obtain estimates for $s_t$. We use intraday transaction data from NYSE Trade and Quote (TAQ) data from 1993–1998 and Institute for the Study of Security Markets (ISSM) data from 1991–1992. We exclude stock-months with less than 200 transactions. We use trades and quotes during regular hours and discard overnight price changes. For all transaction, we only include transactions with positive prices, positive sizes, and positive bid and ask prices with bid minus ask being positive and less than 25% of the mid quote. To make the measure comparable across stocks and over time, we normalize the standard deviation of the pricing error by the standard deviation of the log transaction price and use this ratio as an inverse measure of pricing efficiency, i.e., $PricingError = \sigma(s_t)/\sigma(p_t)$. We construct the pricing error measure at a monthly frequency.

To test the effect of the implementation of EDGAR on stock pricing efficiency, we run the following regression:

$$InverseEfficiency_{i,m} = c_i + c_m + \theta_1 \times Post\text{-}EDGAR_{i,m} + \gamma \mathbf{X}_{i,m-1} + \varepsilon_{i,q}, \tag{4}$$

where $InverseEfficiency_{i,m}$ is one of the three inverse measures of information efficiency for stock $i$ in month $m$; $Post\text{-}EDGAR_{i,m}$ is an indicator set to zero before the stock becomes subject to mandatory EDGAR filing and one afterward; $c_i$ and $c_m$ are firm and month fixed effects, respectively; and $\mathbf{X}_{i,q-1}$ is the same set of firm characteristics used in Eq. (1). We cluster standard errors by firm and by month. If EDGAR implementation increases pricing efficiency, we expect a positive and significant coefficient on the post-EDGAR indicator.

The results, reported in Table 10, show that the coefficient on the post-EDGAR indicator is negative and statistically significant across all six specifications, suggesting that EDGAR implementation leads to more efficient stock pricing. The economic magnitude

is nontrivial: for example, since the mean (standard deviation) of pricing error is 0.131 (0.122), model 6 shows that pricing error decreases by 6.1% (6.6%) relative to its mean (standard deviation) after the implementation of EDGAR. This finding is consistent with a positive effect of EDGAR implementation on informational efficiency. Thus, the implementation of EDGAR not only increases information production by corporate outsiders, but also leads to more efficient stock pricing.

[Insert Table 10 about here]

# 5 Conclusion

Modern information technologies have greatly facilitated the dissemination of information in financial markets. In this paper, we investigate the impact of internet dissemination of corporate disclosures on information production by corporate outsiders, namely individual investors and financial analysts. Using the staggered implementation of the EDGAR system in 1993–1996 as a shock to information dissemination technologies, we find evidence that greater information dissemination facilitated by modern information technologies increases information production by these two types of market participants. Specifically, trades by individual investors in a stock become more informative about future stock returns after the stock becomes subject to mandatory filing on EDGAR. This effect is driven primarily by investors who have access to the internet. As for financial analysts, we find that both the amount and accuracy of information produced by sell-side analysts increase following the EDGAR implementation. Also, market responses to analyst revisions become stronger after firms start to file electronically on EDGAR. Furthermore, stock pricing efficiency improves after a firm becomes an EDGAR filer. Overall, these results suggest that advances in information technology that facilitate greater and broader information dissemination improve information production and stock pricing efficiency.

This paper contributes to our understanding of the effects of modern information technologies on financial markets. Our findings suggest that regulations that aim at promoting the availability of corporate information to a broad base of investors in real time at low costs are likely to enhance the resource allocation role of financial markets by increasing the supply of information by corporate outsiders. Given the profound effects of modern information technologies on stock pricing efficiencies, future research should investigate whether and, if so, how information technologies influence the real decisions of firms.

**Appendix A:** Timetable for Implementation of EDGAR Division of Corporation Finance Filings

April 26, 1993: Phase-in of Group CF–01.

July 19, 1993: Phase-in of Group CF–02.

October 4, 1993: Phase-in of Group CF–03.

December 6, 1993: Phase-in of Group CF–04.

Mid-1994: Final EDGAR rules and phase-in schedule are adopted.

August 1994: Phase-in of all remaining registrants begins in accordance with the final phase-in schedule, commencing with Group CF–05.

November 1994: Phase-in of Group CF–06

May 1995: Phase-in of Group CF–07

August 1995: Phase-in of Group CF–08

November 1995: Phase-in of Group CF–09

May 1996: Phase-in of Group CF–10. All registrants not previously phased in become subject to mandated electronic filing.

# References


[1] Amador, Manuel, and Pierre-Olivier Weill, 2010, Learning from prices: Public communication and welfare, *Journal of Political Economy* 118, 866–907.

[2] Angrist, Joshua D., and Jörn-Steffen Pischke, 2008, *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press, Princeton, NJ.

[3] Asthana, Sharad, and Steven Balsam, 2001, The effect of EDGAR on the market reaction to 10-K filings, *Journal of Accounting and Public Policy* 20, 349–372.

[4] Asthana, Sharad, Steven Balsam, and Srinivasan Sankaraguruswamy, 2004, Differential response of small versus large investors to 10-K filings on EDGAR, *The Accounting Review* 79, 571–589.

[5] Back, Kerry, H. Henry Cao, and Gregory A. Willard, 2000, Imperfect competition among informed traders, *Journal of Finance* 55, 2117–2155.

[6] Barber, Brad, and Terrance Odean, 2000, Trading is hazardous to your wealth: The common stock investment performance of individual investors, *Journal of Finance* 55, 773–806.

[7] Barber, Brad, and Terrance Odean, 2001, The Internet and the investor, *Journal of Economic Perspectives* 15, 41–54.

[8] Barber, Brad, and Terrance Odean, 2008, All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors, *Review of Financial Studies* 21, 785–818.

[9] Barlevy, Gadi, and Pietro Veronesi, 2000, Information acquisition in financial markets, *Review of Economic Studies* 67, 79–90.

[10] Bhushan, Ravi., 1989, Firm characteristics and analyst following, *Journal of Accounting and Economics* 11, 255–274.

[11] Blankespoor, Elizabeth Gregory S. Miller, and Hal D. White, 2014, Initial evidence on the market impact of the XBRL mandate, *Review of Accounting Studies* 19, 1468–1503.

[12] Boehmer, Ekkehart, and Eric K. Kelley, 2009, Institutional investors and the informational efficiency of prices, *Review of Financial Studies* 22, 3563–3594.

[13] Chen, Qi, Itay Goldstein, and Wei Jiang, 2007, Price informativeness and investment sensitivity to stock price, *Review of Financial Studies* 20, 619–650.

[14] Clement, Michael B., and Senyo Y. Tse, 2003 Do investors respond to analysts' forecast revisions as if forecast accuracy is all that matters? *Accounting Review* 78, 227–249.

[15] Coval, Joshua D., David A. Hirshleifer, and Tyler Shumway, 2005, Can individual investors beat the market? Working paper.

[16] D'Avolio, Gene, Efi Gildor, and Andrei Shleifer, 2002, Technology, information production, and market efficiency, In *Economic Policy for the Information Economy: A Symposium Sponsored by the Federal Reserve Bank of Kansas City.* Kansas City, MO.

[17] Dong, Yi, Oliver Z. Li, Yupeng Lin, and Chenkai Ni, 2016, Does information-processing cost affect firm-specific information acquisition? Evidence from XBRL adoption, *Journal of Financial and Quantitative Analysis* 51, 435–462.

[18] Dougal, Casey, Joseph Engelberg, Diego García, and Christopher A. Parsons, 2012, Journalists and the stock market, *Review of Financial Studies* 25, 639–679.

[19] Drake, Michael. S., Darren T. Roulstone, and Jacob R. Thornock, 2012, Investor information demand: Evidence from Google searches around earnings announcements, *Journal of Accounting Research* 50, 1001–1040.

[20] Dugast, Jérôme, and Thierry Foucalt, 2017, Data abundance and asset price informativeness, *Journal of Financial Economics*, forthcoming.

[21] Durnev, Artyom, Randall Morck, Bernard Yeung, and Paul Zarowin, 2003, Does greater firm-specific return variation mean more or less informed stock pricing? *Journal of Accounting Research* 41, 797–836.

[22] Engelberg, Joseph, and Christopher A. Parsons, 2011, The causal impact of media in financial markets, *Journal of Finance* 66, 67–97.

[23] Goldstein, Itay, and Liyan Yang, 2017, Information disclosure in financial markets, *Annual Reviews of Financial Economics*, forthcoming.

[24] Greenstein, Shane, 1998, Universal service in the digital age: The commercialization and geography of US internet access, NBER Working Paper 6453.

[25] Grossman, Sanford J., and Joseph E. Stiglitz, 1980, On the impossibility of informationaly efficient markets, *American Economic Review* 70, 393–408.

[26] Hasbrouck, Joel, 1993, Assessing the quality of a security market: A new approach to transaction-cost measurement, *Review of Financial Studies* 6, 191–212.

[27] Healy, Paul M, and Krishna G Palepu, 2001, Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature, *Journal of Accounting and Economics* 31, 405–440.

[28] Holden, Craig W., and Avanidhar Subrahmanyam, 1992, Long-lived private information and imperfect competition, *Journal of Finance* 47, 247–270.

[29] Ivković, Zoran, and Scott Weisbenner, 2005, Local does as local is: Information content of the geography of individual investors' common stock investments, *Journal of Finance* 60, 267–306.

[30] Ivković, Zoran, Clemens Sialm, and Scott Weisbenner, 2008, Portfolio concentration and the performance of individual investors. *Journal of Financial and Quantitative Analysis* 43, 613–656.

[31] Kaniel, Ron, Shuming Liu, Gideon Saar, and Sheridan Titman, 2012, Individual investor trading and return patterns around earnings announcements, *Journal of Finance* 67, 639–680.

[32] Kelley, Eric K., and Paul C. Tetlock, 2013, How wise are crowds? Insights from retail orders and stock returns, *Journal of Finance* 68, 1229–1265.

[33] Kelley, Eric K., and Paul C. Tetlock, 2017, Retail short selling and stock prices, *Review of Financial Studies* 30, 801–834.

[34] Kim, Oliver, and Robert E. Verrecchia, 1994, Market liquidity and volume around earnings announcements, *Journal of Accounting and Economics* 17, 41–67.

[35] Lang, Mark. H., Karl V. Lins, and Darius P. Miller, 2003, ADRs, analysts, and accuracy: Does cross listing in the United States improve a firm's information environment and increase market value?, *Journal of Accounting Research* 41, 317–345.

[36] Lang, Mark H., and Russell J. Lundholm, 1996, Corporate disclosure policy and analyst behavior, *Accounting Review* 71, 467–492.

[37] Lo, Andrew W., and A. Craig MacKinlay, 1988, Stock market prices do not follow random walks: Evidence from a simple specification test, *Review of Financial Studies* 1, 41–66.

[38] Morck, Randall, Bernard Yeung, and Wayne W. Yu, 2000, The information content of stock markets: Why do emerging markets have synchronous stock price movements?, *Journal of Financial Economics* 58: 215–260.

[39] Morris, Stephen, and Hyun Song Shin, 2002, Social value of public information, *American Economic Review* 92, 1521–1534.

[40] *New York Times*, 1982, S.E.C. Data: Difficult hunt, May 19.

[41] *New York Times*, 1993, Plan opens more data to public, October 22.

[42] O'Brien, Patricia C., and Ravi Bhushan, 1990, Analyst following and institutional ownership, *Journal of Accounting Research* 28, 55–76.

[43] Peress, Joel, 2014, The media and the diffusion of information in financial Markets: Evidence from newspaper strikes, *Journal of Finance* 69, 2007–2043.

[44] Petersen, Mitchell A., 2009, Estimating standard errors in finance panel data sets: Comparing approaches, *Review of Financial Studies* 22, 435–480.

[45] Simon, Herbert A., 1997, Designing organizations for an information-rich world, In D. M. Lamberton (ed.), *The Economics of Communication and Information*. Cheltenham, UK: Edward Elgar.

[46] Shapiro, Carl, and Hal R. Varian, 1999, *Information rules: A strategic guide to the network economy*, Harvard Business Press, Boston, MA.

[47] Shiller, Robert, 2006, *Irrational Exuberance*, Crown Business Publishing, Danvers, MA.

[48] Verrecchia, Robert E., 1982, Information acquisition in a noisy rational expectations economy, *Econometrica* 50, 1415–1430.

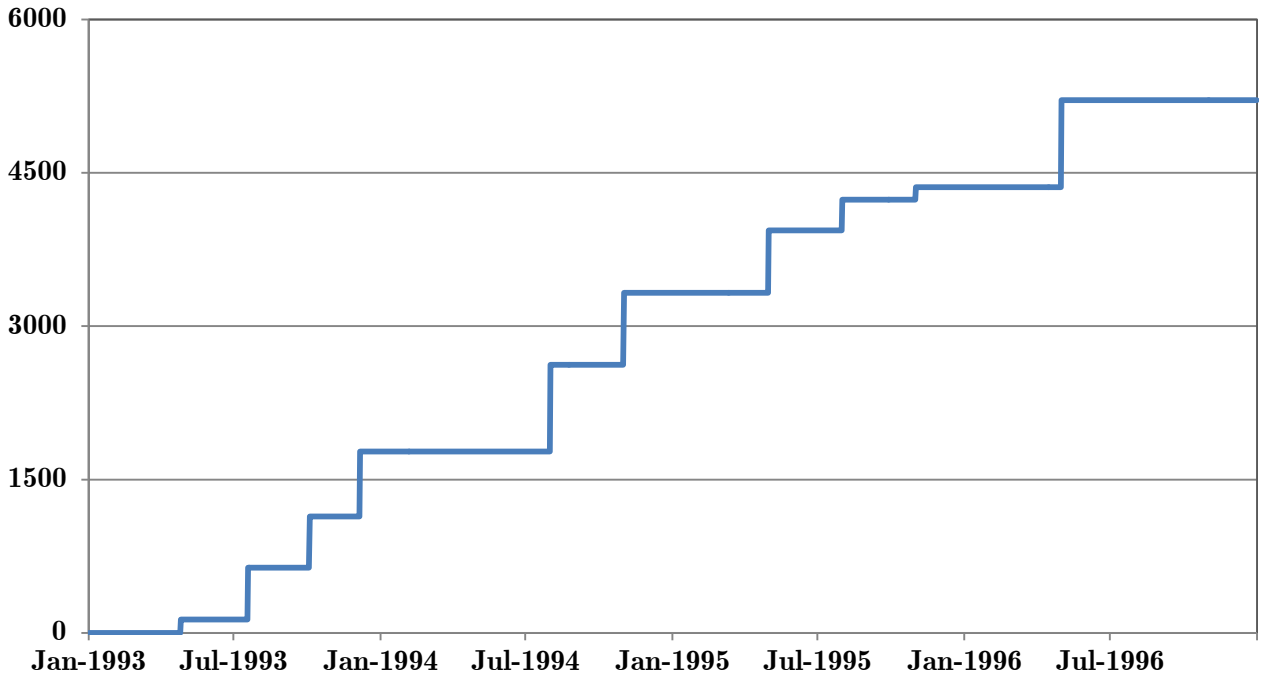


**Figure 1. Staggered implementation of mandatory filing through EDGAR**

This figure plots the number of firms that are subject to mandatory filing through EDGAR during the period from January 1993 through December 1996.

**Table 1: Summary statistics**

This table reports the summary statistics for the individual trading sample (Panel A) and analyst sample (Panel B). *Post-EDGAR* is an indicator that equals one after a firm-quarter becomes a mandatory EDGAR filer and zero otherwise. *Net buying*$_{[+1,\ +20]}$ is the number of shares bought minus the number of shares sold by individual investors as a fraction of the number of shares outstanding during a 20-day window after an earnings announcement. *Net buying by internet users*$_{[+1,\ +20]}$ and *Net buying by non-users*$_{[+1,\ +20]}$ are similarly defined for individual investors with access to the internet and those without, respectively. We identify an investor as an internet user if she placed a trade through the internet in the past. *AbnReturn*$_{[+21,\ +146]}$ and *AbnReturn*$_{[+21,\ +272]}$ are cumulative DGTW-characteristics adjusted returns during a 6- and 12-month window starting from the 21st day after an earnings announcement, respectively. *# of analysts* is the number of analysts making quarterly earnings forecasts in the I/B/E/S database for a stock in a given quarter. *Forecast accuracy* is negative of the absolute value of the difference between the actual earnings per share and the median analyst forecast normalized by stock price (following Lang, Lins, and Miller, 2003). *Revision* is the difference between two consecutive quarterly earnings forecasts of an analyst for the same stock-quarter scaled by stock price (following Clement and Tse, 2003). *CAR*$_{[-1,\ +1]}$ is the cumulative DGTW characteristics-adjusted returns during a three-day window around an earnings forecast revision by an analyst. *Total assets* is the book value of assets of the firm. *Book-to-market* is the book value of common equity divided by the market value of common equity. *Prior stock return* is the buy-and-hold stock return during the past 12 months skipping the most recent month. *ROA* is the ratio of income before extraordinary items to book value of assets. *Book leverage* is the ratio of the book value of total debt to the book value of total assets. *Asset tangibility* is the ratio of net property, plant, and equipment to total assets. *Sales growth* is the percentage change in quarterly sales from four quarters earlier to the current quarter. *CapEx/TA* is the ratio of capital expenditure to total assets. *R&D/TA* is the ratio of R&D expenses to total assets. *Institutional ownership* is the number of shares held by institutional investors as a fraction of the number of shares outstanding. All variables are winsorized at the 0.1% and 99.9% levels in order to minimize the effect of outliers.

Panel A: Summary statistics for the individual trading sample

| | # of obs | Mean | Standard deviation | P10 | Median | P90 |
|---|---|---|---|---|---|---|
| *Main variables* | | | | | | |
| Post-EDGAR | 29,355 | 0.431 | 0.495 | 0.000 | 0.000 | 1.000 |
| Net buying$_{[+1,\ +20]}$ | 29,355 | 0.034% | 3.561% | -1.702% | -0.019% | 1.771% |
| Net buying by internet users$_{[+1,\ +20]}$ | 29,355 | 0.008% | 1.316% | -0.203% | 0.000% | 0.202% |
| Net buying by non-users $_{[+1,\ +20]}$ | 29,355 | 0.025% | 2.955% | -1.421% | 0.000% | 1.451% |
| AbnReturn$_{[+21,\ +146]}$ | 29,355 | 1.194% | 32.053% | -29.901% | -0.055% | 32.326% |
| AbnReturn$_{[+21,\ +272]}$ | 29,355 | 2.897% | 49.574% | -41.622% | 0.651% | 48.238% |
| *Control variables* | | | | | | |
| Total assets ($ mil) | 29,355 | 3,055.680 | 12,070.430 | 24.647 | 258.121 | 5,585.520 |
| Book-to-Market | 29,355 | 0.665 | 0.665 | 0.186 | 0.548 | 1.232 |
| Prior stock return | 29,355 | 0.305 | 0.844 | -0.298 | 0.155 | 0.989 |
| ROA | 29,355 | 0.026 | 0.148 | -0.062 | 0.043 | 0.130 |
| Book leverage | 29,355 | 0.495 | 0.238 | 0.186 | 0.490 | 0.833 |
| Asset tangibility | 29,355 | 0.296 | 0.231 | 0.034 | 0.239 | 0.667 |
| Sales growth | 29,355 | 0.251 | 2.323 | -0.093 | 0.100 | 0.486 |
| CapEx | 29,355 | 0.081 | 0.255 | 0.005 | 0.054 | 0.162 |
| R&D | 29,355 | 0.050 | 0.123 | 0.000 | 0.000 | 0.154 |
| Institutional ownership | 29,355 | 0.393 | 0.224 | 0.088 | 0.389 | 0.692 |

Panel B: Summary statistics for the analyst sample

| | # of obs | Mean | Standard deviation | P10 | Median | P90 |
|---|---|---|---|---|---|---|
| *Main variables* | | | | | | |
| Post-EDGAR | 103,935 | 0.512 | 0.500 | 0.000 | 1.000 | 1.000 |
| # of analysts | 103,935 | 2.488 | 3.917 | 0.000 | 1.000 | 8.000 |
| Forecast accuracy | 56,495 | -0.009 | 0.079 | -0.017 | -0.002 | 0.000 |
| Revision | 358,434 | -0.184% | 0.723% | -0.642% | -0.042% | 0.193% |
| CAR$_{[-1,\ +1]}$ | 358,434 | -0.232% | 4.999% | -4.790% | -0.146% | 4.549% |
| *Control variables* | | | | | | |
| Total assets ($ mil) | 103,935 | 2,368.320 | 12,143.080 | 12.675 | 152.707 | 3,630.230 |
| Book-to-Market | 103,935 | 0.722 | 23.128 | 0.160 | 0.576 | 1.368 |
| Prior stock return | 103,935 | 0.234 | 0.772 | -0.363 | 0.119 | 0.838 |
| ROA | 103,935 | -0.001 | 0.191 | -0.123 | 0.032 | 0.115 |
| Book leverage | 103,935 | 0.539 | 0.308 | 0.193 | 0.517 | 0.909 |
| Asset tangibility | 103,935 | 0.292 | 0.239 | 0.024 | 0.231 | 0.678 |
| Sales growth | 103,935 | 0.356 | 7.398 | -0.124 | 0.086 | 0.472 |
| CapEx | 103,935 | 0.083 | 0.943 | 0.002 | 0.048 | 0.158 |
| R&D | 103,935 | 0.045 | 0.273 | 0.000 | 0.000 | 0.133 |
| Institutional ownership | 103,935 | 0.316 | 0.235 | 0.029 | 0.281 | 0.658 |

**Table 2: Staggered implementation of EDGAR and the informativeness of trades by individual investors**

This table reports regression analysis of the impact of EDGAR on the informativeness of individual investors' trades about subsequent stock returns. The dependent variables are cumulative DGTW-characteristics adjusted returns during a 6- or 12-month window starting from the 21st day after an earnings announcement. Panel A uses net buying by all individual investors in the sample; Panel B decomposes the net buying measure into net buying by internet users and that by non-users; and Panel C interacts the main variable with an indicator for opaque firms. We identify an investor as an internet user if she placed a trade through the internet in the past. *Opaque* is an indicator that equals one if the firm has no analyst coverage and the market capitalization of the firm is in the bottom quartile. Analyst coverage and market cap are measured as of the month-end immediately before the release of the rules regarding EDGAR implementation, i.e., January 31, 1993. All variables are defined in Table 1. Numbers in parentheses are *t*-statistics based on standard errors two-way clustered by firm and quarter. Significance at the 10% (*), 5% (**), or 1% level (***) is indicated. Numbers in square brackets are *p*-values for the null that the coefficients on the two interaction terms are equal.

Panel A: All investors

| Dependent = | $\text{AbnReturn}_{[+21,\ +146]}$ | | $\text{AbnReturn}_{[+21,\ +272]}$ | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Net buying$_{[+1,\ +20]}$ × Post-EDGAR | 0.436 | 0.469 | 0.415 | 0.463 |
| | (2.57)** | (2.93)*** | (1.91)* | (2.34)** |
| Net buying$_{[+1,\ +20]}$ | 0.066 | 0.068 | 0.096 | 0.102 |
| | (1.13) | (1.14) | (0.95) | (1.06) |
| Post-EDGAR | -0.038 | -0.036 | -0.089 | -0.084 |
| | (3.79)*** | (3.74)*** | (6.07)*** | (5.78)*** |
| Log(Total assets) | | -0.128 | | -0.184 |
| | | (9.58)*** | | (9.74)*** |
| Book-to-Market | | 0.049 | | 0.081 |
| | | (2.96)*** | | (2.92)*** |
| Prior stock return | | -0.045 | | -0.067 |
| | | (4.49)*** | | (4.36)*** |
| ROA | | -0.008 | | -0.022 |
| | | (0.20) | | (0.45) |
| Book leverage | | 0.231 | | 0.354 |
| | | (6.33)*** | | (5.85)*** |
| Asset tangibility | | 0.146 | | 0.207 |
| | | (2.39)** | | (2.31)** |
| Sales growth | | -0.001 | | 0.000 |
| | | (0.28) | | (0.17) |
| CapEx | | -0.007 | | -0.019 |
| | | (0.62) | | (1.06) |
| R&D | | -0.089 | | -0.168 |
| | | (2.38)** | | (3.39)*** |
| Institutional ownership | | -0.236 | | -0.391 |
| | | (5.20)*** | | (5.51)*** |
| Firm FEs | Yes | Yes | Yes | Yes |
| Quarter FEs | Yes | Yes | Yes | Yes |
| # of observations | 29,056 | 29,056 | 29,056 | 29,056 |
| Adj. R-squared | 0.12 | 0.16 | 0.26 | 0.31 |

Panel B: Internet users vs. non-users

| Dependent = | $\text{AbnReturn}_{[+21,\ +146]}$ | | $\text{AbnReturn}_{[+21,\ +272]}$ | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Net buying by internet users × Post-EDGAR ($b_1$) | 0.894 | 0.972 | 1.140 | 1.257 |
| | (2.06)** | (2.35)** | (2.75)*** | (3.38)*** |
| Net buying by non-users × Post-EDGAR ($b_2$) | 0.263 | 0.312 | 0.163 | 0.237 |
| | (1.19) | (1.48) | (0.54) | (0.83) |
| Net buying by internet users | 0.154 | 0.068 | 0.130 | -0.003 |
| | (0.91) | (0.41) | (0.62) | (0.02) |
| Net buying by non-users | 0.071 | 0.087 | 0.116 | 0.144 |
| | (0.99) | (1.21) | (0.93) | (1.22) |
| Post-EDGAR | -0.038 | -0.036 | -0.089 | -0.084 |
| | (3.80)*** | (3.74)*** | (6.09)*** | (5.79)*** |
| Firm controls | No | Yes | No | Yes |
| Firm FEs | Yes | Yes | Yes | Yes |
| Quarter FEs | Yes | Yes | Yes | Yes |
| $p$-value for $b_1 = b_2$ | [0.22] | [0.18] | [0.06]* | [0.03]** |
| # of observations | 29,056 | 29,056 | 29,056 | 29,056 |
| Adj. R-squared | 0.14 | 0.18 | 0.25 | 0.30 |

Panel C: High vs. low information asymmetry stocks

| Dependent = | $\text{AbnReturn}_{[+21,\ +146]}$ | | $\text{AbnReturn}_{[+21,\ +272]}$ | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Net $\text{buying}_{[+1,\ +20]}$ × Post-EDGAR × Opaque | 1.074 | 1.060 | 0.958 | 0.947 |
| | (3.29)*** | (3.40)*** | (2.71)*** | (2.68)*** |
| Net $\text{buying}_{[+1,\ +20]}$ × Opaque | 0.047 | 0.079 | 0.074 | 0.117 |
| | (0.26) | (0.47) | (0.30) | (0.50) |
| Post-EDGAR × Opaque | -0.538 | -0.562 | -0.699 | -0.747 |
| | (3.42)*** | (3.52)*** | (3.65)*** | (3.75)*** |
| Net $\text{buying}_{[+1,\ +20]}$ × Post-EDGAR | -0.046 | -0.026 | -0.130 | -0.099 |
| | (2.33)** | (1.42) | (3.60)*** | (3.00)*** |
| Net $\text{buying}_{[+1,\ +20]}$ | 0.273 | 0.287 | 0.356 | 0.385 |
| | (3.38)*** | (3.60)*** | (2.95)*** | (3.46)*** |
| Post-EDGAR | -0.034 | -0.033 | -0.075 | -0.074 |
| | (3.34)*** | (3.41)*** | (5.18)*** | (4.95)*** |
| Firm controls | No | Yes | No | Yes |
| Firm FEs | Yes | Yes | Yes | Yes |
| Quarter FEs | Yes | Yes | Yes | Yes |
| # of observations | 29,056 | 29,056 | 29,056 | 29,056 |
| Adj. R-squared | 0.12 | 0.17 | 0.26 | 0.31 |

**Table 3: Staggered implementation of EDGAR and sell-side analyst research**

This table reports regression analysis of the impact of EDGAR on analyst coverage and analysts' forecast accuracy. The dependent variable in the first (last) two columns is the number of analysts (forecast accuracy). All variables are defined in Table 1. Numbers in parentheses are *t*-statistics based on standard errors two-way clustered by firm and quarter. Significance at the 10% (*), 5% (**), or 1% level (***) is indicated.

| Dependent = | # of analysts | | Forecast accuracy | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Post-EDGAR | 0.224 | 0.234 | 0.142 | 0.119 |
| | (3.07)*** | (3.48)*** | (2.70)** | (2.62)** |
| Log(Total assets) | | 0.835 | | 0.153 |
| | | (13.80)*** | | (2.19)** |
| Book-to-Market | | -0.000 | | -0.004 |
| | | (2.42)** | | (1.46) |
| Prior stock return | | 0.009 | | 0.442 |
| | | (0.69) | | (11.16)*** |
| ROA | | 0.121 | | 1.353 |
| | | (1.65) | | (5.38)*** |
| Book leverage | | -0.322 | | -0.387 |
| | | (4.30)*** | | (2.59)** |
| Asset tangibility | | 0.705 | | -0.327 |
| | | (3.87)*** | | (1.22) |
| Sales growth | | 0.000 | | 0.001 |
| | | (1.16) | | (1.30) |
| CapEx | | 0.006 | | 0.092 |
| | | (1.40) | | (2.19)** |
| R&D | | 0.002 | | 0.542 |
| | | (0.10) | | (2.60)** |
| Institutional ownership | | 2.023 | | 0.990 |
| | | (12.01)*** | | (5.66)*** |
| Firm FEs | Yes | Yes | Yes | Yes |
| Quarter FEs | Yes | Yes | Yes | Yes |
| # of observations | 103,872 | 103,872 | 56,333 | 56,333 |
| Adj. R-squared | 0.81 | 0.82 | 0.29 | 0.31 |

**Table 4: Staggered implementation of EDGAR and market responses to analysts' forecast revisions**

This table reports regression analysis of the impact of EDGAR on the informativeness of analysts' forecast revisions. The dependent variable is the cumulative DGTW characteristics-adjusted returns during a three-day window around analyst revisions. The unit of observation is a revision event. All variables are defined in Table 1. Numbers in parentheses are *t*-statistics based on standard errors three-way clustered by firm, analyst, and quarter. Significance at the 10% (*), 5% (**), or 1% level (***) is indicated.

| Dependent = | Revision $CAR_{[-1,+1]}$ | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Revision × Post-EDGAR | 0.634 | 0.637 | 0.656 | 0.573 |
| | (4.52)*** | (4.69)*** | (4.72)*** | (4.22)*** |
| Revision | 0.357 | 0.246 | 0.222 | 0.248 |
| | (4.72)*** | (3.47)*** | (2.94)*** | (3.13)*** |
| Post-EDGAR | 0.001 | 0.002 | 0.002 | |
| | (1.11) | (2.45)** | (2.71)** | |
| Log(Total assets) | | -0.005 | -0.003 | |
| | | (4.44)*** | (2.87)*** | |
| Book-to-Market | | 0.005 | 0.005 | |
| | | (4.28)*** | (4.05)*** | |
| Prior stock return | | 0.009 | 0.009 | |
| | | (11.67)*** | (10.98)*** | |
| ROA | | -0.008 | -0.008 | |
| | | (1.77)* | (1.66) | |
| Book leverage | | 0.000 | 0.001 | |
| | | (0.08) | (0.27) | |
| Asset tangibility | | -0.009 | -0.010 | |
| | | (2.11)** | (2.40)** | |
| Sales growth | | -0.000 | -0.000 | |
| | | (0.14) | (0.60) | |
| CapEx | | -0.001 | -0.003 | |
| | | (0.51) | (1.28) | |
| R&D | | 0.004 | 0.003 | |
| | | (0.82) | (0.59) | |
| Institutional ownership | | -0.006 | -0.006 | |
| | | (2.55)** | (2.36)** | |
| Firm FEs | Yes | Yes | Yes | No |
| Analyst FEs | Yes | Yes | No | No |
| Quarter FEs | Yes | Yes | No | No |
| Firm × quarter FEs | No | No | No | Yes |
| Analyst × quarter FEs | No | No | Yes | Yes |
| # of observations | 358,071 | 354,844 | 349,104 | 342,815 |
| Adj. R-squared | 0.06 | 0.07 | 0.07 | 0.10 |

**Table 5: Controlling for group-specific time trends**

This table reports regression analysis of the impact of EDGAR on various information production measures after adding controls for group-specific time trends. The dependent variables are cumulative DGTW-characteristics adjusted returns during a 6-month window starting from the 21st day after an earnings announcement, the number of analysts, forecast accuracy, and the cumulative DGTW characteristics-adjusted returns during a three-day window around analyst revisions, respectively. All variables are defined in Table 1. Numbers in parentheses in the first three columns are *t*-statistics based on standard errors two-way clustered by firm and quarter, and those in the last column are three-way clustered by firm, quarter, and analyst. Significance at the 10% (*), 5% (**), or 1% level (***) is indicated.

| Dependent = | AbnReturn $_{[+21,\ +146]}$ (1) | # of analysts (2) | Forecast accuracy (3) | Revision CAR$_{[-1,\ +1]}$ (4) |
|---|---|---|---|---|
| Net buying$_{[+1,\ +20]}$ × Post-EDGAR | 0.540 | | | |
| | (3.06)*** | | | |
| Net buying$_{[+1,\ +20]}$ | 0.527 | | | |
| | (2.68)*** | | | |
| Post-EDGAR | -0.032 | 0.236 | 0.125 | |
| | (3.17)*** | (3.37)*** | (2.77)*** | |
| Revision × Post-EDGAR | | | | 0.600 |
| | | | | (4.64)*** |
| Revision | | | | 0.267 |
| | | | | (1.57) |
| Firm controls | Yes | Yes | Yes | No |
| Group × time trends | Yes | Yes | Yes | No |
| Net buying × group × time trends | Yes | N/A | N/A | N/A |
| Revision × group × time trends | N/A | N/A | N/A | Yes |
| Firm FEs | Yes | Yes | Yes | No |
| Quarter FEs | Yes | Yes | Yes | No |
| Firm × quarter FEs | No | No | No | Yes |
| Analyst × quarter FEs | N/A | N/A | N/A | Yes |
| Clustering SEs | Firm, quarter | Firm, quarter | Firm, quarter | Firm, analyst, quarter |
| # of observations | 29,056 | 103,872 | 56,333 | 342,815 |
| Adj. R-squared | 0.25 | 0.82 | 0.31 | 0.10 |

Table 6: Ease of access to EDGAR filings

This table reports regression analysis of the impact of EDGAR on various information production measures when we partition the post-EDGAR period for the first four groups into two periods based on the ease of access to EDGAR filings. Specifically, we redefine the *Post-EDGAR* indicator for the first four groups to take the value of one if the firm-quarter is after January 17, 1994 (when the filings became available to internet users without additional charges) and zero otherwise. *Interim* is an indicator variable that takes the value of one if the firm-quarter falls in the interim period, i.e., the time from the starting date of mandated electronic filing to EDGAR to January 16, 1994, for the first four groups of companies and zero otherwise. The dependent variables are cumulative DGTW-characteristics adjusted returns during a 6-month window starting from the 21st day after an earnings announcement, the number of analysts, forecast accuracy, and the cumulative DGTW characteristics-adjusted returns during a three-day window around analyst revisions, respectively. All variables are defined in Table 1. Numbers in parentheses in the first three columns are *t*-statistics based on standard errors two-way clustered by firm and quarter, and those in the last column are three-way clustered by firm, quarter, and analyst. Significance at the 10% (*), 5% (**), or 1% level (***) is indicated.

| Dependent = | AbnReturn $_{[+21,\ +146]}$ (1) | # of analysts (2) | Forecast accuracy (3) | Revision CAR$_{[-1,\ +1]}$ (4) |
|---|---|---|---|---|
| Net buying$_{[+1,\ +20]}$ × Post-EDGAR | 0.473 | | | |
| | (2.89)*** | | | |
| Net buying$_{[+1,\ +20]}$ × Interim | 0.322 | | | |
| | (0.89) | | | |
| Net buying$_{[+1,\ +20]}$ | 0.069 | | | |
| | (1.15) | | | |
| Post-EDGAR | -0.055 | 0.241 | 0.142 | |
| | (5.37)*** | (3.41)*** | (2.39)** | |
| Interim | 0.011 | 0.205 | 0.030 | |
| | (0.96) | (1.11) | (0.53) | |
| Revision × Post-EDGAR | | | | 0.673 |
| | | | | (4.70)*** |
| Revision × Interim | | | | 0.025 |
| | | | | (0.23) |
| Revision | | | | 0.358 |
| | | | | (4.78)*** |
| Firm controls | Yes | Yes | Yes | No |
| Group × time trends | Yes | Yes | Yes | No |
| Net buying × group × time trends | Yes | N/A | N/A | N/A |
| Revision × group × time trends | N/A | N/A | N/A | Yes |
| Firm FEs | Yes | Yes | Yes | No |
| Quarter FEs | Yes | Yes | Yes | No |
| Firm × quarter FEs | No | No | No | Yes |
| Analyst × quarter FEs | N/A | N/A | N/A | Yes |
| Clustering SEs | Firm, quarter | Firm, quarter | Firm, quarter | Firm, analyst, quarter |
| # of observations | 29,056 | 103,872 | 56,333 | 358,071 |
| Adj. R-squared | 0.30 | 0.82 | 0.31 | 0.06 |

**Table 7: Placebo tests**

This table reports regression analysis of information production activities around pseudo EDGAR implementations. We define pseudo-events as occurring 24 months prior to the actual implementation. The dependent variables are cumulative DGTW-characteristics adjusted returns during a 6-month window starting from the 21st day after an earnings announcement, the number of analysts, forecast accuracy, and the cumulative DGTW characteristics-adjusted returns during a three-day window around analyst revisions, respectively. "*Post-EDGAR*" is an indicator that equals one for firm-quarters that are in the two-year window after the pseudo-event date and zero for firm-quarters that are in the two-year window immediately before the pseudo-event date. All variables are defined in Table 1. Numbers in parentheses in the first three columns are *t*-statistics based on standard errors two-way clustered by firm and quarter, and those in the last column are three-way clustered by firm, quarter, and analyst. Significance at the 10% (*), 5% (**), or 1% level (***) is indicated.

| Dependent = | AbnReturn$_{[+21,\ +146]}$ (1) | # of analysts (2) | Forecast accuracy (3) | Revision CAR$_{[-1,\ +1]}$ (4) |
|---|---|---|---|---|
| Net buying$_{[+1,\ +20]}$ × "Post-EDGAR" | 0.166 | | | |
| | (1.27) | | | |
| Net buying$_{[+1,\ +20]}$ | 0.064 | | | |
| | (0.76) | | | |
| "Post-EDGAR" | 0.010 | 0.055 | 0.002 | |
| | (0.88) | (0.36) | (0.89) | |
| Revision × "Post-EDGAR" | | | | 0.016 |
| | | | | (0.28) |
| Revision | | | | 0.002 |
| | | | | (0.08) |
| Firm controls | Yes | Yes | Yes | No |
| Firm FEs | Yes | Yes | Yes | No |
| Quarter FEs | Yes | Yes | Yes | No |
| Firm × quarter FEs | No | No | No | Yes |
| Analyst × quarter FEs | N/A | N/A | N/A | Yes |
| Clustering SEs | Firm, quarter | Firm, quarter | Firm, quarter | Firm, analyst, quarter |
| # of observations | 18,354 | 61,045 | 31,292 | 123,828 |
| Adj. R-squared | 0.19 | 0.83 | 0.10 | 0.57 |

**Table 8: Propensity-score matching**

This table reports propensity-score matching diagnostics and the diff-in-diff tests using the propensity-score matched sample. Panel A compares the mean of the matching variables between treatment (i.e., firms that switch from a non-filer to an EDGAR filer in a month) and control (i.e., firms that remain as a filer or non-filer in a 36-month window around the month under consideration) stocks before and after matching. We use the predicted probabilities, or propensity scores, from a logit estimation and perform a one-to-one nearest-neighbor match with replacement. Panel B reports the diff-in-diff tests of the impact of EDGAR on information production using the propensity-score matched sample. All variables are defined in Table 1. Numbers in parentheses are *t*-statistics based on standard errors two-way clustered by firm and quarter. Significance at the 10% (*), 5% (**), or 1% level (***) is indicated.

Panel A: Pre and post-matching characteristics

| Variable | Pre-matching | | | Post-matching | | |
|---|---|---|---|---|---|---|
| | Treatment (1) | Control (2) | Difference (1) – (2) | Treatment (3) | Control (4) | Difference (3) – (4) |
| Log(Total assets) | 5.171 | 5.300 | -0.129 | 5.174 | 5.322 | -0.148 |
| Book-to-Market | 0.676 | 0.607 | 0.069 | 0.673 | 0.709 | -0.036 |
| Prior stock return | 0.204 | 0.196 | 0.008 | 0.224 | 0.228 | -0.004 |
| ROA | -0.014 | -0.015 | 0.001 | -0.016 | -0.023 | 0.007 |
| Book leverage | 0.547 | 0.534 | 0.013 | 0.549 | 0.558 | -0.009 |
| Asset tangibility | 0.290 | 0.294 | -0.004 | 0.290 | 0.298 | -0.009 |
| Sales growth | 0.164 | 0.254 | -0.090 | 0.168 | 0.182 | -0.014 |
| CapEx | 0.066 | 0.079 | -0.012* | 0.066 | 0.066 | 0.000 |
| R&D | 0.038 | 0.044 | -0.006 | 0.038 | 0.035 | 0.003 |
| Institutional ownership | 0.312 | 0.326 | -0.014 | 0.312 | 0.321 | -0.009 |

Panel B: Diff-in-diff tests using the propensity-score matched sample

| | Informativeness of individual trades (1) | # of analysts (2) | Forecast accuracy (3) | Market responses to revisions (4) |
|---|---|---|---|---|
| DiD estimate | 0.673 | 0.369 | 0.183 | 0.468 |
| | (2.47)** | (2.62)** | (1.72)* | (2.00)** |

**Table 9: Excluding Group CF–01 firms**

This table reports regression analysis of the impact of EDGAR on various information production measures after excluding firms assigned to Group CF–01 (which consists mostly of transitional filers) on the phase-in schedule. The dependent variables are cumulative DGTW-characteristics adjusted returns during a 6-month window starting from the 21st day after an earnings announcement, the number of analysts, forecast accuracy, and the cumulative DGTW characteristics-adjusted returns during a three-day window around analyst revisions, respectively. All variables are defined in Table 1. Numbers in parentheses in the first three columns are *t*-statistics based on standard errors two-way clustered by firm and quarter, and those in the last column are three-way clustered by firm, quarter, and analyst. Significance at the 10% (*), 5% (**), or 1% level (***) is indicated.

| Dependent = | AbnReturn $_{[+21,\ +146]}$ (1) | # of analysts (2) | Forecast accuracy (3) | Revision CAR$_{[-1,\ +1]}$ (4) |
|---|---|---|---|---|
| Net buying$_{[+1,\ +20]}$ × Post-EDGAR | 0.461 | | | |
| | (2.84)*** | | | |
| Net buying$_{[+1,\ +20]}$ | 0.065 | | | |
| | (1.08) | | | |
| Post-EDGAR | -0.038 | 0.248 | 0.127 | |
| | (3.69)*** | (3.00)*** | (2.72)*** | |
| Revision × Post-EDGAR | | | | 0.559 |
| | | | | (4.25)*** |
| Revision | | | | 0.253 |
| | | | | (3.46)*** |
| Firm controls | Yes | Yes | Yes | No |
| Firm FEs | Yes | Yes | Yes | No |
| Quarter FEs | Yes | Yes | Yes | No |
| Firm × quarter FEs | No | No | No | Yes |
| Analyst × quarter FEs | N/A | N/A | N/A | Yes |
| Clustering SEs | Firm, quarter | Firm, quarter | Firm, quarter | Firm, analyst, quarter |
| # of observations | 27,693 | 100,652 | 53,539 | 314,426 |
| Adj. R-squared | 0.16 | 0.81 | 0.31 | 0.10 |

**Table 10: Staggered implementation of EDGAR and stock price efficiency**

This table reports regression analysis of the impact of EDGAR on stock price efficiency. The dependent variable is stock price synchronicity (i.e., R-square), the absolute value of stock return autocorrelation, and the standard deviation of the pricing error divided by the standard deviation of the log transaction price, respectively. The unit of observation is a firm-month. All variables are defined in Table 1. Numbers in parentheses are *t*-statistics based on standard errors two-way clustered by firm and month. Significance at the 10% (*), 5% (**), or 1% level (***) is indicated.

| Dependent = | Price synchronicity | | Abs(Stock return autocorrelation) | | Pricing error | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Post-EDGAR | -0.008 | -0.008 | -0.010 | -0.010 | -0.008 | -0.007 |
| | (2.25)** | (2.36)** | (5.41)*** | (5.36)*** | (2.92)*** | (2.83)*** |
| Log(Total assets) | | 0.024 | | -0.013 | | -0.019 |
| | | (6.79)*** | | (9.10)*** | | (8.84)*** |
| Book-to-Market | | -0.023 | | 0.011 | | 0.020 |
| | | (9.42)*** | | (8.07)*** | | (7.92)*** |
| Prior stock return | | 0.025 | | -0.021 | | -0.030 |
| | | (15.98)*** | | (20.68)*** | | (24.51)*** |
| ROA | | -0.009 | | -0.006 | | -0.012 |
| | | (1.20) | | (1.67)* | | (1.96)* |
| Book leverage | | -0.054 | | 0.009 | | 0.028 |
| | | (7.50)*** | | (2.35)** | | (4.89)*** |
| Asset tangibility | | -0.022 | | 0.008 | | -0.004 |
| | | (1.48) | | (1.15) | | (0.41) |
| Sales growth | | 0.004 | | -0.001 | | -0.002 |
| | | (2.82)*** | | (0.72) | | (1.50) |
| CapEx | | 0.036 | | -0.025 | | -0.030 |
| | | (2.52)** | | (3.96)*** | | (3.11)*** |
| R&D | | 0.012 | | -0.008 | | -0.014 |
| | | (1.07) | | (0.87) | | (1.02) |
| Institutional ownership | | 0.099 | | -0.040 | | -0.095 |
| | | (9.99)*** | | (7.83)*** | | (13.05)*** |
| Firm FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Month FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| # of observations | 337,489 | 337,489 | 337,716 | 337,716 | 158,731 | 158,731 |
| Adj. R-squared | 0.64 | 0.65 | 0.17 | 0.18 | 0.39 | 0.42 |