



THE HONG KONG
POLYTECHNIC UNIVERSITY

香港理工大學

Pao Yue-kong Library

包玉剛圖書館

Copyright Undertaking

This thesis is protected by copyright, with all rights reserved.

By reading and using the thesis, the reader understands and agrees to the following terms:

1. The reader will abide by the rules and legal ordinances governing copyright regarding the use of the thesis.
2. The reader will use the thesis for the purpose of research or private study only and not for distribution or further reproduction or any other purpose.
3. The reader agrees to indemnify and hold the University harmless from and against any loss, damage, cost, liability or expenses arising from copyright infringement or unauthorized usage.

IMPORTANT

If you have reasons to believe that any materials in this thesis are deemed not suitable to be distributed in this form, or a copyright owner having difficulty with the material being included in our database, please contact lbsys@polyu.edu.hk providing details. The Library will look into your claim and consider taking remedial action upon receipt of the written requests.

**INFORMATION ACQUISITION FROM SEC
FILINGS AND EARNINGS MANAGEMENT**

JING JIAO

PhD

The Hong Kong Polytechnic University

2019

**The Hong Kong Polytechnic University
School of Accounting and Finance**

**Information Acquisition from SEC Filings
and Earnings Management**

Jing Jiao

A thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

December 2018

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

_____ (Signed)

JING Jiao _____ (Name of student)

ABSTRACT

The literature suggests that investors' information acquisition may have two opposite effects on earnings management. First, greater information acquisition enables investors to better ascertain a firm's operational and financial conditions and thereby to better monitor earnings management. Second, extensive information acquisition, especially the acquisition of interim reports, can induce manager myopia if investors focus on short-term performance. Using the number of downloads of 10-K, 10-Q and 8-K filings to proxy for information acquisition, I find that downloads are negatively related to earnings management, consistent with the monitoring view. In further analyses, I find that this negative relation is stronger when downloads are concentrated among fewer IP addresses, consistent with concentrated information acquisition having a stronger monitoring effect. I also find that the relation is stronger for firms that are more likely to experience agency problems, e.g., firms with weaker external and internal monitoring and firms with incentives to manage earnings. I contribute to the extant research by providing evidence of the influence that investors' information acquisition has on corporate reporting behavior.

Keywords: Information Acquisition, SEC Filings, Earnings Management

ACKNOWLEDGMENTS

Firstly, I would like to express my sincere gratitude to my chief supervisor Prof. Jeffrey NG for the continuous support of my Ph.D. study. In many times, I was so moved by his warm encouragement when I was discouraged by difficulties and obstacles, his great efforts in making good study arrangements on behalf of me, and his patient guidance in helping me with research. I could not have imagined having a better supervisor for my Ph.D. study.

Next, a very special gratitude goes to my former chief supervisor Prof. Nancy Su. She is my first teacher to initiate me into research in my Ph.D. study. I still clearly remember Nancy's great patience when I asked numerous naïve questions in my early years. Her advices on research, career as well as my life have been invaluable.

In addition, profound gratitude goes to my co-supervisor Prof. Agnes Cheng. She provided many valuable resources, such as teaching training workshops, languages classes specifically related to accounting and finance areas, programming training and numerous other visible and invisible supports. I also would like to acknowledge the help from my co-supervisor Dr. Bin Miao, the insightful comments from my BoE Chair, Dr. Yangyang Chen, and two external examiners, Dr. Xin Wang and Dr. Zili Zhuang.

I am also hugely appreciative to Dr. Yuan Ji who offered the coauthoring opportunity, conducted pleasant chat about life and shared valuable experience with me.

Finally, I would like to thank my parents and my husband for their great understanding and continuous supports over the years. A special thanks to my best friend in Ph.D. program, Jiajia Fu for her helps and encouragement. I also thank all my friends and classmates for their company over the years.

TABLE OF CONTENTS

CERTIFICATIE OF ORIGINALITY	i
ABSTRACT.....	ii
ACKNOWLEDGMENTS	iii
CHAPTER 1 INTRODUCTION	6
CHAPTER 2 HYPOTHESIS DEVELOPMENT	16
CHAPTER 3 DATA AND VARIABLES	20
3.1. Sample.....	20
3.2. Information acquisition measures	21
3.3. Earnings management measures	22
3.4. Empirical methodology.....	22
3.5. Descriptive statistics	23
CHAPTER 4 EMPIRICAL RESULTS.....	25
4.1. Baseline regression	25
4.2. Endogeneity tests	26
4.2.1. Firm fixed effects and additional control variables.....	27
4.2.2. Difference-in-difference methodology.....	29
4.2.3. Instrumental variable regressions.....	32
4.3. Robustness tests	35
CHAPTER 5 CROSS-SECTIONAL TESTS	37
5.1. Download concentration	37
5.2. Internal and external monitoring.....	39
5.3. Incentives	41
CHAPTER 6 CONCLUSIONS	44
APPENDIX VARIABLE DEFINITION	46
REFERENCES	49

LIST OF TABLES

Table 1 Summary Statistics	55
Table 2 Industry and Year Distribution of Downloads	56
Table 3 Main Regressions.....	57
Table 4 Endogeneity Issues	58
Table 5 Alternative Measures	61
Table 6 Download Concentration	63
Table 7 Monitoring	64
Table 8 Incentives	65

CHAPTER 1 INTRODUCTION

The financial accounting literature has a long tradition of studying the determinants and economic consequences of corporate disclosure. These studies focus on either the accounting information disclosed or the disclosure quality. However, because there are the costs to acquiring and processing information (e.g., Verrecchia, 1982; Gabaix, Laibson, Moloche, and Weinberg, 2006) as well as limited investor attention (e.g., Chakrabarty and Moulton, 2012; Ben-Rephael, Da, and Israelsen, 2017), the amount of information disclosed might not reflect the amount of information acquired by investors. Over the last several decades, developments in information technology have largely facilitated information acquisition and analysis. As technologies improve, the Securities and Exchange Commission (SEC) continues to update regulations to improve the efficiency of the information acquisition process. For example, in 1993 the SEC started to require electronic filing on the Electronic Data Gathering, Analysis and Retrieval (EDGAR) system, which vastly decreases the time and expense needed to acquire disclosure. More recently, the SEC requires companies to provide financial data using the eXtensible Business Reporting Language, which also reduces acquisition costs (SEC, 2009; Bhattacharya, Cho, and Kim, 2018; Kim, Li, and Liu, 2019).

While much of the prior accounting literature focuses on disclosure, more recently, the literature has started to pay more attention to information acquisition. For instance, a strand of the literature documents investors' information acquisition activities via various venues, including SEC filings in the EDGAR system (Drake, Roulstone, and Thornock, 2015, 2016),

Google search (Drake, Roulstone, and Thornock, 2012), Yahoo Finance (Lawrence, Ryans, and Sun, 2017), Wikipedia (Xu and Zhang, 2013) and the Food and Drug Administration (Gargano, Rossi, and Wermers, 2017). Another strand of the literature examines the information acquisition activities of different users, including auditors (Drake, Lamoreaux, Quinn, and Thornock, 2019), regulators (Stice-Lawrence, 2017), the Internal Revenue Service (Bozanic, Hoopes, Thornock, and Williams, 2017) and the Federal Reserve (Li, Lind, Ramesh, and Shen, 2017). Additionally, more recent literature also investigates the predictability of information acquisition for future earnings and returns (Bartov, Faurel, Mohanram, 2017; Tang, 2018) as well as its impact on market liquidity (Blankespoor, Miller, and White, 2013) and the efficiency of market reaction to information (Drake et al., 2015). However, few studies examine the impact of information acquisition on corporate disclosure. While firms are providers of information disclosure, it is not clear whether and how users' usage behavior influences corporate disclosure. My paper aims to fill this gap by studying the "feedback effect" of corporate disclosure.

In this study, I investigate the relation between information acquisition and firms' earnings management activities. This inquiry is motivated by the differing views that underlie this relation. The two competing views can be briefly described as follows. The monitoring view predicts that more information acquisition makes investors better informed and thus better able to monitor firms' activities, which decreases managers' likelihood of engaging in earnings management activities. A greater number of downloads of SEC filings can be generated by more investors

downloading the filings or by a smaller number of investors downloading more filings. If more investors download a firm's SEC filings, such high investor attention would discourage managers' opportunistic behaviors and thus decrease earnings management. Prior literature shows that investor inattention increases managers' opportunistic behavior (Kempf, Manconi, and Spalt, 2016), while the attention of professional investors such as analysts can serve as an external monitor for managers (e.g., Yu, 2008; Cheng and Subramanyam, 2008; Chen, Harford and Lin, 2015; Kim, Lu, and Yu, 2018; Ellul and Panayides, 2018). Prior literature also shows that different types of filings are all informative (e.g., Griffin, 2003; Callen, Livnat, and Segal, 2006; Lerman and Livnat, 2010; Doyle and Magilke, 2013; Bird and Karolyi, 2016). The acquisition of different types of information from different filings enables investors to gain a better understanding of a firm's operational and financial conditions and increases the effectiveness and timeliness of investor monitoring.¹ Accordingly, the monitoring view suggests that information acquisition is negatively associated with earnings management.

However, the pressure view predicts that more information acquisition can exert undue pressure on managers and increase earnings management. Greater acquisition of performance information may indicate that investors place high emphasis on short-term performance. This is especially true when investors frequently request interim or current reports,

¹ For example, the comprehensive information in a firm's Form 10-K gives investors company information as well as annual performance and corporate governance information. Form 10-Q enables deep insight into a business's current performance and predicts future annual earnings. Various reportable events in Form 8-K give investors the most timely information. For example, events related to asset disposition may indicate whether managers are choosing to sell assets with unrealized gains for earnings management purposes.

which contain information related to short-term performance. Prior literature finds that when a firm changes its reporting frequency from annual or semiannual to quarterly reporting, managers decrease investments (Kraft, Vashishtha, and Venkatachalam, 2018) and engage in earnings management (Ernstberger, Link, Stich, and Vogler, 2017). If the provision of quarterly reporting induces managers' myopia, investors acquiring more information from quarterly reporting would have a similar impact. Furthermore, investor attention itself can exert pressure on managers. A manager who does not meet investors' expectations is more likely to be penalized through cuts to compensation or a forced departure if more investor attention is associated with higher investor engagement in monitoring. Prior literature finds that transitory institutional investors lead to earnings management activities such as R&D investment cuts (Bushee, 1998; Matsumoto, 2002) and that pressure from analysts undermines innovative activities (He and Tian, 2013), promotes real earnings management activities (Irani and Oesch, 2016), and increases the likelihood of beating or meeting analyst forecasts (Huang, Pereira and Wang, 2017). Hence, based on the pressure view, the acquisition of SEC filings is expected to be positively associated with earnings management.

To investigate the relationship between information acquisition and earnings management, I study investors' downloads of SEC filings, including Forms 10-K, 10-Q and 8-K, from EDGAR as a proxy for information acquisition.² Specifically, I construct an aggregate yearly measure by summing the downloads of the last year's 10-K and the current

² For the remainder of the paper, I use the terms "information acquisition" and "download" interchangeably. Both of them refer to users retrieving documents from the SEC EDGAR system.

year's 10-Qs and 8-Ks filings within a year. I focus on these three filing types for several reasons. First, the credibility of SEC filings is much higher than that of information from other sources (e.g., the internet or social media). Investors generally assess information based on its credibility (Thayer, 2011). Second, the information in these three types of SEC filings is most likely related to firm performance and thus to earnings management activities. Form 10-K gives the most comprehensive company information, while Form 10-Q supplements a firm's annual report by showing its interim performance and Form 8-K presents timely information on material events that significantly impact the firm. These filings' different characteristics determine the extent to which investors can obtain the comprehensive and timely information about a firm. Finally, of all filing types, these three types of filings have the highest download rates. It is thus interesting and important to examine how investors' use of these filings influences earnings management.³

Relying on a sample from 2003 to 2015, I find that firms with more downloads have lower discretionary accruals, consistent with the monitoring role of information acquisition. I also find that the impact of downloads is significant even after controlling for the number of filings which is a proxy for the volume of the information disclosure. This finding suggests that in addition to information disclosure, information acquisition has an incremental effect on reducing earnings management.

³ While other types of SEC filings may be helpful for detecting earnings management, this paper mainly examines the impact of filings that are more representative of investors' acquisition of firms' general financial related information. Which filing is the most important for monitoring earnings management is not my focus.

Nevertheless, the main findings might suffer from several endogeneity problems. For example, firms with better accounting quality are likely to attract more investors and thereby will have more frequent downloads of their SEC filings, leading to a reverse causality bias. It is also possible that the existence of omitted variables influences the number of downloads as well as earnings management, leading to an omitted variable bias. Therefore, to hopefully strengthen causal inference, I conduct several tests to mitigate endogeneity concerns. First, I include firm fixed effects in the regressions to control for unobserved time-invariant firm characteristics and continue to find a negative association between information acquisition and earnings management. I also include additional control variables, including media coverage, the number of questions asked in conference calls, CEO turnover, and the uncovering of misstatements, all of which might drive the number of downloads. I find that the effect of information acquisition is significant after including these variables.

Second, I implement a quasi-natural experimental research design that exploits an exogenous shock to information acquisition. Specifically, I rely on the implementation of a new regulation for 8-K filing on August 23, 2004 as a shock.⁴ This regulation increases the number of reportable events that are filed in Form 8-K and accelerates the filing date. Investors can acquire more information about material events in a timely manner and are thereby more likely to download 8-K filings. Thus, the implementation of this new regulation provides a shock to the number of 8-K filing

⁴ Final rule: Additional Form 8-K Disclosure Requirements and Acceleration of Filing Date, release no. 33-8400, release no. 34-49424 (August 23 2004).

downloads.⁵ Using a difference-in-difference methodology, I find that firms that are more exposed to this regulation reduce earnings management more than those that are less exposed.

Next, I now rely on the instrumental variable approach to examine whether information acquisition has an effect on earnings management. I use the navigability of the Investor Relations tab on a company's website homepage as the instrument. A firm's website is another major source, other than the SEC's EDGAR, for various SEC filings. I expect the number of downloads of SEC filings to be smaller for firms where this tab is more navigable, because easier navigation helps investors more easily find and download SEC filings from the company website. Using a two-stage least squares regression with the above instrument, I find that more information acquisition reduces earnings management.

Additionally, I conduct several robustness tests to determine whether my baseline results are robust to alternative measures in the main regressions. First, I use alternative dependent variables. I find my results are robust to using performance-matched discretionary accruals as well as two model-free measures, which indicates whether a firm's earnings meet or beat analysts' earnings forecasts. I then consider alternative measures of the downloads of SEC filings. Using the three-day abnormal returns [-1, +1] around a filing date to differentiate between whether a filing is good and bad news for investors, I find that downloading filings that relate both positive and negative news can decrease earnings management. I also differentiate between the effect of financial and non-financial related news. Similarly,

⁵ I provide evidence in section 4.2.2 that this regulation indeed increases the number of downloads.

downloading both financial and non-financial news is helpful in keeping investors abreast of the firm's operational and other activities, allowing better monitoring.

I further my analyses by implementing several cross-sectional tests. First, using the number of IP addresses and the download concentration ratio to proxy for the download concentration, I find that the effect of downloads on earnings management is stronger when the downloads are concentrated among fewer IP addresses. This finding suggests that more concentrated, as opposed to more dispersed, attention, has a stronger monitoring effect.

Second, there are external and internal monitoring mechanisms that can also play similar roles in monitoring earnings management, such as external auditors (Becker, DeFond, Jiambalvo, and Subramanyam, 1998; Markelevich and Rosner, 2013), analysts (Yu, 2008), institutional investors (Bushee, 1998), and internal control procedures (Doyle, Ge, and McVay, 2007; Ashbaugh-Skaife, Collins, Kinney Jr, and LaFond, 2008; Chan, Farrell, and Lee, 2008). I expect the monitoring effect via information acquisition from SEC filings to be stronger if a firm has weak external and internal monitoring mechanisms. Consistent with this expectation, I find the effect of downloads to be more pronounced when a firm is audited by a non-big 4 auditor, is covered by fewer analysts, has lower institutional ownership, or has internal control weaknesses over financial reporting.

Finally, prior literature finds that the need to raise capital (Cohen and Zarowin, 2010) and managers who are early in the career (Bergstresser and Philippon 2006; Ali and Zhang 2015) create incentives to manage

earnings. I find that the monitoring effect of investors' information acquisition on earnings management is more pronounced for firms that have a seasoned equity offering in the following year and for those with a CEO who is in the early years of their service. This finding is consistent with the monitoring effect of information acquisition being stronger when there are incentives to manage earnings.

My analysis contributes to the extant research in several ways. First, my paper contributes to the literature related to information acquisition. Investors are time- and resource-constrained, which leads to an imbalance between the actual information acquired by investors and the information disclosed by firms. Prior literature investigates the determinants of information acquisition from various sources (Xu and Zhang, 2013; Drake et al., 2012, 2016; Gargano et al., 2016; Lawrence et al., 2017) and of different users (Bozanic et al., 2017; Li et al., 2017; Stice-Lawrence, 2017; Drake et al., 2019). A few studies also investigate the capital market consequences of information acquisition (Drake et al., 2015; Bartov et al., 2017; Tang, 2018). However, there is little empirical evidence on the impact of information acquisition on corporate disclosure. My study contributes to this vein of the literature by showing that information acquisition has a monitoring effect on opportunistic corporate disclosure. In doing so, I also add to the recent literature on how feedback from capital market participants affects corporate disclosure (Zuo, 2016).

Second, this paper contributes to the earnings management literature by documenting a source of monitoring. Prior literature finds that institutional investors and analysts have a monitoring effect on earnings

management (e.g., Koh, 2007; Irani and Oesch, 2016; Huang et al., 2017). While investors take other actions to monitor earnings management (e.g., Dou, Hope, Thomas, and Zou, 2014), information acquisition is one of the most basic factors determining whether investors are able to monitor earnings management. This study shows that acquiring more information is the basis by which investors, as outsiders, understand and monitor a firm.

Lastly, the paper also contributes to the literature related to SEC filings. Prior literature offers evidence of the market response to information disclosed in SEC filings (e.g., Griffin, 2003; Callen et al., 2006; Lerman and Livnat, 2010; Zhao, 2017; McMullin, Miller, and Twedt, 2019). However, given the large variation in information acquisition across firms, it is not clear whether and how the acquisition of SEC filings would influence firms' behaviors. Although investors and analysts can acquire information through conferences and private meetings, the results in my paper suggest that SEC filings are crucial to helping investors better understand a firm and thereby mitigating its managers' opportunistic behaviors.

The rest of this paper is organized as follows. Section 2 reviews prior research and develops the main hypothesis. Section 3 describes the data, variables and summary statistics. Section 4 reports the main findings and the tests that address endogeneity issues, while Section 5 presents results from the cross-sectional analyses. Section 6 concludes.

CHAPTER 2 HYPOTHESIS DEVELOPMENT

The SEC has long required that firms make public disclosures. An extensive accounting literature focuses on firms' choices about what information to supply and its quality and the information content (e.g., Botosan, 1997; Leuz and Verrecchia, 2000; Lang and Lundholm, 2000). Recently, researchers have turned to investors' information acquisition. For example, Drake et al. (2012) investigate the factors that influence investors' demand for public information via Google searches. Drake et al. (2015, 2016) and Shevlin and Thornock (2015) examine the determinants and capital market consequences of information acquisition from SEC filings. Bozanic et al. (2017) study the factors that influence the US Internal Revenue Service's (IRS) information acquisition from SEC filings. In addition to studies documenting the determinants of information acquisition, several studies also investigate the capital market impact of information acquisition. For example, Blankespoor et al. (2013) examine the impact of information dissemination via Twitter on market liquidity and find a positive relation, consistent with investors' greater information acquisition having an impact on capital markets. Drake et al. (2015) find that information acquisition is positively associated with the market's initial reaction to the news and negatively associated with post-announcement return drift. However, no study has considered how investors' information acquisition influences corporate behavior. In this paper, I examine the relation between information acquisition and earnings management.

Extant research presents competing predictions on the relation between information acquisition and earnings management. Based on the

monitoring view, more information acquisition could enable investors to better monitor firms' activities, which would make managers less likely to engage in earnings management activities. Bushman and Smith (2001) highlight that the use of financial accounting information can promote efficiency in the governance processes of corporations. I conjecture that enhanced monitoring from the downloads of SEC filings can be from two sources. First, broadly speaking, more information acquisition can be associated with more investor attention. Prior literature finds that analyst attention decreases earnings management (Yu, 2008), default risk (Cheng and Subramanyam, 2008), and agency conflicts (Chen, Harford, and Lin, 2015), while investor inattention increases managers' opportunistic behaviors (Kempf et al., 2016). For example, constructing an investor inattention measure by exploiting shocks to unrelated portions of institutional investors' portfolios, Kempf et al. (2016) find that firms that receive low investor attention are more likely to engage in opportunistic behavior, such as an opportunistically timed CEO stock option grant, dividend cut, or a reduction in the CEO turnover rate. To the extent that information acquisition from SEC filings reflects the attention investors pay to a firm and improve corporate governance, it is expected to be negatively associated with earnings management.

Second, more information acquisition from SEC filings can be associated with investors' downloads of more types of filings, which results in investors having a more comprehensive understanding a firm. The SEC requires various filings, including 10-K, 10-Q and 8-K filings. All these filings are informative (e.g., Griffin, 2003; Callen et al., 2006; Lerman and

Livnat, 2010; Doyle and Magilke, 2013; Bird and Karolyi, 2016). At the same time, each of these filings has a different focus on a specific type of information. For example, the 10-K filing provides the most comprehensive summary of a company's performance and other information. The 10-Q filing shows less detailed but more timely financial performance. The 8-K filing contains information about material events which have significantly impacted a firm. Acquiring more information from the different types of SEC filings enables investors to assemble a more complete picture of a firm's conditions and thereby improves monitoring. Accordingly, the monitoring role played by investors suggests that the number of SEC filing downloads negatively affects earnings management.

In contrast to the monitoring view, another line of research argues for a pressure effect, which suggests that information acquisition can increase earnings management. Greater acquisition of performance reports, especially interim reports, may indicate investors' heavy focus on short-term performance. Using the transition of U.S. firms from annual to semi-annual and then to quarterly reporting over the period from 1950 to 1970, Kraft et al. (2018) find that increased reporting frequency is associated with an economically large decline in investments. Similarly, using a European Union setting, where the reporting frequency increased with the introduction of a mandate to issue Interim Management Statements on a quarterly basis, Ernstberger et al. (2017) find an increase in real activities manipulations. Based on the above evidence, it is possible that greater acquisition of interim reports would also increase manager myopia because of investors' focus on short-term performance. Prior literature documents evidence of

such performance pressure on managers from investors and analysts. For example, Bushee (1998) finds that transitory institutional investors pressure managers to engage in myopic behaviors, such as cutting R&D investments. Matsumoto (2002) finds that managers manage earnings upward to avoid negative earnings surprises. Prior literature also provides evidence of such pressure from analysts. For example, He and Tian (2013) also document that pressure from greater analyst coverage undermines a firm's innovative activities. Irani and Oesch (2016) find that to meet analysts' expectation, firms engage in real earnings management when analyst coverage is high. Huang et al. (2017) also find that firms with high analyst coverage are more likely to beat or meet earnings benchmarks. Accordingly, based on the pressure framework, information acquisition from SEC filing requests is expected to be positively associated with earnings management.

In light of these conflicting views, the relation between information acquisition and earnings management remains an empirical issue. The monitoring view predicts a negative relation while the pressure view envisages a positive one. No significant relation is also possible if the two effects cancel each other out. Given the conflicting views on the effect of information acquisition on earnings management, I treat my main hypothesis as ultimately an empirical question and state it non-directionally as follows.

H1: Information acquisition is associated with earnings management.

CHAPTER 3 DATA AND VARIABLES

3.1. Sample

The data on information acquisition from SEC filings comes from the SEC; its EDGAR system logs all search traffic.⁶ My sample consists of all firms between 2003 and 2015. The sample period begins in 2003 because the EDGAR server logs start on January 1 of that year. Each observation in the raw data contains information, including each visitor's IP address, the date and time a request was made, the firm's Central Index Key (CIK), and the accession numbers of each filing. I requested the raw data from the EDGAR log website using Python. I match the CIK in the EDGAR to COMPUSTAT to identify the firms. To identify the filings, I use the accession number to the Master Index File from the SEC website. In this paper, I assume that the acquired information is actually used by investors. However, downloads of SEC filings by automated web-crawlers may go unread and thus have a limited impact on earnings management. Following prior literature (e.g., Lee, Ma, and Wang, 2015; Drake et al., 2015), I clean the data by eliminating robot-generated requests. Specifically, I delete the requests from IP addresses that access more than five filings per minute or 1,000 filings per day.

The data described above is then merged with data from other sources. I obtain analyst coverage data from the I/B/E/S database (I/B/E/S) and institutional ownership data from the Thomson Reuters Institutional (13f) Holdings. Data related to CEOs comes from Execucomp. I also use

⁶ <https://www.sec.gov/data/edgar-log-file-data-set>.

seasonal equity offerings data from Thomson SDC database and internal control data from AuditAnalytics.

3.2. Information acquisition measures

In this paper, I use investors' downloads of SEC filings to capture their financial information acquisition. I mainly focus on downloads of the previous year's annual financial report (10-K filings) and the current year's quarterly financial reports (10-Q filings) and current reports (8-K filings). Form 10-K is an annual report required by the SEC. It gives the most comprehensive and reliable summary of a company's financial performance and other important information. It also provides the most comparable earnings number to the current year's earnings. Therefore, it appears that the information in 10-Ks enables investors to form a basic assessment of a company, which allows them to better monitor it. While the quarterly reports filed in Form 10-Q are less detailed and generally less reliable than the annual reports are, they are better at informing investors about recent financial performance. Form 8-K is one of the most common filed forms with the SEC. A company must file a Current Report on Form 8-K within four business days, so it helps investors obtain timely information on a firm's material changes that may have a significant impact on its operations or financial conditions. Overall, the three types of filings all aid in informing investors and thus are all useful in monitoring a firm.

Accordingly, I construct an aggregate measure to proxy for investors' entire acquisition of financial and non-financial information that has a significant impact on the firm's financial performance. Specifically, I add the number of downloads of the previous year's 10-K filing, three of the

current year's 10-Qs filings and all of the current year's 8-Ks filings together. By adding all the downloads together, this measure is able to capture the degree of attention investors give to a firm.

3.3. Earnings management measures

I use the modified Jones model (Dechow, Sloan, and Sweeney, 1995) to estimate accrual-based earnings management. The model is estimated for each Fama-French 48 industry and year with more than 20 observations, as follows:

$$Accruals_t = \alpha_1 + \alpha_2(\Delta REV_t - \Delta AR_t) + \alpha_3 PPE_t + \varepsilon_t, \quad (1)$$

where t indexes the year; *Accruals* is income before extraordinary items minus operating cash flows; ΔREV is the change in sales; ΔAR is the change in accounts receivables; and *PPE* is gross property, plant and equipment. All variables are scaled by lagged total assets. The discretionary accruals are the residuals from the above regression. The results are robust to estimating equation (1) by SIC industry classification. In Section 4.3, I also use alternative earnings management measures.

3.4. Empirical methodology

I examine the effect of information acquisition on earnings management following the model below:

$$\begin{aligned} Discretionary\ Accruals = & \beta_0 + \beta_1 Downloads + \beta_2 Filings + \beta_3 Size + \beta_4 BM + \\ & \beta_5 Leverage + \beta_6 IO + \beta_7 Analyst + \beta_8 ROA + \beta_9 CFO + \beta_{10} Lagged_Accruals + \\ & \beta_{11} Lagged_NOA + \beta_{12} Age + \beta_{13} Growth + \varepsilon. \end{aligned} \quad (2)$$

The dependent variable is discretionary accruals estimated based on the description in Section 3.3. The variable of interest is *Downloads*, which is used to capture investors' information acquisition from SEC filings in the

EDGAR system. The control variables consist of three categories. The first comprises firm characteristics that may affect managers' incentives to manage earnings, including firm size (*Size*), book-to-market ratio (*BM*), firm leverage (*Leverage*), firm age (*Age*), financial performance (*ROA* and *CFO*) and firm growth (*Growth*). Next, I control for total accruals (*Lagged_Accruals*) and net operating assets (*Lagged_NOA*) at the beginning of year *t*, which reflects the extent of previous earnings management (Barton and Simko, 2002). Finally, I also control for institutional ownership (*IO*) and analyst following (*Analyst*) which significantly impact the number of downloads. The t-statistics are computed using standard errors that are robust to heteroskedasticity and clustering at the firm level. Industry (at the 2-digit SIC level) and year fixed effects are included.

3.5. Descriptive statistics

Table 1 presents summary statistics for all the key variables. The sample size is 40,117 firm-year observations with non-missing control variables from 2003 to 2015, representing 6,354 firms. Since the raw value of the number of downloads is highly skewed, I use the natural logarithm of downloads. The mean of *Downloads* is 7.344, which corresponds to 2,793.72 downloads of a firm's three types of SEC filings in a year. The standard deviation of *Downloads* is 1.023, showing a deviation of downloads among firms. The mean of *Filings* is 16.090, implying that on average a firm would file approximately 16 10-Ks, 10-Qs, 8-Ks and their amended filings in a year. The descriptive statistics of the other variables are similar to those in prior studies.

[Insert Table 1 here]

Table 2 demonstrates the industry and year averages for the downloads. As shown in Panel A, manufacturing firms account for the largest group in my sample. Firms in the agriculture, forestry and fishing industries have the highest number of downloads while those in the public administration and other non-classifiable industries have the lowest number. The variation in the number of downloads does not seem to be very large across industries. Table 2, Panel B shows the sample distribution and the average number of downloads over the years. The table shows an increasing trend in the number of downloads, suggesting that recent years have seen investors increasingly use SEC filings.

[Insert Table 2 here]

CHAPTER 4 EMPIRICAL RESULTS

In this section, I test the hypothesis that investors' information acquisition is associated with earnings management. Section 4.1 examines the primary relation and Section 4.2 describes tests intended to address endogeneity issues, while Section 4.3 reports tests using alternative measures to enhance robustness.

4.1. Baseline regression

I begin by performing a multivariate regression analysis on the relation between information acquisition and earnings management. The baseline regression results are reported in Table 3. To provide some comparisons, Columns (1) and (2) present the results from tests for the effect of information acquisition (*Downloads*) and disclosure (*Filings*) on earnings management, respectively. In Column (1), the coefficient on *Downloads* is -0.021 (t-statistic=-6.38), implying that firms with more downloads have lower discretionary accruals. This finding suggests that there is a negative relation between information acquisition and earnings management, which is consistent with the monitoring role of investors' information acquisition from SEC filings. In Column (2), the coefficient on *Filings* is -0.001(t-statistic=-3.24). Both coefficients are negative and significant at the one percent significance level. Above evidence suggests that both information acquisition and disclosure can influence investors' monitoring role on firms' earnings management activities.

Column (3), which is the main regression for this paper, presents the results from the same regression models as those in Columns (1)-(2), but with *Filings* included as an additional control variable. As shown in the

table, after including *Downloads* and *Filings* simultaneously in the regression, the coefficient on *Downloads* remains significant (t-statistic=-5.65), while the coefficient on *Filings* becomes insignificant (t-statistic=0.05). This finding shows that information acquisition has an incremental effect on earnings management activities. To investigate the effect of information acquisition in addition to information disclosure, *Filings* is controlled in all following analyses.

The results for the control variables are largely consistent with the recent literature. Specifically, discretionary accruals are positively related to size, book-to-market, institutional investor ownership, ROA and firm age, and negatively related to leverage, analyst following and operating cash flow. The coefficients on *Lagged_Accruals*, *Lagged_NOA* and *Growth* are insignificant. Untabulated results show that the largest variance inflation factor (VIF) is 3.71, mitigating the multicollinearity concern.

Overall, I provide evidence supporting the main hypothesis that high information acquisition from SEC filings decreases discretionary accruals, which is consistent with the monitoring role of information acquisition on earnings management behavior.

[Insert Table 3 here]

4.2. Endogeneity tests

In the baseline regression, I use the most straightforward way to examine how information acquisition from SEC filings influences earnings management by regressing discretionary accruals on the number of downloads. However, the regression results are likely to suffer from endogeneity issues, such as omitted variable and reverse causality issues.

For example, it is possible that investors require more information because of omitted firm characteristics that are related to earnings management behavior. With regard to the reverse causality issue, a negative association between information acquisition and earnings management may reflect the fact that firms with better accounting quality are more likely to attract more investors, thus leading to more downloads of SEC filings, as opposed to the reverse causal impact of information acquisition on earnings management. Therefore, in this section, I conduct several tests that address endogeneity concerns.

4.2.1. Firm fixed effects and additional control variables

In the baseline regressions, I include industry and year fixed effects because different industries' characteristics can lead to investors' different information needs. For example, a firm in the high tech industry is more complex to analyze than one in the retail industry, so investors have a stronger motivation to acquire information about a high tech firm in order to monitor its business. I thus include industry fixed effects to control for variations in downloads and discretionary accruals across sectors of the economy. However, time-invariant firm characteristics can also affect the level of information acquisition, as well as the association between information acquisition and earnings management. Therefore, to offer more evidence of the causality relationship, in this section I include firm fixed effects to control for unobserved heterogeneity across firms, which leaves only within-firm variation to estimate the effect.

The results are presented in the Column (1) of Panel A of Table 4 and are qualitatively the same as those in my main regressions. Specifically,

the coefficient on *Downloads* remains negative and statistically significant (t-statistic=-4.67). This evidence suggests that my core findings persist after including firm fixed effects.

Next, I examine whether my results are robust to the inclusion of several additional control variables. First, I consider alternative information acquisition channels. Investors can acquire information not only from the SEC EDGAR system but also other sources such as the media. Using media data from RavenPack, I construct a variable of the number of news articles published about a firm in a fiscal year, *Media*, and include it as a control variable. Column (2) of Panel A of Table 4 presents the results. As shown in the table, the coefficient on *Downloads* remains negative and statistically significant (t-statistic=-2.65). Second, I re-estimate the regression after including the number of questions asked in conference calls. To the extent that many downloads of SEC filings follow earnings announcements, the negative association between downloads and earnings management may only reflect investors' attention to earnings announcements. Thus, I control for the number of questions asked in the conference calls. As shown in Column (3) of Panel A of Table 4, the coefficient on *Downloads* is still statistically significant (t-statistic=-2.46). Moreover, I control for CEO turnover in the regression. A firm hiring a new CEO may experience business reconstruction as well as attract investors' attention. The results in Column (4) show that after controlling for CEO turnover, the coefficient on *Downloads* is still statistically significant (t-statistic=-2.39). Finally, I include the uncovering of a prior misstatement as a control variable. When a firm's prior misstatement is uncovered, investors may pay more attention to

this firm, which increases the downloads of SEC filings. As shown in Column (5), the coefficient on *Downloads* remains negative and statistically significant (t-statistic=-5.45).

Overall, these results suggest that my findings are not materially sensitive to time-invariant firm characteristics, alternative channels of information acquisition as well as certain events that can trigger a short-term increase in downloads of SEC filings.

4.2.2. Difference-in-difference methodology

To better address endogeneity concerns, in this section I explore the effects of a positive exogenous shock to a firm's information acquisition due to a change in regulations for 8-K filing and employ a difference-in-difference methodology.

On August 23, 2004, the SEC implemented a new regulation for 8-K filing, which sought to provide investors with better and timelier information about significant events. Under this new rule, the number of reportable events under the form increased to 22. The SEC also shortened the filing deadline for most items. Firms need to file Form 8-K within four business days after the occurrence of a triggering event. The expanded disclosure of events allows investor to acquire, in a timely way, material information instead of waiting for the next periodic report. As documented in Carter and Soo (1999), timely filings contain valuable information content; delayed filings do not. With the acceleration of the filing date, investors are able to increasingly depend on current reports to gain a better understanding of a firm.

Because of the increased number of reportable events, firms must file more 8-Ks and investors are thereby able to download more of these forms. Relying on this exogenous shock, I employ a difference-in-difference methodology to gain more insight into causality. Specifically, I focus on a short window covering the effective date of the SEC's new regulation. For the pre-shock year, I use data from the latest fiscal year that ended before the regulation became effective. For the post-shock year, I use data from the fiscal year after the regulation became effective. The sudden change in requests due to the regulation change is a more exogenous test of the causality relation. In addition, the results from the longer window may be influenced by other confounding events, such as passage of the Sarbanes Oxley Act.

If the new rule is effective in providing investors with timely information and thereby helps them to play a better monitoring role, it is expected that the firms most affected by the regulation will exhibit the greatest decrease in earnings management. Therefore, I determine the treatment and control samples based on the extent to which the new regulation is likely to impact a firm. Specifically, I calculate the number of 8-K filings that contain the item "Other Events" and define these as voluntarily filed 8-Ks. One major change in the new regulation is that firms must file more 8-Ks based on an expanded list of reportable events. If a firm provided voluntary 8-K disclosure before the rule change took effect, it is less likely to be strongly influenced by the new regulation. In contrast, if a firm was reluctant to provide voluntary 8-K disclosure, it is more likely to be impacted by the regulation. Therefore, a firm is in the treatment group if

the number of its voluntarily filed 8-Ks is below the sample median; if the number is above the median, it is in the control group.

Unlike the main regressions in Table 3, which use the total downloads of 10-K, 10-Q and 8-K filings, here I use the downloads of 8-K filings in the difference-in-difference specification. I first test the validity of using the change in regulation mentioned above as a shock to the downloads of 8-K filings. Panel B1 of Table 4 presents the results. As shown in the table, the number of downloads increases for both the control and treatment groups. Specifically, the average number of downloads from 304.45 (112.76) to 412.86 (226.34) for the control (treatment) group. It suggests that after the change takes effect, the treatment group experiences a 100.73% increase in downloads while the control group only has a 35.61% increase. This finding is consistent with my prediction that the change in the Form 8-K disclosure rule has a significant impact on the treatment firms with regard to the downloads of their 8-K filings.

Table 4, Panel B2 reports the results from the test of the difference-in-difference analysis. As shown in the table, the coefficient on the interaction item, *Treat*Post*, is negative and statistically significant with a value of -0.022 and a t-statistic of -2.40, suggesting that compared to the control group, after the change in regulation the treatment group has lower discretionary accruals. This finding is consistent with my prediction of a negative relation between information acquisition and earnings management. It provides support for the notion that when investors are more active in acquiring information about a firm, its managers are less willing and/or able

to engage in earnings management. Thus, an important moral hazard problem is mitigated.

4.2.3. Instrumental variable regressions

My third test, which addresses potential endogeneity issues, employs an instrumental variable technique. This approach relies on the notion that the instrument variable is a significant determinant of information acquisition from SEC filings and arguably it does not have a direct effect on earnings management activities.

I select the navigability of the Investor Relation (or Investor) tab on a firm's official website homepage as the instrumental variable. Specifically, I define a dummy variable, *Investor_Relation*, which is equal to one if: (1) the firm's Investor Relation tab is placed at the top of the homepage or (2) the Investor Relation tab is a secondary tab below either About Us or the Company main tab, which is placed at the top of homepage, and zero otherwise.

The navigability of the Investor Relation tab meets the relevance condition of an instrumental variable. Company websites are another major source, other than SEC EDGAR, for various SEC filings. It is well known that the vast majority of firms provide different document formats (e.g., pdf, excel, html etc.) of SEC filings on their websites.⁷ Since firms tend to disclose a comprehensive set of financial information together on a subpage of the website, investors can always navigate to the firm's SEC filings by clicking on the Investor Relation tab. Navigability reflects the ease with which website users can move around to find needed information. Therefore,

⁷ Using a sample of S&P 500 firms, I find 100% of them provide SEC filings on the website.

the prominence of the Investor Relation tab on the homepage determines how easily an investor can acquire the SEC filings. Prior literature finds that a website's navigability is an important determinant of user visits (e.g., Rosen and Purinton, 2004; Tarafdar and Zhang, 2008; Al-Qeisi, Dennis, Alamanos, and Jayawardhena, 2014). Accordingly, I expect that when that tab is on the top of the homepage, there would be more downloads of SEC filings from the company website and thus fewer downloads in the SEC EDGAR system.

Further, the navigability of the Investor Relation tab is unlikely related to firms' earnings management activities. Prior research shows that the decision and content of Internet financial reporting is influenced by various factors, such as firm size (Marston and Polei, 2004), information asymmetry (e.g., Ettredge, Richardson, and Scholz, 2002) and corporate governance (e.g., Kelton and Yang, 2008). However, given that all S&P 500 firms provide SEC filings on their company website and have an Investor Relation tab on the homepage,⁸ the location of this tab is less likely to be determined by the above factors. Instead, as the starting point for most user visits, homepage design is largely influenced by the facilitation of e-commerce, marketing purposes and aesthetics (e.g., Pandir and Knight, 2006; Tarafdar and Zhang, 2008). For example, luxury homepage design tends to use darker background colors and a substantial reduction in the number elements (Kluge, Königsfeld, Fassnacht, and Mitschke, 2013). More importantly, the design of a website homepage is unlikely to frequently change over the years, while earnings management depends on the current

⁸ This finding is based on my data collection in May 2019.

year's performance. Thus, the navigability of the Investor Relation tab meets the exclusion restriction for an instrumental variable.

In May 2019, I hand collected the data of the location of the Investor Relation tab on firms' homepages for S&P 500 firms. I then match this data to the earnings management measures. In this regression, I focus on the most recent three-year period of the main sample because it is more likely that the website configuration will not be dramatically different from 2019. The results are robust to using a five-year or longer period.

Panel C of Table 4 presents the results. In Column (1), the results of the first-stage regression are reported. As shown in the table, the coefficient on *IV* is negative and statistically significant at the 1% level (t-statistics=-3.73). This finding is consistent with my prediction that when the Investor Relation tab is more prominently located, it is easier to find SEC filings on company website, investors would download fewer filings on the SEC EDGAR system. The results of the second-stage regression are reported in Column (2). I find that the coefficient on *Instrumented Downloads* is positive and statistically significant with a t-statistic of 2.02, indicating that firms with a lower predicted number of downloads have a higher level of discretionary accruals, which is consistent with the main results presented in Table 3.⁹ I caution that the analysis in this section relies on the validity of the navigability of the Investor Relation tab as an instrument. The analysis is intended to provide some supplementary evidence about the effect of downloads on earnings management. One limitation of the instrument is that the website configuration during the data collection period might be

⁹ The partial F-statistic is statistically significant (F-statistic=13.89), which helps confirm that a weak instrument variable is not a concern. The Wu-Hausman test (F-statistic=6.65) rejects the null hypothesis that *Downloads* is exogenous.

different from what it was during the sample period. These results should be interpreted with caution.

[Insert Table 4 here]

4.3. Robustness tests

In this section, I conduct several robustness tests to check whether my baseline results are robust to alternative measures. Specifically, I consider alternative dependent variables and different construction of downloads variables.

First, I re-estimate the regressions after replacing the dependent variable, which is the earnings management measure based on the modified Jones model (Dechow et al., 1995), in the main regressions. In this section, I use the measure of performance-matched discretionary accruals in Kothari, Leone, and Wasley (2005), as well as the model-free measures that are based on whether a firm's earnings meet or beat analysts' forecasts. The results are reported in Table 5, Panel A. In Column (1), the dependent variable is performance-matched discretionary accruals. The coefficient on *Downloads* is negative and statistically significant at the one percent level (t-statistic=-3.14), consistent with the finding in the main regression. In Column (2), the dependent variable is *Meet*, defined as an indicator equal to one if the difference between the actual earnings per share and the consensus analyst forecast of earnings per share is in the range of [0, 1 cent]. In Column (3), the dependent variable is *Beat*, defined as an indicator equal to one if a firm's actual earnings exceed the consensus analyst forecast by more than one cent. I use a logistic model to perform the regressions in Columns (2) and (3). As shown in the table, the coefficients on *Downloads*

are negative and highly statistically significant in both columns (t-statistics=-5.02 and -4.76 respectively), suggesting that firms with filings that are downloaded more frequently are less likely to meet or beat consensus analyst forecasts. These results support the notion that when investors are more active in acquiring information about the firm, firms are less likely to manage earnings upward and thereby meet or beat analysts' forecast.

Second, I consider an alternative measure of SEC filing downloads. I separate the total downloads into those with positive or negative news based on whether the three-day abnormal returns [-1, +1] around a filing date of the news is positive or negative. The results are shown in Column (1) of Panel B of Table 5. For both *Pos_Downloads* and *Neg_Downloads*, the coefficients are negative and significant, suggesting that downloading filings related to both positive and negative news can help investors better understand and thus monitor a firm. In addition, I examine whether investors' downloads of filings related to non-financial or financial news have a monitoring effect on earnings management. As shown in Table 5, Panel B, Column (2), the coefficients on *Downloads_NonFin* and *Downloads_Fin* are significant. This evidence implies that non-financial news is also helpful in informing investors about the firm's operational and other activities, so they can better monitor it. One explanation is that non-financial information indicates potential influences on future performance.

Overall, above evidence suggests that my results are robust to alternative measures.

[Insert Table 5 here]

CHAPTER 5 CROSS-SECTIONAL TESTS

While H1 examines the average effect of information acquisition on earnings management, this effect could vary across subsets of firms. To address this variation, I propose several additional hypotheses. Specifically, I focus on the downloads' characteristics, the forces that can monitor a firm's earnings management behavior and managers' incentives to manage earnings.

5.1. Download concentration

The main findings in the baseline regressions are consistent with the monitoring role of information acquisition. Prior literature finds that compared to dispersed shareholders, concentrated shareholders can better monitor management. For example, Shleifer and Vishny (1986) propose that concentrated ownership by outside shareholders enhances firm value by increasing monitoring. Intuitively, more concentrated attention on the corporate reports is likely to result in a more intense monitoring of potential agency problems at the firm, especially if one source of better monitoring comes from a more comprehensive understanding of the firm's activities from downloading of various downloads by the firm (as discussed in my development of H1). Hartzell and Starks (2003) find that institutional concentrated concentration associates with higher pay-for-performance sensitivity and lower CEO pay, which is also consistent with the monitoring role served by shareholders. Similarly, Burns, Kedia, and Lipson (2010) find that concentration of holdings reduces financial misreporting. Above evidence suggests that firms are better monitored when ownership is concentrated among a small group of shareholders. In my setting, when a

firm's SEC filings are downloaded by a small group of investors rather than a dispersed investor base, these downloads are likely to relate to sophisticated investors who have time and resources to study multiple SEC filings. Sophisticated investors are able to monitor the firm. In addition, these investors can obtain a more comprehensive understanding of the firm via the study of multiple filings and thus better monitor it. Therefore, I expect the monitoring effect of downloading SEC filings on earnings management to be stronger when the downloads are concentrated among fewer people. This leads to my second hypothesis, stated in alternate form:

H2: The negative association between information acquisition and earnings management is stronger for firms with more concentrated downloads.

I focus on two download concentration measures, the number of IP addresses (*Fewer IP Number*) and the download concentration ratio (*Download HHI*). *Fewer IP Number* is a dummy variable equal to one if the number of IP addresses is below the sample median and zero otherwise. An IP address is counted if it requests at least one filing over the course of a year. The fewer number of IP addresses, the more concentrated the downloads. *Download HHI* is a dummy variable equal to one if a firm's download concentration ratio is above the sample median and zero otherwise. Download concentration ratio is computed as the sum of the square of an IP's downloads out of the total downloads for each firm. In this case, higher values indicate more concentrated downloads. Hence, I expect the importance of information acquisition to earnings management to be

larger for firms that have fewer IP addresses or more concentrated downloads.

Table 6 presents the results. In Column (1), the coefficient on the interaction of *Downloads*Concentration*, in which concentration is *Fewer IP Number*, is negative and statistically significant with a t-statistic of -4.38, implying that when the filings are less broadly downloaded, the effect of information acquisition on earnings management is stronger. In Column (2), the coefficient on the interaction of *Downloads*Concentration*, in which concentration is measured by *Download HHI*, is negative and statistically significant with a t-statistic of -2.39, suggesting that the effect of information acquisition is stronger when the filing downloads are concentrated among a smaller group of investors. This evidence is consistent with the prediction that an investor can better monitor a firm when there is more concentrated attention.

[Insert Table 6 here]

5.2. Internal and external monitoring

Next, I consider how alternative monitoring mechanisms influence the relation between information acquisition and earnings management. Prior literature presents evidence that firms are less likely to manage earnings if they are under certain monitoring mechanisms, either internal or external. For example, external auditors (Becker et al., 1998; Markelevich and Rosner, 2013), analysts (Yu, 2008), and institutional investors (Bushee, 1998) can play a monitoring role in reducing managers' opportunistic behaviors. Internal control procedures also have significant impact on firm' earnings management activities (Doyle et al., 2007; Ashbaugh-Skaife et al.,

2008; Chan et al., 2008). Since information acquisition reflects investor monitoring, its effect should be weaker if a firm is already subject to other strong monitoring mechanisms. Therefore, I expect the monitoring effect of downloading SEC filings on earnings management to be stronger when monitoring is weak. This leads to my third hypothesis, stated in alternate form:

H3: The negative association between information acquisition and earnings management is stronger for firms with weak monitoring.

To examine the influence of monitoring on the association between information acquisition and earnings management, I first examine how external monitoring influences this association. I expect the effect of information acquisition to be stronger when there is less external monitoring of a firm. I use three proxies, including whether a firm's auditor is a non-big 4 auditor (*NonBig4*), whether the firm is covered by fewer numbers of analysts (*Fewer Analysts*), and whether it has lower institutional ownership (*Lower Institutional Ownership*). *NonBig4* is an indicator equal to one if a firm is audited by a non-big 4 auditor and zero otherwise. *Fewer Analysts* is an indicator equal to one if a firm's analyst following is below the sample median and zero otherwise. *Lower Institutional Ownership* is a dummy variable equal to one if a firm's institutional ownership is below the sample median and zero otherwise. Auditor data comes from Compustat, analyst coverage data is from the I/B/E/S database (I/B/E/S), and institutional ownership data is from the Thomson Reuters Institutional (13f) Holdings. I also examine how an internal control weakness (*IPCW*) influences this association. Firms with an internal control weakness are likely to have a

weak internal monitoring force. I expect to observe that the importance of information acquisition to earnings management is larger for those with an internal control weakness. *IPCW* is a binary variable equal to one if a firm has an internal control weakness in a year and zero otherwise. Internal control weakness data comes from AuditAnalytics.

Table 7 presents the results. In Column (1), the coefficient on the interaction item, *Downloads*Weak_Monitor*, in which monitoring is proxied by an indicator of whether a firm is audited by a non-big 4 auditor, is negative and statistically significant with a t-statistic of -5.27. *Weak_Monitor* is proxied by non-big 4 auditors, fewer analysts, lower institutional ownership and an internal control weakness in Columns (2)-(4), respectively. The coefficients on the interaction item, *Downloads*Weak_Monitor*, are statistically significant in all columns. This suggests that when there are weaker external monitoring forces, the information acquisition effect is more pronounced, consistent with the prediction in H3.

[Insert Table 7 here]

5.3. Incentives

Next, I examine the impact of incentives to manage earnings on the relation between information acquisition and earnings management. Prior literature provides evidence of various incentives to manage earnings. The need to raise financing is one such incentive. Cohen and Zarowin (2010) examine the relation between earnings management and financing needs and present evidence of a positive relation. Managers' compensation and career concerns also impact earnings management. Bergstresser and Philippon

(2006) show that a firm has higher earnings management when the value of stock and option holdings account for a significant part of CEOs' potential total compensation. Ali and Zhang (2015) focus on CEO tenure and find that CEOs tend to overstate firm earnings more in the early years of their tenure to favorably influence the market's perception of their ability. If information acquisition can help investors to better monitor a firm's earnings management activities, firms with incentives to manage earnings are more likely to be constrained by monitoring. Therefore, I expect the effect of information acquisition on earnings management to be stronger for firms with incentives to manage earnings. This leads to my final hypothesis, stated in alternate form:

H4: The negative association between information acquisition and earnings management is stronger for firms with a greater incentive to manage earnings.

I use firms' choice of a seasonal equity offering (*SEO*) in the following year to proxy for the financing incentive and CEO tenure (*CEO_Tenure*) to proxy for the CEO's incentives to manage earnings. I obtain data about seasonal equity offerings from Thomson SDC database and CEO tenure data from Execucomp. Following prior literature (e.g., Bergstresser and Philippon 2006; Cohen and Zarowin, 2010; Ali and Zhang 2015), *SEO* is a dummy variable equal to one if a firm has a seasoned equity offering in the following year. *CEO Tenure* is an indicator equal to one if a firm's CEO is in the first three years of her tenure and zero otherwise. I expect the effect of information acquisition to be stronger for firms with an

SEO and those with CEOs who are took up that position within the past three years.

As shown in Column (1) of Table 8, the coefficient on the interaction term of *Downloads*Incentives* is negative and significant with a t-statistic of -1.75. This result demonstrates that the monitoring effect of investors' information acquisition on earnings management is more pronounced for firms with a stronger incentive to manipulate earnings. In Column (2), where incentive is proxied by CEO tenure, the coefficient is negative and significant with a t-statistic of -1.83, suggesting that the effect of information acquisition is stronger for CEOs in the first three years of their tenure. Overall, the results are consistent with my final hypothesis.

[Insert Table 8 here]

CHAPTER 6 CONCLUSIONS

In this study, I examine the association between information acquisition from SEC filings and earnings management. A tension underlies this association. On one hand, greater information acquisition enables investors to better ascertain a firm's operational and financial conditions and thereby to monitor its earnings management activities. On the other hand, extensive information acquisition, especially of interim and current reports, can induce manager myopia because of investors' focus on short-term performance. I investigate this issue using an aggregate measure of the number of downloads of 10-K, 10-Q and 8-K filings because the credibility of SEC filings is much higher than that of other information sources and because these three types of filings are both the most frequently downloaded and are most likely to be related to firm performance and thus earnings management activities.

I find that downloads are negatively related to earnings management, consistent with the monitoring role of information acquisition. To address endogeneity issues, I add firm fixed effects and additional control variables and employ a difference-in-difference design, an instrumental variables technique and alternative specifications. The results show that my findings are robust. While each method above has its limitations in fully ruling out endogeneity concerns, the results of these tests allow me to more confidently draw a causal inference between information acquisition and earnings management.

In a cross-sectional analysis, I find that the impact of information acquisition on earnings management is stronger when the downloads of

SEC filings are concentrated among fewer IP addresses. Moreover, the importance of information acquisition is larger in firms subject to less external and internal monitoring. Specifically, investors' information acquisition plays a larger role when a firm is audited by a non-big 4 auditor, is covered by fewer analysts, has lower institutional ownership, or has an internal control weakness. Finally, I also find the influence of information acquisition to be more pronounced for firms that have a seasoned equity offering in the coming year or those that have a new CEO.

APPENDIX VARIABLE DEFINITION

Main variables	
<i>Downloads</i>	The natural logarithm of one plus the number of downloads of the previous year's 10-K and the current year's 10-Q and 8-K filings. Downloads of the previous year's 10-K are calculated from their filing dates to three months after the fiscal year end. Downloads of this year's 10-Qs are calculated from each file's filing date to three months after the fiscal year end. Downloads of this year's 8-Ks are calculated from each 8-K's filing date to three months after the fiscal year end.
<i>Pos_Downloads</i>	The natural logarithm of one plus the number of downloads of the previous year's 10-K and the current year's 10-Q and 8-K filings if the stock market has a positive response to the new filing in the three day window [-1, +1] around the filing date.
<i>Neg_Downloads</i>	The natural logarithm of one plus the number of downloads of the previous year's 10-K and the current year's 10-Q and 8-K filings if the stock market has a negative response to the new filing in the three day window [-1, +1] around the filing date.
<i>Downloads_financial</i>	The natural logarithm of one plus the number of downloads of 8-K filings containing item 2.02 as well as 10-K and 10-Q filings.
<i>Downloads_Nonfinancial</i>	The natural logarithm of one plus the number of downloads of 8-K filings, excluding those that contain item 2.02.
Other variables(in alphabetical order)	
<i>Age</i>	The natural logarithm of firm age.
<i>Analyst</i>	The number of analysts following a firm in a fiscal year.
<i>Beat</i>	One if a firm's actual earnings exceed the consensus analyst forecast by more than one cent.
<i>BM</i>	Book value of equity divided by the market value of equity.
<i>CEO_Tenure</i>	A dummy variable equal to one if the CEO has served a firm for three years or less, zero otherwise.

<i>CEO_Turnover</i>	A dummy variable equal to one if a firm has a new CEO in a year and zero otherwise.
<i>CFO</i>	Cash flow from operations in year t scaled by total assets at the beginning of year t.
<i>Discretionary Accruals</i>	Residuals from the Modified Jones model (Dechow et al., 1995) or following the performance-matched approach in Kothari et al. (2005).
<i>Download HHI</i>	A dummy variable equal to one if a firm's download concentration ratio is above the sample median and zero otherwise. Download-number-based concentration ratio for all IP addresses, calculated as: $\text{Download HHI} = \sum_{j=1}^J \left(\frac{\text{Downloads}_{jit}}{\text{Downloads}_{it}} \right)^2$ where Downloads_{it} represents firm i 's total downloads in year t, and Downloads_{jit} represents firm i 's downloads from IP address j in year t.
<i>Fewer Analyst</i>	A dummy variable equal to one if a firm's analyst following is below the sample median and zero otherwise.
<i>Fewer Institutional Ownership</i>	A dummy variable equal to one if a firm's institutional ownership is below the sample median and zero otherwise.
<i>Filings</i>	The number of Form 10-K, 10-Q and 8-K filed by each firm within each fiscal year.
<i>Growth</i>	Sales growth defined as the difference between sales in year t minus sales in year t-1 divided by sales in year t-1.
<i>Lagged_Accruals</i>	Total accruals in year t-1 scaled by total assets at the beginning of the year.
<i>Lagged_NOA</i>	Net operating assets at the beginning of year t, defined as shareholders' equity less cash and marketable securities, plus total debt, deflated by sales.
<i>Leverage</i>	The sum of the book value of long-term debt plus the book value of current liabilities, divided by the book value of assets.
<i>IO</i>	The percentage of stocks held by institutional investors.
<i>IPCW</i>	An indicator variable equal to one if a firm has an internal control weakness in a year and zero otherwise.
<i>Fewer IP Number</i>	A dummy variable equal to one if the number of IP addresses is below the sample median and zero otherwise.

	An IP address is counted if it downloaded any one of a firm's previous year's 10-K or the current year's 10-Q or 8-K filings.
<i>Media</i>	The natural logarithm of one plus the number of all news articles released for each firm within each fiscal year.
<i>Meet</i>	One if the difference between the actual earnings per share and the consensus analyst forecast of earnings per share is in the range of [0, 1 cent].
<i>NonBig4</i>	An indicator variable equal to one if a firm is audited by a non-big4 auditor and zero otherwise.
<i>Questions</i>	The natural logarithm of one plus the number of questions that are asked in the conference calls in a year.
<i>Restate_Uncover</i>	A dummy variable equal to one if a restatement is filed in a year and zero otherwise.
<i>ROA</i>	Return on assets, calculated as income before extraordinary items divided by the beginning-of-year book value of total assets.
<i>SEO</i>	An indicator variable set equal to one if the firm announces a seasoned equity offering in the next year, 0 otherwise, based on Thomson SDC data.
<i>Size</i>	The natural logarithm of the market value of equity.

REFERENCES

- Ali, A., & Zhang, W. (2015). CEO tenure and earnings management. *Journal of Accounting and Economics*, 59(1), 60-79.
- Al-Qeisi, K., Dennis, C., Alamanos, E., & Jayawardhena, C. (2014). Website design quality and usage behavior: Unified theory of acceptance and use of technology. *Journal of Business Research*, 67(11), 2282-2290.
- Ashbaugh-Skaife, H., Collins, D. W., Kinney Jr, W. R., & LaFond, R. (2008). The effect of SOX internal control deficiencies and their remediation on accrual quality. *The Accounting Review*, 83(1), 217-250.
- Barton, J., & Simko, P. J. (2002). The balance sheet as an earnings management constraint. *The Accounting Review*, 77(s-1), 1-27.
- Bartov, E., Faurel, L., & Mohanram, P. S. (2017). Can Twitter help predict firm-level earnings and stock returns? *The Accounting Review*, 93(3), 25-57.
- Becker, C. L., DeFond, M. L., Jiambalvo, J., & Subramanyam, K. R. (1998). The effect of audit quality on earnings management. *Contemporary Accounting Research*, 15(1), 1-24.
- Ben-Rephael, A., Da, Z., & Israelsen, R. D. (2017). It depends on where you search: Institutional investor attention and underreaction to news. *The Review of Financial Studies*, 30(9), 3009-3047.
- Bergstresser, D., & Philippon, T. (2006). CEO incentives and earnings management. *Journal of Financial Economics*, 80(3), 511-529.
- Bhattacharya, N., Cho, Y. J., & Kim, J. B. (2018). Leveling the playing field between large and small institutions: evidence from the SEC's XBRL mandate. *The Accounting Review*, 93(5), 51-71.
- Bird, A., & Karolyi, S. A. (2016). Do institutional investors demand public disclosure? *The Review of Financial Studies*, 29(12), 3245-3277.
- Blankespoor, E., Miller, G. S., & White, H. D. (2013). The role of dissemination in market liquidity: Evidence from firms' use of Twitter. *The Accounting Review*, 89(1), 79-112.
- Botosan, C. A. (1997). Disclosure level and the cost of equity capital. *The Accounting Review*, 72(3), 323-349.
- Bozanic, Z., Hoopes, J. L., Thornock, J. R., & Williams, B. M. (2017). IRS attention. *Journal of Accounting Research*, 55(1), 79-114.
- Burns, N., Kedia, S., & Lipson, M. (2010). Institutional ownership and monitoring: Evidence from financial misreporting. *Journal of Corporate Finance*, 16(4), 443-455.

- Bushee, B. J. (1998). The influence of institutional investors on myopic R&D investment behavior. *The Accounting Review*, 73(3), 305-333.
- Bushman, R. M., & Smith, A. J. (2001). Financial accounting information and corporate governance. *Journal of Accounting and Economics*, 32(1-3), 237-333.
- Callen, J. L., Livnat, J., & Segal, D. (2006). The information content of SEC filings and information environment: A variance decomposition analysis. *The Accounting Review*, 81(5), 1017-1043.
- Carter, M. E., & Soo, B. S. (1999). The relevance of Form 8-K reports. *Journal of Accounting Research*, 37(1), 119-132.
- Chakrabarty, B., & Moulton, P. C. (2012). Earnings announcements and attention constraints: The role of market design. *Journal of Accounting and Economics*, 53(3), 612-634.
- Chan, K. C., Farrell, B., & Lee, P. (2008). Earnings management of firms reporting material internal control weaknesses under Section 404 of the Sarbanes-Oxley Act. *Auditing: A Journal of Practice & Theory*, 27(2), 161-179.
- Chen, T., Harford, J., & Lin, C. (2015). Do analysts matter for governance? Evidence from natural experiments. *Journal of Financial Economics*, 115(2), 383-410.
- Cheng, M., & Subramanyam, K. R. (2008). Analyst following and credit ratings. *Contemporary Accounting Research*, 25(4), 1007-1044.
- Cohen, D. A., & Zarowin, P. (2010). Accrual-based and real earnings management activities around seasoned equity offerings. *Journal of Accounting and Economics*, 50(1), 2-19.
- Dechow, P. M., Sloan, R. G., & Sweeney, A. P. (1995). Detecting earnings management. *The Accounting Review*, 70(2), 193-225.
- Dou, Y., Hope, O. K., Thomas, W. B., & Zou, Y. (2014). Blockholder exit threats and financial reporting quality. *Contemporary Accounting Research*, 35(2), 1004-1028.
- Doyle, J. T., Ge, W., & McVay, S. (2007). Accruals quality and internal control over financial reporting. *The Accounting Review*, 82(5), 1141-1170.
- Doyle, J. T., & Magilke, M. J. (2013). Decision usefulness and accelerated filing deadlines. *Journal of Accounting Research*, 51(3), 549-581.
- Drake, M. S., Lamoreaux, P. T., Quinn, P. J., & Thornock, J. R. (2019). Auditor benchmarking of client disclosures. *Review of Accounting Studies*, 24(2), 393-425.
- Drake, M. S., Roulstone, D. T., & Thornock, J. R. (2012). Investor information demand: Evidence from Google searches around

- earnings announcements. *Journal of Accounting Research*, 50(4), 1001-1040.
- Drake, M. S., Roulstone, D. T., & Thornock, J. R. (2015). The determinants and consequences of information acquisition via EDGAR. *Contemporary Accounting Research*, 32(3), 1128-1161.
- Drake, M. S., Roulstone, D. T., & Thornock, J. R. (2016). The usefulness of historical accounting reports. *Journal of Accounting and Economics*, 61(2-3), 448-464.
- Ellul, A., & Panayides, M. (2018). Do financial analysts restrain insiders' informational advantage? *Journal of Financial and Quantitative Analysis*, 53(1), 203-241.
- Ernstberger, J., Link, B., Stich, M., & Vogler, O. (2017). The real effects of mandatory quarterly reporting. *The Accounting Review*, 92(5), 33-60.
- Ettredge, M., Richardson, V. J., & Scholz, S. (2002). Dissemination of information for investors at corporate Web sites. *Journal of Accounting and Public Policy*, 21(4-5), 357-369.
- Gabaix, X., Laibson, D., Moloche, G., & Weinberg, S. (2006). Costly information acquisition: Experimental analysis of a boundedly rational model. *American Economic Review*, 96(4), 1043-1068.
- Gargano, A., Rossi, A. G., & Wermers, R. (2017). The freedom of information act and the race toward information acquisition. *The Review of Financial Studies*, 30(6), 2179-2228.
- Griffin, P. A. (2003). Got information? Investor response to Form 10-K and Form 10-Q EDGAR filings. *Review of Accounting Studies*, 8(4), 433-460.
- He, J. J., & Tian, X. (2013). The dark side of analyst coverage: The case of innovation. *Journal of Financial Economics*, 109(3), 856-878.
- Hartzell, J. C., & Starks, L. T. (2003). Institutional investors and executive compensation. *The Journal of Finance*, 58(6), 2351-2374.
- Huang, S. X., Pereira, R., & Wang, C. (2017). Analyst coverage and the likelihood of meeting or beating analyst earnings forecasts. *Contemporary Accounting Research*, 34(2), 871-899.
- Irani, R. M., & Oesch, D. (2016). Analyst coverage and real earnings management: Quasi-experimental evidence. *Journal of Financial and Quantitative Analysis*, 51(2), 589-627.
- Kelton, A. S., & Yang, Y. W. (2008). The impact of corporate governance on Internet financial reporting. *Journal of Accounting and Public Policy*, 27(1), 62-87.

- Kempf, E., Manconi, A., & Spalt, O. (2016). Distracted shareholders and corporate actions. *The Review of Financial Studies*, 30(5), 1660-1695.
- Kim, J. B., Li, B., & Liu, Z. (2019). Information - Processing Costs and Breadth of Ownership. *Contemporary Accounting Research*, Forthcoming.
- Kim, J. B., Lu, L. Y., & Yu, Y. (2018). Analyst coverage and expected crash risk: Evidence from exogenous changes in analyst coverage. *The Accounting Review*. Forthcoming.
- Kluge, P. N., Königsfeld, J. A., Fassnacht, M., & Mitschke, F. (2013). Luxury web atmospherics: An examination of homepage design. *International Journal of Retail & Distribution Management*, 41(11/12), 901-916.
- Koh, P. S. (2007). Institutional investor type, earnings management and benchmark beaters. *Journal of Accounting and Public Policy*, 26(3), 267-299.
- Kothari, S. P., Leone, A. J., & Wasley, C. E. (2005). Performance matched discretionary accrual measures. *Journal of Accounting and Economics*, 39(1), 163-197.
- Kraft, A. G., Vashishtha, R., & Venkatachalam, M. (2018). Frequent financial reporting and managerial myopia. *The Accounting Review*, 93(2), 249-275.
- Lang, M. H., & Lundholm, R. J. (2000). Voluntary disclosure and equity offerings: Reducing information asymmetry or hyping the stock? *Contemporary Accounting Research*, 17(4), 623-662.
- Lawrence, A., Ryans, J. P., & Sun, E. Y. (2017). Investor demand for sell-side research. *The Accounting Review*, 92(2), 123-149.
- Lee, C. M., Ma, P., & Wang, C. C. (2015). Search-based peer firms: Aggregating investor perceptions through internet co-searches. *Journal of Financial Economics*, 116(2), 410-431.
- Lerman, A., & Livnat, J. (2010). The new Form 8-K disclosures. *Review of Accounting Studies*, 15(4), 752-778.
- Leuz, C., & Verrecchia, R. E. (2000). The economic consequences of increased disclosure. *Journal of Accounting Research*, 38, 91-124.
- Li, E., Lind, G., Ramesh, K., & Shen, M. (2017). The Federal Reserve's Reliance on SEC Filings. Working Paper.
- Matsumoto, D. A. (2002). Management's incentives to avoid negative earnings surprises. *The Accounting Review*, 77(3), 483-514.
- Markelevich, A., & Rosner, R. L. (2013). Auditor fees and fraud firms. *Contemporary Accounting Research*, 30(4), 1590-1625.

- Marston, C., & Polei, A. (2004). Corporate reporting on the Internet by German companies. *International Journal of Accounting Information Systems*, 5(3), 285-311.
- McMullin, J. L., Miller, B. P., & Twedt, B. J. (2019). Increased mandated disclosure frequency and price formation: Evidence from the 8-k expansion regulation. *Review of Accounting Studies*, 24(1), 1–33.
- Pandir, M., & Knight, J. (2006). Homepage aesthetics: The search for preference factors and the challenges of subjectivity. *Interacting with Computers*, 18(6), 1351-1370.
- Rosen, D. E., & Purinton, E. (2004). Website design: Viewing the web as a cognitive landscape. *Journal of Business Research*, 57(7), 787-794.
- Securities and Exchange Commission. (2009). Interactive data to improve financial reporting. Retrieved from: <https://www.sec.gov/rules/final/2009/33-9002.pdf>.
- Shevlin, T., & Thornock, J. (2015). Market (in) attention and the strategic scheduling and timing of earnings announcements. *Journal of Accounting and Economics*, 60(1), 36-55.
- Shleifer, A., & Vishny, R. W. (1986). Large shareholders and corporate control. *Journal of Political Economy*, 94(3), 461-488.
- Stice-Lawrence, L. (2017). Too small to regulate: International evidence on regulatory monitoring. Doctoral Dissertation, the University of North Carolina at Chapel Hill.
- Tang, V. W. (2018). Wisdom of crowds: Cross-sectional variation in the informativeness of third-party-generated product information on Twitter. *Journal of Accounting Research*, 56(3), 989-1034.
- Tarafdar, M., & Zhang, J. (2008). Determinants of reach and loyalty—A study of website performance and implications for website design. *Journal of Computer Information Systems*, 48(2), 16-24.
- Thayer, J. (2011). Determinants of investors' information acquisition: Credibility and confirmation. *The Accounting Review*, 86(1), 1-22.
- Verrecchia, R. E. (1982). Information acquisition in a noisy rational expectations economy. *Econometrica*, 50(6), 1415-1430.
- Xu, S. X., & Zhang, X. Q. (2013). Impact of Wikipedia on market information environment: Evidence on management disclosure and investor reaction. *MIS Quarterly*, 37(4), 1043-1068.
- Yu, F. F. (2008). Analyst coverage and earnings management. *Journal of Financial Economics*, 88(2), 245-271.
- Zhao, X. (2017). Does information intensity matter for stock returns? Evidence from Form 8-K filings. *Management Science*, 63(5), 1382-1404.

Zuo, L. (2016). The informational feedback effect of stock prices on management forecasts. *Journal of Accounting and Economics*, 61(2-3), 391-413.

Table 1 Summary Statistics

This table reports the descriptive statistics for all variables in the main regression. All continuous variables are winsorized at 1% and 99%. Variable definitions can be found in the appendix.

	N	Mean	Std. dev.	P25	P50	P75
<i>Discretionary Accruals</i>	40,117	-0.027	0.344	-0.070	-0.002	0.069
<i>Downloads</i>	40,117	7.344	1.023	6.613	7.330	8.039
<i>Filings</i>	40,117	16.090	6.492	12.000	15.000	19.000
<i>Size</i>	40,117	5.787	2.288	4.196	5.877	7.361
<i>BM</i>	40,117	0.396	1.355	0.205	0.411	0.714
<i>Leverage</i>	40,117	0.667	1.334	0.305	0.492	0.682
<i>IO</i>	40,117	0.401	0.360	0.000	0.378	0.759
<i>Analyst</i>	40,117	7.790	8.665	1.000	5.000	12.000
<i>ROA</i>	40,117	-0.144	0.965	-0.086	0.028	0.081
<i>CFO</i>	40,117	-0.002	0.405	-0.006	0.077	0.141
<i>Lagged_Accruals</i>	40,117	-0.137	0.551	-0.116	-0.058	-0.018
<i>Lagged_NOA</i>	40,117	0.620	3.648	0.231	0.480	0.888
<i>Age</i>	40,117	2.745	0.673	2.303	2.773	3.296
<i>Growth</i>	40,117	0.179	0.808	-0.040	0.070	0.207

Table 2 Industry and Year Distribution of Downloads

This table presents the mean value of downloads by industry and year.

Panel A: Industry distribution

	Obs.	<i>Downloads</i>
Agriculture, Forestry, And Fishing (SIC code 0100-0999)	23	7.683
Mining (SIC code 1000-1499)	2,572	7.542
Construction (SIC code 1500-1799)	412	7.525
Manufacturing (SIC code 2000-3999)	20,785	7.293
Transportation, Communications, Electric (SIC code 4000-4899)	2,598	7.582
Wholesale Trade (SIC code 5000-5199)	1,588	7.250
Retail Trade (SIC code 5200-5999)	2,929	7.663
Services (SIC code 7000-8999)	8,847	7.261
Public Administration and non-classifiable (SIC code 9100-9999)	363	7.014
Total	40,117	7.346

Panel B: Year distribution

Year	Obs.	<i>Downloads</i>
2003	2,739	6.667
2004	3,383	6.874
2005	3,224	6.146
2006	3,143	6.499
2007	3,066	6.894
2008	3,089	7.240
2009	3,247	7.562
2010	3,136	7.735
2011	3,043	7.865
2012	3,000	7.780
2013	3,021	8.095
2014	3,022	8.074
2015	3,004	8.156
Total	40,117	7.346

Table 3 Main Regressions

This table reports the regression results from tests of the effect of information acquisition on earnings management. All variables are defined in the appendix. Industry and year fixed effects are included. The t-values in the parentheses are based on heteroscedasticity robust standard errors clustered by firm. ***, **, and * denote significance at the 1%, 5% and 10% levels, respectively.

	<i>Discretionary Accruals</i>		
	(1)	(2)	(3)
<i>Downloads</i>	-0.021*** (-6.38)		-0.022*** (-5.65)
<i>Filings</i>		-0.001*** (-3.24)	0.000 (0.05)
<i>Size</i>	0.009*** (5.82)	0.006*** (3.92)	0.009*** (5.79)
<i>BM</i>	0.005*** (2.90)	0.005*** (3.27)	0.005*** (2.90)
<i>Leverage</i>	-0.029*** (-4.66)	-0.029*** (-4.66)	-0.029*** (-4.67)
<i>IO</i>	0.005 (1.27)	0.006* (1.67)	0.005 (1.27)
<i>Analyst</i>	-0.001*** (-2.58)	-0.001*** (-5.21)	-0.001** (-2.57)
<i>ROA</i>	0.323*** (15.46)	0.323*** (15.49)	0.323*** (15.46)
<i>CFO</i>	-0.339*** (-12.79)	-0.338*** (-12.80)	-0.339*** (-12.80)
<i>Lagged_Accruals</i>	0.013 (0.77)	0.013 (0.78)	0.013 (0.77)
<i>Lagged_NOA</i>	-0.001 (-0.68)	-0.001 (-0.73)	-0.001 (-0.68)
<i>Age</i>	0.032*** (10.27)	0.031*** (9.93)	0.032*** (10.32)
<i>Growth</i>	0.005 (0.87)	0.005 (0.91)	0.005 (0.87)
<i>Constant</i>	0.052** (2.40)	-0.056*** (-4.95)	0.052** (2.28)
Industry Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Observations	40,117	40,117	40,117
Adj. R-squared	0.5233	0.5225	0.5233

Table 4 Endogeneity Issues

The table presents the regression results from tests using several methodologies used to address endogeneity issues. Panel A reports the results for the regressions that include firm and year fixed effects as well as additional control variables. Panel B presents the results for regressions that use the difference-in-difference method. Panel C shows the results from tests of the instrumental variable regressions. The t-values in the parentheses are based on heteroscedasticity robust standard errors clustered by firm. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Panel A: Firm fixed effects and additional controls

	<i>Discretionary Accruals</i>				
	(1)	(2)	(3)	(4)	(5)
<i>Downloads</i>	-0.023*** (-4.67)	-0.013*** (-2.65)	-0.013** (-2.46)	-0.008** (-2.39)	-0.021*** (-5.45)
<i>Media</i>		0.001 (1.05)			
<i>Questions</i>			0.002 (0.63)		
<i>CEO_Turnover</i>				-0.003 (-1.11)	
<i>Restate_Uncover</i>					-0.017*** (-3.88)
<i>Filings</i>	0.000 (0.95)	0.001** (1.98)	0.000 (0.29)	0.000 (1.32)	0.000 (0.26)
<i>Size</i>	0.012*** (4.20)	0.016*** (5.12)	0.001 (0.26)	-0.001 (-0.77)	0.009*** (5.67)
<i>BM</i>	0.003 (1.51)	0.001 (0.25)	0.007*** (2.90)	0.002* (1.65)	0.005*** (2.93)
<i>Leverage</i>	-0.030*** (-3.57)	-0.029 (-1.02)	0.030*** (3.40)	0.029*** (3.56)	-0.029*** (-4.66)
<i>IO</i>	-0.003 (-0.68)	-0.004 (-1.14)	-0.009 (-1.64)	-0.003 (-0.75)	0.005 (1.29)
<i>Analyst</i>	-0.001** (-2.06)	-0.001*** (-4.17)	-0.000* (-1.82)	-0.000** (-2.07)	-0.001*** (-2.73)
<i>ROA</i>	0.338*** (13.05)	0.592*** (6.69)	0.863*** (30.37)	0.913*** (34.62)	0.323*** (15.47)
<i>CFO</i>	-0.490*** (-14.61)	-0.719*** (-13.99)	-0.804*** (-29.15)	-0.822*** (-30.70)	-0.339*** (-12.80)
<i>Lagged_Accruals</i>	-0.049** (-2.42)	-0.070* (-1.70)	-0.060*** (-4.07)	-0.053*** (-4.47)	0.013 (0.76)
<i>Lagged_NOA</i>	-0.001 (-0.83)	-0.001 (-0.73)	0.001 (1.54)	0.001 (1.13)	-0.001 (-0.69)
<i>Age</i>	0.009 (0.71)	0.020** (2.33)	0.012*** (3.80)	0.010*** (4.01)	0.032*** (10.29)
<i>Growth</i>	0.018*** (2.63)	0.013*** (3.69)	0.009* (1.74)	0.008 (1.42)	0.005 (0.86)
<i>Constant</i>	0.087* (1.83)	-0.012 (-0.36)	0.097*** (3.68)	0.079*** (4.49)	0.049** (2.13)
Firm Fixed Effects	Yes				
Industry Fixed Effects		Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	40,117	29,920	10,829	16,439	40,117
Adj. R-squared	0.6875	0.6235	0.4488	0.3955	0.5235

Panel B1: Validity test

	Control	Treat
Before	304.45	112.76
Post	412.86	226.34

Panel B2: Difference-in-differences specification

	<i>Discretionary Accruals</i>
<i>Treat*Post</i>	-0.022** (-2.40)
<i>Post</i>	0.007 (1.02)
<i>Treat</i>	0.011 (1.63)
<i>Filings</i>	0.000 (0.25)
<i>Size</i>	0.001 (0.50)
<i>BM</i>	0.002 (0.85)
<i>Leverage</i>	0.002 (0.23)
<i>IO</i>	0.012** (2.01)
<i>Analyst</i>	-0.001 (-1.45)
<i>ROA</i>	0.508*** (18.23)
<i>CFO</i>	-0.547*** (-11.89)
<i>Lagged_Accruals</i>	0.025 (1.14)
<i>Lagged_NOA</i>	-0.007*** (-3.75)
<i>Age</i>	0.033*** (6.58)
<i>Growth</i>	0.023*** (3.45)
<i>Constant</i>	-0.066*** (-3.46)
Industry Fixed Effects	Yes
Year Fixed Effects	Yes
Observations	8,200
Adj. R-squared	0.6308

Panel C: Instrumental variable analysis

	<i>Downloads</i>	<i>Discretionary Accruals</i>
	(1)	(2)
<i>IV</i>	-0.199*** (-3.73)	
<i>Instrumented Downloads</i>		0.077** (2.02)
<i>Filings</i>	0.019*** (6.11)	-0.002* (-1.76)
<i>Size</i>	0.432*** (14.66)	-0.046*** (-2.78)
<i>BM</i>	0.500*** (5.17)	-0.025 (-0.75)
<i>Leverage</i>	0.288** (1.98)	0.003 (0.11)
<i>IO</i>	-0.061 (-0.76)	0.007 (0.45)
<i>Analyst</i>	0.014*** (4.72)	-0.001 (-1.52)
<i>ROA</i>	0.274 (0.77)	-0.753*** (-9.58)
<i>CFO</i>	-0.331 (-1.38)	-0.108 (-1.35)
<i>Lagged_Accruals</i>	-0.984*** (-2.85)	1.041*** (11.97)
<i>Lagged_NOA</i>	-0.024 (-0.52)	0.003 (0.35)
<i>Age</i>	0.063 (1.65)	0.010 (1.57)
<i>Growth</i>	0.012 (0.10)	0.017 (0.62)
<i>Constant</i>	4.024*** (14.96)	-0.183 (-1.12)
Industry Fixed Effects	Yes	Yes
Year Fixed Effects	Yes	Yes
Observations	954	954
Adj. R-squared	0.7386	0.2984

Table 5 Alternative Measures

The table presents the regression results for various robustness checks. In Panel A, alternative earnings management measures are used, including discretionary accruals following the performance-matched approach and two model-free measures that indicate whether a firm's earnings meet or beat analysts' forecasts. In Panel B, downloads of filings are divided into downloads of positive or negative news based on whether a filing has a positive or negative three-day abnormal returns [-1, +1] around a filing date. The t-values in the parentheses are based on heteroscedasticity robust standard errors clustered by firm. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Panel A: Alternative earnings management measures

	<i>Performance-Matched Discretionary Accruals</i>	<i>Meet</i>	<i>Beat</i>
	(1)	(2)	(3)
<i>Downloads</i>	-0.027*** (-3.14)	-0.139*** (-5.02)	-0.096*** (-4.76)
<i>Filings</i>	0.000 (0.37)	-0.007*** (-3.30)	-0.000 (-0.11)
<i>Size</i>	-0.000 (-0.04)	0.073*** (6.36)	0.126*** (13.94)
<i>BM</i>	0.001 (0.45)	-0.048*** (-4.20)	0.004 (0.34)
<i>Leverage</i>	-0.035*** (-2.60)	-0.120** (-2.24)	0.021 (0.88)
<i>IO</i>	-0.007 (-0.91)	-0.058* (-1.77)	0.154*** (6.02)
<i>Analyst</i>	-0.001** (-1.99)	0.009*** (4.46)	0.003* (1.78)
<i>ROA</i>	0.361*** (10.65)	0.095 (1.35)	0.186*** (2.71)
<i>CFO</i>	-0.633*** (-12.71)	-0.168* (-1.75)	0.095 (1.23)
<i>Lagged_Accruals</i>	-0.148*** (-4.13)	-0.143** (-2.38)	-0.169*** (-3.18)
<i>Lagged_NOA</i>	-0.006** (-1.97)	0.004 (0.99)	-0.011*** (-3.36)
<i>Age</i>	0.036* (1.75)	-0.028 (-1.47)	-0.048*** (-3.27)
<i>Growth</i>	0.026*** (3.60)	-0.012 (-0.75)	0.033*** (2.92)
<i>Constant</i>	0.152* (1.88)	-0.509* (-1.88)	-0.824*** (-4.40)
Industry Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Observations	40,117	31,092	31,108
Adj. R-squared	0.3555	0.0331	0.0372

Panel B: Types of downloads

	<i>Discretionary Accruals</i>	
	(1)	(2)
<i>Downloads_Pos</i>	-0.005*** (-3.62)	
<i>Downloads_Neg</i>	-0.007*** (-4.32)	
<i>Downloads_NonFin</i>		-0.006** (-2.10)
<i>Downloads_Fin</i>		-0.006*** (-2.67)
<i>Filings</i>	0.000 (0.92)	0.001** (2.14)
<i>Size</i>	0.009*** (3.15)	0.011*** (4.85)
<i>BM</i>	0.005*** (3.47)	-0.001 (-0.68)
<i>Leverage</i>	0.012 (1.39)	-0.043*** (-2.63)
<i>IO</i>	0.001 (0.30)	-0.005 (-1.31)
<i>Analyst</i>	-0.001*** (-4.89)	-0.001*** (-3.77)
<i>ROA</i>	0.704*** (8.68)	0.518*** (9.73)
<i>CFO</i>	-0.723*** (-14.37)	-0.651*** (-16.35)
<i>Lagged_Accruals</i>	-0.014 (-0.38)	-0.056** (-2.24)
<i>Lagged_NOA</i>	-0.001 (-0.51)	-0.002 (-1.31)
<i>Age</i>	0.011*** (5.46)	0.010 (1.09)
<i>Growth</i>	0.006* (1.86)	0.017*** (4.83)
<i>Constant</i>	0.034** (2.51)	0.013 (0.46)
Industry Fixed Effects	Yes	Yes
Year Fixed Effects	Yes	Yes
Observations	28,862	34,660
Adj. R-squared	0.4166	0.6067

Table 6 Download Concentration

The table presents the regression results of cross-sectional tests that consider the download concentration of IP addresses. The t-values in the parentheses are based on heteroscedasticity robust standard errors clustered by firm. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

	<i>Discretionary Accruals</i>	
	(1) <i>Fewer IP Number</i>	(2) <i>Download HHI</i>
<i>Downloads*Concentration</i>	-0.282** (-2.39)	-0.015*** (-4.38)
<i>Downloads</i>	-0.020*** (-4.93)	-0.016*** (-3.94)
<i>Concentration</i>	1.777** (2.33)	0.101*** (4.32)
<i>Filings</i>	0.000 (0.30)	0.000 (0.37)
<i>Size</i>	0.009*** (5.78)	0.009*** (5.81)
<i>BM</i>	0.005*** (2.94)	0.005*** (2.95)
<i>Leverage</i>	-0.029*** (-4.67)	-0.029*** (-4.68)
<i>IO</i>	0.006 (1.47)	0.006 (1.56)
<i>Analyst</i>	-0.001*** (-3.00)	-0.001*** (-3.36)
<i>ROA</i>	0.323*** (15.46)	0.323*** (15.47)
<i>CFO</i>	-0.339*** (-12.80)	-0.339*** (-12.80)
<i>Lagged_Accruals</i>	0.013 (0.76)	0.013 (0.76)
<i>Lagged_NOA</i>	-0.001 (-0.69)	-0.001 (-0.68)
<i>Age</i>	0.032*** (9.96)	0.031*** (9.92)
<i>Growth</i>	0.005 (0.88)	0.005 (0.90)
<i>Constant</i>	0.040 (1.64)	0.011 (0.45)
Industry Fixed Effects	Yes	Yes
Year Fixed Effects	Yes	Yes
Observations	40,117	40,117
Adj. R-squared	0.5234	0.5235

Table 7 Monitoring

This table reports results from a cross-sectional test of whether the effect of downloads on earnings management is more pronounced for firms with weak monitoring. The t-values in the parentheses are based on heteroscedasticity robust standard errors clustered by firm. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

	<i>Discretionary Accruals</i>			
	(1)	(2)	(3)	(4)
	<i>NonBig4</i>	<i>Fewer Analysts</i>	<i>Lower Institutional Ownership</i>	<i>IPCW</i>
<i>Downloads* Weak_Monitor</i>	-0.021*** (-5.27)	-0.009*** (-3.14)	-0.008*** (-3.22)	-0.006* (-1.90)
<i>Downloads</i>	-0.036*** (-6.44)	-0.020*** (-5.97)	-0.018*** (-5.02)	-0.011*** (-4.03)
<i>Weak_Monitor</i>	0.139*** (5.02)	0.065*** (3.32)	0.055*** (3.11)	0.040* (1.70)
<i>Filings</i>	0.000 (0.36)	0.000 (0.53)	0.000 (0.25)	0.000* (1.65)
<i>Size</i>	0.008*** (5.06)	0.007*** (5.04)	0.009*** (6.02)	0.005*** (3.67)
<i>BM</i>	0.005*** (3.09)	0.005*** (2.94)	0.005*** (2.88)	0.004*** (3.01)
<i>Leverage</i>	-0.030*** (-4.74)	-0.029*** (-4.72)	-0.029*** (-4.68)	0.026*** (5.01)
<i>IO</i>	0.005 (1.31)	0.006 (1.53)		-0.001 (-0.41)
<i>Analyst</i>	-0.001*** (-4.16)		-0.001*** (-3.10)	-0.001*** (-5.33)
<i>ROA</i>	0.323*** (15.45)	0.323*** (15.46)	0.323*** (15.46)	0.803*** (24.76)
<i>CFO</i>	-0.339*** (-12.80)	-0.339*** (-12.79)	-0.339*** (-12.78)	-0.733*** (-23.17)
<i>Lagged_Accruals</i>	0.012 (0.74)	0.013 (0.77)	0.013 (0.77)	-0.048** (-2.53)
<i>Lagged_NOA</i>	-0.001 (-0.69)	-0.001 (-0.70)	-0.001 (-0.69)	0.001 (1.53)
<i>Age</i>	0.031*** (9.72)	0.032*** (9.98)	0.032*** (10.08)	0.010*** (4.94)
<i>Growth</i>	0.005 (0.95)	0.005 (0.91)	0.005 (0.88)	0.002 (0.70)
<i>Constant</i>	0.157*** (4.28)	0.036** (1.98)	0.024 (1.18)	0.047*** (3.15)
Industry Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	40,114	40,117	40,117	25,963
Adj. R-squared	0.5240	0.5233	0.5234	0.5027

Table 8 Incentives

This table reports results from a cross-sectional test of whether the effect of downloads on earnings management is more pronounced for firms with higher incentives to manage earnings. The t-values in the parentheses are based on heteroscedasticity robust standard errors clustered by firm. ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

	(1)	(2)
	<i>SEO</i>	<i>CEO_Tenure</i>
<i>Downloads*Incentives</i>	-0.004*	-0.004*
	(-1.75)	(-1.83)
<i>Downloads</i>	-0.013***	-0.007**
	(-3.06)	(-2.04)
<i>Incentives</i>	0.081***	0.031*
	(4.93)	(1.93)
<i>Filings</i>	-0.000	0.000
	(-0.04)	(1.29)
<i>Size</i>	0.004**	-0.001
	(2.42)	(-0.70)
<i>BM</i>	0.004***	0.003**
	(2.71)	(1.99)
<i>Leverage</i>	-0.028***	0.025***
	(-4.47)	(3.19)
<i>IO</i>	0.004	-0.003
	(1.00)	(-0.97)
<i>Analyst</i>	-0.000	-0.000**
	(-1.32)	(-2.28)
<i>ROA</i>	0.322***	0.890***
	(15.52)	(30.18)
<i>CFO</i>	-0.343***	-0.790***
	(-13.05)	(-25.84)
<i>Lagged_Accruals</i>	0.017	-0.050***
	(1.02)	(-4.32)
<i>Lagged_NOA</i>	-0.000	0.001
	(-0.33)	(0.82)
<i>Age</i>	0.033***	0.011***
	(10.74)	(4.28)
<i>Growth</i>	0.002	0.010*
	(0.36)	(1.88)
<i>Constant</i>	-0.022	0.067***
	(-0.85)	(3.72)
Industry Fixed Effects	Yes	Yes
Year Fixed Effects	Yes	Yes
Observations	40,114	16,777
Adj. R-squared	0.5282	0.3908