Title
Essays in Corporate Finance

Author
Liu, Yibin

Publication Date
2021

UNIVERSITY OF CALIFORNIA SAN DIEGO

Essays in Corporate Finance

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Economics

by

Yibin Liu

Committee in charge:

Professor Joseph E. Engelberg, Co-Chair
Professor Roger H. Gordon, Co-Chair
Professor Jeremy Bertomeu
Professor Eric Floyd
Professor William Mullins
Professor Charles D. Sprenger
Professor Allan G. Timmermann

2021

The Dissertation of Yibin Liu is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2021

DEDICATION

To my parents and grandparents.

EPIGRAPH

Knowledge and action are one inseparable unit
(zhi xing he yi)
—Wang Yangming

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I am extremely grateful to my dissertation committee members: Joseph Engelberg (Co-Chair), Roger Gordon (Co-Chair), Jeremy Bertomeu, Eric Floyd, William Mullins, Charles Sprenger, Allan Timmermann.

I am also grateful to many faculty members and students for all of their generous help and thoughtful feedback - particularly Snehal Banerjee, Julie Cullen, Edwige Cheynel, Alexis Toda, Valerie Ramey, Peicong (Keri) Hu, Jun Chen, Ying Liang, Xiao Ma, Linyan Zhu, Nobuhiko Nakazawa, Wenbin Wu, and Yuting Bai.

Lastly, I would like to thank several anonymous reviewers, and many participants at seminars and conferences.

Chapter 1, in part, is currently being prepared for submission for publication of the material. Liu, Yibin. The dissertation author was the primary investigator and author of this material.

Chapter 2, in part, has been submitted for publication of the material. Bertomeu, Jeremy. Hu, Peicong (Keri). Liu, Yibin. The dissertation author was one of the primary investigators and authors of this paper.

Chapter 3, in part, is currently being prepared for submission for publication of the material. Chen, Jun. Liu, Yibin. The dissertation author was one of the primary investigators and authors of this paper.

VITA

2015	B.A. in Economics (with Honors) and Mathematics, *summa cum laude*, Franklin and Marshall College
2021	Ph. D. in Economics, University of California San Diego

ABSTRACT OF THE DISSERTATION

Essays in Corporate Finance

by

Yibin Liu

Doctor of Philosophy in Economics

University of California San Diego, 2021

Professor Joseph E. Engelberg, Co-Chair
Professor Roger H. Gordon, Co-Chair

This dissertation addresses several questions in corporate finance. A common thread is the study of stock market investors' processing of disclosure by public firms. The first chapter studies the effect of public scrutiny on financial misreporting. I exploit the staggered implementation of the EDGAR system, which provides all investors with free and instant access to financial reports. Firms phased into EDGAR received higher public scrutiny and stronger stock market reaction to earnings announcements. A plausibly exogenous increase in public scrutiny incentivizes firms to substitute between different methods of earnings management. Moreover, the increase in public scrutiny impacts firms differently depending on the ex-ante level of scrutiny that firms already have, consistent with theoretical predictions.

The second chapter models investors' allocation of attention to financial disclosures and its impact on firms' voluntary disclosure. We jointly solve investors' optimal allocation of limited attention and managers' choice to disclose their privately observed signals (e.g., forecasts of future earnings). We predict an inverse-U-shaped relation between firms' likelihood of disclosure and investor attention, supported by our empirical tests.

The third chapter also studies investors' reactions to financial reports. In particular, we examine whether earnings management by manipulating firms distorts investors' response to financial reports by (similar) non-manipulating firms. We exploit a unique setting in China's stock market that de-lists firms if they report consecutive negative annual earnings. We find that non-manipulating firms suffer from significant adverse capital market effects, resulting from investors' distrust.

Chapter 1

Going Digital: The Causal Effect of Information Technology on Financial Reporting

Abstract

This paper studies the causal effect of the implementation of the EDGAR system on financial reporting. Ten groups of public firms were phased-into the EDGAR system from 1993 to 1996, which provides exogenous variations in information acquisition cost. Firms that phased-into EDGAR: 1) substitute away from accruals to real earnings management, highlighting a significant unintended consequence of EDGAR; 2) have an increase in earnings response coefficient which boosts up the marginal benefit of earnings management; 3) firms with low (high) ex-ante public scrutiny increase (decrease) overall earnings management, consistent with the prediction of the inverse-U relation between ex-ante public scrutiny and reporting bias by Samuels, Taylor, and Verrecchia (2020).

1.1 Introduction

In recent decades, the cost of acquiring corporate information in the U.S. has been continuously declining. Before the launch of the Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system in 1993, investors needed to pay a hefty subscription fee to companies that retrieve corporate filings for them, and the turnaround would usually take days (*The Washington Post*, 1993). The EDGAR system has been revolutionary in providing investors with free and instant corporate information.[1] However, providing investors with easier access to corporate information might affect managers' trade-offs in making operational and financial reporting decisions. One possible dark side of more transparency is that managers may be incentivized to reallocate resources to areas that are measured and scrutinized by investors, which can be socially sub-optimal.

My paper focuses on managers' trade-offs in biasing financial reports after a dramatic decrease in investors' cost of acquiring corporate filings, facilitated by the staggered implementation of the EDGAR system from 1993 to 1996. The answer to my research question is

[1] In addition to EDGAR, investors nowadays can directly access financial information announced by public firms on Twitter. Jung et al. (2017) find that almost half of all S&P 1500 firms manage their own Twitter accounts to enhance investors' awareness of their public releases by 2015.

of paramount interest to securities regulators at the U.S. SEC and other countries worldwide. My findings may inform securities regulators on trade-offs between making public firms more transparent and inducing managers to truthfully report their firms' performance.

The EDGAR system dramatically lowers the information acquisition cost and increases public scrutiny of companies. Before EDGAR, public firms transmitted all corporate filings in paper to the SEC which were stored in public reference rooms. It was both time-consuming and expensive to acquire corporate filings.[2] The advent of EDGAR made corporate filings freely and instantly accessible online. The average daily visits to the EDGAR website exceeded 267,000 in 1996 (*The SEC*, 1996).[3] The heightened scrutiny of corporate filings translates into more informative trades by retail investors and more accurate forecasts by financial analysts (Gao and Huang, 2020).

Most crucially for identification, ten groups of public firms were phased-into the EDGAR system from 1993 to 1996, which provides exogenous variations in information acquisition cost.[4] This novel archival setting offers an ideal set of counterfactuals for how earnings management would have changed across time in the absence of changes in information acquisition cost. Consequently, I can disentangle the effect of information acquisition cost from other unobservable and time-varying determinants of earnings management.

The seminal theory on reporting bias by Fischer and Verrecchia (2000) and followed-up theories shed light on how information acquisition cost affects earnings management.[5] Extending Fischer and Verrecchia (2000), Ewert and Wagenhofer (2005) model accruals and real earnings

[2]Investors need to contact companies which specialize in retrieving documents and pay a base rate of $32 for a copy of 10-K, $16 for a 10-Q, and $25 for an annual report (*The Washington Post*, 1993).

[3]As a comparison, Yahoo!, one of the ten most websites at the time, had daily visits around 500, 000 in 1995 (*New York Times*, 1995). Please see section 1.2.2 for more details on how actively the public uses EDGAR for information acquisition.

[4]Please see section 1.2.2 for more details on the staggered implementation of EDGAR.

[5]In a nutshell, managers in the model choose the amount of bias to add to a privately-observed earnings signal. Stock prices are determined by investors' rational expectations, which are based on managers' reported earnings. Intuitively, managers choose a reporting bias that is decreasing in the cost of bias and is increasing in the marginal benefit of inflating earnings. Heightened public scrutiny of financial reports increases the chance of detecting bias and thus the cost of biasing reports. On the other hand, higher stock price response to accounting earnings boosts up the marginal benefit of inflating earnings (Ferri et al., 2018).

management as substitutes in biasing earnings.[6] Managers choose the amount of accruals and real earnings management based on their relative costs. An increase in the relative cost of accruals management would lead managers to use less accrual-based but more real earnings management.

Heightened public scrutiny of corporate filings through EDGAR increases the likelihood of detecting earnings management in general. However, for firms on EDGAR, the increase in the cost of accruals earnings management is probably higher than that of real earnings management for two reasons. Compared to accruals earnings management, it is much more challenging to uncover real earnings management, which are deviations from optimal business decisions.[7] Secondly, even if investors can detect real earnings management, it is very difficult for them to take legal actions against management since real operational changes are protected under the ”business judgment rule” (e.g., Lo (2008), Shon and Yan (2015), Heater et al. (2017), etc).[8] Hence, it is reasonable to assume that being phased-into EDGAR has a more substantial impact on the cost of accrual-based than that of real earnings management. With this additional assumption, Ewert and Wagenhofer (2005) predict a decrease in accruals earnings management and an increase in real earnings management.

I exploit the staggered implementation of EDGAR to empirically test the prediction by Ewert and Wagenhofer (2005). Compared to control firms, firms phased-into EDGAR have lower absolute discretionary accruals (0.6% of lagged assets) and higher negative discretionary accruals (0.7% of lagged assets), implying lower accruals earnings management. On the other hand, firms significantly overproduce inventory to lower their costs of goods sold after they are on EDGAR, resulting in higher abnormal production costs (0.8% of lagged assets). Taken together, empirical evidence that firms shift away from accruals to real earnings management is

[6] Empirical evidence on the substitutability between accrual and real earnings management can be found in papers such as Zang (2012), Cohen and Zarowin (2010a), etc.

[7] The inherent uncertainty in business environments makes it hard to determine what optimal business decisions should have been without real earnings management (Lo, 2008).

[8] On the other hand, accruals earnings management requires managerial estimates that are not related to operational decisions and thus are not protected under the ”business judgment rule” (Heater et al., 2017). Consequently, accruals earnings management has been a focus of shareholder litigations and poses high litigation risks for managers.

consistent with the prediction by Ewert and Wagenhofer (2005).

Turning back to Ewert and Wagenhofer (2005), the substitution between accruals and real earnings management is endogenous. The intuition is that: higher relative cost of accruals reduces the use of accruals manipulation, which boosts up earnings' quality as well as value relevance. Consequently, managers are incentivized by the higher marginal benefit of earnings management to engage in more real earnings management. Empirically, it is crucial to present evidence on a significant increase in the marginal benefit of earnings management after firms became EDGAR filers.

I document a significant rise in stock price response to accounting earnings (earnings response coefficient (ERC)) after firm were phased-into EDGAR. As shown in Table 3.3, the change in ERC is 0.7 to 1.4 higher for firms on EDGAR compared to control firms going from pre- to post-EDGAR periods.[9] This increase in ERC (0.7-1.4) is comparable in magnitude to major accounting events.[10] Evidence on a higher marginal benefit of real earnings management is not only consistent with Ewert and Wagenhofer (2005) but also provides a reasonable explanation of why managers engage in costly real earnings management.

Lastly, Samuels et al. (2020a) examine the impact of ex-ante public scrutiny on reporting bias. In their model, heightened public scrutiny increases the chance of detection and thus the cost of biasing reports. On the other hand, higher public scrutiny also increases the weight investors place on accounting earnings in valuing firms. As illustrated in Figure 1.1, Samuels et al. (2020a) predict an inverse-U relation between reporting bias and public scrutiny.[11] An essential empirical implication of the inverse-U shaped relation is that: after an exogenous increase in public scrutiny such as being phased-into EDGAR, firms with low ex-ante public scrutiny (the red circle) are predicted to increase their reporting bias. In contrast, for firms with high ex-ante public scrutiny (the black triangle), Samuels et al. (2020a) would predict either a decline or an insignificant change in reporting bias.

[9]This result is robust to adding firm-level controls, fiscal year-quarter fixed effects, and firm fixed effects.

[10]For example, ERC increases by 0.8 around the introduction of the PCAOB inspection regime (Gipper et al., 2019) and by 0.925 after enhanced disclosure of executive pay (Ferri et al., 2018).

[11]In other words, reporting bias first increases in public scrutiny, peaks, and then drops in the end.

To test the prediction by Samuels et al. (2020a), I examine the heterogeneous impacts of being phased-into EDGAR on firms with low versus high ex-ante public scrutiny. I follow Gao and Huang (2020) in defining firms with low ex-ante public scrutiny as those with no analyst coverage and market capitalization below the median value.[12]

My results on the heterogeneous impacts of EDGAR are presented in Table 1.8 & 1.9[13] and summarized in Table 1.11 Panel B. Compared to control firms, firms with low ex-ante public scrutiny increase both absolute and positive discretionary accruals and have substantially higher abnormal production costs after they became EDGAR filers. In contrast, firms with high ex-ante public scrutiny reduce their accruals earnings management with lower absolute accruals and higher negative accruals and have insignificant changes in real earnings management after they were on EDGAR. In summary, consistent with the inverse-U relation predicted by Samuels et al. (2020a), firms with low (high) ex-ante public scrutiny increase (decrease) their overall earnings management in response to an exogenous increase in public scrutiny.

My study has examined the EDGAR system, a modern information technology that revolutionized investors' acquisition of corporate information. But more generally, my paper speaks to a hard choice that all securities regulators around the world need to make in designing regulations that either enhance investors' access to information or mandate firms to be more transparent. My findings shed light on a crucial trade-off between making corporate outsiders more informed and distorting corporate insiders' incentives. Managers might be incentivized to window dress their reports and deviate from optimal operational and reporting decisions, catering to corporate outsiders' attention and preferences. In this respect, my results point to the dark side of more transparency and potentially designing current and future disclosure regulations, especially in the age of big data and AI which has been fundamentally transforming investors' information processing.

[12] My results are robust to two alternatives definitions of low ex-ante public scrutiny. Please see section 1.4.4 for more details.

[13] *Post-EDGAR* measures the difference in changes in earnings management between firms with high ex-ante public scrutiny and control firms. I add up *Low-scrutiny*EDGAR* and *Post-EDGAR* to obtain the difference between firms with low ex-ante public scrutiny and control firms.

I perform a battery of robustness checks in Section 1.4.5. Crucial for my generalized difference-in-differences design, I test the parallel trends assumption by analyzing the dynamic treatment effect of being phased-into EDGAR on earnings management. The results support the parallel trends assumption. Moreover, I conduct a falsification test by creating a pseudo-event occurring two years *after* each firm's actual phase-in year. Thirdly, I find that the increase in abnormal production costs is concentrated in manufacturing firms. Fourthly, I exclude firms that volunteered to file electronically on the pilot EDGAR system. Fifthly, I include Industry×Year fixed effects in my specifications to absorb industry-specific shocks and include phase-in group specific time trends as additional controls. Lastly, I repeat my empirical tests with two alternative measures of discretionary accruals. My estimated results continue to go through after the robustness checks.

The rest of the paper is organized as follows: Section 2 discusses my contribution to the literature and also presents institutional details on the staggered implementation of the EDGAR system. Section 3 discusses sample selection and summary statistics. Section 4 describes my empirical results and robustness checks. Section 5 concludes.

1.2 Literature Review and Institutional Details

1.2.1 Literature Review

My paper contributes to a growing strand of literature studying the impact of information cost on earnings management, which is reviewed by Blankespoor et al. (2020a).[14] High information acquisition/processing costs are often assumed to prevent investors from scrutinizing and detecting earnings management (e.g., Dechow et al. (2010), Kim et al. (2019)). Consistent with this assumption, Lo et al. (2017) and Niessner (2015) find evidence that managers manipulate processing costs to hide misreporting. However, as formalized by Samuels et al. (2020a), higher public scrutiny of firms not only facilitates the detection of earnings management but also increases the weight investors place on accounting earnings which increases the marginal benefit

[14] As mentioned in Blankespoor et al. (2020a), "research is just beginning to examine the effects of disclosure processing costs on [...] corporate actions."

of reporting bias. Hence, Samuels et al. (2020a) predict an inverse-U relation between ex-ante public scrutiny and reporting bias. I contribute to this literature by providing the first evidence on the impact of an exogenous decrease in information acquisition cost on managers' trade-off of accrual-based and real earnings management. Furthermore, motivated by Ewert and Wagenhofer (2005), I provide further evidence on an endogenous increase in the marginal benefit of real earnings management. Lastly, I empirically test and find evidence consistent with the inverse-U relation predicted by Samuels et al. (2020a).

Secondly, my paper joins an emerging literature studying the impact of EDGAR by exploiting its staggered implementation. Early studies on the EDGAR system generally treat the launch of EDGAR as a one-time event and document significant market reactions to 10-K/Q filings on EDGAR.[15] Recent papers, using the staggered implementation for identification, have mainly focused on investors and financial analysts.[16] However, given the well-documented impact of EDGAR on corporate outsiders, evidence on how corporate insiders behave after their firms were phased-into EDGAR has been scarce. The only other paper that focuses on managers, to the best of my knowledge, is Goldstein et al. (2020) which find an increase in corporate investment (due to better equity financing) but a decrease in investment to price sensitivity (due to reduced managerial learning from prices) for firms on EDGAR. My paper is the first to study the causal effect of mandatory EDGAR filings on managers' earnings management decisions.

[15] For instance, Qi et al. (2000) find that 10-K reports on EDGAR contain incremental information useful for firm valuation. Asthana and Balsam (2001) find that price and trading volume react significantly to filings on EDGAR. Griffin (2003) reports that the absolute value of the excess return is higher immediately after filings dates from 1996 to 2001. Asthana et al. (2004) document an increase (no change) in the volume of small trades (large investors) for firms that filed 10-K on EDGAR for the first time. Li and Ramesh (2009) show that the stock market reacts significantly to 10-Q/QSB/KSB reports that release earnings information for the first time, as well as 10-K reports that are filed around the calendar quarter-ends from 1996 to 2006.

[16] Compared to control firms, firms phased-into EDGAR have: reduced investor disagreement (Chang et al., 2020a), lower information asymmetry between managers and investors but higher information asymmetry between more- and less- sophisticated investors (Gomez, 2020), lower home bias (Emery and Gulen, 2019), more informative retail investor trades (Gao and Huang, 2020), more information production by analysts (Gao and Huang, 2020), and constrained strategic analyst behavior (Chang et al., 2020b).

1.2.2 Institutional Details on the EDGAR system

Dissemination of Corporate Filings Before EDGAR

In the 1980s, the SEC's public reference rooms were the ultimate repository for stock analysts, lawyers, investment bankers, and investors to gain information on public companies (*New York Times*, 1982). Firms registered with the SEC transmitted paper copies of all corporate filings to the SEC.[17] To acquire a copy of 10 K/Q, an investor would call one of a dozen companies which maintain an army of professionals in the public reference rooms specializing in retrieving documents.[18] The professionals would then chase down a microfiche, read it with a computer, and then print out copies for clients. In terms of pricing, *New York Times* (1982) reported that a page costs 35-90 cents. Fast forward to 1993 right before EDGAR, *The Washington Post* (1993) reported that Disclosure Inc. of Bethesda, a contractor of the SEC, operated the SEC public reference rooms and handled all requests for filing information. Disclosure Inc. charges a base rate of $32 for copies of 10-K filings, $16 for 10-Q filings, and $25 for annual reports.[19]

The Staggered Implementation of EDGAR

The development of the EDGAR system by the SEC consists mainly of two stages: a pilot system commencing in 1984 and a fully operational system starting in 1993. A group of approximately 150 companies volunteered to participate in the pilot system, including AT&T, Exxon, General Motors, IBM, and other major corporations. After the success of the pilot system, the SEC proceeded with developing a fully operational EDGAR system. On February

[17] Paper copies were first reviewed by the SEC staff and then stored in three reference rooms for public viewing in Washington DC, New York, and Chicago.

[18] The public reference rooms were difficult to navigate even for professionals whose job was to retrieve files quickly for clients, letting alone investors who may want to visit the public reference rooms. The 15 reference room staff members did not have time to help since they were preoccupied with sorting and filing about 160,000 documents, responding to 15,000 written requests a year, and also answering as many as 400 phone inquiries per day (*New York Times*, 1982). A professional with Disclosure Inc said "It's just incredible the number of problems you can run into trying to find something you need." and Director of research for Charles E. Simon & Company, another professional research firm, put it simply, "The place can be a zoo." (*New York Times*, 1982).

[19] There are other ways to obtain 10-K/Qs than the SEC's public reference rooms. Public firms mail a paper copy of their annual reports to shareholders. Non-shareholders, such as institutional and retail investors and financial analysts, may write to firms requesting a copy of 10-K/Q. However, the fact that a dozen companies made a living on retrieving documents suggests that direct requests to firms were unlikely to be an effective way of acquiring information.

23, 1993, the SEC released a mandatory phase-in schedule for firms to transmit corporate filings electronically to the EDGAR system. Public firms were categorized into ten groups, and each group was phased-into EDGAR every three to six months over four years. The first four groups were phased-in in 1993: Group 1 (April)[20], Group 2 (July), Group 3 (Oct.), and Group 4 (Dec.). Two groups were phased-in in 1994: Group 5 (Aug.) and Group 6 (Nov.). Another three groups were in 1995: Group 7 (May), Group 8 (Aug.), and Group 9 (Nov.). The staggered implementation was completed when the last group (10) became EDGAR filers in May 1996.

Does EDGAR increase public scrutiny of corporate filings?

As mentioned in the SEC's annual report (*The SEC*, 1996), investors can access "10K/Q and all other corporate filings instantly on home computer screens". They can "display current comparative price-earning, yield, and other data on securities; instantly refine such lists by industry, size, markets and other criteria". Moreover, a substantial fraction of investors had access to the internet in the 1990s.[21]

Furthermore, I provide direct evidence on the number of visits to the EDGAR website. As reported in the 1996 annual report by the SEC (*The SEC*, 1996):

> During the first full year of operation, the [EDGAR] system was heavily accessed,[...] **Average daily connections** exceeded **267,000** and daily data volume downloaded averaged over 10,500,000 bytes.[...] The SEC's home page has become one of the most popular government sites on the World Wide Web.

As a comparison, Yahoo! was one of the ten most-visited websites and had an average daily visit of 500,000 in 1995 (*New York Times*, 1995). Gao and Huang (2020) document that individual investors likely represent over 24.45% of the total number of requests and 31.39%

[20]Firms that volunteered to file electronically on the pilot EDGAR system were assigned to phase-in group one. Firms may volunteer to join the pilot EDGAR system for strategic reasons. My results are unaffected if I exclude the first phase-in group. Please see section 1.4.5 for more details.

[21]11.4% of the U.S. households owned a personal computer with a modem as reported in the 1994 Current Population Survey. The commercial access to the Internet costs as little as two dollars per hour (*Wall Street Journal*, 1992). *New York Times* (1993) further highlighted that "many college students may now obtain Internet access as part of their tuition costs and many businesses buy a high-speed Internet connection [...] permits employees to share unlimited access to the network." The actual percentage of households with internet access is probably much higher than 11.4%.

of the total amount of data requested.[22] More importantly, the large amount of visits to the EDGAR system leads to more informative trades by retail investors and more accurate analysts' forecasts (Gao and Huang, 2020). All together, EDGAR has been greatly facilitating the public's scrutinizing and researching of corporate filings by providing free and instant access online.

How did the SEC assign the phase-in groups?

When the SEC released the Request for Proposal (RFP) for a fully operational EDGAR system in 1987, the RFP requested that contractors propose phase-in schedules by criteria such as company size, industry, or dissemination market interest (*The SEC*, 1987). Chang et al. (2020b) receive confirmation from Scott Bauguess, then Acting Chief Economist of the SEC, that phase-in assignments were decided solely based on firm size. This phase-in criterion comes at no surprise since larger firms tend to have better technological facilities to transition from paper to electronic filings on EDGAR. Important for identification, the assignment of firms to different phase-in groups is random conditional on firm size. I explicitly control for firm size throughout all of my estimations. Furthermore, I test for parallel pre-trends across treatment and control groups and use a placebo event occurring two years after each firm's phase-year.[23] These robustness checks further alleviate the concern that my results may be driven by unobservable differences across phase-in groups.

Can Firms Switch to a Different Phase-in Group?

As stated in the SEC Release No.33-6977, firms can request to the SEC to switch to a different phase-in group. The SEC would only permit firms to change phase-in dates if they indeed face technical difficulties in filing electronically. Around 3% of all firms started electronic filing on a different date than what was specified in the original SEC release (Gao and Huang, 2020).

[22] Gao and Huang (2020) obtain data on the number of requests to the EDGAR system from New York University (https://town.hall.org/govt/tuttle/stats_edgar_domain_073095.html), which provides a breakdown of visits to the EDGAR website by domain names during the week ending July 30, 1995.

[23] Please see section 1.4.5 for more details.

Since firms may delay or accelerate their electronic filings for strategic reasons, I use pre-specified rather than actual phase-in dates.[24] However, using pre-specified phase-in dates introduces measurement errors into my variable of interest. Hence, the estimated coefficient is attenuated towards zero and provides a lower bound of the true effect of EDGAR on earnings management.

1.3 Data and Summary Statistics

I obtain the phase-in schedule from Appendix B of the SEC Release No. 33-6977 (released on February 23, 1993), which provides company name, the Central Index Key (CIK), phase-in group number (CF 01 to 10), and phase-in date for each group. Next, I match firms on the SEC's phase-in schedule with COMPUSTAT using CIK and company name. There are 5,913 firms on the SEC phase-in schedule that have financial information in COMPUSTAT as of 12/31/1992.[25] Furthermore, I obtain data on earnings forecasts from Institutional Brokers' Estimate System (I/B/E/S) and data on institutional ownership from Thomson Reuters. Similar to Gao and Huang (2020) and Gomez (2020), my main sample period starts in 1991 (two years before the first phase-in group) and ends in 1998 (2 years after the last phase-in group). As seen from Table B.2, my main final sample consists of 20,385 firm-year observations and 3,048 unique firms.[26]

For discretionary accruals, I follow Cohen et al. (2008a) in constructing absolute discretionary accruals (*Abs. DA*), positive discretionary accruals (*Pos. DA*), and negative discretionary accruals (*Neg. DA*).[27] In addition, I follow Roychowdhury (2006) in constructing industry-year expectation models for 1) abnormal cash flow from operations (*CFO*), 2) abnormal production

[24] Using actual phase-in dates can potentially bias my estimates. The actual phase-in dates may be correlated with firms' unobservable incentives to manage earnings.

[25] Firms that went public after the release of the phase-in schedule were excluded since these firms may have strategically chosen the timing of their IPO in response to the staggered implementation of EDGAR.

[26] Please see Table B.2 for detailed steps in sample selection (Panel A) and industry decomposition (Panel B).

[27] As in Cohen et al. (2008a), *Abs. DA* is calculated as the absolute value of discretionary accruals (*DA*) from the modified Jones model (Jones, 1991) following Dechow et al. (1995). *Pos. DA* (*Neg. DA*) is equal to discretionary accruals (*DA*) for firm-years with positive (negative) *DA*, and set to zero otherwise. Please see Appendix A.1 for more details on discretionary accruals calculations.

costs (*Prod.*), 3) abnormal discretionary expenditure (*Disc.)*. The residuals from industry-year expectation models are then used as proxies for real earnings management.[28]

Since my earnings management proxies are all constructed at annual frequency, I re-group the ten phase-in groups into four different treatment groups based on when their 10-Ks were filed on EDGAR. The first treatment group consists of the first four phase-in groups that were phased-in in 1993. For firms in the first treatment group, 1993 10-K was the first 10-K they filed on EDGAR. Hence, $Post\text{-}EDGAR_{i,t}$ equals one for firms in the first treatment group in 1993 and thereafter.[29]

I eliminate firms in the regulated industries (SIC codes between 4400 and 5000) and banks and financial institutions (SIC between 6000 to 6500) since their disclosure requirements and accounting rules are significantly different. All continuous variables are winsorized at 1% and 99% percentile. Table B.3 presents the summary statistics of my main sample. The mean (standard deviation) values of absolute, positive, and negative discretionary accruals are 0.096 (0.120), 0.045 (0.085), and -0.050 (0.100). For real earnings management proxies, the mean (standard deviation) values of abnormal cash flow from operations, abnormal production costs, and abnormal discretionary expenditure are 0.010 (0.173), -0.026 (0.241), and 0.027 (0.326), respectively.

1.4 Empirical Results

1.4.1 Impact of EDGAR on Earnings Management

My identification strategy exploits the staggered implementation of the EDGAR system over different phase-in groups. The first goal is to evaluate the impact of mandatory electronic corporate filings on public firms' earnings management. To do that, I estimate the following equation:

$$EM_{i,t} = c_i + c_t + \beta * Post\text{-}EDGAR_{i,t} + Controls_{i,t} + \varepsilon_{i,t}, \tag{1.1}$$

[28]Please also see Appendix A.1 for more details on the construction of proxies for real earnings management.

[29]Similarly, $Post\text{-}EDGAR_{i,t}$ equals one for firms in the second treatment group in 1994 and thereafter which are Group CF 05 and 06 in the SEC's schedule. Lastly, $Post\text{-}EDGAR_{i,t}$ equals one for the third treatment group (CF 07-09) in 1995 and the fourth group (CF 10) in 1996 and thereafter.

where i indexes firm and t indexes year. The dependent variable is a proxy for either discretionary accruals or real earnings management depending on the specification. $Post\text{-}EDGAR_{i,t}$ equals to one when firm i is subject to the mandatory filing on EDGAR in year t and stays one afterwards. I control for fixed idiosyncratic firm earnings management choices with firm fixed effects (c_i) and time-related effects with year fixed effects (c_t). Standard errors are clustered by firm to account for potential transitory shocks that are correlated across time for a specific firm.

The coefficient of interest, β, is identified from plausibly exogenous time variations in when different groups of firms were required to file electronically on EDGAR. β captures the difference between the change in phased-in firms' level of earnings management and the change in control firms' earnings management from pre- to post- EDGAR periods.

Furthermore, I control for firm-level factors of earnings management. More specifically, I control for firm size by including the natural log of lagged total assets (e.g., Dechow (1994); Dechow and Dichev (2002)). To control for growth opportunities across firms, I include the Market-to-Book ratio (market capitalization/book equity) and also Sales Growth (change in sales/lagged sales). Moreover, I include Long-term Leverage (ratio of long-term liabilities to total assets) following DeFond and Jiambalvo (1994). As in Kothari et al. (2005), I control for potential confounding correlation between cash flows and accruals by including the ratio of cash flow from operations to lagged total assets. I also include an indicator that equals one if a firm's financials are audited by a big 4/5 auditor to control for auditor quality (Becker et al., 1998). The other controls included are return on assets (ROA calculated as income before extraordinary items/total assets), Interest Coverage Ratio (interest expense/income before extraordinary items), and an indicator for negative income (Loss).

Results on Discretionary Accruals

As shown from Table 1.4 Panel A, absolute discretionary accruals go down by 0.6% to 1.3% of lagged total assets once a firm starts filing on EDGAR. The estimated coefficient is stable and statistically significant across three different specifications: (1) the univariate specification (column 1), (2) with both firm and year fixed effects (column 2), (3) with both fixed effects and time-varying firm-level controls (column 3).[30]

A significant drop in absolute discretionary accruals suggests that firms engage in less accrual-based earnings management (Cohen et al., 2008a). However, it remains unclear whether the decline in absolute discretionary accruals is due to a decrease in positive discretionary accruals, or an increase in negative discretionary accruals, or both. Results from Table 1.4 Panel B report an insignificant change in positive discretionary accruals, whereas Table 1.4 Panel C shows a significant increase in negative discretionary accruals by 0.7% to 0.9% of lagged total assets after firms become EDGAR filers. An increase in negative discretionary accruals indicates that managers reduce cookie jar accounting, which refers to income decreasing accruals in the current period for opportunities to boost future earnings (Healy and Wahlen, 1999).

Results on Real Earnings Management

Table 3.2 presents results on the impact of filing electronically on EDGAR on three proxies of real earnings management. There is no significant change in abnormal cash flow from operations (column 1) or abnormal discretionary expenditure (column 2). However, firms increase abnormal production costs by 0.8% to 1% of lagged assets (columns 3 to 5) when they are on EDGAR compared to control firms.[31] In other words, firms overproduce products to lower the cost of goods sold (COGS), which inflates their reported profitability in the current period.

Trade-off Between Accrual-based and Real Earnings Management

As summarized in Table 1.11 Panel A, firms substitute away from accrual-based towards real earnings management once their filings are freely accessible on EDGAR. More specifically, firms significantly reduce their use of cookie jar accounting and overproduce products at the same time.

[30] Furthermore, results on absolute discretionary accruals are robust to using two alternative measures of discretionary accruals, accounting for industry-level shocks, and phase-in group-specific time-trends. Please see Section 1.4.5 Additional Robustness Checks for more details.

[31] The estimated coefficient of abnormal production costs is robust and stable across three different specifications: the univariate specification (column 3), with both firm and year fixed effects (column 4), and with both fixed effects, and time-varying firm-level controls (column 5).

My empirical findings are consistent with the theoretical prediction by Ewert and Wagenhofer (2005). In their model, managers choose the amount of accrual versus real earnings management depending on their relative costs. More public scrutiny of corporate filings on EDGAR increases the chance of detecting earnings management overall. Compared with accruals earnings management, real earnings management are more difficult for investors to uncover and take successful legal actions against (Lo, 2008). Consequently, a relatively higher cost in managing accruals will result in lower accruals earnings management and higher real earnings management for firms phased-into EDGAR.

1.4.2 Why do managers engage in costly real earnings management?

Since real management activities impose real costs on firms, it is crucial to present evidence on why managers engage in higher costly real earnings management after their firms became EDGAR filers. The intuition in Ewert and Wagenhofer (2005) is that the higher relative cost of accruals reduces the use of accruals manipulation, making accounting earnings more value relevant. Consequently, a higher marginal benefit of inflating earnings incentivizes managers to take on more real earnings management.

I document a significant jump in stock price response to accounting earnings (earnings response coefficient (ERC)) after firms were phased-into EDGAR. As shown in Table 3.3, the change in ERC is 0.7 to 1.4 higher for firms phased in EDGAR compared to control firms going from pre- to post-EDGAR periods, which is estimated from the following generalized difference-in-differences design similar to Ferri et al. (2018):

$$\begin{aligned} CAR_{i,t} = c_i + c_t + \beta_1 UE_{i,t} * Post\text{-}EDGAR_{i,t} + \beta_2 UE_{i,t} + \beta_3 Post\text{-}EDGAR_{i,t} \\ + Controls + UE_{i,t} * Controls + \varepsilon_{i,t}, \quad (1.2) \end{aligned}$$

where $CAR_{i,t}$ is 3-day market-adjusted stock return around quarterly earnings announcements dates. $UE_{i,t}$ is unexpected earnings calculated as the difference between actual quarterly earnings per share (EPS) and median of one-quarter-ahead analysts' quarterly EPS forecasts scaled by stock price two days before the earnings announcement. *Post-EDGAR*$_{i,t}$ is an indicator variable which equals to one for those quarterly earnings announcements by firm i that occurred after firm i became an EDGAR filer. I control for fixed idiosyncratic stock price reaction to earnings with firm fixed effects and time-related effects with fiscal year-quarter fixed effects. Standard errors are clustered by earnings announcement dates to account for potential cross-sectional dependencies among firms announcing earnings on the same day (DellaVigna and Pollet (2009), Hirshleifer et al. (2009), and Ferri et al. (2018)). Following the literature on ERCs (e.g., Collins and Kothari (1989a), Easton and Zmijewski (1989), and Ferri et al. (2018)), I include both time-varying firm-level controls and their interaction terms with UE_{it}: firm size, Market-to-book ratio, Long-term Leverage, a loss indicator (Loss), and analysts' forecast dispersion.[32]

The coefficient of interest, β_1, measures the difference between the change in phased-in firms' ERCs and the change in control firms' ERCs.[33] The estimated β_1 ranges from 0.7 to 1.4 and remains statistically significant across five different specifications.[34] Furthermore, the estimated increase in ERC (0.7-1.4) brought by EDGAR is comparable in magnitude to changes in ERCs around major accounting events.[35] The substantial rise in stock price response to accounting earnings is not only consistent with the prediction by Ewert and Wagenhofer (2005) but also sheds light on managers' incentives in engaging in costly real earnings management.

[32]Dispersion is calculated as the difference between the highest and lowest analyst forecasts, scaled by stock price from two days before the earnings announcement.

[33]Table 3.3 column 1 reports the results of regressing CAR on UE in the pooled sample without either controls or any fixed effects. The estimated coefficient of UE in the pooled sample is 1.467, a reassuring result broadly in line with ERC estimates in the literature (Kothari (2001)).

[34]Column 2 the uni-variate specification, column 3 with both firm-level controls and their respective interaction terms with UE_{it}, column 4 with fiscal year-quarter fixed effects, SIC (4-digit) industry fixed effects, and controls, column 5 with firm fixed effects, fiscal year-quarter fixed effects, and controls, lastly column 6 as a robustness check using mean analyst forecast instead of median analyst forecast to calculate UE_{it}.

[35]For example, Gipper et al. (2019) report an increase of about 0.8 in the annual ERC around the introduction of the PCAOB inspection regime. Ferri et al. (2018) document an improvement of ERC of 0.925 for firms with enhanced executive pay disclosures mandated by the SEC in 2006.

1.4.3 Additional Outcome

Meet or Beat Analyst Forecasts

A natural follow-up question is whether firms are more likely to meet or beat analysts' forecasts. Given that firms on EDGAR substitute away from accruals to real earnings management, the net effect of earnings management on meeting earnings targets is ambiguous. I run the following linear probability model to access whether firms that phased-into EDGAR are more likely to meet or beat analysts' forecasts:

$$MeetorBeat_{i,t} = c_i + c_t + \beta * Post\text{-}EDGAR_{i,t} + Controls_{i,t} + \varepsilon_{i,t},$$

where i indexes firm and t indexes fiscal year-quarter.[36] $MeetorBeat_{i,t}$ is a dummy variable that equals to one if a firm i meets or beats analysts' expectations of earnings per share by 1) zero or one cent; 2) zero, one, or two cents (e.g., Burgstahler and Eames (2006), Cheong and Thomas (2018), Heater et al. (2017), etc). Table 1.7 shows that firms phased-into EDGAR are not significantly more likely to meet or beat analysts' forecasts, suggesting that the decrease in accruals earnings management might have offset the increase in real earnings management in meeting earnings targets.

1.4.4 Heterogeneous impact of EDGAR as a test of Samuels et al. (2020a)

Samuels et al. (2020a) examine how the ex-ante level of public scrutiny affects reporting bias. There are two countervailing effects of ex-ante public scrutiny on reporting bias. Heightened public scrutiny facilitates investors' monitoring of managers and thus deters them from biasing reports. Secondly, as public scrutiny increases, investors weigh accounting earnings more heavily in valuing the firm, which increases managers' marginal benefit of inflating one extra dollar.

The two countervailing effects combined result in an inverse-U relation between public scrutiny and reporting bias. As illustrated in Figure 1.1, reporting bias is first increasing in the

[36] c_i is firm fixed effect and c_t is fiscal year-quarter fixed effect.

level of ex-ante public scrutiny, peaks, and goes down eventually. An exogenous increase in public scrutiny is predicted to have heterogeneous impacts on firms with low versus high ex-ante public scrutiny. As EDGAR increases investors' scrutiny of corporate filings, firms with low ex-ante public scrutiny (the red circle) are predicted to have a higher reporting bias, whereas firms with high ex-ante public scrutiny (the black triangle) have a lower (or an insignificant change in) reporting bias after they are phased-into EDGAR compared to control firms.

I use analyst coverage, market capitalization, and institutional ownership to proxy for firm-level ex-ante public scrutiny following Gao and Huang (2020) and Gomez (2020). Since the theory by Samuels et al. (2020a) requires the level of public scrutiny to be *ex-ante* (i.e., before managers' choice of bias), all of my proxies are measured at the end of 1992 which is the year before the phase-in of the first group (1993). I define firms with low ex-ante public scrutiny as those with no analyst coverage and market capitalization below the median value. For robustness, I use two alternative definitions for low ex-ante public scrutiny: 1) if a firm has no institutional ownership and market capitalization below the median value; 2) if a firm has no analyst coverage and institutional ownership below the median value.

Heterogeneous impact of EDGAR on discretionary accruals

Table 1.8 reports how filing electronically on EDGAR affects firms with high versus low ex-ante public scrutiny differently. More precisely, *Post-EDGAR* measures the difference in changes in discretionary accruals between firms with high ex-ante public scrutiny and control firms. *Low-scrutiny* EDGAR* captures the difference between firms with low and firms with high ex-ante public scrutiny. Lastly, I add up *Low-scrutiny*EDGAR* and *Post-EDGAR* to obtain the difference in changes of discretionary accruals between firms with low ex-ante public scrutiny and control firms.

Table 1.8 columns 1 to 3 show a significant decrease in absolute discretionary accruals for firms with high ex-ante public scrutiny (*Post-EDGAR* < 0) after they were phased-into EDGAR across all three definitions of public scrutiny. In contrast, firms with low ex-ante public scrutiny experience an increase in absolute discretionary accruals after they were on EDGAR compared to control firms (*Low-scrutiny*EDGAR* + *Post-EDGAR* > 0).[37] Turning to columns 4 to 6, I do not find a significant change in positive discretionary accruals for firms with high ex-ante public scrutiny. In contrast, there is a substantial increase in positive discretionary accruals for firms with low ex-ante public scrutiny.[38] Lastly, I document a significant increase in negative discretionary accruals for firms with high ex-ante public scrutiny and no significant change in negative discretionary accruals for firms with low ex-ante public scrutiny.[39]

Heterogeneous impact of EDGAR on real earnings management

Table 1.9 shows the differential impact of EDGAR implementation on real earnings management for firms with high versus low ex-ante public scrutiny. Similar to analysis on discretionary accruals above, *Post-EDGAR (Low-scrutiny*EDGAR + Post-EDGAR)* measures the difference in changes of real-activity based earnings management between firms with high (low) ex-ante public scrutiny and control firms.

Across all three definitions of ex-ante public scrutiny, I do not find a significant change in abnormal cash flow from operations (columns 1 to 3) or abnormal discretionary expenditure (columns 7 to 9) for firms either with high or low ex-ante public scrutiny after they were on EDGAR. However, firms with low ex-ante public scrutiny have a substantial rise in abnormal production costs compared to control firms whereas there is no significant change for high ex-ante public scrutiny firms.[40] In other words, the increase in abnormal production costs in the whole sample is mainly driven by firms with low ex-ante public scrutiny.

[37] The caveat is that the *Low-scrutiny*EDGAR* + *Post-EDGAR* is not statistically significant. This concern is alleviated in the sub-sample analysis presented in Table 1.10 Panel A. Using only firms with high ex-ante public scrutiny, column 2 shows a statistically significant drop in absolute discretionary accruals after firms were on EDGAR compared to controls firms.

[38] As I show in the sub-sample analysis in Table 1.10 Panel A, column 3 (4) using only firms with low (high) ex-ante public scrutiny reports a statistically significant increase (an insignificant change) in positive discretionary accruals compared to control firms.

[39] The results for negative discretionary accruals are confirmed by sub-sample analysis reported by Table 1.10 Panel A columns 5 and 6.

[40] In addition, what I report in Table 1.9 is confirmed by the sub-sample analysis in Table 1.10 Panel B.

Discussion

Table 1.11 Panel B summarizes the heterogeneous impacts of EDGAR on firms with high and low ex-ante public scrutiny. For firms with low ex-ante public scrutiny, there is a significant increase both accrual-based (both positive and absolute discretionary accruals) and real activity earnings management (abnormal production costs) after they are on EDGAR. On the other hand, firms with high ex-ante public scrutiny have a decrease in accrual-based earnings management and no change in real based earnings management.

In short, after an exogenous increase in ex-ante public scrutiny, firms with low (high) ex-ante public scrutiny increase (decrease) their overall earnings management activities. My results are consistent with an inverse-U relation between ex-ante public scrutiny and reporting bias predicted by Samuels et al. (2020a).[41]

1.4.5 Additional Robustness Checks

Testing the parallel trends assumption

The parallel trends assumption is critical for my generalized difference-in-differences design. To test the assumption, I analyze the dynamic effect of mandatory filings on EDGAR in the years before and after the actual phase-in date using the following specification:[42]

$$EM_{i,t} = c_i + c_t + \sum_{l=-4, l\neq -1}^{l=4} \beta_l * EDGAR_{i,l} + Controls_{i,t} + \varepsilon_{i,t}. \tag{1.3}$$

The key variables of interest are a set of 8 indicator variables $EDGAR_{i,l}$, which indicates

[41] My test complements the set of empirical tests in Samuels et al. (2020a). One of their tests exploits the mandatory EDGAR reporting of Form 4 (managers' equity trades) in 2003. Their findings on the effect of EDGAR dissemination of managers' equity trades are consistent with the prediction of their model.

My test shall be interpreted with caution since it assumes that the increase in public scrutiny is reasonably small. The exact magnitude of the increase in public scrutiny brought by EDGAR is unobservable. If indeed EDGAR leads to a small increase in public scrutiny, the EDGAR setting is ideal for testing Samuels et al. (2020a). However, if the increase of public scrutiny for firms on EDGAR is so large that it passes the inflection point, I may not find a significant rise in reporting bias for firms with low ex-ante public scrutiny even if predictions by Samuels et al. (2020a) are correct. My test is a joint test of the magnitude of the increase in public scrutiny and the theory by Samuels et al. (2020a).

[42] This design to test parallel trends assumption has also been adopted by papers including Christensen et al. (2016), Christensen et al. (2017), Duguay et al. (2019), Samuels (2020), Samuels et al. (2020a), etc.

the relative year around when each firm's 10-K was first available on EDGAR. l goes from -4 to $+4$ with the year immediately preceding each firm's phase-in year ($EDGAR_{i,l=-1}$) omitted from the regression and its coefficient set as zero.[43]

Figure 1.2 presents the dynamic effect of filing electronically on EDGAR on absolute (sub-figure 1.2a) and negative (sub-figure 1.2b) discretionary accruals. Consistent with the parallel trends assumption, the estimated coefficients of the years before the actual phase-in year (i.e., the four dots to the left of the dashed line) are not statistically significant for either absolute or negative discretionary accruals. Similarly, Figure **??** shows results that are consistent with the parallel trends assumption for abnormal production costs.[44] Taken together, trends in earnings management across mandatory EDGAR filers and control firms are not significantly different before firms' phase-in years. The impact of EDGAR on earnings management only materializes when firms become EDGAR filers (i.e., the dots to the right of the dashed line). Combined with the fact that the assignment of firms to phase-in groups is random conditional on firm size (Chang et al., 2020b), my estimated effects are unlikely to be driven by unobservable firm characteristics that correlate with both the timing of the EDGAR implementation and earnings management.

Falsification Test

In addition to testing for parallel trends *before* the phase-in year, I conduct a falsification test by creating a pseudo-event occurring two years *after* each firm's actual phase-in year. As shown in Table 1.12 Panel A, there is neither statistically nor economically significant change in earnings management around the pseudo-event. For example, the estimated coefficients on the *pseudo Post-EDGAR* indicator are 0.000377, -0.00184, -0.00112 for negative, absolute

[43] $EDGAR_{i,l=-4}$ ($EDGAR_{i,l=+4}$) corresponds to four or more years before (after) when each firm's 10-K was first available on EDGAR. As an example: suppose firm i's first 10-K on EDGAR is its 1994 10-K. Since my sample for dynamic treatment effect starts from 1989 (4 years before the first phase-in group in 1993) to 2000 (4 year after the last phase-in group in 1996), $EDGAR_{i,l=-4}$ is set to 1 for firm i in year 1989 and 1990; $EDGAR_{i,l=-3} = 1$ for year 1991; $EDGAR_{i,l=-2} = 1$ for year 1992; $EDGAR_{i,l=-1} = 1$ for year 1993 (the omitted dummy); $EDGAR_{i,l=0} = 1$ for year 1994 (the phase-in year); $EDGAR_{i,l=1} = 1$ for year 1995; $EDGAR_{i,l=2} = 1$ for year 1996; $EDGAR_{i,l=3} = 1$ for year 1997; $EDGAR_{i,l=+4} = 1$ for year 1998, 1999, and 2000.

[44] For completeness, I also estimate the dynamic treatment effect for the remaining outcome variables: positive discretionary accruals (sub-figure B.1c); abnormal cash flow from operations (sub-figure **??**); and abnormal discretionary expenditure (sub-figure 1.4c).

discretionary accruals, and abnormal production costs, whereas results using the true EDGAR indicator are: 0.00674, −0.00603, 0.00829, respectively.

Excluding phase-in group one

As mention in section 1.2.2 above, over one hundred firms volunteered to file electronically on the pilot EDGAR system. The volunteer firms were assigned to phase-in group one in the phase-in schedule. To lessen the concern that my estimated results might be affected by these firms, I repeat my empirical tests after excluding the first phase-in group. Table 1.12 Panel B presents results estimated after excluding phase-in group one. The effect of mandatory EDGAR filings on both accrual and real earnings management continues to go through, and the magnitude of the estimated coefficients is qualitatively unchanged. For comparisons, the new (old) estimated coefficients for negative, absolute discretionary accruals, and abnormal production costs are 0.00598 (0.00674), −0.00496 (−0.00603), 0.00852 (0.00829).

Manufacturing versus Non-manufacturing Firms

My documented higher abnormal production costs (relative to sales) after firms become EDGAR filers can result from both overproduction of inventory and price discounts (Roychowdhury, 2006). While price discounts can be used by both manufacturing and non-manufacturing firms, overproduction as a means to manage earnings is only available to manufacturing firms. I do not find significant evidence that firms offer more price discounts (i.e., abnormal cash flow from operations not significantly different) after they become EDGAR filers. Consequently, the documented jump in abnormal production costs has to come from manufacturing firms overproducing inventory.

Empirically, I expect to find that manufacturing (non-manufacturing) firms have significantly higher (insignificant) abnormal production costs after they become EDGAR filers. Following Cohen et al. (2008a), I define manufacturing firms as those with two-digit SIC code between 20 and 39. As seen from Table 1.12 Panel C, non-manufacturing firms do not have higher abnormal production costs (column 1), whereas manufacturing firms have significantly higher abnormal production costs after they became EDGAR filers (columns 2 and 3).[45]

Controlling for Group-specific Time Trends & Industry×Year Fixed Effects

One threat to my identification strategy is that the estimated coefficient of the variable of interest might be capturing differential time trends across phase-in groups. To counter this threat, I include phase-in group-specific time trends as additional controls. Therefore, the impact of EDGAR on earnings management is identified as each group's deviation from pre-existing group-specific time trends. If my estimated effects are not driven by differential time trends, the inclusion of such time trends will not change either the statistical significance or economic magnitude of the estimated effects. Furthermore, I include Industry×Year fixed effects to absorb industry-specific unobservable shocks. Results from Table 1.12 Panel D show that the estimated effects continue to go through.

Alternative Measures of Discretionary Accruals

Table 1.12 Panel E reports results using two alternative measures of discretionary accruals. Columns 1 and 2 show results using the first alternative in which $\Delta REV_{i,t}/Assets_{i,t-1}$ is replaced with $(\Delta REV_{i,t} - \Delta AR_{i,t})/Assets_{i,t-1}$ in the first stage regression of estimating the modified Jones model following Cohen et al. (2008a).[46] Columns 3 and 4 present results using performance-matched discretionary accruals following Kothari et al. (2005).[47] The estimated results using these two alternative measures remain consistent with what was reported in the paper before.

[45]Results from column 1 are estimated using only non-manufacturing firms, column 2 using only manufacturing firms, column 3 using both manufacturing and non-manufacturing firms. *Post-EDGAR* measures the difference in abnormal production costs between either non-manufacturing (column 1) or manufacturing firms (column 2) and control firms. *Manu* is a dummy variable that equals to one for manufacturing firms. *Manu* × *Post-EDGAR* measures the difference in abnormal production costs between manufacturing and non-manufacturing firms.

[46]Please see Appendix A.1 for more details.

[47]Each sample firm is matched with another firm that is from the same fiscal year industry and has the closest return on assets as the given firm. The performance-matched discretionary accruals are then computed as each sample firm's discretionary accruals minus the discretionary accruals of the matched firm.

1.5 Conclusion

This paper studies the causal impact of information acquisition cost on managers' decisions to manage earnings. For identification, the staggered implementation of the EDGAR system from 1993 to 1996 provides exogenous time variations in when information acquisition cost is lowered for ten phase-in groups.

My empirical tests are motivated by seminal theories on reporting bias that originate from Fischer and Verrecchia (2000). Consistent with the trade-off theory of accruals and real earnings management by Ewert and Wagenhofer (2005), I find that firms substitute away from accruals to real earnings management. This result highlights a significant unintended consequence of the EDGAR system on managers' incentives even though the EDGAR system has been revolutionary in providing timely corporate information to the general public. To shed light on managers' incentives in engaging in costly real earnings management, I document a significant increase in the marginal benefit of real earnings management after firms were phased-into EDGAR. Lastly, the heterogeneous impacts of mandatory EDGAR filings on firms with high versus low ex-ante public scrutiny support the inverse-U relation between public scrutiny and reporting bias predicted by Samuels et al. (2020a).

Acknowledgements

This chapter, in part, is currently being prepared for submission for publication of the material. Liu, Yibin. The dissertation author was the primary investigator and author of this material.

I am grateful for the insightful comments and suggestions from Chyhe Becker (the SEC), Jeremy Bertomeu, Stefano Cascino, Julie Cullen, Frank Ecker, Joseph Engelberg (Co-Chair), Eric Floyd, Roger Gordon (Co-Chair), Peicong (Keri) Hu, Yuping Jia, Chuchu Liang (Discussant), Ying Liang, William Mullins, Martin Nienhaus (Discussant), Veronika Pool, Delphine Samuels, Guoman She, Laurence van Lent, Yanhua (Sunny) Yang (Discussant), Yachang Zeng, Ning Zhang, and Ruishen Zhang as well as participants at NUS, CUHK, HKU, HK Baptist, Tilburg,

Bocconi, 2020 American Accounting Association Annual Meeting, 2020 Hawai'i Accounting Research Conference, 2019 Financial Management Association Doctoral Student Consortium & Job Market Paper Presentation Session, the 2nd Emerging Scholars in Accounting Conference 2019 at Frankfurt School of Finance & Management, and Duke Accounting Theory Summer School 2019 (Poster Session).

Table 1.1. Variable Definitions

Variables	Definitions	Sources
Dependent Variables		
Abs. DA	The absolute value of abnormal accruals, following Cohen et al. (2008a)	Compustat
Pos./Neg. DA	The positive/negative value of abnormal accruals, following Cohen et al. (2008a)	Compustat
CFO	Abnormal cash flow from operations, following Roychowdhury (2006)	Compustat
Prod.	Abnormal production costs, following Roychowdhury (2006)	Compustat
Disc.	Abnormal discretionary expenditure, following Roychowdhury (2006)	Compustat
$CAR_{[-1.+1]}$	3-day market-adjusted stock return around quarterly earnings announcements dates	CRSP
Variables of Interest		
Post-EDGAR	Indicator Variable: equal to one if a firm is on EDGAR and thereafter	SEC release
Unexpected Earnings	Difference between actual quarterly EPS and median of one-quarter-ahead analysts' quarterly EPS forecasts scaled by stock price two days before the earnings announcement	I/B/E/S
Control Variables		
Interest Cov. Ratio	Interest expense / Income before extraordinary items	Compustat
Oper. Cash Flows	Cash flow from operations / Lagged total assets	Compustat
Long-term Lev.	Long-term liabilities / Total assets	Compustat
Sales Growth	Change in sales / Lagged sales	Compustat
Size	Natural log of lagged total assets	Compustat
Market-to-Book	Market capitalization / Book value of equity	Compustat
ROA	Return on Assets (Income before extraordinary items / Total assets)	Compustat
Loss	Indicator if income before extraordinary items is negative	Compustat
Big 4/5 Auditor	Indicator if a firm is audited by a Big 4/5 auditor	Compustat

Table 1.2. Sample Selection and Industry Composition

Panel A: Sample Selection Procedure

	Details	Firm-Year	Firms
Step 1	Firms on SEC's phase-in schedule and have financial information on COMPUSTAT as of 12/31/1992		5,913
	a) : sample period 1991-1998	39,386	
Step 2	Exclude firms with missing values of proxies for discretionary accrual and real earnings management	(10,528)	(1,607)
Step 3	Exclude observations without controls for main regressions	(5,760)	(891)
Step 4	Exclude firms from financial (SIC 6000-6900) and utility industries (SIC 4900-4949)	(2,713)	(367)
	Total	20,385	3,048

Panel B: 2-Digit SIC Industry Composition

2-digit SIC	Industry	No.	%
01-09	Agricultural and Forestry	52	0.26
10-19	Mining, Oil and Gas, and Others	1,673	8.21
20-27	Food, Printing and Publishing	2,391	11.73
28-29	Chemicals, Petroleum and Coal, Rubber and Plastics	1,931	9.47
30-39	Metal, Machinery and Equipment	7,949	38.99
50-59	Wholesale and Retail	2,447	11.93
70-79	Business Services, Auto Repair and Recreation	2,533	12.00
80-89	Health, Engineering and Management Service	944	4.63
99	Others	465	2.28
Total		20,385	100

Table 1.3. Summary Statistics

	# of Obs	Mean	Median	SD	P25	P75
Post-EDGAR	20,385	0.590	1.000	0.492	0.000	1.000
Abs. DA	20,385	0.096	0.057	0.120	0.025	0.117
Pos. DA	20,385	0.045	0.001	0.085	0.000	0.056
Neg. DA	20,385	-0.050	0.000	0.100	-0.057	0.000
Abnormal CFO	20,385	0.010	0.025	0.173	-0.048	0.092
Abnormal Production Costs	20,385	-0.026	-0.026	0.241	-0.153	0.095
Abnormal Discretionary Expenditure	20,385	0.027	-0.009	0.326	-0.129	0.142
Cumulative Abnormal Return	16,114	0.003	0.003	0.064	-0.028	0.035
Unexpected Earnings	16,114	-0.002	0.000	0.007	-0.001	0.001
Interest Cov. Ratio	20,385	0.391	0.069	2.283	-0.042	0.410
Oper. Cash Flows	20,385	0.039	0.070	0.202	-0.000	0.131
Long-term Lev.	20,385	0.174	0.122	0.189	0.012	0.274
Sales Growth	20,385	0.144	0.072	0.475	-0.037	0.211
Size	20,385	4.427	4.353	2.112	2.925	5.838
Market-to-Book	20,385	2.599	1.791	4.432	1.011	3.158
Return on Asset	20,385	-0.041	0.033	0.305	-0.039	0.078
Loss	20,385	0.321	0.000	0.467	0.000	1.000
Big 4/5 Auditor	20,385	0.732	1.000	0.443	0.000	1.000

Note: this table reports summary statistics for variables used in my main analysis. All continuous variables have been winsorized at 1% and 99%. Please see Table B.1 for detailed definitions of each variable.

Table 1.4. Changes in Discretionary Accruals after Firms Phased-into EDGAR

This table shows how discretionary accruals change after firms became EDGAR filers compared to control firms. The coefficients are estimated from the following specification:

$$EM_{i,t} = c_i + c_t + \beta * Post\text{-}EDGAR_{i,t} + Controls_{i,t} + \varepsilon_{i,t}$$

where i indexes firm and t indexes year. Firm and year fixed effects are captured by c_i and c_t, respectively. $Post - EDGAR_{i,t}$ equals one when a firm i started filing on EDGAR in year t and thereafter. The sample includes all firms on the SEC's phase-in list that can be matched to COMPUSTAT with available information from 1991 to 1998. All standard errors are clustered at firm-level. t statistics in parentheses. * indicates statistical significance at the 10% level,** at the 5% level, and *** at the 1% level. **Results on absolute, positive, and negative discretionary accruals are reported in Panel A, B, and C, respectively.**

Panel A: Absolute Discretionary Accruals

	(1) Abs. DA	(2) Abs. DA	(3) Abs. DA
Post-EDGAR	-0.0126*** (-6.89)	-0.00635** (-2.45)	-0.00603** (-2.32)
Interest Cov. Ratio			-0.00102*** (-3.07)
Oper. Cash Flows			-0.0391*** (-2.94)
Long-term Lev.			0.000154 (0.02)
Sales Growth			0.0250*** (7.61)
Size			-0.00505 (-1.58)
Market-to-Book			0.00100*** (3.09)
ROA			-0.103*** (-11.68)
Loss			-0.00460 (-1.46)
Big 4/5 Auditor			-0.00903* (-1.83)
Year FE		✓	✓
Firm FE		✓	✓
Controls			✓
Observations	20385	20385	20385
Adjusted R^2	0.002	0.331	0.372

Table 1.4. Changes in Discretionary Accruals after Firms Phased-into EDGAR (Continued)

Panel B: Positive Discretionary Accruals

	(1) Pos. DA	(2) Pos. DA	(3) Pos. DA	(4) Pos. DA
Post-EDGAR	-0.000583	0.00294	0.00271	0.00111
	(-0.43)	(1.20)	(1.39)	(0.55)
Interest Cov. Ratio			-0.00216***	-0.00210***
			(-8.75)	(-8.81)
Oper. Cash Flows			-0.269***	-0.270***
			(-30.79)	(-30.60)
Long-term Lev.			0.0162**	0.0144**
			(2.47)	(2.21)
Sales Growth			-0.00143	-0.00110
			(-0.69)	(-0.54)
Size			-0.0149***	-0.0168***
			(-6.19)	(-6.84)
Market-to-Book			-0.000122	-0.000140
			(-0.56)	(-0.64)
ROA			0.137***	0.138***
			(20.49)	(20.75)
Loss			-0.0348***	-0.0364***
			(-15.95)	(-16.70)
Big 4/5 Auditor			-0.00494	-0.00521
			(-1.37)	(-1.46)
Year FE		✓	✓	
Firm FE		✓	✓	✓
Controls			✓	✓
Ind×Year FE				✓
Group Trends				✓
Observations	20385	20385	20385	20385
Adjusted R^2	0.000	0.188	0.470	0.478

Table 1.4. Changes in Discretionary Accruals after Firms Phased-into EDGAR (Continued)
Panel C: Negative Discretionary Accruals

	(1) Neg. DA	(2) Neg. DA	(3) Neg. DA
Post-EDGAR	0.00845*** (5.05)	0.00855*** (2.90)	0.00674*** (3.06)
Interest Cov. Ratio			-0.0010*** (-4.77)
Oper. Cash Flows			-0.241*** (-21.16)
Long-term Lev.			0.0168** (2.03)
Sales Growth			-0.0279*** (-9.45)
Size			-0.0109*** (-3.73)
Market-to-Book			-0.0012*** (-4.57)
ROA			0.252*** (30.77)
Loss			-0.0294*** (-12.82)
Big 4/5 Auditor			0.00416 (1.26)
Year FE		✓	✓
Firm FE		✓	✓
Controls			✓
Observations	20385	20385	20385
Adjusted R^2	0.001	0.262	0.601

Table 1.5. Changes in Real Earnings Management after Firms Phased-into EDGAR

	(1) CFO	(2) Disc.	(3) Prod.	(4) Prod.	(5) Prod.
Post-EDGAR	0.000340	-0.00195	0.00979***	0.00811**	0.00829**
	(0.13)	(-0.39)	(3.02)	(2.09)	(2.06)
Interest Cov. Ratio	-0.002***	0.000972**			0.00170***
	(-8.57)	(1.99)			(3.77)
Oper. Cash Flows	0.727***	-0.165***			-0.183***
	(51.86)	(-8.26)			(-13.57)
Long-term Lev.	0.00166	-0.046**			0.0107
	(0.17)	(-2.19)			(0.73)
Sales Growth	-0.008**	0.175***			0.0138***
	(-2.03)	(22.94)			(2.75)
Size	-0.029***	0.0449***			0.0424***
	(-7.09)	(5.77)			(8.49)
Market-to-Book	-0.00018	0.00102			0.000653
	(-0.50)	(1.60)			(1.52)
ROA	-0.053***	-0.125***			-0.057***
	(-5.85)	(-7.88)			(-6.18)
Loss	-0.039***	0.0270***			0.0272***
	(-13.39)	(4.94)			(6.71)
Big 4/5 Auditor	-0.00701	0.0235***			-0.0103
	(-1.56)	(2.72)			(-1.51)
Year FE	✓	✓		✓	✓
Firm FE	✓	✓		✓	✓
Controls	✓	✓			✓
Observations	20385	20385	20385	20385	20385
Adjusted R^2	0.781	0.786	0.002	0.695	0.734

This table shows how real earnings management changes after firms became EDGAR filers compared to control firms. The coefficients are estimated from the following specification:

$$EM_{i,t} = c_i + c_t + \beta * Post\text{-}EDGAR_{i,t} + Controls_{i,t} + \varepsilon_{i,t}$$

where i indexes firm and t indexes year. Firm and year fixed effects are captured by c_i and c_t, respectively. $Post-EDGAR_{i,t}$ equals one when a firm i started filing on EDGAR in year t and thereafter. The dependent variables are proxies for real earnings management as in Roychowdhury (2006). The proxies include abnormal cash flow from operations (*CFO*), abnormal production costs (*Prod.*), and abnormal discretionary expenditure (*Disc.*). The sample includes all firms on the SEC's phase-in list that can be matched to COMPUSTAT with available information from 1991 to 1998. All standard errors are clustered at firm-level. t statistics in parentheses. * indicates statistical significance at the 10% level,** at the 5% level, and *** at the 1% level.

Table 1.6. Change in Earnings Response Coefficient after Firms Phased-into EDGAR

	CAR (3-day)					
	(1)	(2)	(3)	(4)	(5)	(6)
UE	1.467***	1.164***	1.385***	1.296***	1.439***	
	(15.45)	(9.32)	(4.56)	(4.19)	(3.93)	
$UE * Post\text{-}EDGAR$		1.345***	0.849***	0.836***	0.731**	
		(5.03)	(3.12)	(3.03)	(2.47)	
$UE(mean)$						1.492***
						(4.08)
$UE(mean) * Post\text{-}EDGAR$						0.705**
						(2.38)
Firm-level controls			✓	✓	✓	✓
UE*Firm-level controls			✓	✓	✓	✓
Fiscal Year-Quarter FE				✓	✓	✓
SIC (4-digit) FE				✓		
Firm FE					✓	✓
Observations	16114	16114	16114	16114	16114	16114
Adjusted R^2	0.027	0.030	0.034	0.039	0.062	0.063

Note: this table presents results of testing whether the marginal benefit of earnings management goes up for firms phased-into EDGAR compared to control firms. The marginal benefit to managers is captured by earnings response coefficient which is the estimated β_1 from the following specification:

$$CAR_{i,t} = c_i + c_t + \beta_1 UE_{i,t} * Post\text{-}EDGAR_{i,t} + \beta_2 UE_{i,t} + \beta_3 Post\text{-}EDGAR_{i,t} + Controls + UE_{i,t} * Controls + \varepsilon_{i,t}$$

where $CAR_{i,t}$ is 3-day market-adjusted stock return around quarterly earnings announcements dates. $UE_{i,t}$ is unexpected earnings calculated as the difference between actual quarterly EPS and median of one-quarter-ahead analysts' quarterly EPS forecasts scaled by stock price two days before the earnings announcement. $Post-EDGAR_{i,t}$ is an indicator variable which equals to one for those quarterly earnings announcements by firm i that occurred after firm i was on EDGAR. The sample period goes from 1993 (first phase-in group) to 1996 (last phase-in group) and only firm-quarter observations with non-missing full sets of controls are included. t statistics (in the parenthesis) are calculated based on robust standard errors clustered at the earnings announcement date level. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Column (1): regressing CAR on UE in the pooled sample without either controls or any fixed effects; Column (2): uni-variate specification with an interaction term between UE and $Post-EDGAR$; Column (3): with both firm-level controls and their respective interaction terms with UE_{it}; Column (4:) with fiscal year-quarter fixed effects, SIC(4-digit) industry fixed effects, and controls; Column (5): with firm fixed effects, fiscal year-quarter fixed effects, and controls; Column (6): as a robustness check using mean analyst forecast instead of median analyst forecast to calculate UE_{it}.

Table 1.7. Additional Outcome: Meet or Beat Analysts' Forecasts

	(1) Meet or Beat (1 cent)	(2) Meet or Beat (2 cents)
Post-EDGAR	0.0229	0.0209
	(1.61)	(1.48)
Year FE	✓	✓
Firm FE	✓	✓
Controls	✓	✓
Observations	16114	16114
Adjusted R^2	0.138	0.149

Note: this table presents results on whether firms are more likely to meet or beat analysts' forecasts after they become EDGAR filers. I estimate the following equation:

$$MeetorBeat_{i,t} = c_i + c_t + \beta * Post\text{-}EDGAR_{i,t} + Controls_{i,t} + \varepsilon_{i,t},$$

where i indexes firm and t indexes fiscal year-quarter. c_i is firm fixed effect and c_t is fiscal year-quarter fixed effect. $MeetorBeat_{i,t}$ is a dummy variable that equals to one if a firm i meets or beats analyst expectations of earnings per share by 1) zero or one cent; 2) by zero, one, or two cents (e.g., Burgstahler and Eames (2006), Cheong and Thomas (2018), Heater et al. (2017), etc).

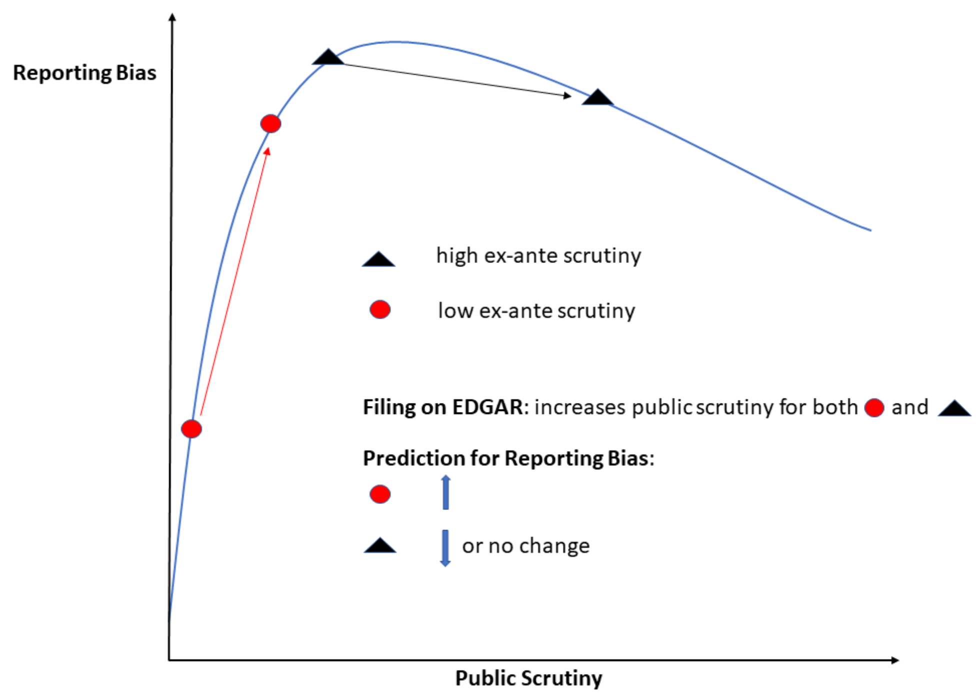


Figure 1.1. Testing the Inverse-U Relation Between Ex-ante Public Scrutiny and Reporting Bias (Samuels et al., 2020a)

Note: Figure 1.1 illustrates the inverse-U relation between ex-ante public scrutiny and reported bias predicted by Samuels et al. (2020a). For firms with low ex-ante public scrutiny (the red dot), an increase in public scrutiny will increase the equilibrium level of reporting bias for them, as illustrated by the shift of the red dot to the right. In contrast, firms with high ex-ante public scrutiny (the black triangle) are predicted to have a decrease or (an insignificant change) in reporting bias resulting from the same exogenous increase in public scrutiny.

Table 1.8. Changes in Discretionary Accruals: High versus Low ex-ante Public Scrutiny

	(1) Abs. DA	(2) Abs. DA	(3) Abs. DA	(4) Pos. DA	(5) Pos. DA	(6) Pos. DA	(7) Neg. DA	(8) Neg. DA	(9) Neg. DA
Post-EDGAR	-0.00633*	-0.00794**	-0.00630*	0.000558	0.000448	0.00101	0.00847***	0.00606**	0.00660**
	(-1.76)	(-2.20)	(-1.78)	(0.25)	(0.20)	(0.50)	(3.36)	(2.23)	(2.29)
Low-scrutiny*EDGAR	0.0163**	0.0164***	0.0154**	0.0121***	0.00999***	0.00858**	-0.00533*	-0.00473*	-0.00565*
	(2.49)	(3.16)	(2.26)	(2.82)	(2.68)	(2.41)	(-1.81)	(-1.65)	(-1.81)
Low-scr*EDGAR+Post-EDGAR	0.0100	0.00850	0.00908	0.0126***	0.0104***	0.00959***	0.00313	0.00133	0.000950
	(1.49)	(1.54)	(1.29)	(2.88)	(2.65)	(2.58)	(0.89)	(0.38)	(0.25)
Interest Cov. Ratio	-0.00110***	-0.00109***	-0.00111***	-0.00217***	-0.00272***	-0.00216***	-0.00118***	-0.00136***	-0.00194***
	(-2.95)	(-2.91)	(-2.97)	(-8.24)	(-10.88)	(-9.88)	(-4.79)	(-5.84)	(-7.33)
Oper. Cash Flows	-0.0434**	-0.0434**	-0.0430**	-0.277***	-0.202***	-0.165***	-0.235***	-0.154***	-0.188***
	(-2.43)	(-2.43)	(-2.41)	(-27.15)	(-16.30)	(-15.73)	(-17.67)	(-10.80)	(-11.85)
Long-term Lev.	-0.00252	-0.00230	-0.00235	0.0100	0.00224	0.00741	0.0176*	0.0123	0.0129
	(-0.21)	(-0.19)	(-0.20)	(1.33)	(0.32)	(1.20)	(1.85)	(1.34)	(1.29)
Sales Growth	0.0288***	0.0288***	0.0289***	-0.00407	0.00588**	0.00836***	-0.0304***	-0.0225***	-0.0207***
	(6.38)	(6.41)	(6.43)	(-1.64)	(2.42)	(3.64)	(-8.49)	(-6.59)	(-5.99)
Size	-0.00500	-0.00496	-0.00529	-0.0139***	-0.00694***	-0.00722***	-0.00899***	-0.00728**	-0.00780**
	(-1.26)	(-1.24)	(-1.33)	(-4.96)	(-2.73)	(-3.22)	(-2.70)	(-2.14)	(-2.11)
Market-to-Book	0.00126***	0.00128***	0.00127***	0.0000818	0.000148	0.000217	-0.00113***	-0.000867***	-0.000891***
	(3.12)	(3.18)	(3.15)	(0.32)	(0.64)	(1.06)	(-3.90)	(-2.81)	(-2.78)
ROA	-0.115***	-0.115***	-0.115***	0.130***	0.0840***	0.0694***	0.242***	0.196***	0.212***
	(-9.84)	(-9.84)	(-9.84)	(15.77)	(10.40)	(10.04)	(25.64)	(17.68)	(17.61)
Loss	-0.00628	-0.00628	-0.00640*	-0.0356***	-0.0497***	-0.0380***	-0.0306***	-0.0342***	-0.0465***
	(-1.64)	(-1.64)	(-1.67)	(-14.19)	(-21.48)	(-18.29)	(-11.79)	(-12.91)	(-16.28)
Big 4/5 Auditor	-0.00862	-0.00863	-0.00895	-0.00340	-0.00281	-0.00411	0.00298	0.00270	0.00246
	(-1.43)	(-1.43)	(-1.48)	(-0.80)	(-0.74)	(-1.16)	(0.79)	(0.61)	(0.52)
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	20385	20385	20385	20385	20385	20385	20385	20385	20385
Adjusted R^2	0.387	0.373	0.375	0.474	0.477	0.489	0.611	0.626	0.615

Note: this table presents how the effect of the EDGAR implementation on discretionary accruals varies across firms with high versus low ex-ante public scrutiny.

Table 1.9. Changes in Real Earnings Management: High versus Low ex-ante Public Scrutiny

	(1) CFO	(2) CFO	(3) CFO	(4) Prod.	(5) Prod.	(6) Prod.	(7) Disc.	(8) Disc.	(9) Disc.
Post-EDGAR	-0.00161	-0.00187	-0.00201	0.00567	0.00357	0.00230	0.000148	0.00798	-0.000129
	(-0.55)	(-0.64)	(-0.69)	(1.17)	(0.72)	(0.46)	(0.03)	(1.37)	(-0.02)
Low-scrutiny*EDGAR	-0.000347	0.00165	0.00285	0.0229**	0.0224***	0.0204***	0.000758	0.000847	0.00230
	(-0.06)	(0.22)	(0.38)	(2.42)	(3.01)	(3.05)	(0.06)	(0.34)	(0.17)
Low-scrutiny*EDGAR + Post-EDGAR	-0.00196	-0.000223	0.000841	0.0285***	0.0260***	0.0227***	0.000906	0.00883	0.002171
	(-0.34)	(-0.03)	(0.11)	(3.00)	(3.47)	(3.39)	(0.22)	(0.66)	(0.16)
Interest Cov. Ratio	-0.00222***	-0.00222***	-0.00223***	0.00155***	0.00157***	0.00158***	0.00111**	0.00113**	0.00132**
	(-7.33)	(-7.35)	(-7.35)	(3.11)	(3.15)	(3.15)	(2.06)	(2.09)	(2.40)
Oper. Cash Flows	0.734***	0.734***	0.735***	-0.186***	-0.186***	-0.186***	-0.152***	-0.152***	-0.162***
	(42.79)	(42.81)	(42.82)	(-11.72)	(-11.73)	(-11.75)	(-6.58)	(-6.56)	(-6.63)
Long-term Lev.	-0.00136	-0.00138	-0.00138	0.0167	0.0170	0.0179	-0.0367	-0.0364	-0.0295
	(-0.12)	(-0.12)	(-0.12)	(1.00)	(1.01)	(1.06)	(-1.52)	(-1.51)	(-1.22)
Sales Growth	-0.00455	-0.00458	-0.00459	0.0121**	0.0121**	0.0122**	0.179***	0.179***	0.178***
	(-0.99)	(-0.99)	(-1.00)	(2.00)	(2.01)	(2.01)	(18.80)	(18.77)	(18.49)
Size	-0.0241***	-0.0240***	-0.0240***	0.0395***	0.0395***	0.0391***	0.0416***	0.0413***	0.0389***
	(-5.18)	(-5.17)	(-5.16)	(7.25)	(7.24)	(7.14)	(5.04)	(5.02)	(4.66)
Market-to-Book	-0.000205	-0.000204	-0.000204	0.000933*	0.000964*	0.000975*	0.000867	0.000870	0.00113
	(-0.49)	(-0.49)	(-0.49)	(1.88)	(1.94)	(1.96)	(1.16)	(1.16)	(1.51)
ROA	-0.0476***	-0.0476***	-0.0476***	-0.0589***	-0.0587***	-0.0584***	-0.116***	-0.116***	-0.105***
	(-4.24)	(-4.24)	(-4.24)	(-5.35)	(-5.33)	(-5.31)	(-6.36)	(-6.36)	(-5.45)
Loss	-0.0349***	-0.0349***	-0.0349***	0.0244***	0.0243***	0.0243***	0.0291***	0.0293***	0.0304***
	(-9.90)	(-9.93)	(-9.93)	(5.39)	(5.39)	(5.35)	(4.76)	(4.80)	(4.86)
Big 4/5 Auditor	-0.00698	-0.00689	-0.00683	-0.00954	-0.00959	-0.00998	0.0222**	0.0221**	0.0252**
	(-1.37)	(-1.36)	(-1.34)	(-1.24)	(-1.24)	(-1.29)	(2.28)	(2.27)	(2.54)
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	20385	20385	20385	20385	20385	20385	20385	20385	20385
Adjusted R^2	0.784	0.789	0.791	0.748	0.755	0.746	0.791	0.789	0.794

Note: this table presents how the effect of the EDGAR implementation on real earnings management varies across firms with high versus low ex-ante public scrutiny.

Table 1.10. Changes in Accrual and Real Earnings Management: Sub-sample Analysis

Panel A: Changes in Discretionary Accruals

	(1) Abs. DA	(2) Abs. DA	(3) Pos. DA	(4) Pos. DA	(5) Neg. DA	(6) Neg. DA
Post-EDGAR	0.0178* (1.67)	-0.00599* (-1.84)	0.0167** (2.05)	0.000898 (0.41)	0.00251 (0.52)	0.00512** (2.25)
Year FE	✓	✓	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Observations	6485	13900	6485	13900	6485	13900
Sub-sample	low-scrutiny	high-scrutiny	low-scrutiny	high-scrutiny	low-scrutiny	high-scrutiny
Adjusted R^2	0.310	0.333	0.338	0.477	0.503	0.603

Note: this table presents the sub-samples analysis of the impact of EDGAR on discretionary accruals. Column 1,3, and 5 (2, 4, and 6) are estimated using firms with low (high) ex-ante public scrutiny only. Firms with low ex-ante public scrutiny are defined as those with no analyst coverage and with a market capitalization below the median value. All standard errors clustered at firm-level. All standard errors clustered at firm-level. t statistics in parentheses. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Panel B: Changes in Real Earnings Management

	(1) CFO	(2) CFO	(3) Prod.	(4) Prod.	(5) Disc.	(6) Disc.
Post-EDGAR	-0.00665 (-0.67)	-0.00178 (-0.62)	0.0192** (2.31)	0.00454 (1.01)	-0.00815 (-0.43)	0.00146 (0.28)
Year FE	✓	✓	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Observations	6485	13900	6485	13900	6485	13900
Sub-sample	low-scrutiny	high-scrutiny	low-scrutiny	high-scrutiny	low-scrutiny	high-scrutiny
Adjusted R^2	0.764	0.794	0.641	0.820	0.703	0.830

Note: this table presents the sub-samples analysis of the impact of EDGAR on real earnings management activities. Column 1,3, and 5 (2, 4, and 6) are estimated using firms with low (high) ex-ante public scrutiny only. Firms with low ex-ante public scrutiny are defined as those with no analyst coverage and with a market capitalization below the median value. All standard errors clustered at firm-level. t statistics in parentheses. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 1.11. Summary of Main Results

Panel A: Baseline Results

Discretionary Accruals			**Real Earnings Management**		
Abs. DA	Pos. DA	Neg. DA	CFO	Prod.	Disc.
↓	0	↑	0	↑	0

Firms shift from accrual earnings management (Abs. DA ↓) to real earnings management (Prod. ↑)

Panel B: Heterogeneous Effects: High versus Low ex-ante Public Scrutiny

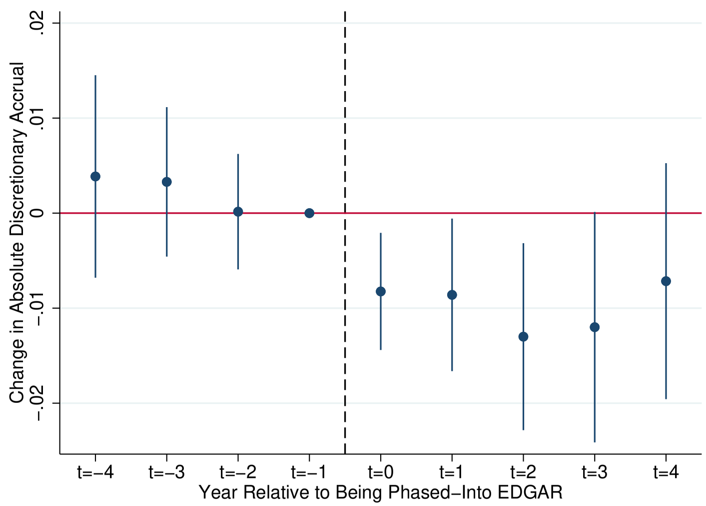
Discretionary Accruals			**Real Earnings Management**		
Abs. DA	Pos. DA	Neg. DA	CFO	Prod.	Disc.
↓	0	↑	0	0	0

High ex-ante public scrutiny firms: ↓ in overall earnings management

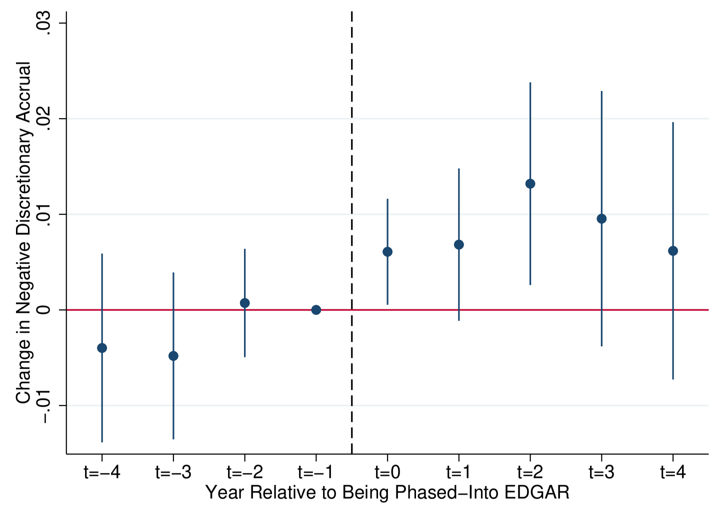
↑	↑	0	0	↑	0

Low ex-ante public scrutiny firms: ↑ in overall earnings management

Note: this table summarizes the estimated impact of being phased-into EDGAR on earnings management. Panel A shows that firms substitute away from accrual-based towards real earnings management once they start filing on EDGAR. Panel B presents results on the heterogeneous impact of EDGAR on firms with low versus high ex-ante public scrutiny. For firms with low ex-ante public scrutiny, there is a significant increase both accrual-based (both positive and absolute discretionary accruals) and real activity earnings management (abnormal production costs) after they are on EDGAR. On the other hand, firms with high ex-ante public scrutiny have a decrease in accrual-based earnings management and no change in real based earnings management.



(a) Absolute Discretionary Accrual



(b) Negative Discretionary Accrual

Figure 1.2. Dynamic Treatment Effect of EDGAR on Absolute and Negative Discretionary Accruals

Note: this figure presents the dynamic treatment effect of mandatory EDGAR filings on absolute and negative discretionary accruals in the years around when each firm's 10-K was first available on EDGAR using the following specification:

$$EM_{i,t} = c_i + c_t + \sum_{l=-4, l\neq -1}^{l=4} \beta_l * EDGAR_{i,l} + Controls_{i,t} + \varepsilon_{i,t}.$$

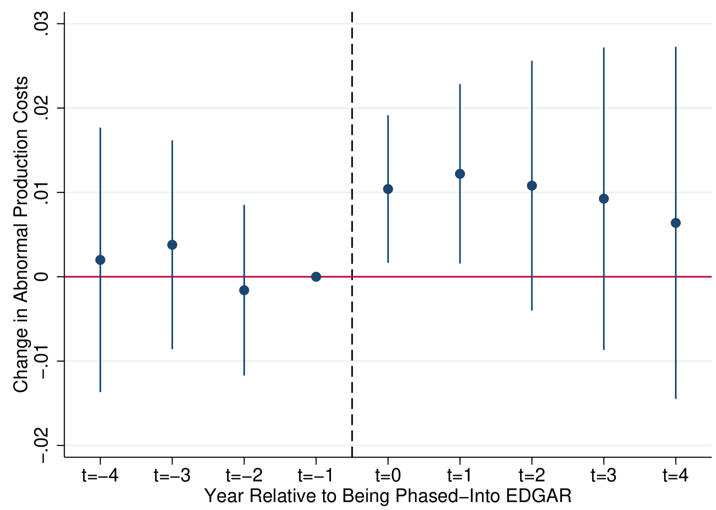
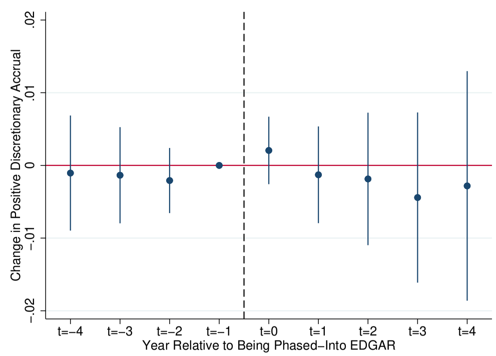


Figure 1.3. Dynamic Treatment Effect of EDGAR on Abnormal Production Costs

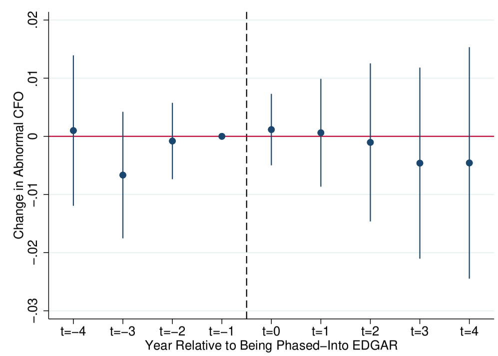
Note: this figure presents the dynamic treatment effect of EDGAR on abnormal production costs in the years around when each firm's 10-K was first available on EDGAR using the following specification:

$$EM_{i,t} = c_i + c_t + \sum_{l=-4,l\neq-1}^{l=4} \beta_l * EDGAR_{i,l} + Controls_{i,t} + \varepsilon_{i,t}$$

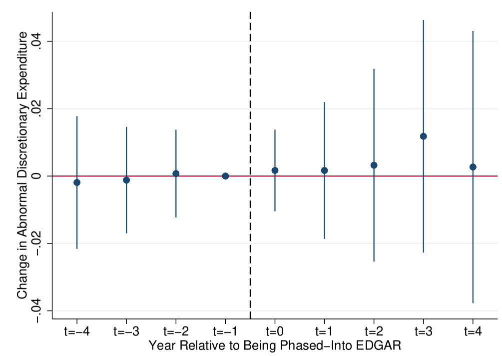
The key variables of interest are a set of 8 indicator variables $EDGAR_{i,l}$, which indicates the relative year around when each firm's 10-K was first available on EDGAR. l goes from -4 to $+4$ with the year immediately preceding each firm's phase-in year ($EDGAR_{i,l=-1}$) omitted from the regression and its coefficient set as zero. $EDGAR_{i,l=-4}$ ($EDGAR_{i,l=+4}$) corresponds to four or more years before (after) the phase-in year. The estimated coefficients of $EDGAR_{i,l}$ for abnormal production costs, and the 95% confidence intervals are displayed. The full set of control variables, as well as both firm and year fixed effects, are included. Standard errors are clustered at firm level.



(a) Positive Discretionary Accruals



(b) Abnormal Cash Flow from Operations



(c) Abnormal Discretionary Expenditure

Figure 1.4. Additional Dynamic Treatment Effect Graphs

Note: this figure presents the dynamic treatment effect of mandatory EDGAR filings on positive discretionary accruals, abnormal cash flow from operations, and abnormal discretionary expenditure in the years around when each firm's 10-K was first available on EDGAR using the following specification:

$$EM_{i,t} = c_i + c_t + \sum_{l=-4, l\neq -1}^{l=4} \beta_l * EDGAR_{i,l} + Controls_{i,t} + \varepsilon_{i,t}.$$

The key variables of interest are a set of 8 indicator variables $EDGAR_{i,l}$, which indicates the relative year around when each firm's 10-K was first available on EDGAR. l goes from -4 to $+4$ with the year immediately preceding each firm's phase-in year ($EDGAR_{i,l=-1}$) omitted from the regression and its coefficient set as 0. $EDGAR_{i,l=-4}$ ($EDGAR_{i,l=+4}$) corresponds to four or more years before (after) the phase-in year. The estimated coefficients of $EDGAR_{i,l}$, and the 95% confidence intervals are displayed in each sub-figure. The full set of control variables, as well as both firm and year fixed effects, are included. Standard errors are clustered at firm level. Across these three sub-figures, there is no significant impact of EDGAR implementation on the specified outcome variables.

Table 1.12. Additional Robustness Checks

Panel A: Falsification Test

	(1) Neg. DA	(2) Abs. DA	(3) Prod.
Pseudo Post-EDGAR	0.000377	-0.00184	-0.00112
	(0.16)	(-0.56)	(-0.26)
Year FE	✓	✓	✓
Firm FE	✓	✓	✓
Controls	✓	✓	✓
Observations	20385	20385	20385
Adjusted R^2	0.594	0.353	0.727

Note: this table presents results from a falsification test using a pseudo-event occurring two years *after* each firm's actual phase-in year.

Panel B: Excluding phase-in group one

	(1) Neg. DA	(2) Abs. DA	(3) Prod.
Post-EDGAR	0.00598***	-0.00496**	0.00852*
	(2.65)	(-2.08)	(1.96)
Year FE	✓	✓	✓
Firm FE	✓	✓	✓
Controls	✓	✓	✓
Observations	19976	19976	19976
Adjusted R^2	0.604	0.373	0.737

Note: this table presents results estimated after excluding phase-in Group one, which are companies that volunteered to file electronically on EDGAR during its pilot stage before EDGAR's official launch starting 1993. The effect of mandatory filing on EDGAR on both accrual and real earnings management continues to go through, and the magnitude of the estimated coefficients is qualitatively unchanged.

Table 1.12. Additional Robustness Checks (Continued)
Panel C: Abnormal Production Costs: Manufacturing vs Non-manufacturing Firms

	(1) Prod.	(2) Prod.	(3) Prod.
Post-EDGAR	-0.00222	0.0123***	0.000163
	(-0.34)	(2.62)	(0.03)
Manu $\times$ ***EDGAR***			0.00919**
			(2.28)
Year FE	✓	✓	✓
Firm FE	✓	✓	✓
Controls	✓	✓	✓
Observations	8114	12271	20385
Sample	Only Non-Manu	Only Manu	Manu+ Non-Manu
Adjusted R^2	0.737	0.745	0.741

Note: following Cohen et al. (2008a), manufacturing firms are defined as those with two-digit SIC code falling between 20 and 39. Results from column 1 are estimated using only non-manufacturing firms, column 2 using only manufacturing firms, column 3 using both manufacturing and non-manufacturing firms. *Post-EDGAR* measures the difference in abnormal production costs between either non-manufacturing (column 1) or manufacturing firms (column 2) and control firms. *Manu* is a dummy variable that equals to one for manufacturing firms. *Manu* $\times$ *Post-EDGAR* measures the difference in abnormal production costs between manufacturing and non-manufacturing firms.

Panel D: Controlling for Group-specific Time Trends & Industry $\times$ Year FE

	(1) Neg. DA	(2) Abs. DA	(3) Prod.
Post-EDGAR	0.00619***	-0.00517*	0.00760**
	(2.71)	(-1.83)	(2.31)
Firm FE	✓	✓	✓
Firm-level Controls	✓	✓	✓
Industry $\times$ Year FE	✓	✓	✓
Group Time Trends	✓	✓	✓
Observations	20385	20385	20385
Adjusted R^2	0.624	0.383	0.757

Note: this table presents results on the impact of EDGAR on negative, absolute discretionary accruals, and abnormal production costs estimated from specifications that include: 1) phase-in group-specific time trends which alleviate the concern that the estimated coefficients before might be capturing differential time trends across phase-in groups; 2) Industry $\times$ Year fixed effects to absorb industry-level unobservable shocks. Results show that the estimated effects of EDGAR on earnings management are not driven by different time trends across the phase-in groups or industry-specific shocks.

Table 1.12. Additional Robustness Checks (Continued)

Panel E: Alternative Measures of Discretionary Accruals

	Alternative 1		**Alternative 2**	
	(1) Neg. DA	(2) Abs. DA	(3) Neg. DA	(4) Abs. DA
Post-EDGAR	0.00606***	-0.00545**	0.00415**	-0.00463**
	(2.72)	(-2.03)	(2.49)	(-2.23)
Year FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Observations	20385	20385	20385	20385
Adjusted R^2	0.604	0.343	0.617	0.365

Note: this table reports results using two alternative measures of discretionary accruals. Columns 1 and 2 show results using the first alternative in which $\Delta REV_{i,t}/Assets_{i,t-1}$ is replaced with $(\Delta REV_{i,t} - \Delta AR_{i,t})/Assets_{i,t-1}$ in the first stage regression of estimating the modified Jones model following Cohen et al. (2008a). Please see Appendix A.1 for more details. Columns 3 and 4 present results using performance-matched discretionary accruals following Kothari et al. (2005). Each sample firm is matched with another firm that is from the same fiscal year industry and has the closest return on assets as the given firm. The performance-matched discretionary accruals are then computed as each sample firm's discretionary accruals minus the discretionary accruals of the matched firm. The estimated results using these two alternative measures of discretionary accruals remain consistent with what was reported before.

Chapter 2

Disclosure and Investor Rational Inattention

Abstract

Investors have a finite capacity to organize all information they receive from financial disclosures. Under rational inattention, we show that investor attention capacity affects the probability of disclosure. In the model, an informed firm makes a strategic voluntary disclosure subject to proprietary costs (Verrecchia 1983) or uncertainty about information endowment (Dye 1985) and investors optimally allocate their attention as a function of their conjectures about the disclosure strategy. Our main result is that the probability of disclosure is inverse U-shaped in investor attention: for low levels of attention, more attention facilitates communication and increases disclosure; for high levels of attention, more attention better identifies, and therefore deters, unfavorable voluntary disclosure. We provide preliminary empirical evidence that the relationship between investor attention and management forecast is concave, using institutional ownership as a proxy for investor attention.

2.1 Introduction

In a standard model with rational investors using all public information, economic agents use all available sources of information to make optimal decisions. Challenges to the theory have been widely documented and call for renewed interest in theories where investors cannot fully process the rich and diverse information released to the market (Blankespoor et al. 2020b). Prior research focuses on behavioral models where some agents use a mis-calibrated model when updating their beliefs. For example, inattentive investors may be unaware of certain sources of information (Hirshleifer and Teoh 2003) leading to persistent misvaluations of accounting numbers. This approach can explain a variety of observable features in the financial market (Banerjee and Kremer 2010; Barberis and Thaler 2003; Daniel et al. 2002; Hirshleifer et al. 2004).

In this study, we explore a close cousin of behavioral models, known as “rational” inattention, and examine its implication in the context of disclosure theory. As in behavioral models, investors subject to rational inattention cannot correctly process all public information;

however, in this approach, investors are cognizant of the limitation and treat attention as a capacity constraint that can be allocated efficiently, see, e.g., Sims (2003), Veldkamp (2011) and Maćkowiak et al. (2018). The main purpose of this approach is to discipline the model so that the allocation of attention endogenously responds to the qualities of the information. This is of particular interest in voluntary disclosure theory because (a) disclosures are strategic and, therefore, choices over what information to disclose respond to how investors allocate their attention, (b) in comparative statics that affect the disclosure process, investors will presumably re-adjust their attention toward signals that are more informative about fundamentals.

In the model, a firm makes a disclosure subject to disclosure costs (Verrecchia 1983) or uncertainty about information endowment (Dye 1985), with an objective to increase market prices. We deviate from the standard model by assuming that investors cannot price the firm using all the information contained in the disclosure but have a finite capacity to mentally represent information. Specifically, we use a model of rational inattention that maintains the (partitional) structure of disclosure games and such that investors can only recall a finite number of messages or memory, see Gray and Neuhoff (1998) or chapter 4 in Rubinstein (1998). Investors program how to classify disclosures or non-disclosures in this finite memory as a function of their expectations of the disclosure process. Inattention affects the non-disclosure price and the price for the marginal discloser which, in turn, affects the disclosure threshold away from the fully rational model. Our main contribution is to jointly solve for the allocation of attention and the frequency and nature of disclosures in this framework.

Inattention has two countervailing effects on incentives to disclose. First, inattentive investors respond less to public information and, therefore, weaken the link between price and disclosure which, all other things equal, will reduce voluntary disclosure. Second, inattention will increase price reaction to the lowest disclosed information (or marginal type) because inattentive investors may inaccurately classify this disclosed signal with more favorable states. This increases incentives to disclose at the marginal discloser. Combining both forces, we determine that the link between attention capacity and voluntary disclosure is inverse U-shaped.

Disclosure first increases for very low levels of attention in which inattention is an impairment to communication, and then decreases as more attention reduces price reaction to unfavorable news. In particular, for sufficiently high levels of attention, firms always disclose less when subject to more attention.

We develop supplementary theoretical results that offer novel testable implications linking proxies of attention capacity and disclosure frictions. We show that disclosure frictions affect whether disclosure increases or decreases in inattention. The model explicitly captures how attention is differentially allocated for changes in disclosure frictions. In environments where frictions are higher and most unfavorable events are unreported, attention is reallocated so that investors price firms more accurately conditional on disclosure. This implication differs from standard disclosure theory in which disclosures, when they occur, reflect the private information of the firm. In extensions, we find that inattention may reduce incentives to acquire private information and, in the multi-period model of Einhorn and Ziv (2008), attention is reallocated as a function of past disclosures. Further analyses with the normal distribution also reveal, as intuitive, that attention is more concentrated toward more likely disclosures near the mode of the distribution.

We develop a simple empirical application, which examines the relation between likelihood of management forecast and investor attention proxied by institutional ownership.[1] This application does not intend to be a complete test of the theory but offers preliminary evidence on the main prediction of our study. In univariate analyses, we sort firms into both deciles and quintiles based on institutional ownership measured immediately before management forecasts. We find that the likelihood of managers' making a forecast is increasing in the first 4 (8) quantiles (deciles) of institutional ownership, and drops in the 5^{th} (9^{th} and 10^{th}) quintile (deciles). We also estimate polynomial ordinary least squares (OLS) and logistic (Logit) models that include both linear and squared term of institutional ownership as well as industry and year fixed effects

[1] Our results are also robust to using an alternative measure of institutional ownership that adjusts for long-term strategic institutional investors who may have lower incentives to acquire and process management forecasts (Ali et al., 2008; Miao et al., 2016).

and firm-level controls. We find that the linear term of institutional ownership is significantly positive while the squared term significantly negative, which lends preliminary support to our theoretical prediction of a hump shape relation between disclosure and investor attention.

Understanding inattention is a critical, and yet not fully understood, topic in accounting research with much to be discovered as to how inattention shapes financial communication. Our results speak to defining tests over one of the three categories of attention in the framework of Blankespoor et al. (2020b). They decompose attention in three mental processes: awareness ("knowledge of the existence of a disclosure"), acquisition ("extraction of the signal from the disclosure") and integration ("mapping of the signal into firm value"). Uncertainty about information endowment (Dye 1985) is mathematically equivalent to awareness in models such as ours, because whether the firm cannot disclose because it is uninformed or discloses but its message is not received implies the same belief structure. As intuitive, lower awareness unambiguously increases strategic non-disclosure. Our main result is about acquisition, given that investors extract and simplify information from reports, possibly confounding multiple reports as a coarse message. Acquisition, we show, implies a non-monotonic link between acquisition capacity constraints and disclosure. Left for further research, our model does not capture integration because investors in our model always correctly form an expectation about value from an extracted (coarse) signal.

Financial communication has increased over time, facilitated by the free and instant access to corporate filings on the EDGAR system (Liu 2020), the dissemination by the financial press and, more recently, the implementation of machine-readable eXtensible Business Reporting Language (XBRL) in financial statements and footnotes (Blankespoor 2019). With the growth in online communication, financial communication now takes the form of an extensive documentation of conference calls (webcasts and transcripts), a wide net of unstructured disclosures (Blankespoor et al. 2014), or Google searches (Da et al. 2011). This information overload is unlikely to be met with increased investor time, creating a need to understand how inattention may pose limits on how public information is reflected into price. At a conceptual level, full development of the theory will explain that more accounting information or footnotes, on their own, does not increase market efficiency if it is not organized with the proper means of delivery and with better financial education.

In practice, we also observe that many companies which garner high levels of investor attention do not necessarily choose forthcoming levels of disclosure, even though their market leadership and quality of information systems make a proprietary cost or information endowment explanation somewhat less persuasive. Companies such as Alphabet, Facebook, Tesla, or Groupon are frequently noted in the financial press to be less than forthcoming and unpredictable in their financial communications and sometimes openly note an unwillingness to report. For example, the CEO of Tesla Elon Musk comments in a 2018 email to employees that "Being public also subjects us to the quarterly earnings cycle that puts enormous pressure on Tesla to make decisions that may be right for a given quarter, but not necessarily right for the long term." Our model provides one channel that may explain this pattern, noting that firms with a very high level of attention may disclose less.

The model also has implications about the role of regulators in facilitating access to information, for example, via the better organization of financial communications and accounting numbers (e.g., the XBRL mandate or structuring of accounting numbers in the income statement). It is generally assumed that increasing investor attention would benefit communication. We show here that a small amount of inattention starting from a fully rational market will always increase disclosure. Hence, we argue, more broadly, that increasing attention may come with a trade-off and reduce incentives by firms to disclose information voluntarily. This echoes long-standing concerns by firms to have greater control over their reporting process.

Our theoretical analysis contributes to a growing literature in accounting, discussing the role of attention in understanding financial communications (Blankespoor et al. 2020b). While linking to this literature in its entirety is far beyond our scope, we note below a few studies that closely relate to our results.

Extending the model of misreporting of Chen et al. (2007), Chen et al. (2017) develop a model of bilateral "it takes two to tango" model of costly attention, in which firms make a disclosure clarity choice and investors make an attention choice revealing the existence of manipulated numbers. As in our paper, the choice of attention by investors is a function of the communication strategy made by the firm. However, their model and focus are quite different from ours. In their model, the choice of clarity is part of a signaling game which jointly affects investors' attention and which projects are financed. They show how additional mandatory disclosure can change the outcome of the game from a separating equilibrium in which investment decisions are efficient, to a pooling equilibrium in which firms choose low clarity. In other words, our primary focus in this paper is whether more investor attention can reduce communication; their focus, by contrast, is whether more mandatory disclosure may discourage joint efforts to communicate.

While there is an extensive literature in economics and finance considering rational inattention (Gabaix, 2019; Sims, 2010; Veldkamp, 2011), this type of approach is relatively novel in accounting. Two recent studies model attention in terms of an entropy constraint, bounding the amount of information that can be transferred from public signals. Jiang and Yang (2017) consider a game in which a privately-informed but impatient firm seeks to maximize proceeds from issuing equity. In this type of model, absent an accounting system, the firm must reduce its equity to signal its type. When the information released by the accounting system is subject to entropy constraint, they show that different accounting reports must always prescribe different lower bounds akin to a conservative reporting system which identifies the lowest possible outcomes. This result emerges in their study because the signaling inefficiency increases in the distance from the lower bound.[2]

To our knowledge, the only other study specifically focusing on inattention and disclosure is by Lu (2019). His primary focus is on the effect of investor inattention on aggregation in

[2] While inattention is a special case of behavioral cognition constraint, there are other studies in the literature that focus on other types of behavioral effects which impact the response of a sender to information, e.g., ambiguity aversion (Budanova et al. 2020; Caskey 2009), disagreement (Banerjee 2011; Bloomfield and Fischer 2011), non-monetary investment preferences (Friedman and Heinle 2016) or the self-fulfilling anticipation of a price bubble (Fischer et al. 2016). Similar to rational inattention, these models can be jointly interpreted as deviations from the prediction of a traditional rational model and a behavioral assumption about how players *optimally* solve the game.

financial statements. In his model, the firm may use an aggregated signal or supplement the signal with disaggregated details, in an economy subject to strategic complementarities. He shows how additional details in this environment can lead investors subject to inattention to over-emphasize certain details that are privately, but not socially, desirable; on the other hand, removing details can aggravate coordination failures by coordinating all investors on the same simplified (but correlated) signals. A key difference between this approach and ours is that we model attention to the realization of a signal, while his model focuses on attention to particular subcomponents of the information.

While our model features truthful communication by the firm and is not a cheap talk game, our approach using a partitional (imprecise) model of investor attention draws heavily from the methods in the cheap talk literature (Crawford and Sobel 1982; Farrell and Rabin 1996; Stocken 2013). Within this literature, Fischer and Stocken (2001) show that more informed senders may decrease the receivers' information through its effect on the sender's partition. Likewise, in our model, more investor attention, which (presumably) should increase communication, may change the disclosure strategy of the sender and reduce effective communication. Other studies such as Stocken (2000), Morgan and Stocken (2003), Kumar et al. (2012), Bertomeu and Marinovic (2016) or Liang et al. (2018) provide applications of cheap talk in models of financial communication.

Lastly, our model aims to show that factors that intuitively increase communication may, in the context of a strategic game between sender and receiver, imply a (testable) non-monotonic relation between communication and disclosure and, as such, rationalize mixed empirical results. We briefly note several recent studies below that suggest an hump shape relationship between characteristics of disclosure and various frictions. Fang et al. (2017) show theoretically and empirically that the response of earnings to restatements is concave in the prevalence of restatements in an industry, if both the noise in the reporting process and the cost of manipulation are driven by a common characteristic. Samuels et al. (2020b) consider the effect of public scrutiny, admittedly a reduction to obstacles to communication, on misreporting.

Noting that scrutiny increases market response to disclosure, hence, payoffs to misreporting, they show and test that misreporting is inverse U-shaped in public scrutiny. In the context of voluntary disclosures, Kim et al. (2020) show that characteristics of the business increasing both the probability of receiving private information and the cost of publicly revealing this information can explain the non-linear relationships between disclosure and characteristics found empirically. Aghamolla et al. (2019) document evidence that the relationship between disclosure and earnings is, contrary to standard models, inverse U-shaped. They show that, in equilibrium, high-ability managers counter-signal by withholding guidance.

2.2 The Model

2.2.1 Assumptions

The model features an owner-manager ("the firm") and boundedly rational investors: investors in our model have a finite capacity to save and recall messages. The firm generates an expected cash flow $\tilde{v}$ with realizations v drawn from a probability distribution with mean μ and full support on an interval normalized to $[0,1]$, and probability density function $f(.)$. As in Dye (1985), Jung and Kwon (1988) and Beyer and Dye (2020), there is a probability $p \in (0,1]$ that the firm observes v. Then, the firm can disclose $d \in \{\text{"}ND\text{"}, \text{"}s\text{"}\}$ where "*ND*" stands for non-disclosure and "*s*" stands for truthful disclosure. When the firm does not observe the signal, it has no means to credibly convey it is uninformed and must disclose $d =$ "*ND*." As in Verrecchia (1983), disclosure involves a cost which reduces the surplus of the owner by $c \geqslant 0$. The objective of the firm is to maximize the market price $P(d)$ minus disclosure costs. For all results stated in the formal analysis, we require the existence of a friction, i.e., if $c = 0$, the probability of information endowment p must be strictly less than one.

In traditional voluntary disclosure models, investors form expectations using all information contained in the disclosure $P(d) = \mathbb{E}(\tilde{v}|d)$. That is, all the informational content of d can be processed by investors to predict v. We develop here an extension of this model in which d is observable subject to capacity constraints to classify, recall, and use information. Specifically, investors can only remember $I > 1$ different messages, where I is their capacity to process information. This representation follows what Gray and Neuhoff (1998) refer as a quantization of the information into a finite number of bits (see example below) and, for our purpose, offers a model of inattention that meshes well with discrete features of voluntary disclosure equilibria. We define investors' information as a partition $\{A_i\}_{i=1}^{I}$ of the message space $[0,1] \cup$ "ND", i.e., such that $\cup_{i=1}^{I} A_i =$ "ND" $\cup [0,1]$ and $A_i \cap A_j = \varnothing$ for any $i \neq j$. The partition corresponds to information sets in decision theory and means that investors cannot distinguish between disclosures located in the same information set A_i. Importantly, while I is an exogenous measure of investors' attention capacity, the choice of the partition will be made endogenous. As I becomes large, the ability of investors to distinguish messages converges to the traditional model with fully-rational prices.[3]

Example: Consider the following machine representation of investors' information processing. The disclosure must be classified using a finite memory capacity that must be encoded into memory bits (a number equal to 0 or 1). If investors have only one bit of capacity, they can only distinguish between two information signals, or $I = 2$. With two bits, investors can classify information as 00, 01, 10 or 11, corresponding to $I = 4$. More generally, with b bits of memory, the corresponding number of elements in the partition is $I = 2^b$; vice-versa, a value of I corresponds to a memory of $[\ln I / \ln 2]$ bits (ignoring integer constraints).

For information sets $\{A_i\}_{i=1}^{I}$, the market price forms as the best estimate of v conditional on this coarse understanding of the disclosure. Then, the market price forms based on this

[3]We represent the set of investors as a single investor subject to bounded rationality, in the sense of the firm making a take-it-or-leave-it offer to a boundedly rational investor. In practice, however, the market may feature multiple investors and, in these settings, we could think about the optimal partition for a set of investors as the intersection of individual partitions using $I = nI'$ as the set of message separated by the market as a whole if n investors can each distinguish between I' messages.

partition, i.e.,

$$P(i|D(\tilde{v}) \in A_i) \equiv \mathbb{E}[\tilde{v}|D(\tilde{v}) \in A_i], \tag{2.1}$$

where $D(\cdot)$ is the anticipated disclosure strategy as a function of v and has c.d.f. $G(.)$.

We further restrict the analysis to (intuitive) partitions in which investors' information sets preserve the ordinal ranking of cash flows.[4] Given that no-disclosure must lead to the worst prior in this type of model, we assign the no-disclosure event to the first element of the partition A_1 and denote the associated price, in short-hand, by $P(i) \equiv P(i|D(\tilde{v}) \in A_i)$. A formal definition is given below.

Definition 1. *A partition* $\{A_i\}_{i=1}^{I}$ *is monotonic if there exists an increasing sequence* $\{a_i\}_{i=1}^{I-1}$ *given* a_0 *and* a_I *such that: (a)* $A_1 = \{ND\} \cup [a_0, a_1)$*, (b) for each* $i \in [2, I]$*,* $A_i = [a_{i-1}, a_i)$*.*

Since we focus exclusively on monotonic partitions, the information set will now be represented as a sequence $\{a_i\}_{i=1}^{I-1}$. Investors are aware of the capacity constraint and choose $\{a_i\}_{i=1}^{I-1}$ in the best possible manner to make their inference correct. To capture (in reduced-form) a penalty for incorrect inferences, we assume that investors face an ex-ante quadratic loss function

$$L(D) = \underbrace{p\int_0^1 (v - P(i|D(v) \in A_i))^2 f(v)dv}_{\text{Loss when firm receives a signal}} + \underbrace{(1-p)\int_0^1 (v - P(1))^2 f(v)dv}_{\text{Loss when firm does not receive a signal}}, \tag{2.2}$$

where $f(v)$ is the probability density of the cash flow v. This preference can also be interpreted as the receiver matching the state, e.g., Chakraborty and Harbaugh (2010). We might also interpret the partition as an analyst or financial expert (receiver) obtaining the signal and mapping it into a recommendation understood by investor as a coarse message, under the assumption that the expert is evaluated more favorably when the message is more precise.

Example (cont.): Although the memory of the machine is limited to b bits, it can be programmed in advance to process information in a certain manner, leading to the encoding of

[4] In the case of uniform distributions, we show in Section 2.3 that the optimal information set is in the form of a monotonic partition. However, this may not be the case in general.

the various disclosed messages into different sequences of zeros and ones. For example, when observing no disclosure, the machine may encode it as a sequence of zeros (A_1 in the model). Note also that the machine is perfectly able to recognize the initial message it needs to encode, but its information storage capacity is bounded.

The number of different elements of the partition $I > 1$ is an exogenous parameter capturing investors' capacity constraints. Although investors are limited in their ability to process disclosures, they are entirely rational in terms of (a) understanding the limitation, (b) anticipating the equilibrium disclosure strategy, (c) making rational choices about which events they should classify more precisely.

2.2.2 Equilibrium

The timeline of the model is as follows: simultaneously, investors choose their information sets $\{a_i\}_{i=1}^{I-1}$ and the firm chooses the disclosure policy with t denoting the minimum disclosed cash flow when informed (aka, disclosure cutoff). Then the message is sent and payoffs realize.

Definition 2. *An equilibrium Γ is given by a disclosure cutoff $t \in [0,1]$, where $D(v) =$ "ND" if the firm gets no signal or $v < t$ and $D(v) = v$ if $v \geqslant t$, and an investor partition $\{a_i\}_{i=1}^{I-1}$ such that:*

1. *For any v, the firm discloses optimally given the anticipated investor attention:*

$$P(i|D(v) \in A_i) - c \cdot 1_{D(v) \neq ND} = \max\{P(1), P(i|v \in A_i) - c\}.$$

2. *Conditional on the anticipated disclosure policy $D(\cdot)$, investors set their attention optimally:*

$$\{a_i\}_{i=1}^{I-1} \in \arg\min\{p \cdot \int_0^1 (v - P(D(v) \in A_i))^2 f(v) dv + (1-p) \cdot \int_0^1 (v - P(1))^2 f(v) dv\}.$$

The notion of partitional information structure represents a natural restriction about how investors process information (Chiba and Leong 2013; Dworczak and Martini 2019; Ivanov 2010a,b; Kolotilin and Li 2019; Kolotilin and Zapechelnyuk 2019; Krishna and Morgan 2001). We focus on the most-informative equilibrium to model the maximum feasible level of communication. For simplicity, we state a definition below in terms of the equilibrium that maximizes the probability of disclosure, hereafter maximal equilibrium. It can be shown that a maximal equilibrium minimizes pricing error.[5]

Definition 3. *An equilibrium is maximal if there is no other monotonic equilibrium with a strictly lower disclosure cutoff t.*

As is common in communication games, there can be equilibria with the same beliefs and payoffs (hence, equivalent) but using different messages. In Definition 4 below, we say that two equilibria are equivalent under these circumstances and, in the rest of our analysis, do not distinguish between equilibria in the same equivalence class.

Definition 4. *Two equilibria Γ and Γ' are equivalent if*

$$\int_{d\in A_i} \mathbb{E}(\tilde{v}|D(\tilde{v})=d)dG(d) = \int_{d\in A_i} \mathbb{E}(\tilde{v}|D'(\tilde{v})=d)dG(d)$$

and, if $c>0$, $\{v : D(v) = \text{“}ND\text{”}\} = \{v : D'(v) = \text{“}ND\text{”}\}$.

The next Lemma provides an intuitive application of this definition. For any equilibrium with $t \neq a_1$, no disclosure is ever made below the disclosure threshold t and prices are constant for any disclosure below a_1. Hence, for any equilibrium with $t \neq a_1$, there exists an equivalent equilibrium with $t' = a_1' = \max(t, a_1)$, such that the upper bound of the first information set coincide. Equipped with this observation, we set the upper bound of the first element A_1 of the partition equal to the disclosure threshold, i.e., $t = a_1$, in later analyses.

Lemma 1. *For any equilibrium Γ, there exists an equivalent equilibrium Γ' such that $a_1' = t'$.*

[5] While our model does not involve cheap talk (i.e., disclosures are truthful), this property is common in many communication equilibria with partitional signals; see Fischer and Stocken (2001) for another example. Other studies such as Hart et al. (2017) and Rappoport (2020) focus on receiver-preferred equilibria.

2.2.3 No strategic withholding benchmark

We solve a benchmark in which the manager is non-strategic and always discloses when receiving information. Investors locate each element of the partition $A_i = [a_{i-1}, a_i)$ to minimize the pricing error:

$$(K_0) : \{a_i\}_{i=1}^{I-1} \in \arg\min_{\{\hat{a}_i\}_{i=0}^{I}} \{p \sum_{i=2}^{I} \int_{\hat{a}_{i-1}}^{\hat{a}_i} (v - \mathbb{E}[\tilde{v}|\hat{a}_{i-1} \leqslant \tilde{v} < \hat{a}_i])^2 f(v)dv$$
$$+ (1-p)\int_{\hat{a}_1}^{1} (v - P(1))^2 f(v)dv + \int_{\hat{a}_0}^{\hat{a}_1} (v - P(1))^2 f(v)dv\},$$
$$\text{s.t. } P(1) = \frac{pF(\hat{a}_1)\mathbb{E}(\tilde{v}|\hat{a}_0 \leqslant \tilde{v} < \hat{a}_1) + (1-p)\mathbb{E}(\tilde{v})}{pF(\hat{a}_1) + (1-p)}, \hat{a}_0 = 0, \hat{a}_I = 1. \tag{2.3}$$

Lemma 2. *A solution $\{a_i^\dagger\}$ to program (K_0) satisfies*

$$a_i^\dagger = \frac{\mathbb{E}[\tilde{v}|a_i^\dagger \leqslant \tilde{v} < a_{i+1}^\dagger] + \mathbb{E}[\tilde{v}|a_{i-1}^\dagger \leqslant \tilde{v} < a_i^\dagger]}{2} \tag{2.4}$$

for $i = 2, \ldots, I-1$.

The cutoffs chosen for $a_i^\dagger$ $(i = 2, \ldots, I-1)$ can be reinterpreted as equalizing pricing errors in any two contiguous elements of the partition at each side of a_i, that is:[6]

$$-(\mathbb{E}[\tilde{v}|a_{i-1}^\dagger \leqslant \tilde{v} < a_i^\dagger] - a_i^\dagger)^2 = -(\mathbb{E}[\tilde{v}|a_i^\dagger \leqslant \tilde{v} < a_{i+1}^\dagger] - a_i^\dagger)^2.$$

For the first cutoff $a_1^\dagger$, the conditional expectation is slightly different because the message $A_1 = [a_0^\dagger, a_1^\dagger]$ may also be the result of not receiving information. Adapting equation (2.4), the

[6] This characterization draws an interesting analogy to Equation (1) in Morgan and Stocken (2003) in which a sender cares about a weighted average of price and accuracy. The accuracy component in their model implies, expectedly, a very similar condition which equates the errors across information sets. The price incentive implies that the pricing error must be increasing in price while, by contrast, the pricing error is constant in our model and the number of elements in the partition is exogenously specified in terms of the degree of bounded rationality.

first cutoff $a_1^\dagger$ is given by[7]

$$\frac{\partial \Delta_1}{\partial a_1}|_{a_1=a_1^\dagger} = 0, \tag{2.5}$$

where

$$g\Delta_1 = p\int_{a_1}^{a_2^\dagger}(v-\mathbb{E}[\tilde{v}|a_1 \leqslant \tilde{v} < a_2^\dagger])^2 f(v)dv + (1-p)\int_{a_1}^{1}(v-P(1))^2 f(v)dv$$
$$+\int_0^{a_1}(v-P(1))^2 f(v)dv \tag{2.6}$$

includes the terms in (K_0) that depend on a_1. Naturally, when the firm always receives information $p=1$, equation (2.5) simplifies to equation (2.4) evaluated at $I=1$, that is,

$$a_i^\dagger = \frac{\mathbb{E}[\tilde{v}|a_1^\dagger \leqslant \tilde{v} < a_2^\dagger] + \mathbb{E}[\tilde{v}|0 \leqslant \tilde{v} < a_1^\dagger]}{2}.$$

The following technical assumption guarantees that this solution to program (K_0) is unique, which is similar to the "Monotonicity" condition in Crawford and Sobel (1982).

Assumption 1 (Monotonicity I)**.** *For $I \geqslant 1$, if two sequences $a \equiv \{a_i\}_{i=0}^I$ and $a' \equiv \{a_i'\}_{i=0}^I$ satisfy Equations (2.4) and (2.5) with $a_{I-1} < a_{I-1}' < a_I = a_I'$, then $a_i < a_i'$ for all $0 \leqslant i \leqslant I-1$.*[8]

2.2.4 Discussion

The model of information classification is designed to reflect investors' inability to process all relevant information. While this model has an intuitive interpretation in terms of reducing the message space, it is also technically convenient in the special context of disclosure theory: disclosure equilibria partition the state space into a disclosure and a non-disclosure region.

[7] We maintain in the benchmark the assumption that uninformed firms must be classified in A_1 because the main role of this preliminary is to help state the solution to the problem with strategic withholding. Naturally, investors could do even better by classifying non-disclosures with disclosures near the unconditional mean; however, this type of solution would not be feasible with strategic withholding because the non-disclosure message must always generate the lowest posterior.

[8] Consider the solution to (2.3) when $\tilde{v}$ is Uniform and rearrange terms $a_{i+1}^\dagger = 2a_i^\dagger - a_{i-1}^\dagger$ implying that Monotonicity I is satisfied for the case of the Uniform distribution.

Hence, bounded rationality works to coarsen the state space further but otherwise maintains the partitional structure of the communication game. Below, we discuss some of the key assumptions in this setting.

(i) A different approach is to model attention capacity in terms of a maximal reduction in entropy (Sims, 2003). This criterion may alter the nature of the game because, with a bound in (differential) entropy, investors will never be able to rule out any state with certainty regardless of a disclosure or non-disclosure - thus, implying a setting perceived by investors as noisy disclosure and no longer has a partitional nature, see, e.g., Van Nieuwerburgh and Veldkamp (2010), Jiang and Yang (2017) or Lu (2019). To our knowledge, the properties of voluntary disclosure games when investors have entropy constraints have not yet been worked out but present interesting research opportunities in this area. Sims (2003) also discusses finite codes as a foundation for entropy (p. 668-669), noting that entropy can be recovered as the information recovered from a finite code observed over a continuous time. This formulation suggests that a finite code may represent a single disclosure event, while entropy may reflect the information collected over a given time horizon composed of many disclosure events.

(ii) In the baseline model, we use I as a measure of the collective ability of investors to distinguish messages: for example, in the form of the intersection of all partitions chosen by each individual investor as it would be efficient for investors to choose non-overlapping information sets. Other interpretations may feature institutional aspects of information providers in which the message is discrete. For example, financial auditors issue an unqualified, qualified or adverse opinion; rating agencies rate debt issues on a scale; stock analysts issue a stock recommendation within a scale. This is also true for quality certifications outside of financial reporting (Dranove and Jin 2010): restaurants, hospitals and movies may receive qualitative grades. This type of coarse partition may be desirable if small investors or consumers have limited ability to process more complex (or continuous) message spaces.

(iii) We present the analysis in terms of investor-driven capacity constraints but a different model may involve manager-driven capacity constraints if, say, the manager can only use I separate messages when disclosing to the market. If the firm has some means to pre-commit itself to a level of disclosure (Aghamolla et al. 2019; Heinle and Verrecchia 2016; Suijs and Wielhouwer 2019), the firm will be better off committing to complete inattention to reduce disclosure costs. However, if the firm cannot credibly commit to attention, it can be readily verified that the maximal equilibrium in the manager attention model will coincide with the baseline investor attention model. Hence, I can also be thought of as the maximum feasible attention by investors and the firm.

2.3 Uniform Payoffs

We lay out the intuitions in the context of $\tilde{v}$ being uniformly distributed and the only friction is a non-zero cost $c > 0$ of disclosure. As we will show next, this specification captures the main trade-offs of the model in closed-form. Another interesting property of the uniform model is that it can be formally shown that the information sets formed by investors must be a monotone partition (thus demonstrating that monotonic partitions do not seem pathological in simple settings), as we claim next.

Proposition 1. *When $\tilde{v}$ is uniformly distributed, all equilibrium information structures induce monotone partitions on the state space.*

The intuition for Proposition 1 is that the investors can always reduce the pricing error by modifying a non-monotone partition. Two prior studies, by Bergemann et al. (2012) and Kos (2012), prove this property using the single-crossing properties of cheap talk with an upper bound on the number of possible messages. Information sets have this form in our model but for different reasons: there is no single-crossing property and disclosures are verifiable; instead, the interval structure are selected because they minimize the pricing error of an uninformed receiver.

Next, we derive equilibrium in this game. Absent strategic withholding, investors optimally separate the state space in intervals of equal length, so that

$$a_i^\dagger = i/I. \tag{2.7}$$

We need to verify if this (ideal) information structure is feasible when the firm can strategically withhold. Specifically, for this to be sustainable in an equilibrium, the firm must report $v \geqslant a_1^\dagger$, that is,

$$\underbrace{\mathbb{E}(\tilde{v}|\tilde{v} \leqslant a_1^\dagger)}_{=a_1^\dagger/2} \leqslant \underbrace{\mathbb{E}(\tilde{v}|a_1^\dagger \leqslant \tilde{v} \leqslant a_2^\dagger)}_{=(a_1^\dagger+a_2^\dagger)/2} - c.$$

Reinjecting the values of $a_i^\dagger$ from (2.7), this partition is feasible as long as incentives to strategically withhold are not too high, that is, if $c \leqslant 1/I$. Intuitively, when the friction is small, the pooling over low strategic types in the non-disclosure region A_1 required by the voluntary disclosure game is less than the pooling directly caused by investor inattention.

Suppose next that $c > 1/I$. Then, the disclosure threshold $t = a_1$ must be set strictly higher than $a_1^\dagger$. The optimal information structure for investors having $I-1$ messages to learn about the remaining state space $[t,1]$ is to set, likewise to (2.7), $I-1$ intervals of equal size on $[t,1]$, i.e., for any $i \geqslant 2$,

$$a_i = t + (1-t)\frac{i-1}{I-1}. \tag{2.8}$$

The maximal equilibrium will prescribe setting t as low as possible, which should involve making the firm exactly indifferent between withholding and disclosing when $v = t$, that is

$$\underbrace{\mathbb{E}(\tilde{v}|\tilde{v} \leqslant t)}_{=t/2} = \underbrace{\mathbb{E}(\tilde{v}|t \leqslant \tilde{v} \leqslant a_2)}_{=(t+a_2)/2} - c,$$

which simplifies to $a_2 = 2c$. Note that a threshold $2c$ would be the disclosure threshold in a fully rational model, but since the threshold here is $t = a_1 < a_2 = 2c$, we know that, in this case, there is more disclosure with rational inattention: put differently, inattentive investors induce the firm to disclose information that would have been withheld if investors had infinite attention capacity.

Using Equation (2.8) to recover a_2 and solving for t readily yields the following equilibrium.

Proposition 2. *The maximal equilibrium is given as follows:*

(i) If $c \leqslant 1/I$, $t = 1/I$ and $a_i = i/I$.

(ii) If $c > 1/I$, $a_1 = t = \frac{2c(I-1)-1}{I-2}$ and, for $i > 1$, $a_i = \frac{2c(I-i)+i-2}{I-2}$.

This equilibrium has two core properties illustrating how inattention affects the characteristics of the voluntary disclosure equilibrium.

First, in classic disclosure models, investors are fully attentive to all disclosures and, therefore, the uncertainty that may remain after a disclosure event is not affected by strategic behavior. In the inattention model, by contrast, a higher disclosure cost implies a higher threshold t. This, in turn, implies that disclosures above t receive more attention and lead to more accurate prices. Put differently, as attention is optimally allocated, investors trade off more inaccuracy due to non-disclosure with more *accurate* pricing conditional on disclosure. As in standard disclosure models, the withholding region is (weakly) the least precise but the inattention model predicts an inverse relationship between the frequency of disclosure and the degree of attention to each disclosure.

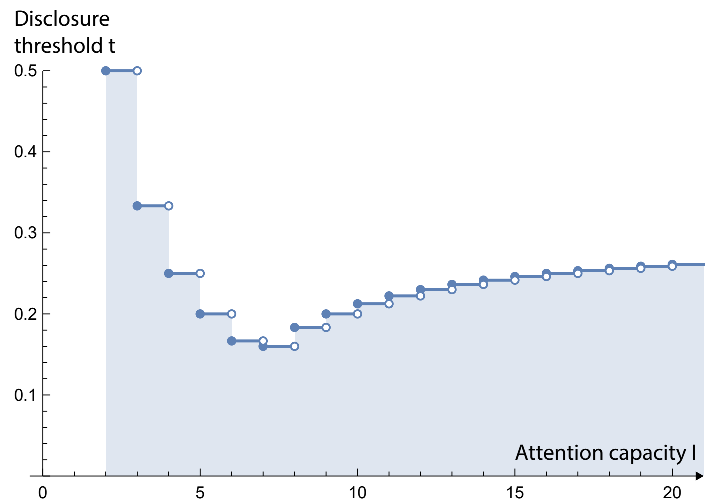


Figure 2.1. Disclosure threshold and attention capacity ($c = .15$)

Second, the equilibrium has a central comparative static that ties how the degree of inattention affects the probability of disclosure. We illustrate the trade-off in Figure 2.1 by varying the degree of inattention. At the maximal level of inattention ($I = 2$), $t = 1/2$ means that only above-average outcomes are disclosed. As the degree of attention increases (up to $I = 1/c$), the cutoff t decreases: intuitively, the partition of the message space becomes more precise as the market becomes more attentive. We refer to this first part of the trade-off as the "informativeness effect" of attention. As the degree of attention increases further (from $I = 1/c$ onward), the cutoff point t *increases*. The intuition for this region is better obtained by considering a decrease in inattention: when investors are inattentive, they classify incorrectly the marginal discloser as a better firm $[a_1, a_2)$, leading to more incentives to disclose; we refer to this second part of the trade-off as the "marginal discloser effect" of attention. The disclosure threshold, i.e., the probability of non-disclosure, is plotted as a function of the degree of attention in Figure 2.1.

To explain the non-monotonicity further, Figure 2.2 illustrates the change in cutoffs when $I = 2, 3, 4$ for $c = 1/3$. Up to $I = 3$, investors are implementing their ideal message space with three signals (i.e., subdividing the message space in three equal intervals). The voluntary disclosure problem does not affect the determination of the cutoff t and, as a result, the informativeness effect must dominate as the precision of all intervals increases. Starting at $I = 4$, however, the voluntary disclosure equilibrium prescribes $t > 1/I$ and there is a loss in precision due to strategic withholding. Then, the marginal discloser effect dominates as incentives to disclose decrease with more attention. In summary, the relationship between attention and the probability of disclosure $1-t$ is inverse U-shaped, with the maximal probability of disclosure achieved at $I = [1/c]$ or $I = [1/c]+1$.

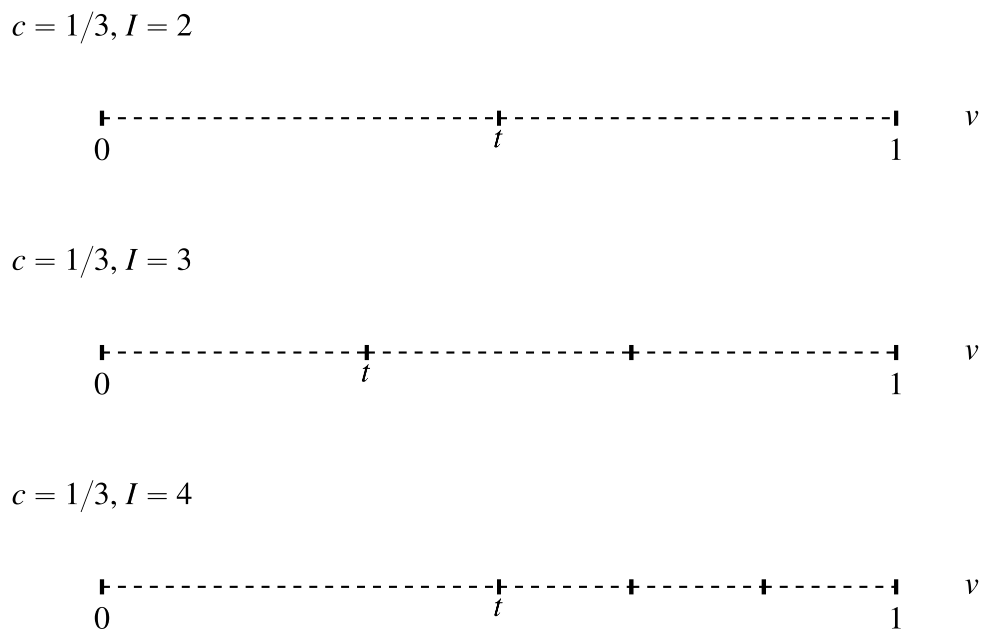


Figure 2.2. Disclosure cutoffs

2.4 General Analysis

2.4.1 Preliminaries

We prove next these results in the general model, lifting the assumption that $\tilde{v}$ is uniform and allowing for a non-zero probability $1-p>0$ of not receiving information. The next statement formally demonstrates that the maximal equilibrium minimizes the pricing error.

Proposition 3. *The equilibrium with the lowest disclosure cutoff gives investors the highest expected payoff over all equilibria that induce interval partitions.*

Given that investors make an equilibrium conjecture about the threshold t, we can rewrite the objective function of investors as a program $K(t)$ which consists of choosing all but the first element of the partition to minimize pricing error over *disclosed* values:

$$K(t) : \{a_i\}_{i=2}^{I-1} \in \arg \min_{\{\hat{a}_i\}_{i=2}^{I-1}} \sum_{i=2}^{I} \int_{\hat{a}_{i-1}}^{\hat{a}_i} (v - \hat{P}(i))^2 f(v) dv, \tag{2.9}$$

$$\text{s.t. } \hat{a}_1 = t,\ \hat{a}_I = 1, \text{ and } \hat{P}(i) = \mathbb{E}[\tilde{v}|\hat{a}_{i-1} \leqslant \tilde{v} < \hat{a}_i].$$

Lemma 3 characterizes an optimal choice a_i in the above program. The proof is identical to Lemma 2, hence omitted, except that the information sets are optimized starting at the disclosure threshold t and over $I-1$ intervals.

Lemma 3. *For any $i \geqslant 2$,*

$$a_i = \frac{\mathbb{E}[\tilde{v}|a_i \leqslant \tilde{v} < a_{i+1}] + \mathbb{E}[\tilde{v}|a_{i-1} \leqslant \tilde{v} < a_i]}{2}. \tag{2.10}$$

2.4.2 Main Analysis

We are now equipped to characterize a solution of the model. We proceed in two simple steps that closely follow the argument in the uniform model.

First, in what follows next, we show that the maximal equilibrium coincides with the benchmark partition $\{a_i^\dagger\}$ if $A_1^\dagger = [a_0^\dagger, a_1^\dagger) \cup \{ND\}$ can be sustained as the withholding region. To verify this, it must be verified that values in the next information set $v \in A_2^\dagger = [a_1^\dagger, a_2^\dagger)$ would not be strategically withheld. That is,

$$\mathbb{E}(\tilde{v}|\tilde{v} \in (a_1^\dagger, a_2^\dagger)) - c \geqslant \frac{pF(a_1^\dagger)\mathbb{E}(\tilde{v}|\tilde{v} \leqslant a_1^\dagger) + (1-p)\mathbb{E}(\tilde{v})}{pF(a_1^\dagger) + (1-p)}, \tag{2.11}$$

where the right-hand side is the non-disclosure price in Jung and Kwon (1988).

When condition (2.11) is satisfied, which means that the disclosure threshold with fully rational investors is lesser or equal than $a_1^\dagger$, investors will respond to any $t < a_1^\dagger$ by increasing the cutoff of the first information set to $a_1^\dagger$. The firm will, of course, respond by increasing the disclosure threshold to $t = a_1^\dagger$, implying a maximal equilibrium that coincides with the model without disclosure frictions and similar to Proposition 2 (i).

Second, suppose that Equation (2.11) is not satisfied, in which case the partition preferred by the investor is too fine and would encourage the manager to withhold some $v > a_1^\dagger$. Recall then that the maximal equilibrium is the equilibrium with the smallest disclosure cutoff $t = a_1$, which involves a choice of $a_1 = t > a_1^\dagger$ binding the withholding constraint:

$$\mathbb{E}(\tilde{v}|\tilde{v} \in [a_1, a_2)) - c = \frac{pF(a_1)\mathbb{E}(\tilde{v}|\tilde{v} \leqslant a_1) + (1-p)\mathbb{E}(\tilde{v})}{pF(a_1) + (1-p)}, \tag{2.12}$$

corresponding to Proposition 2 (ii) where the withholding region A_1 is driven by the binding incentive constraint on the cutoff. The next theorem summarizes these observations and is the main result of our study.

Theorem 1. *Let $\overline{p}$ be the probability cutoff such that (2.11) is met at equality.*

(i) If Equation (2.11) is satisfied (i.e., $p \geqslant \overline{p}$), $\{a_i^\dagger\}$ is the maximal equilibrium and the manager discloses when informed with $v \geqslant a_1^\dagger$;

(ii) Otherwise, the equilibrium disclosure cutoff t is strictly greater than $a_1^\dagger$ and the maximal equilibrium $\{a_i\}$ satisfies equations (2.10) and (2.12).

To summarize Theorem 1, investors will try to set their ideal information sets $\{a_i^\dagger\}$. But this is only feasible if there are limited incentives to withhold $A_2^\dagger$ - which, in turn, requires the market to be sufficiently skeptical after a non-disclosure to decrease their beliefs when observing $A_1^\dagger$. As is well-known in this type of model, this can only occur if firms are expected to be informed, i.e., likely to be strategically withholding, when reporting in $A_1^\dagger$. When firms are likely to be uninformed, this equilibrium is no longer sustainable because firms with $v \in A_2^\dagger$ will be better-off withholding (and pretend to be uninformed). Then, the disclosure cutoff must increase to $t = a_1 > a_1^\dagger$ to satisfy the indifference condition of the marginal discloser (2.12). The remaining cutoffs $\{a_i\}$ for $i \geqslant 2$ are then set according to (2.10) to minimize pricing errors over the disclosure region $[a_1, 1]$.

For reasons similar to Crawford and Sobel (1982) and the assumed monotonicity condition in (1), it is possible for the necessary conditions in (2.10) and (2.12) to have multiple solutions because while these are second-order sequences subject to two boundary points, $a_0 = 0$ and $a_I = 1$, the equilibrium equations are non-linear. While these seem to describe pathological cases, we formally show below that the Monotonicity condition can be adapted to the current setting so that these conditions are necessary and sufficient.

Assumption 2 (Monotonicity II)**.** *Given $c > 0$ and $0 < p < 1$, for $I \geqslant 1$, if two sequences $a \equiv \{a_i\}_{i=0}^I$ and $a' \equiv \{a_i'\}_{i=0}^I$ satisfy Equations (2.10) and (2.12) with equality such that $a_{I-1} < a'_{I-1} < a_I = a'_I$, then $a_i < a'_i$ for all $0 \leqslant i \leqslant I-1$.*

The monotonicity assumption guarantees that, as for the case with low disclosure frictions, a solution exists and is unique.

Corollary 1. *Suppose that Assumption 2 holds. Then, if Equation (2.11) is not satisfied (i.e., $p < \overline{p}$), the maximal equilibrium is uniquely given by the solution to (2.10) and (2.12).*

The main results will continue to hold with an arbitrary, e.g., unbounded, support as long as we adjust the boundary conditions a_0 and a_I when solving for (2.10)-(2.12) in Theorem 1. Below in Figure 2.3, we plot a numerical example with the normal distribution, comparing the optimal partition under complete information (dashed) versus under strategic withholding (solid). The equilibrium features a large strategic withholding region followed by compressed information sets in the disclosure region. Note also that, in the example of the normal distribution, investors set more precise information sets near the median distribution over events that have greater likelihood.

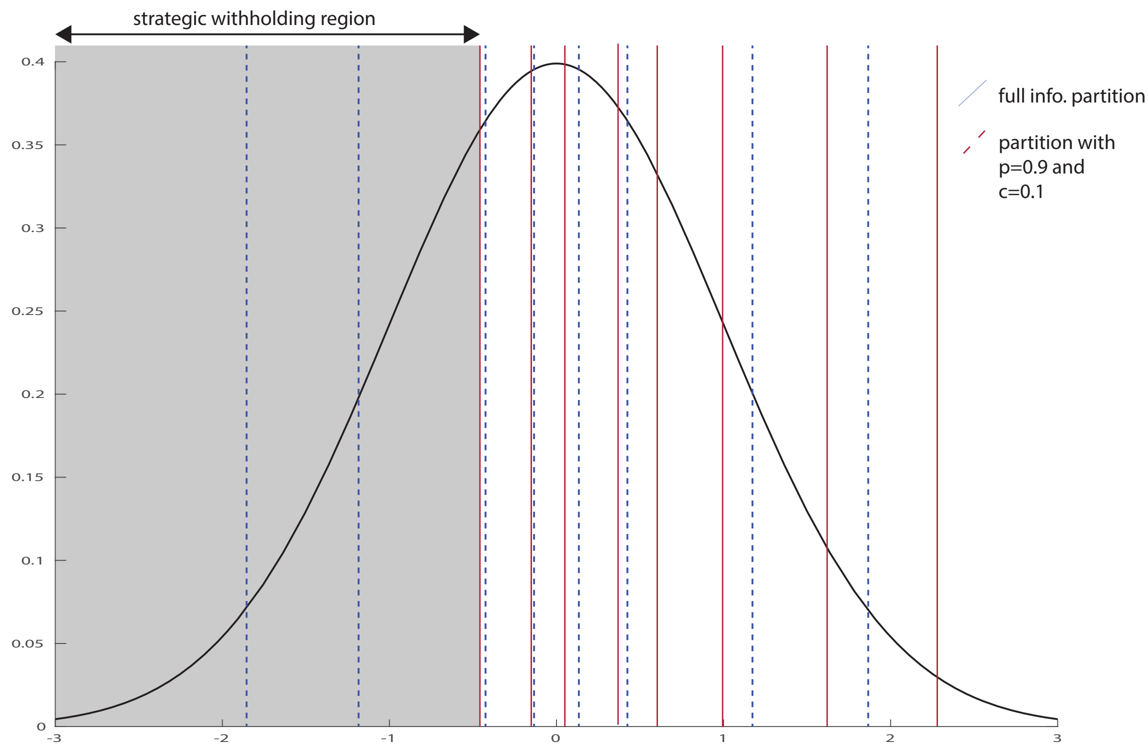


Figure 2.3. Rational Inattention: Normal Distribution

2.4.3 Properties of the equilibrium

We discuss next the key properties of the model, generalizing the observations made in the case of the uniform distribution. To begin with, we demonstrate a few results that establish several standard insights of classic voluntary disclosure models - shown to be preserved with minor adjustments for any degree of inattention.

Proposition 4. *The voluntary disclosure cutoff t increases in the disclosure cost c and decreases in the probability of being informed p.*

When strategic withholding constrains investor learning, the probability of strategic withholding is affected by the friction in a manner similar to the traditional models - even though the disclosure cutoff need not be set at the same location. Interestingly, note that the standard disclosure model would always predict that non-disclosure implies the least precise beliefs. In fact, Dye and Hughes (2018) show a stronger result: in the context of normal distributions, the residual variance of $\tilde{v}$ conditional on non-disclosure is equal to the ex-ante variance $Var(\tilde{v})$. Neither statements need to hold with rational inattention and it may be the case that uncertainty is higher conditional on disclosure than conditional on non-disclosure. The choice of cutoffs $\{a_i\}$ in (2.10) equates the pricing error for the *marginal type* located at $v = a_i$, this needs not hold for the average type in an information set and, therefore, when comparing $Var(\tilde{v}|A_1)$ to $Var(\tilde{v}|A_i)$ with $i \geqslant 2$.

We turn next to new insights unique to the inattention setting.

Proposition 5. *The expected pricing error is increasing in the disclosure cost c and decreasing in the probability of being informed p. As an example, in the special case of uniform cash flows $\tilde{v}$,*

(i) The pricing error conditional on disclosure is decreasing in the disclosure cost c and increasing in the probability of being informed p;

(ii) For sufficiently large cost, the expected pricing error is first strictly decreasing and then strictly increasing in attention capacity I. The pricing error conditional on disclosure is strictly decreasing in attention capacity I for I sufficiently large, i.e., when Inequality (2.11) does not hold.

In Proposition 5, we show how inattention reallocates investors' information sets between disclosure and non-disclosure regions. Apart from the results that are true for any general distributions, there are some interesting comparative statics that hold under uniform distributions. The pricing error conditional on disclosure decreases when a disclosure friction increases, which illustrates the trade-off between less precise non-disclosure and more precise disclosures. Investors who cannot observe well strategically withheld low events pay more attention to fewer disclosed news: in other words, inattention creates an inherent trade-off between frequency of disclosure and (perceived) quality of disclosure.

The next proposition summarizes the key main result from our analysis and demonstrates how disclosure varies as a function of inattention.

Proposition 6. *The disclosure cutoff t is first strictly decreasing and then strictly increasing in the partition size I.*

Proposition 6 generalizes the observations made in the context of the uniform distributions (where all elements of the investor partition *conditional on disclosure* are of equal size) to the case of general distributions. The effect of attention on the probability of disclosure is a simple inverse U-shaped relationship with, first, the probability of disclosure being increasing for low levels of attention and, then, decreasing for high levels of attention. The probability of disclosure is maximal at an interior level of investor attention binding Equation (2.11) and is the point at which the non-disclosure region corresponds exactly to how unconstrained investors would have chosen the lowest element of the partition.

Below, we state an additional result in the context of $c = 0$, i.e., with only uncertainty about information endowment. In this type of model with fully rational investors, Acharya et al. (2011) and Guttman et al. (2014) demonstrate that the equilibrium satisfies the "minimum" principle, that is, minimizes the price conditional on any possible disclosure cutoff. We show below that the minimum principle may be upset in the presence of rational inattention.

Proposition 7. *Suppose $c = 0$. There exists at most a single level of attention I_m such that the minimum principle and, subject to $I \in \mathbb{N}$ being an integer, is not generic, i.e., the set of parameters $p \in (0, 1)$ such that the minimum principle holds has zero mass. If the cutoff t when $I = 2$ is lower than the cutoff in the rational model (Jung and Kwon 1988), the non-disclosure belief is always strictly higher under rational inattention for any $I \geqslant 2$.*

To explain this result, note that the minimum principle is a generalization of the unravelling principle (Milgrom 1981) in the presence of a disclosure friction. The principle relies on the ability of an informed firm to separate (by disclosing) which, in turn, causes skepticism in beliefs following non-disclosure. Reformulated, the minimum principle, just like the unravelling principle, states that any equilibrium features the maximal rationalizable skepticism. Rational inattention works as a constraint on the ability of informed firms to separate, thus reducing the equilibrium skepticism. Counter-intuitively, the higher non-disclosure belief under this constraint implies that strategically withholding firms achieve a higher surplus than under the rational model. In particular, Proposition 7 implies that investor inattention benefits (on average) strategic firms at the expense of firms that were truly uninformed.

2.5 Discussions

2.5.1 Information Acquisition

In a seminal study, Shavell (1994) shows that voluntary disclosure induces excessive information acquisition because informed firms have discretion to strategically disclose. Rational inattention can interact with this effect: as firms cannot "freely" strategically disclose because investors are not allocating enough attention, incentives for excess information acquisition may be muted. We discuss this idea formally below.

As in Shavell (1994), assume that, ex-ante, information has social value (otherwise, any reduction in information acquisition is socially beneficial). Let v be a productivity signal, with density $f(.)$, and let x be an investment. The firm's market value is then given by $vx - \psi(x)$, where $\psi(0) = \psi'(0) = 0$ and $\psi''(\cdot) > 0$. Let $x^*(v)$ be the optimal investment at v. As benchmark, we restate Proposition 5 of Shavell (1994) below.

Lemma 4. *The value of information to firms V exceeds the social value of information V^*.*

We show next that this problem can be mitigated if investors have limited attention.

Proposition 8. *Suppose that there is only acquisition cost and no disclosure cost, i.e., $c = 0$. For **any** finite information capacity, the value of information to firms is less than the full-information case.*

The proof is provided in the appendix that utilizes the minimum principle (Acharya et al., 2011) and the properties of the equilibrium. Proposition 8 demonstrates that inattention reduces incentives to acquire information. This does not mean, however, that inattention necessarily increases social welfare. When information is useful to determine the optimal level of investment, it is socially desirable for investors to understand more information *given that it has been obtained by the firm*. Hence, there is a trade-off between information precision and acquisition cost. If investors are able to allocate more attention to firms' disclosure, the quality of information potentially increases and more informed decisions could be made, which is socially beneficial.

2.5.2 Dynamics

Rational inattention also has multi-period implications. Consider a two-period simplified version of the model by Einhorn and Ziv (2008). The cash flows realized at the end of each period are independent across periods and publicly observed. At the beginning of each period, the firm potentially receives a (noisy) signal s that is equal to the cash flow v in the current period with probability $q(v) > 0$ and a pure error with probability $1 - q(v) > 0$.[9] Assume that the firm cannot distinguish between signals about the cash flow and errors. Let G be the probability distribution of the error that is independent of the cash flow v. The probability that the firm receives a signal in period t is $p_t \in (0,1)$ $(t = 1,2)$, where $p_1 = \lambda$, $p_2 = \lambda_0$ if the firm does not have a signal in period 1; $p_2 = \lambda_1 > \lambda_0$ if the firm has a signal in period 1. The investors' attention capacities are I_1 and I_2, respectively, in periods 1 and 2.

Proposition 9. *Let A_1^1 be the first element of the investor's information set in period 1. Let a_1^2 be the first cutoff that the investor selects in period two. Let t_1 be the disclosure threshold in period one. The cutoff a_1^2 will be lower if the investor does not observe A_1^1 in period one or if the realized cash flow in period one falls below the disclosure threshold t_1 (when the observation about the signal is A_1^1).*

At the end of period 1, investors will update their beliefs about the second-period signal

[9] The probability of signal being informative could potentially depend on the actual state. Our result hold for any function $q(\cdot)$ as long as $q(v) > 0$ for all v.

endowment of the firm to

$$p_2 = \begin{cases} \lambda_1 & \text{if investors does not observe } A_1^1 \text{ in period 1} \\ \phi(v_1) = \frac{1-\lambda}{M}\lambda_0 + (1 - \frac{1-\lambda}{M})\lambda_1 & \text{otherwise,} \end{cases}$$

where $M = (1-\lambda) + \lambda \Pr(s_1 < t_1 | v_1)$. For any information set different from A_1^1, it is certain that the firm received a signal in this period conditional on the disclosure. In turn, this implies that the firm is more likely to be informed again in period 2, and investors will pay more attention to lower disclosures and choose a lower cutoff a_1^2 in period 2 (than if the first-period observation is A_1^1) by Proposition 4.

Similarly, if the realized cash flow v_1 in period 1 is higher than t_1, the probability $\Pr(s_1 < t_1 | v_1)$ that s_1 is lower than the equilibrium threshold t_1 (and the firm then withdraws the low signal) will be lower, because a signal lower than t_1 can only be generated by error. So when forming the belief p_2, there is less weight assigned to the case where the firm conceals a low signal in period 1 (given the observation A_1^1). Then $\phi(v_1)$ will be smaller because $\lambda_1 > \lambda_0$. Hence the investors will pay less attention to lower disclosures and choose a higher a_1^2 in period 2. In summary, if information endowments are correlated, attention will be serially correlated as well and vary over time as a function of disclosures and realized signals.

2.6 Empirical Application

2.6.1 Sample Selection

Our main theoretical prediction is that firm disclosure has an inverse-U shaped relation with investor attention. We develop preliminary evidence about this prediction using management forecast as a proxy for firm disclosure. Management forecasts are voluntary disclosures and managers face substantial uncertainty in making forecasts about future earnings. Moreover, management forecasts are released as part of earnings conference calls which are highly publicized and and discussed by the financial press. Management forecasts generally garner more significant price reactions than most other types of firm disclosures (Beyer et al. 2010).

We present the definitions and sources of our main variables in Table 1 and sample selection procedures in Table 2. We start with all annual earnings announcements made by the U.S. firms for fiscal years ending between January 1^{st}, 2004 and December 31^{st}, 2016 obtained from the Institutional Brokers' Estimate System (I/B/E/S) earnings announcement database.[10] We construct a sample of earnings per share (EPS).[11] Our sample starts with 67,239 firm-year observations and 10,945 unique firms. We only keep observations with non-missing current and prior year earnings announcement dates, which are used to construct a time window for management forecasts.

We merge earnings announcements with management forecasts which are acquired from the I/B/E/S management forecast guidance (CIG) database using I/B/E/S unique tickers and forecast period end dates. We can match 70,198 management forecasts to the I/B/E/S earnings announcement sample. As in Bertomeu et al. (2019), we further require all forecasts to be made after the prior year's earnings announcement date but at least six months before the current period end date, which shrinks the number of annual management forecasts to 28,787. The majority of management forecasts is bundled with earnings announcements and takes place between 10 to 11 months before the current fiscal year end. Since our theory is silent on how frequently managers forecast within a period, we only retain the earliest management forecast for periods with more than one forecast.[12]

[10] Our sample starts from 2004 due to two significant regulatory changes in the U.S. in 2000 and 2002, which have fundamentally changed both managers' incentives to disclose information and process of collecting management forecasts. Since August 2000, Regulation Fair Disclosure (Reg FD) has shut down most private communications between managers and financial analysts. Consequently, Reg FD have increased the frequency of public managerial forecasts. In addition, since July 2002, the Sarbanes-Oxley Act (SOX) has dramatically increased internal controls and management responsibilities. From a data-collection perspective, SOX also requires conference calls to be recorded in transcript form, which allows for much more convenient identification of management forecasts.

[11] Earnings in I/B/E/S are reported as pro-forma earnings calculated under the same accounting principles for both analysts' and managers' forecasts (Bertomeu et al. 2019). We choose to use raw EPS since it is the actual nominal variable being forecasted by managers and analysts and has been kept within a similar range across firms (Cheong and Thomas, 2011). EPS, that have been adjusted for stock splits, are more problematic since its magnitude tends to decline as firms split their shares.

[12] Note that management forecasts in I/B/E/S have already been adjusted for the number of shares. We construct the raw earnings forecasts by multiplying forecasts in I/B/E/S with I/B/E/S adjustment factor, which is recovered using the ratio of raw earnings to adjusted earnings in the I/B/E/S earnings database.

We obtain information on stock prices from CRSP, accounting fundamentals from Compustat, and institutional ownership from Thomson Reuters. After the merge, our sample shrinks to 50,703 firm-year observations, 7,864 unique firms, and 11,451 management forecasts. Lastly, we drop firms that either always or never make a forecast since these firms probably have committed to a fixed forecast policy for reasons out of the scope of our model. Our final sample has a total of 16,508 firm-year observation, 2,583 unique firms, and 7,392 management forecasts. As shown in Table B.3, 45% of all firm-years have management forecasts in our sample. A median firm-year in our sample has institutional ownership of 75%, leverage ratio of 53%, market capitalization of 1.01 billion U.S. dollars, and book-to-market ratio of 53%. 15% of all firm-years report negative earnings, and 68% have an increase in EPS.

2.6.2 Proxies for Investor Attention

A critical empirical challenge in testing our theory is to construct a plausible proxy for investors' attention. The proxy should capture investors' aggregate capacity constraints to extract managers' forecasts. Institutional investors hold more than 70% of the common shares of NYSE/NASDAQ/AMEX stocks as of 2012 (Kempf et al., 2017).[13] Moreover, institutional investors generally have both better skills and stronger incentives than retail investors to proactively acquire and process management forecasts. Hence, our firm-level proxy for investors' capacity should be increasing with the amount of influence institutional investors have on managers. Secondly, since the number of messages (I) investors can remember is set before managers' disclosure in our model, our empirical proxy for investors' capacity should be measured *prior to* managers' forecasts.

With these considerations, we use the percentage of institutional ownership measured immediately before management forecasts as a firm-level proxy for investor attention. Higher institutional ownership correspond to higher capacity by investors to acquire and process man-

[13] Institutional investors interact and communicate with firm managers their demands of disclosure. In contrast, retail investors' demands are much more opaque to managers (Basu et al., 2020).

agers' voluntary disclosure. We will refer to this measure as $Capacity\ (percent)$ for the rest of the paper.

Institutions that hold significant stakes ($> 5\%$) usually have strategic considerations and are less likely to acquire and trade on management forecasts. Following Ali et al. (2008) and Miao et al. (2016), our second measure $Capacity\ (ratio)$ refines the first measure by adjusting for long-term institutional ownership:

$$Capacity\ (ratio) = \frac{Ins - Ins(LT)}{1 - Ins(LT)},$$

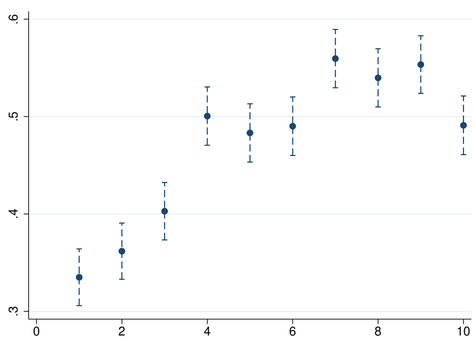
where Ins is the percentage of institutional ownership and $Ins(LT)$ is the percentage shareholdings by institutions that own more than 5% of shares.

Admittedly, our empirical proxies might be related to how likely managers make a forecast for reasons other than what we conjecture in the model. In other words, $Capacity\ (percent)$ and $Capacity\ (ratio)$ might affect management forecasts through channels other than investor attention. For example, institutional investors may demand more voluntary disclosure to balance their portfolio or combine public disclosures with their own private information (Cheynel and Levine 2020). If we fail to find an inverse U-shaped relation between $Capacity$ and management forecast in the data, it could either be: 1) the channel through investor attention predicted by our theory does not exist; 2) the other channels add sufficiently substantial noise into $Capacity$ such that our tests do not have enough power to detect our theoretical channel. Our empirical proxies may capture other determinants of management forecasts, which could be biased against us finding an inverse U-shaped relation between $Capacity$ and management forecasts in the data. However, other determinants of management forecasts have no reason to produce the inverse U-shaped relation on their own.

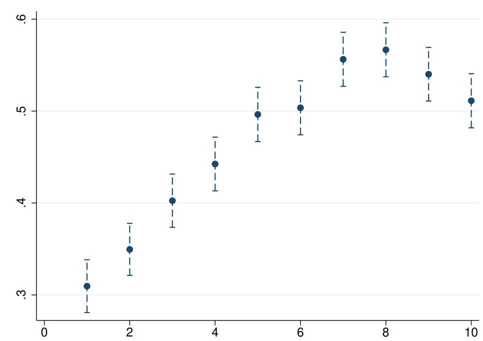
2.6.3 Empirical Analysis

Investor Attention and Management Forecast

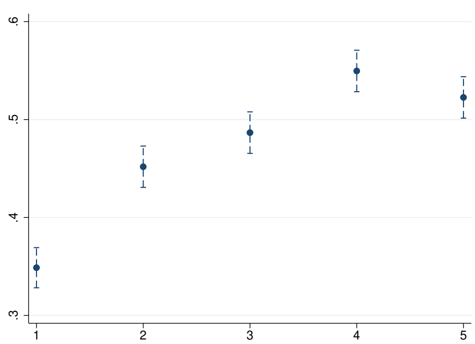
We begin by graphically presenting the relation between investor attention and managers' likelihood of making a forecast. In Figure 2.4, we sort firms into either ten deciles or five quintiles based on $Capacity\ (percent)$ and $Capacity\ (ratio)$ in year $t-1$. Within each decile or quintile, we calculate and report the average probability of managers making a forecast in year t. The 95% confidence intervals are plotted around the mean values for each decile and quintile. Consistent with an inverse U-shape relation predicted by the theory, we find that the likelihood of making a forecast increases in the first 4 (8) quintiles (deciles), and then declines in the 5^{th} (9^{th} and 10^{th}) quintile (deciles).



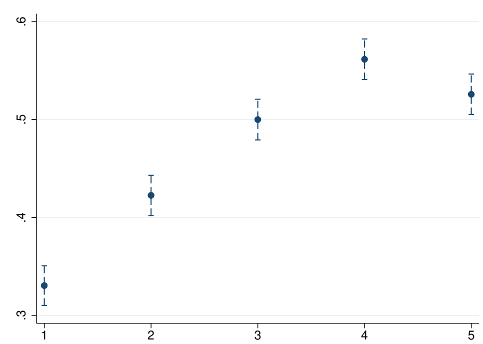
(a) Institutional Ownership ($Capacity\ (percent)$)



(b) Institutional Ownership ($Capacity\ (ratio)$)



(c) Institutional Ownership ($Capacity\ (percent)$)



(d) Institutional Ownership ($Capacity\ (ratio)$)

Figure 2.4. Likelihood of Management Forecast Across Deciles and Quintiles of Institutional Ownership

Note: this figure plots percentage of firms with management forecasts in year t across deciles (sub-figures a and b) and quintiles (sub-figures c and d) of investor attention which is proxied by either $Capacity\ (percent)$ (sub-figures a and c) or $Capacity\ (ratio)$ (sub-figures b and d) measured in year $t-1$. All sub-figures plot the 95% confidence interval around the mean values for each decile or quintile.

We conduct next additional tests to lend further support to the theoretical prediction.

First, we estimate a polynomial regression model which includes both proxies for investors' bounded capacity ($Capacity\ (percent)$ and $Capacity\ (ratio)$) and their respective squared terms to test if investor bounded capacity has a hump-shaped relation with management forecast. The hump-shaped relation will be supported if: 1) the estimated coefficients of $Capacity\ (percent)$ and $Capacity\ (ratio)$ are significantly positive; 2) the estimated coefficients of the squared terms of $Capacity\ (percent)$ and $Capacity\ (ratio)$ are significantly negative. Our polynomial regression model is specified as follow:

$$MF_{i,t} = \alpha_t + \alpha_j + \beta Capacity_{i,t-1} + \gamma Capacity^2_{i,t-1} + Controls_{i,t-1} + \varepsilon_{i,t}, \tag{2.13}$$

where α_t is year fixed effect and α_j 4-digit SIC industry fixed effect. The dependent variable $MF_{i,t}$ equals to one if a firm i makes a forecast on future earnings in year t. The variables of interest are $Capacity_{i,t}$ and $Capacity^2_{i,t-1}$. All independent variables are lagged one period relative to management forecasts. Standard errors are clustered by firm to account for potential transitory shocks that are correlated across time for a specific firm. In addition, to capture firm-level variables that can influence manager's decision to forecast, we control for firm size with *Size*, growth opportunities with $Book\ to\ Market$, leverage with $Leverage\ Ratio$, whether a firm reports negative earnings with *Loss*, whether a firm has an increase in earnings per share with $EPS\ increase$, the absolute value of the change in earnings per share with $Abs.\ EPS\ Change$.[14]

The polynomial regression results are presented in Table B.4. Panel A reports results from estimating our polynomial model with OLS and Panel B results from a Logit regression. For both panels, we show results from the same set of six different specifications. Columns 1 and 2 on both panels estimate a univariate regression. Columns 3 and 4 include both year and industry fixed effects, which control for common time trends and persistent differences across industries, respectively. Lastly, columns 5 and 6 further control for firm-level characteristics. The estimated coefficients of control variables are generally consistent with prior literature on

[14] Please see Table B.1 for more details on variable construction.

management forecasts.[15] The positive estimated coefficients of *EPS Increase*, *Leverage Ratio*, and *Size* suggest that well-performing, highly-levered, and large firms are more likely to issue management forecasts. Besides, the negative estimated coefficients of *Loss*, *Abs EPS Change*, and *Book to Market* imply that firms with poor and volatile financial performances and with fewer growth opportunities are less likely to issue management forecasts.

Consistent with our predicted inverse U-shaped relation, we find that the estimated coefficients of $Capacity_{i,t-1}$ are significantly positive and the coefficients of the squared term - $Capacity^2_{i,t-1}$ are significantly negative across all of our six different specifications. These formal statistical tests, along with patterns in the raw data shown in Figure 2.4, add to the credibility of our primary theoretical prediction.

Investor Attention and Management Forecast using an alternative research design

In addition to the polynomial regressions above, we adopt an alternative research design to lend further support to our theoretical prediction. More precisely, we estimate a spline regression that treats the relation between the likelihood of management forecast and investor attention as piecewise linear. In other words, we estimate a separate slope for each side of a threshold τ of investor attention as follows:

$$MF_{i,t} = \alpha_t + \alpha_j + \beta_1(Capacity_{i,t-1} - \tau < 0) + \beta_2(Capacity_{i,t-1} - \tau \geqslant 0) + Controls_{i,t-1} + \varepsilon_{i,t}.$$

If our theoretical prediction holds, we expect to see that the slope between likelihood of management forecast and institutional ownership to be significantly positive (negative) if institutional ownership is below (above) the threshold τ (i.e., $\beta_1 > 0$ and $\beta_2 < 0$).

However, the major challenge of estimating a spline regression is that we need first to specify the threshold τ, which our theory is silent on. We approach this challenge in two ways. Firstly, by eyeballing Figure 2.4, we conjecture that the threshold is around the 80^{th} percentile of both $Capacity(percent)$ and $Capacity(ratio)$ because the probability of management forecasts

[15] For example: Cheng et al. 2013, Goodman et al. 2014, Li and Yang 2016, Tsang et al. 2019, Guan et al. 2020, Basu et al. (2020), and Abramova et al. (2020), etc.

declines in the 5^{th} (9^{th} and 10^{th}) quintile (deciles). For robustness, we set $\tau = 70^{th}, 75^{th}, 80^{th}, 85^{th}$ percentile of both $Capacity(percent)$ and $Capacity(ratio)$.

Our results estimated from the spline regression are consistent with our inverse U-shaped relation prediction. Table B.5 Panel A reports results using $Capacity(ratio)$ across four pre-specified values of τ and Panel B reports results using $Capacity(percent)$. Across four different thresholds τ and two proxies for investor attention ($Capacity(ratio)$ and $Capacity(percent)$, we find a statistically significant positive slope between management forecast and investor attention for values of investor attention that are below the thresholds τ (i.e.,$Capacity - \tau < 0$). In addition, the slope between management forecast and investor attention for values of investor attention that are above the thresholds τ is significantly negative in all specifications (i.e.,$Capacity - \tau \geqslant 0$).

Our second approach employs the multivariate adaptive regression spline (MARS) method, which simultaneously determines the optimal threshold τ^* and the sign of the slope on either side of τ^*. This statistical method is developed by Friedman (1991) and has been recently applied by Samuels et al. (2020b) to test an inverse-U relation predicted by their model. The primary advantage of MARS is that it does not require a pre-specified threshold τ by researchers. Instead, MARS searches for the optimal threshold τ^*, which minimizes the mean-squared errors of our spline regression model.

Again, our empirical results estimated from the MARS method are consistent with our theoretical prediction and are reported in Table B.5 Panel C. We report results using $Capacity(ratio)$ as a proxy in column 1 and $Capacity(percent)$ in column 2. First, the optimal threshold τ^* that minimizes the mean squared errors of our spline regression model corresponds to 79^{th} percentile of $Capacity(ratio)$ and 81^{th} percentile of $Capacity(percent)$. The optimal threshold τ^* matches and confirms our conjectured τ at around 80^{th} percentile from patterns in the raw data. Second, similar to our results using pre-specified thresholds τ, the slope for values of investor attention that are below (above) the estimated optimal threshold τ^* is significantly positive (negative). In other words, for values of investor attention that are below either the 79^{th} percentile of $Capacity(ratio)$ and 81^{th} percentile of $Capacity(percent)$, an increase in investor attention is associated with a higher likelihood of management forecast. In contrast, for values of investor attention that are above the estimated optimal threshold τ^*, investor attention is negatively associated with managers' likelihood of making a forecast.

Three Types of Institutional Investor Attention and Management Forecast

Our empirical tests above provide robust evidence that the likelihood of management forecast has an inverse U-shaped relation with institutional investor attention. To paint a more granular picture of the roles played by different types of institutional investors, we follow Bushee and Noe (2000) to classify institutional investors into one the three categories: quasi-indexers, transient investors, and dedicated investors.[16]

Similar to our graphical analysis above, we start by plotting the probability of management forecasts across ten deciles of each of the three types of institutional investor ownership. Sub-figure a of figure B.1 sorts firms by quasi-indexers' ownership, sub-figure b by transient investors, and sub-figure c by dedicated investors. The probability of management forecast is positively associated with all three types of institutional ownership for low levels of institutional ownership. In particular, the likelihood of management forecasts is monotonically increasing in the bottom eight deciles of quasi-indexers' ownership. Our result on quasi-indexers is consistent with the finding in the literature that quasi-indexers have a strong preference for management forecasts and firms cater to quasi-indexers' demands.[17] In addition, we document three novel associations. First, the probability of management forecast declines in the 9^{th} and 10^{th} deciles of ownership by quasi-indexers, suggesting that sufficiently high levels of quasi-indexers' owner-

[16] Bushee (1998) and Bushee and Noe (2000) use principal component analysis to construct factors that capture institutional investors' average size of stake in their portfolio firms and degree of portfolio turnover. Similar institutional investors are grouped together into one of the three clusters: dedicated, quasi-indexers, and transient investors. Dedicated investors generally have significant stakes in a small number of firms and hold their stakes for a long period of time. Quasi-indexers consist of passive index funds and active funds that have a diverse portfolio of companies, trade infrequently, and closely benchmark against indexes. Lastly, transient investors trade frequently on a select of firms, and they use short-run strategies (Basu et al., 2020).

[17] Relevant papers in the literature include: Boone and White (2015), Bird and Karolyi (2016), Schoenfeld (2017), Basu et al. (2020), and Abramova et al. (2020), etc). Quasi-indexers generally hold a well-diversified portfolio and hence, face enormous costs in collecting private information on their portfolio firms. In addition, quasi-indexers' tracking strategies limit their ability to trade on private information. Consequently, quasi-indexers demand higher firm transparency with more public disclosures, which reduces the information asymmetry between them and their portfolio firms and lowers the costs of monitoring portfolio firms.

ship reduces managers' incentives to forecast. Second, the probability of management forecast does not respond to changes in transient investors' ownership in the top eight deciles. Third, the probability of management forecast increases in the bottom five deciles of dedicated investors' ownership and declines thereafter, which is a clear inverse U-shaped relation consistent with our theoretical prediction.

Lastly, we provide formal statistical tests on the relation between management forecast and different types of institutional investor ownership. We re-estimate equation (2.13) by replacing $Capacity_{i,t-1}$ with each of the three types of institutional investor ownership at $t-1$. Table B.6 presents the results from our regressions with the full-set of firm controls as well as year and industry fixed effects. The main takeaway is that: while all three types of institutional investor ownership are positively associated with management forecasts, the inverse U-shaped relation is primarily driven by dedicated investor ownership (i.e., the estimated coefficient of $Dedicated^2$ is statistically significant at 1% level and with the highest magnitude).[18]

2.7 Conclusion

Inattention is a complex behavioral constraint that can, in its application to capital markets, restrict how much information is incorporated into price. In this study, we examine how investor inattention affects strategic withholding in a standard model of voluntary disclosure. Inattention is jointly determined with disclosure choices. On the one hand, inattention alters how prices respond to disclosure and can either increase or decrease incentives to withhold. On the other hand, investors allocate their attention as a function of their expectations in the disclosure game. Our primary result is that disclosure first increases and then decreases in investors' attention capacity. We also show how the informativeness of disclosures as perceived by market participants changes as a function of attention capacity and market frictions.

[18]Consistent with the apparent inverse U-shaped relation in the raw data, the estimated coefficient of the linear term *Dedicated* is not statistically significant. This insignificant result is in line with Abramova et al. (2020) that attention by non-passive investors (i.e., investors other than quasi-indexers) does not have a significant impact on whether firms make a forecast.

Our model presents only first steps into the role of inattention, when reading through the lens of an otherwise generic disclosure theory. This presents advantages and disadvantages. The advantage is that the general properties of these models are well-understood with perfect attention. Hence, we can easily observe the incremental effect of inattention in a manner that extends existing insights. A disadvantage is that our model only intends to develop one applied setting of inattention, but disclosure models, on their own, do not aim to represent all forms of communication, in particular regulated and audited financial reports.

Having noted these, many questions are left open for future research in manners that would, likely, not require a model of voluntary disclosure. As an example, further research may consider the role of enforcement and its effect on investor attention. In particular, whether enforcement may allocate attention away from the manipulative activities (Schantl and Wagenhofer 2020). We also do not know the interactions between mandatory and voluntary disclosure (Einhorn 2005) in the context of finite attention capacity. Finally, while our primary purpose has been to present the theory and offer some tentative empirical facts, more empirical tests are required to validate the theory. Inattention, even "rational" inattention, violates semi-strong market efficiency in that not all public information is reflected into price (Fama 1970). There is still disagreement between proponents of the efficient market hypothesis and behavioral finance as to whether such violations is significant enough, especially given that new technologies have increased how to organize massive datasets, while simultaneously allowing for broader use of statistics and machine learning to summarize information.

Acknowledgements

This chapter, in part, has been submitted for publication of the material. Bertomeu, Jeremy. Hu, Peicong (Keri). Liu, Yibin. The dissertation author was one of the primary investigators and authors of this paper.

We thank seminar participants at Financial Accounting and Reporting Section (2021), Hawai'i Accounting Research Conference (2021), Carnegie Mellon University Accounting, Rutgers Accounting, UC Berkeley Accounting, UC San Diego Accounting, the Doctoral Program in Accounting, Reporting, and Taxation (DART) and the Swiss Doctoral Program Network in Accounting Research (DAR) Accounting Theory Seminar, UC San Diego Economics. The authors thank Iván Marinovic, Joel Sobel, and Jonathan Bonham (discussant) for their advice and suggestions.

Chapter 3

The Spillover Effect of Earnings Management: Evidence from China's Stock Market De-listing Policy

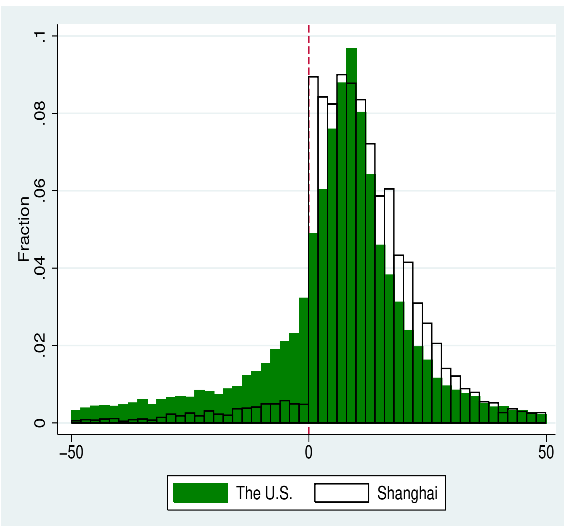
Abstract

We propose and test the spillover effect of earnings management by a set of firms on market reaction to other similar firms. We first document that China's de-listing policy separates public firms into high and low information segments based on their reported earnings. A large proportion of firms in the low information segment are suspects of earnings management, which has a spillover effect on all the other firms in the same segment. We show that investors can not identify which firm has managed earnings in the low information segment. Hence, investors distrust and react less to earnings announcements by all firms in the low information segment. More specifically, firms in the low information segment suffer from lower stock market investors' reaction, lower cumulative abnormal return around earnings announcements, insignificant earnings response coefficient, lower trading liquidity, higher systematic risk, and higher stock price synchronicity. Lastly, we support our proposed spillover effect with causal evidence by studying China's public firms that exogenously shift from high to low information segment due to the U.S. 2007-08 financial crisis.

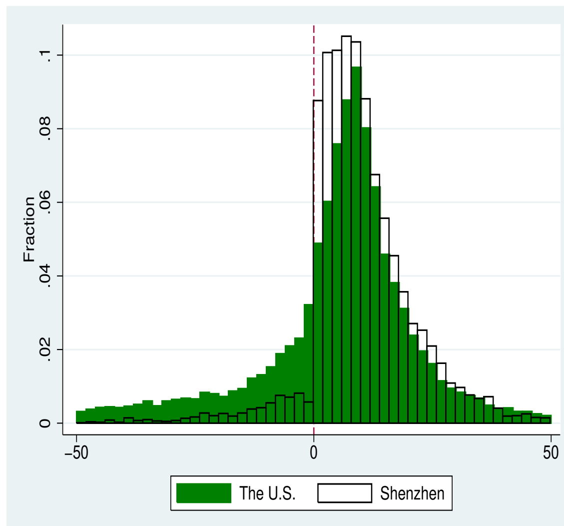
3.1 Introduction

The financial and real consequences of earnings management are central to accounting research. An emerging literature focuses on the effect of earnings management on **peer** firms. Beatty et al. (2013) and Li (2015) show that financial misreporting by prominent firms leads to sub-optimal investments on capital investment, R&D, and advertising by peer firms. Instead of studying peer firms' real decisions, we propose and test a novel inter-firm spillover effect of earnings management in the stock market: does earnings management by manipulating firms distort investors' reaction to financial reports by (other similar) non-manipulating firms?

We take advantage of the unique de-listing policy in China's stock market to answer our research question. China's Securities Regulatory Committee (CSRC, counterpart of the SEC in the U.S.) set the rule in 1998 that public firms would be de-listed if they consecutively reported negative annual earnings. This earnings-based de-listing policy was designed to protect investors from risks imposed by under-performing firms. However, a crucial unintended consequence of China's de-listing policy is that it incentivizes firms to engage in massive earnings management to stay listed when they expect to report negative earnings. Consequently, there is an abnormally large amount of firms in China that report a small and positive earnings compared to firms in the U.S. as shown in Figure 3.1 below. Chinese investors are well aware of both the de-listing policy and what firms have been doing.



(a) Shanghai v.s. the U.S.



(b) Shenzhen v.s. the U.S.

Figure 3.1. Firm ROE Distribution: China v.s. the U.S. (2009-2016)

Note: we pool together all the listed firms in China and the U.S. from 2009 to 2016 and plot their respective ROE distribution. The x-axis is ROE from -50% to 50% and the y-axis is the fraction of firms falling into each 2% ROE bin. China has two major stock exchanges. Figure 3.1a plots all the firms listed in the Shanghai Stock Exchange whereas Figure 3.1b all the firms in the Shenzhen Stock Exchange.

In our research, we define all firms with an annual ROE from 0 to 4% as low information segment since many firms in this segment are suspects of earnings management. Correspondingly, firms with ROE in (7%,+∞) are categorized as high information segment since these firms do not face an imminent pressure of delisting (as shown in figure 3.2). We provide evidence on why we divide China's stock market this way[1]. We also show that firms in the low information segment indeed have higher earnings management than other firms.

[1]As for firms with ROE in (4%, 7%), investors are much less certain whether they have managed their earnings or not. In our analysis, we leave firms with ROE from (4%, 7%) out and directly compare firms in high and low information segments. Our results are robust to changing the ROE threshold from 4% to 3% or 5%, and from 7% to 6% or 8%.

Our major contribution is to show that manipulating firms in the low information segment (ROE from 0 to 4%) have a spillover effect on non-manipulating firms in the same segment. The low information segment consists of a tremendous amount of manipulating firms as evident from the ROE distribution and more importantly non-manipulating firms whose true ROE is from (0,4%). There are a substantial amount of public firms that would have a true ROE from 0 to 4% in both China and the U.S.. According to the statistics compiled by Aswath Damodaran at NYU Stern, the U.S. firms in the industries such as education, advertising, insurance, and green & renewable energy on average report an ROE below 4%. Our own calculations show that Chinese firms from industries such as healthcare, education, entertainment, and technology service have an industry-average ROE below 4% in year 2000-2016. [2]

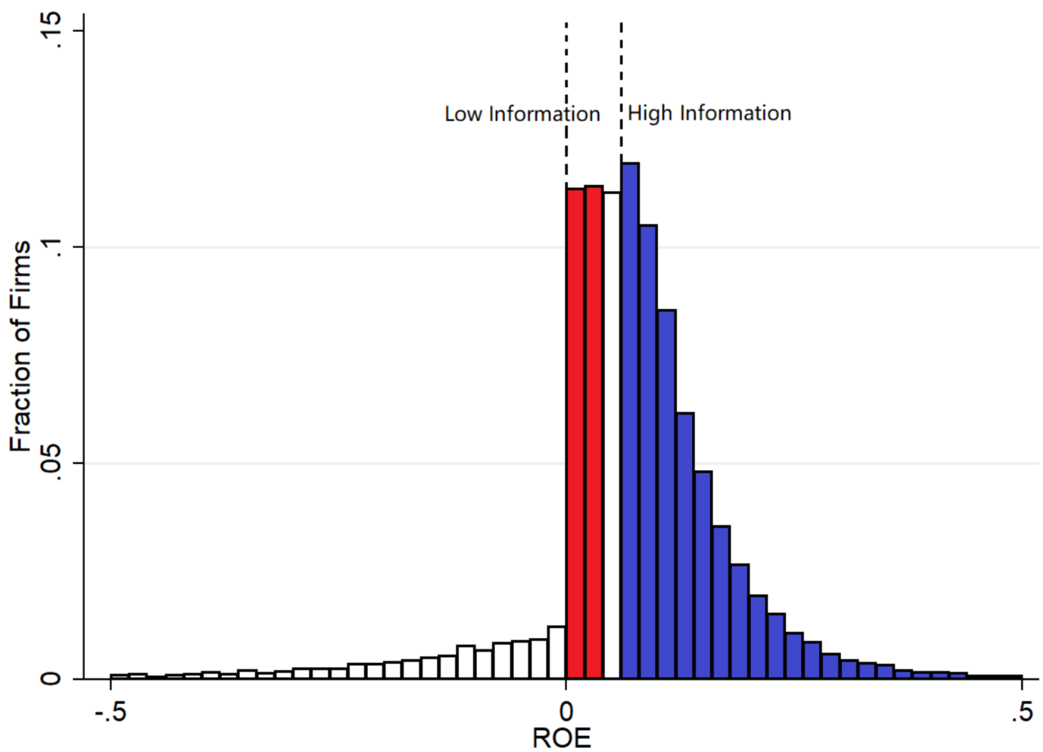


Figure 3.2. Two Information Segments in China

We provide empirical evidence that investors cannot distinguish which firm actually engages in earnings management. First, we divide all firms in low information segment into five

[2]The full list of industries with an average industry ROE below 4% from 2000-2016 in China: A: Agriculture H: Restaurant/Dining M: Technology Service P: Education Q: Healthcare R:Entertainment S: Social Service.

quantiles based on two widely used measures of earnings management, namely discretionary accrual and real earnings management. Investors treat firms in the 1^{st} quantile and 5^{th} quantile indifferently under both measures. Second, we show that investors react indifferently to annual reports by firms that accidentally fall into the low information segment and firms that systematically stay there. More specifically, we show investors cannot distinguish between future stayers and escapers for firms falling into the low information segment this period.

Furthermore, we show that firms in the low information segment suffer from lower stock market investors' reaction and lower cumulative abnormal return around earnings announcements, insignificant earnings response coefficient, lower trading liquidity, higher systematic risk, and higher synchronicity. These results imply that investors distrust the earnings numbers reported by firms in the low information segment. As a result, investors react less to earnings announcements and incorporate less firm-specific information in the stock prices. In other words, stock prices of firms in the low information segment are less informative about firms' fundamentals and co-move more significantly with the overall stock market. Our findings offer a new explanation on the unusually large stock price co-movement among individual stocks in China as documented in Morck et al. (2000a).

We further corroborate our findings with causal evidence. We identify a group of firms that exogenously switch from high information segment to low information segment as a result of the 2007-08 global financial crisis. Comparing this group of firms with firms that had the same magnitude of drop in ROE but stayed in high information segment, we confirm that firms that exogenously fall into the low information segment suffer from adverse effects in the financial market due to investors' distrust.

Our paper provides the very first evidence about the spillover effect of earnings management. More specifically, we show that large scale of earnings management has a negative externality effect on all relevant firms with respect to market reaction and price informativeness. In our setting, earnings management, though implemented by each individual firm, could affect other non-manipulating firms' well-being. Our finding owes to three unique features of China's capital market. First, a large proportion of firms manage their earnings due to the delisting policy. Second, investors are well aware of this situation and it is relatively easy for them to pool firms together based on a fixed accounting number (zero here due to regulation). Third, retail investors take up more than 70 percent of Chinese stock holdings. It is, if not impossible, extremely difficult for them to detect a specific firm's earnings management.

Our paper also contributes to a growing literature on the real and financial effects of market transparency and price informativeness. For nearly all of the prior literatures that studied the effects of market transparency (Levine and Zervos (1996)), a cross-country analysis is implemented to obtain the necessary variation of transparency, which leads to an inevitable edogeneity problem. However, due to the delisting policy, a subset of Chinese firms (those with barely negative earnings) have a huge incentive to manage their earnings, which gives us a significant variation of market transparency in China (high and low information segments as discussed above). Therefore, we are disengaged from all unobservable country fixed effect. Our paper also provides a rare opportunity to study the real value of financial market development. Treating low information segment as developing, less supervised market, and high information segment as developed, more mature market, we offer a rich soil for future within-country studies on financial market development.

Our empirical findings are subject to several caveats. First, although we divide China's stock market into high and low information segments, we have not defined a direct measure of information level either for firms or segments. Rather, we provide evidence of cross-segment variation for a mass of short-term and long-term financial measures. We further show investors cannot detect a specific firm's earnings management in the low information segment. This finding rules out the possibility that our finding merely comes from an average effect of all manipulating firms. Also, the extremely large magnitude of earnings management around 0 in China further guarantees that the two segments defined in our paper have entirely different information transparency. Second, the spillover effect in our paper mainly focuses on the investor/market side. Firms do not further generate a separating equilibrium in the low information segment because of the following reasons: first, their incentive to manage earnings drops dramatically without the danger of delisting; second, we are the very first to document this "two-market" phenomenon and firms may not know the additional capital market consequences at all.

The rest of the paper is organized as follows. Section 2 reviews the literature. Section 3 provides institutional background on China's delisting policy. Section 4 presents summary statistics on the data. We present evidence on firms' earnings management in Section 5. In section 6, we document the existence of two information segments and the financial effects of falling into the low information segment. In section 7, we show that investors treat good and bad firms similarly in the low information segment. In section 8, we pin down a group of firms that exogenously falls into the low information segment as a result of the global financial crisis and present causal evidence on the effects of sliding into the low information segment. Section 9 concludes.

3.2 Literature Review

Our research is closely related to several strands of literature in finance and accounting.

3.2.1 Earnings Management

Our paper is related to a massive accounting literature on earnings management. We proxy for earnings management with both discretionary accrual estimated with modified Jones model as advocated in Dechow et al. (1995) and also with proxies for real earnings management as in Roychowdhury (2006). [3]

3.2.2 Market Transparency

We contribute to the current literature on the real and financial effects of market transparency in the following 3 aspects.

First, for all of the prior literatures that directly study the effects of market transparency (Levine and Zervos, 1996), they use a cross-country analysis to acquire the necessary variation

[3] See Appendix for more details.

of transparency, which leads to an inevitable endogeneity problem. However, due to one specific delisting policy, Chinese firms have a huge incentive to manage their earnings right above 0, which gives us a significant variation of market transparency in China.

Second, even though the real and financial effects of disclosure level have been widely studied, these effects have rarely been investigated in the transparency area. The most important reason here is that prior research mainly focus on individual firm-level disclosure measure. They cannot link disclosure level to market transparency since there is not a systematically biased distribution of disclosure quality inside the market. The connection between individual firm-level disclosure and aggregate market transparency in our paper depends on the dramatically different earnings manipulation incentives across different ROE ranges in Chinese stock market.

Third, unlike the US market, individual investors take up more than 70 percent of Chinese stock holdings. The variation of market transparency comes from the investors' inability to fully detect a specific firm's earnings manipulation. The large percentage of individual investors in China further strengthens the connection between individual firm-level disclosure and aggregate market transparency. Moreover, we sort firms based on their measures of discretionary accruals and real earnings management and observe no evidence for investors' detection in either measurement.

3.2.3 Market Reaction and Price Informativeness

Our paper studies short-term market reactions and long-term price informativeness of firms in the low information segment.

For market reaction measures, we first use the earnings respond coefficient(ERC) following Collins and Kothari (1989b), which basically describes the relationship between cumulative abnormal return and unexpected earnings. The ERC has been widely adopted both in accounting and finance literature. Furthermore, we use two other announcement reaction measures following Pevzner et al. (2015a). One is the abnormal return volatility, which mainly measures the abnormal return volatility during announcement window versus the estimation window. The other is the abnormal trading volume, which is constructed similarly only instead using the trading volume. We expect ERC to be less significant and two abnormal reaction measures to be lower in the low information segment.

For price informativeness measures, we first choose the synchronicity following Morck et al. (2000b). Synchronicity comes from the R^2 in CAPM model and describes the degree that a stock price co-moves with the market index. The higher R^2 is, the less firm-specific information is incorporated into the stock price. We expect the synchronicity to be higher in the low information segment since investors distrust firms' announcement. We also use the factor loadings from CAPM model as an alternative measure. We expect the market β to be higher in the low information segment due to a higher systematic risk and cost of capital.

3.3 Institutional Background on China's Delisting Policy

The delisting policy in China was established in 1998 by the China Securities Regulatory Commission (CSRC). The intention of the policy is to protect unsophisticated retail investors by reminding them of the risk in investing in the stock market. Specifically, the delisting policy mandates that if a publicly-listed firm reports negative accounting earnings in two consecutive years, its stock will be put under *special treatment* status (ST). There are various trading and financing restrictions on ST stocks[4]. If an ST firm reports one more annual loss, it is suspended from trading on the stock exchanges. After a fourth annual loss, the stock will be de-listed from the stock exchange. In total, approximately 100 firms have actually been delisted in China.

The delisting policy has a far-reaching impact on all firms in China. Every firm wants to avoid being put under special treatment status which we refer to as a delisting threat. A delisting threat not only brings stigma to a firm but also strictly restricts firm's financing activities in the capital market. As a result, firms go great length to avoid reporting two consecutive negative annual earnings by engaging in earnings management. We will first show evidence on firms' earnings management and then present the real and financial consequences of the delisting policy.

[4]ST companies' daily stock price movement is restricted to be no more than 5% in either direction. Non-ST stocks' daily price range is restricted to 10% in either direction. ST companies' semi-annual reports must be audited. Furthermore, ST firms cannot raise additional capital from stock market.

3.4 Data and Summary Statistics

Our research utilizes data on stock price and firm-level fundamentals for all listed firms in the U.S. and China from 2009 to 2016. For firms listed in China, we mostly rely on data from the China Stock Market and Accounting Research (CSMAR) database. We obtain data from CSMAR on daily stock return, market return, and announcement dates of annual financial report along with other firm-level variables such as firm size (total assets), return-on-equity (ROE), sales, account receivables, leverage (book debt/total assets), operating and net cash flows, R&D expenditure, advertising, selling, general, and administrative expenses, cost of goods sold, and inventory. We obtain data on stock price, ROE, and announcement dates of annual financial reports for all firms listed in the U.S. from Compustat, CRSP, and the Bloomberg Terminal.

Table 3.1 presents the summary statistics for key variables used in our research. Before each one of our regression analysis, we winsorize all continuous variables at 1st and 99th percentile to mitigate the impact of outliers.

Table 3.1. Summary Statistics For Companies Listed in China 2009-2016

	N	Mean	Std	p25	p50	p75
Abnormal Return Variance	8823	2.05	4.22	0.35	0.82	1.97
Abnormal Trading Volume	6987	1.29	1.09	0.64	1.00	1.58
Log (Firm Size)	8818	21.58	1.18	20.78	21.43	22.19
Firm Leverage	8823	0.46	0.21	0.30	0.46	0.62
Return on Equity	8823	0.10	0.10	0.05	0.09	0.15
Unexpected Earnings	8002	0.00	0.02	0.00	0.01	0.02

3.5 Do Low Information Segment Firms have more Earnings Management?

In this section, we present two pieces of evidence that public firms in the low information segment indeed engage in more earnings management than other firms. First, we plot the histograms of firms' return on equity (ROE) [5] distribution for both China and the U.S. The tremendously high proportion of firms falling into ROE range (0, 4%) in China compared to the U.S. suggests that a large number of Chinese firms engages in massive earnings management to report positive earnings. Second, we present direct evidence that Chinese firms with ROE from 0 to 4% have significantly more real earnings management than other firms.

3.5.1 Firm ROE Distribution Histograms: China V.S. the U.S.

Figure 3.1 plots the ROE distribution histograms for listed firms in China and in the U.S.. We pool together all the listed firms from 2009 to 2016. The x-axis is ROE from -50% to 50% and the y-axis is the fraction of firms falling into each 2% ROE bin. China has two major stock exchanges. Figure 3.1a plots all the firms listed in the Shanghai Stock Exchange whereas Figure 3.1b all the firms in the Shenzhen Stock Exchange.

We find similar patterns across the Shanghai and Shenzhen Stock Exchanges. Comparing the ROE distribution of Chinese firms to the U.S. firms, we make two immediate observations: 1) 18 % firms listed in China report an ROE from 0 to 4% compared to 10% firms in the U.S. 2) the difference between fractions of firms in ROE range (-2%, 0) and (0, 2%) is 8% in China compared with 1.5% in the U.S.. A much higher mass of firms with ROE from (-2%, 0) than firms with (0, 2%) convincingly suggests that firms engages in earnings management to achieve positive earnings.

We further divide Chinese firms into two categories: firms with a positive ROE last year and those with a negative ROE last year. Since the firms only face a de-listing threat after two consecutive years of negative earnings, we expect that firms with negative ROE last year would have a much stronger incentive to manage and to report a positive earnings this year. Consequently, we expect to see a even higher mass of firms with negative ROE last year falling into small and positive ROE range this year, than firms with positive ROE last year. We see exactly what we have expected on Figure 3.3: at least 50% of firms with a negative ROE last year report an ROE from 0 to 4% this year. On the contrary, less than 20% of firms with a positive

[5] Return on Equity (ROE) = Net Earnings/Book Equity

ROE last year reporting a (0, 4%) ROE in the following year. We further run a regression using investor reaction measure as dependent variable and delisting threat indicator (1 if negative last year; 0 if positive last year) as independent variable in the appendix. We confirm that investors do react less to firms with delisting threat.

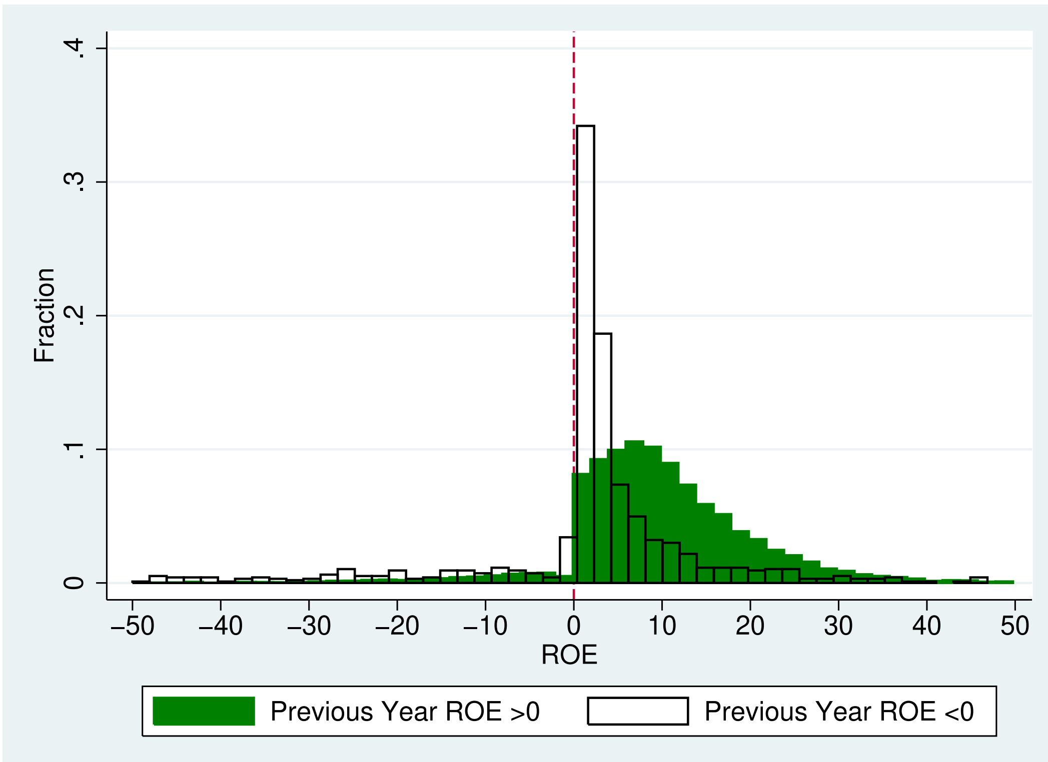


Figure 3.3. Firm ROE Distribution in China: Positive v.s. Negative ROE (Last Year)

To sum up, firms listed in China have a much stronger incentive to manage their earnings than firms in the U.S. due to China's distorted delisting policy. Furthermore, the incentive to report a positive earning is even stronger for firms that had a loss in the previous year in China. Following Chetty et al. (2011), we use a bunching estimator to retrieve the counterfactual ROE distribution without any earnings management. According to Figure C.1, there are approximately 40% of firms in the low information segment (ROE from 0 to 4%) which should have ended up with a negative ROE if they had not engaged in earnings management.

3.5.2 Testing if Low Information Segment Firms have more Earnings Management

Firstly, we calculate both Discretionary Accruals (DA) and Real Earnings Management (RM) for each listed firm in China following Dechow et al. (1995) and Roychowdhury (2006). [6]

Second, we test if firms in the low information segment (ROE range (0,4%)), which are highly suspected of managing their earnings based on ROE distribution histograms, have higher discretionary accruals than other firms by running the following regression:

$$DA_i = \alpha + \beta_1 * 1_{\mathbf{ROE} \in \mathbf{(0,0.04)}} + \beta_i * Controls_i + \varepsilon_i \tag{3.1}$$

where DA_i is the discretionary accrual of firms i. $1_{\mathbf{ROE} \in \mathbf{(0,0.04)}}$ is a dummy variable that equals to 1 if a firm's ROE is in (0, 4%), 0 otherwise. We also include control variables such as firm size, leverage, industry, and year dummies that can explain firms' discretionary accruals.

Our results in Table 3.2 show that firms in the ROE range (0, 4%) do not have a significantly higher discretionary accrual than other firms. There result may due to increasing attention from securities authorities on firms' abnormal accruals. Moreover, as presented in the below, firms in the low information segment tend to overly engage in real earning management, which reduces the usage of discretionary accruals.

We test if firms in the low information segment have higher real earnings management than other firms by running the following regression:

$$RM_i = \alpha + \beta_1 * 1_{\mathbf{ROE} \in \mathbf{(0,0.04)}} + \beta_i * Controls_i + \varepsilon_i \tag{3.2}$$

where $1_{\mathbf{ROE} \in \mathbf{(0,0.04)}}$ is a dummy variable that equals to 1 if a firm's ROE is in (0, 4%), 0 otherwise. We also include control variables such as firm size, leverage, industry, and year dummies that can explain firms' real earnings management.

Our results in Table 3.2 show that real earnings management as a share of last year's total

[6] See Appendix for details.

asset is 3-6% higher for firms with ROE from 0 to 4%. This result lends direct support to our claim that firms in China tend to manage their earnings under the pressure of a delisting threat. Also, since real earnings management has to be conducted throughout the entire accounting year, discretionary accruals (adjusted at the year end) have been seldomly used as a major managing method.

Table 3.2. Earnings Management across Firms: 2009-2016 China

	(1) RM	(2) RM	(3) DA	(4) DA
1_{ROE}∈**(0,0.04)**	0.0622***	0.0342***	-0.000881	0.00110
	(0.00508)	(0.00473)	(0.00262)	(0.00271)
Firm Size		-0.00639***		0.00133
		(0.00243)		(0.00119)
Firm Leverage		0.165***		0.00252
		(0.0132)		(0.00586)
Return on Equity		-0.535***		0.0305***
		(0.0257)		(0.0113)
Observations	12231	12144	12231	12144
Adjusted R^2	0.013	0.131	0.004	0.007

Note: In the parentheses below coefficient estimates are robust standard errors adjusted for heteroskedasticity and firm-level clustering. All continuous variables are winsorized at the 1st and 99th percentile. *, **, and *** denote significance at the 0.1, 0.05, and 0.01 level, respectively. RM stands for real earnings management and DA stands for discretionary accrual.

3.6 Two Information Segments in China's Stock Market

A large fraction of firms with ROE from 0 to 4% are suspects of earnings management. We establish that there are two information segments within China's stock market with event studies focusing on firms' annual earnings announcements. The low information segment consists of firms in ROE (0,4%) in which many of them are suspects of massive earnings management. The high information segment has all the firms with ROE (7%, $+\infty$) which are not under an immediate pressure of delisting. Consequently, firms in the high information segment are much more truthful about their earnings.

We show that firms in the low information segment suffer from lower stock market investors' reaction and lower cumulative abnormal return around the dates of earnings announcements, insignificant earnings response coefficient, lower stock trading liquidity, higher systematic risk, and higher co-movement with the overall stock market. These results imply that investors distrust the earnings numbers reported by all firms in the low information segment. As a result, investors react less to earnings announcements and incorporate less firm-specific information in the stock prices. In other words, stock prices of firms in the low information segment are less informative about firms' fundamentals and co-move more significantly with the overall stock market. Our findings offer a new explanation on the unusually large stock price co-movement among individual stocks in China as documented in Morck et al. (2000a).

3.6.1 Abnormal Stock Return Variance

Abnormal return variance is calculated as the average of the squared market-model adjusted daily returns over the event window (-1, +1), scaled by the stock return variance over the estimation window (-120, -21) (Pevzner et al., 2015b). The market model is estimated over the estimation window (-120, -21). Specifically, firm i's market model adjusted returns on day t during the event window is computed as follows:

$$U_{it} = R_{it} - (\alpha_i + \beta_i R_{mt})$$

where R_{it} is the daily stock return of firm i on day t, R_{mt} is the daily market return on day t, and α_i and β_i are firm i's market model estimates obtained from the estimation window. Stock return variance over the event window (-1, +1) then is calculated as the average of the squared market adjusted return, U_{it}^2. The stock return variance over the estimation window (-120, -21) equals the variance of the residual returns from the firm's market model estimated over the estimation window.

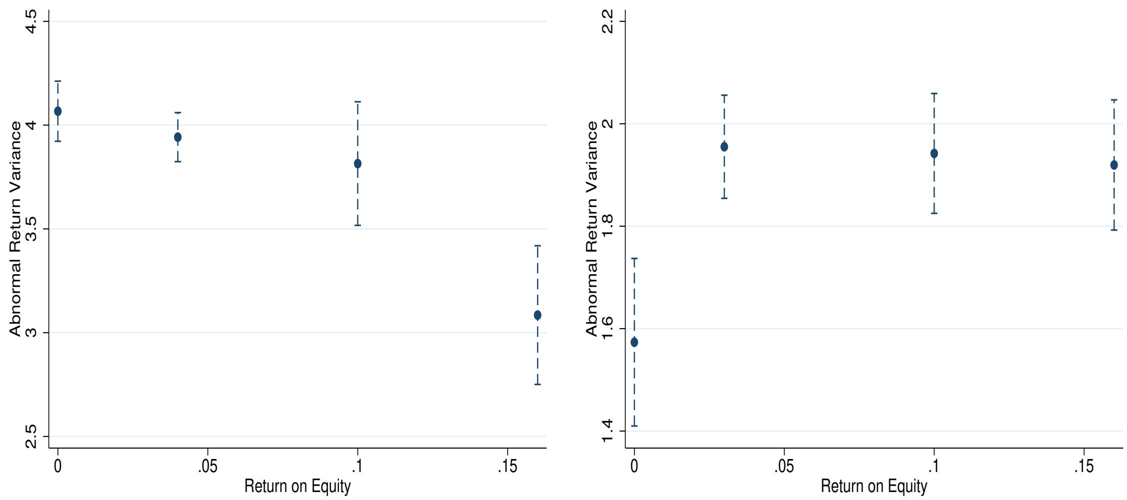
We plot the abnormal return variance on Figure 3.4 for all the listed firms in China and the U.S. from year 2009 to 2016. The X-axis is the firm ROE in percentage and the Y-axis is the level of abnormal return variance. Each dot is the average of abnormal return variance for all firms in the corresponding ROE range. The first dot is the average for firms in ROE from 0 to 4%, second 4-10%, third 10-16% and fourth all firms with a ROE above 16%. The dashed bars are the 1.96 standard error of the mean. Notably, observing a clear pattern in the figure is much stronger than the traditional regression result. We include the regression result as well in the appendix.

Firstly, Figure 3.4 shows that American firms have an average abnormal return variance of 4 [7] which is much higher than the average of 1.9 for Chinese firms. The difference in magnitude indicates that the U.S. stock market is more efficient in incorporating firms' annual earnings news into stock prices than China's.

Secondly, we notice that abnormal return variance of American firms is slightly decreasing with ROE. In contrast, abnormal return variance is significantly positive correlated with ROE for Chinese firms. For now, we do not take a stand on why abnormal return variance is declining with ROE in the U.S.. We are using the firms in the U.S. to illustrate what the correlation between abnormal return variance and ROE would normally look like in a stock market without a delisting policy based solely on firms' earnings. Comparing with the decreasing trend in the U.S., an increasing trend of abnormal return variance in ROE in China seems rather peculiar and is likely related to its delisting policy.

We address potential concerns that the positive correlation between abnormal return variance and ROE in China is a spurious correlation by controlling for covariates such as firm size, leverage, unexpected earnings, industry, and year. Specifically, we filter out the impact of the covariates mentioned above by regressing our firm-level abnormal return variance on those covariates and plot the residual of the abnormal return variance on Figure 3.5. Firms with an ROE from 0 to 4% still have a lower abnormal return variance (residual) compared to other firms. This finding not only supports our hypothesis but also alleviates the endogeneity concerns on what we find on Figure 3.4.

[7] which is similar to what Pevzner et al. report in their paper (Pevzner et al., 2015b)



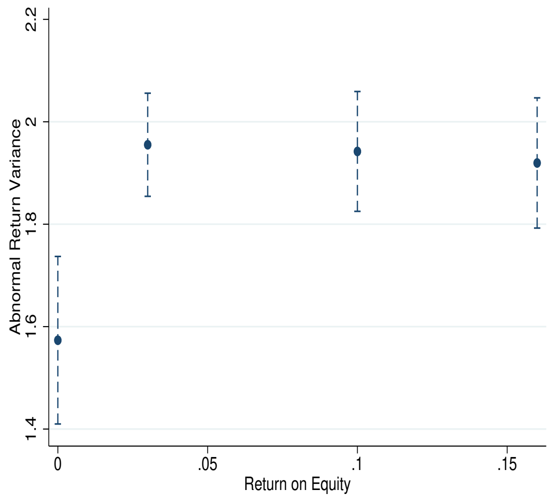
(a) the U.S. **(b)** China

Figure 3.4. Abnormal Return Variance Around Firms' Annual Earnings Announcement: China v.s. the U.S. (2009-2016)

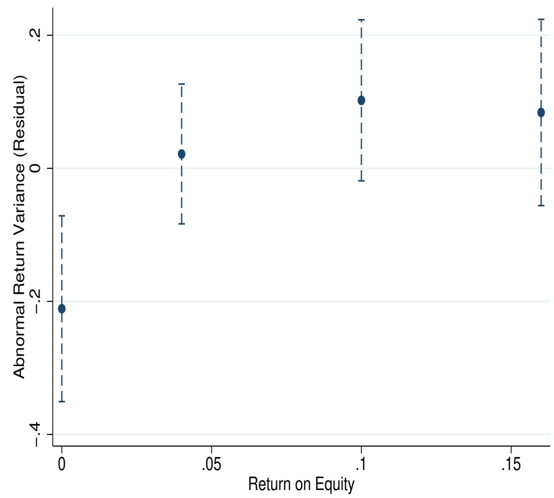
From Figure 3.5, we see that firms in the ROE range of 0 to 4% have an abnormal return variance that is about 0.3 lower than firms with ROE greater than 10%. The difference is statistically significant and is free of impacts of common covariates of stock return variance. The sample average of abnormal return variance is around 1.9, which means that average return variance for a firm when it announces its annual report is 90 % higher than its average return variance in normal times. Firms with ROE from 0 - 4% only have an average abnormal return variance of 1.6 which is 60% higher than normal times. We could define the *extra* return variance brought by earnings announcement as abnormal return variance - 1. We see that normal firms (ROE$>$ 0.1) have an *extra* return variance that is 1.5 times as large as firms with ROE from 0-4%. The magnitude is economically significant and lends support to our hypothesis that investors distrust and react less to earnings reported in the balance sheet of suspicious firms in terms of return variance.

3.6.2 Abnormal Trading Volume

We measure abnormal trading volume by calculating average trading volume over the event window (-1, +1), scaled by the average trading volume over (-120, -21) (Pevzner et al.,



(a) Level of Abnormal Return Variance



(b) Residual of Abnormal Return Variance

Figure 3.5. Abnormal Return Variance Around Firms' Annual Earnings Announcement: China

Note: residual is predicted after regressing abnormal return variance on firm size, leverage, absolute value of unexpected earnings, industry, and year dummies.

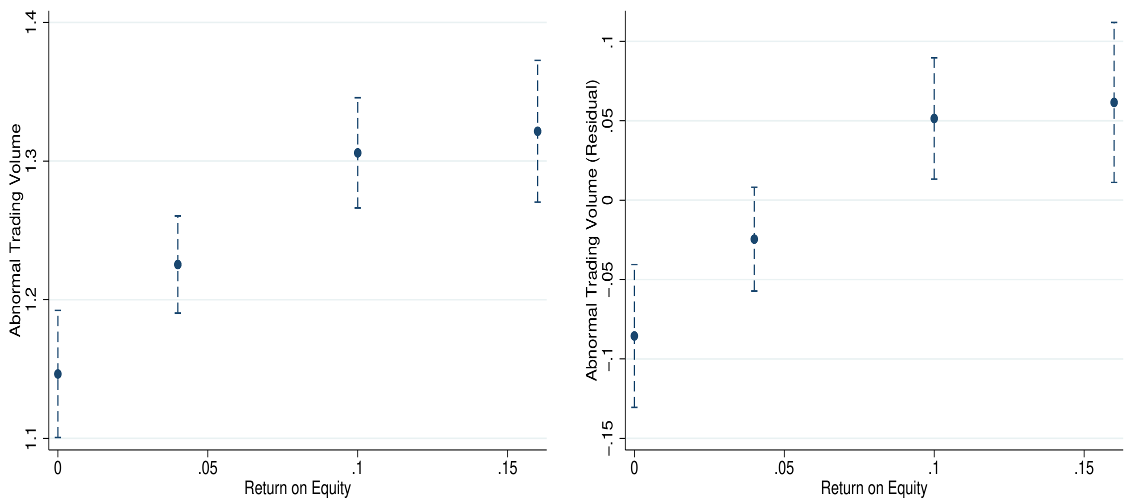
2015b). Trading volume is defined as the number of shares of firm i traded on day t divided by the total number of shares outstanding of firm i on day t.

We plot the abnormal trading volume (residual) [8] on Figure 3.6. The X-axis is the firm ROE in percentage and the Y-axis is the residual of abnormal trading volume. Each dot is the average of abnormal trading volume (residual) for all the firms in the corresponding ROE range. The first dot is the average for firms in ROE from 0 to 4%, second 4-10%, third 10-16%, and fourth all firms with a ROE above 16%. The bars are 1.96 standard errors of the mean.

From Figure 3.6, we see that firms in the ROE range of 0 to 4% have an abnormal trading volume that is 0.15 lower than firms with ROE greater than 10%. The difference is statistically significant and is after controlling for common covariates of trading volume. The sample average of abnormal trading volume is around 1.2, which means that average trading volume for a firm when it announces its annual report is 20 % higher than its average trading volume in normal times. Firms with ROE from 0 - 4% only have an average abnormal trading volume of 1.05 which

[8] We take the residual after regressing abnormal trading volume on firm size, leverage, absolute value of unexpected earnings, industry, and year effects.

is 5% higher than normal times. We could define the *extra* trading volume brought by earnings announcement as abnormal trading volume - 1. We see that normal firms (ROE>0.1) have an *extra* trading volume that is 4 times as large as firms with ROE from 0-4%. The magnitude is economically significant and reinforces our hypothesis that investors discount the earnings numbers reported by suspicious firms and react less accordingly in the stock market.



(a) Level of Abnormal Trading Volume **(b)** Residual of Abnormal Trading Volume

Figure 3.6. Abnormal Trading Volume Around Firms' Annual Earnings Announcement: China

Note: residual is predicted after regressing Abnormal Trading Volume on firm size, leverage, absolute value of unexpected earnings, industry, and year dummies.

3.6.3 Earnings Response Coefficient

We provide further evidence on whether investors discount the earnings of suspicious firms by calculating earnings response coefficient (ERC) for each firm. Suppose that firm A and B report the same and positive unexpected earnings and investors trust firm A's earnings more, we expect that firm A's price increase would be higher than that of firm B's. We estimate the ERC using the following regression:

$$CAR_i = \alpha + \beta_1 * UE_i + \sum_{2}^{i=k} \beta_i * Controls_i + \varepsilon_i \tag{3.3}$$

where CAR_i is the three-day cumulative abnormal return over event window (-1,+1) with 0 denoting the day when the annual earnings announcement is made. UE_i is firm i's unexpected earnings which is defined as actual annual earnings minus the most recent mean analyst forecast, scaled by the most recent stock price. We also include covariates such as: firm size, ROE, leverage, industry, and year dummies.

The estimated coefficient ($\hat{\beta}_1$) of UE_i is the earnings response coefficient and measures how stock prices respond to firms' unexpected earnings. There is extensive empirical finance research documenting that ERC ($\hat{\beta}_1$) should be significantly positive. Stock prices are expected to rise after a positive unexpected earnings. A ERC that is not significantly different from 0 suggests that price response to earnings surprises is sluggish, implying that investors do not believe in the earnings reported by the firms.

Our hypothesis is that investors distrust the earnings reported by suspicious firms' (ROE $\in (0, 4\%)$). Hence, we expect to see a ERC, estimated within the sub-sample of suspicious firms, that is either not significantly different from 0 or smaller than ERC estimated within the sub-sample of normal firms (ROE$>$ 10%). We find exactly what we have expected in Table 3.3. The first column is estimated using the whole sample and we see that ERC is significantly positive which is consistent with the previous literature. A one unit increase in UE results in a 17.4% gain in three-day cumulative abnormal return around earnings announcement. The second column provides strong evidence that investors do not react to unexpected earnings of suspicious firms. The ERC for firms with ROE greater than 4% is positive and statistically significant, indicating that investors do respond to firm-level earnings surprises if they trust what these firms say on their balance sheet.

3.6.4 Price Informativeness

In principal, stock price movements of an individual firm can be decomposed into movements due to market/industry level news and firm-level news (Roll, 1998). Suppose that firm A and B publish the same amount of idiosyncratic news and investors believe that the quality of firm A's news is higher, we expect that the price informativeness of firm A's stock price

Table 3.3. Earnings Response Coefficient Across Sub-samples (2009-2016 China)

	CAR			
	All Firms	ROE ∈(0,0.04)	ROE ∈(0.04,0.1)	ROE ∈(0.1,+∞)
Unexpected Earnings	0.174***	0.0233	0.235***	0.213***
	(0.0485)	(0.131)	(0.0858)	(0.0743)
Return on Equity	0.0184***	-0.00355	-0.00406	0.0400***
	(0.00700)	(0.0997)	(0.0485)	(0.0133)
Firm Size	0.00114**	0.00305**	0.000746	0.00112
	(0.000512)	(0.00127)	(0.000996)	(0.000779)
Firm Leverage	-0.00574**	-0.0132*	-0.00425	-0.00765*
	(0.00290)	(0.00692)	(0.00514)	(0.00453)
Year effect	Yes	Yes	Yes	Yes
Industry effect	Yes	Yes	Yes	Yes
Observations	7403	1188	2593	3382
Adjusted R^2	0.010	0.004	0.006	0.017

Note: in the parentheses below coefficient estimates are robust t-statistics based on standard errors adjusted for heteroskedasticity and firm-level clustering. All continuous variables are winsorized at the 1st and 99th percentile. *, **, and *** denote significance at the 0.1, 0.05, and 0.01 level, respectively.

will be higher since investors are more likely to trade on firm A's idiosyncratic news. Investors are aware that the trustworthiness of annual financial reports for firms with ROE from 0 to 4% is substantially lower than those published by firms with ROE greater than 10%. We further hypothesize that the stock prices of firms with ROE from 0 - 4% contain less idiosyncratic firm-level information and hence shall co-move significantly more with the market.

We test our hypothesis using price non-synchronicity proposed by Roll (1998). Price non-synchronicity basically measures the correlation between a firm's return and a market or industry benchmark. The higher the correlation between a firm's stock return and market return, the less informative stock price is about the company's idiosyncratic news and fundamentals. Papers that adopt this measure include Morck et al. (2000a), Durnev et al. (2003), and Chen et al. (2006). Durnev et al. (2003) show that price non-synchronicity is positively related to the correlation between returns and future earnings at the industry level, which helps to validate it as a measure of informativeness.

Following Jin and Myers (2006); Morck et al. (2000a), we estimate a Capital Asset Pricing Model (CAPM):

$$r_{it} = \alpha_i + \beta_i(r_{mt} - r_{ft}) + \varepsilon_{it} \tag{3.4}$$

separately in the pre-event period (-100, -1) and post-event period (+1, +100) for each individual firm. r_{it} is firm i's return on date t; r_{mt} stock market return on date t; r_{ft} risk-free rate on date t. We define R^2_{diff} as the difference between the R^2 of the CAPM in pre- and post-event period: R^2_{pre}, R^2_{post}. We are plotting on Figure 3.7 the average of R^2_{diff} for four groups of firms based on their ROE: (0, 4%), (4%, 10%), (10%, 16%), (16%, $+\infty$).

In the pre-event period (-100, -1) which corresponds to 4 months to 1 day before the annual earnings announcement of a firm, there are a lot of uncertainties on how the firm performed in the past year and what its earnings would be. Individual stock price comove greatly the overall market due to scarcity of firm-level idiosyncratic news. As soon as firms publish their annual earnings numbers, the uncertainties are largely dissolved and stock prices would reflect more of firms' fundamentals instead of market-wide news such as GDP growth, unemployment, inflation, etc.

What we see on Figure 3.7 is consistent with our reasoning. We see from Figure 3.7 that on average, firms with ROE greater than 10% have a significant drop in R^2 of over 0.02 (3.4-6.7 % of the sample average R^2 (0.3)) from pre-event to post-event period, which is a sign that uncertainties on firms' earnings are dissipated and stock prices reflect more of firms' own fundamentals. However, for firms with ROE from 0 to 4%, they actually experience a significant increase in R^2 of 0.02 (a 6.7% increase of sample mean (0.3)) from pre-event to post-event period. We are not sure how to interpret the increase of R^2. For now, we take it as strong evidence that investors distrust the financial reports published by these firms. In contrast with firms with high quality reports, there are still a lot of uncertainties and speculations on the actual performance of firms reporting a ROE from 0 to 4%.

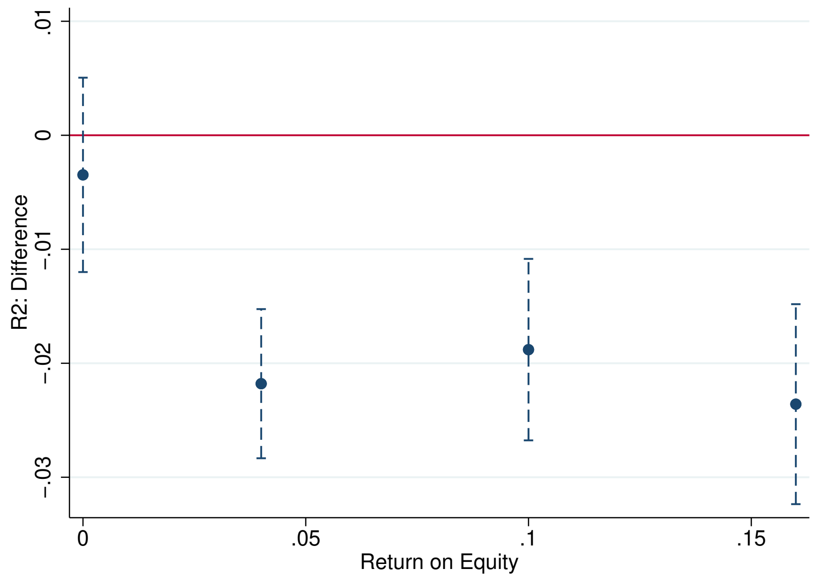


Figure 3.7. Difference in R^2: Pre- and Post- Earnings Announcement

Note: we estimate R^2 pre- and post- annual earnings announcement using event days (-60,-5) and (5,60) respectively. Results similar if controlling for industry and year

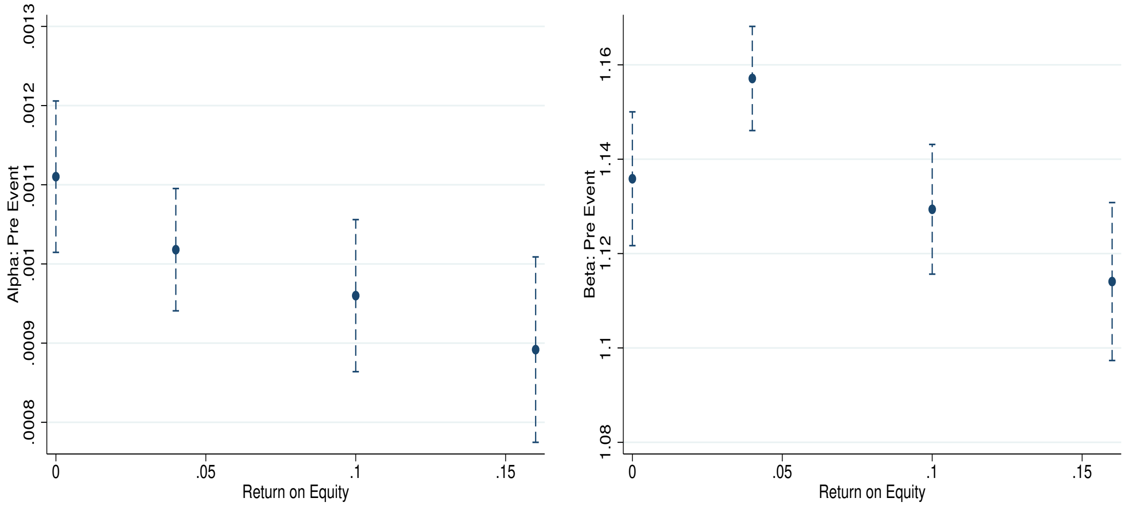
3.6.5 Risk Factor Loadings

We are interested in whether risk factor loadings would be different across different ROE ranges as a consequence of market transparency. We are particularly interested in testing whether firms with low transparency are more exposed to market risk. Following (Jin and Myers, 2006; Morck et al., 2000a), we estimate a Capital Asset Pricing Model (CAPM):

$$r_{it} = \alpha_i + \beta_i(r_{mt} - r_{ft}) + \varepsilon_{it} \tag{3.5}$$

in post-event period (+1, +100) for each individual firm. r_{it} is firm i's return on date t; r_{mt} stock market return on date t; r_{ft} risk-free rate on date t. We obtain $\hat{\alpha}_i$ and $\hat{\beta}_i$ for each firm in the post event period. We then plot on Figure 3.9 the average of $\hat{\alpha}_i$ and $\hat{\beta}_i$ for four groups of firms based on their ROE: (0, 4%), (4%, 10%), (10%, 16%), (16%, $+\infty$).

We see from Figure 3.9 that $\hat{\alpha}$ is stable across ROE groups. Since α measures the mispricing of an individual stock based on CAPM, we conclude that firms with low transparency

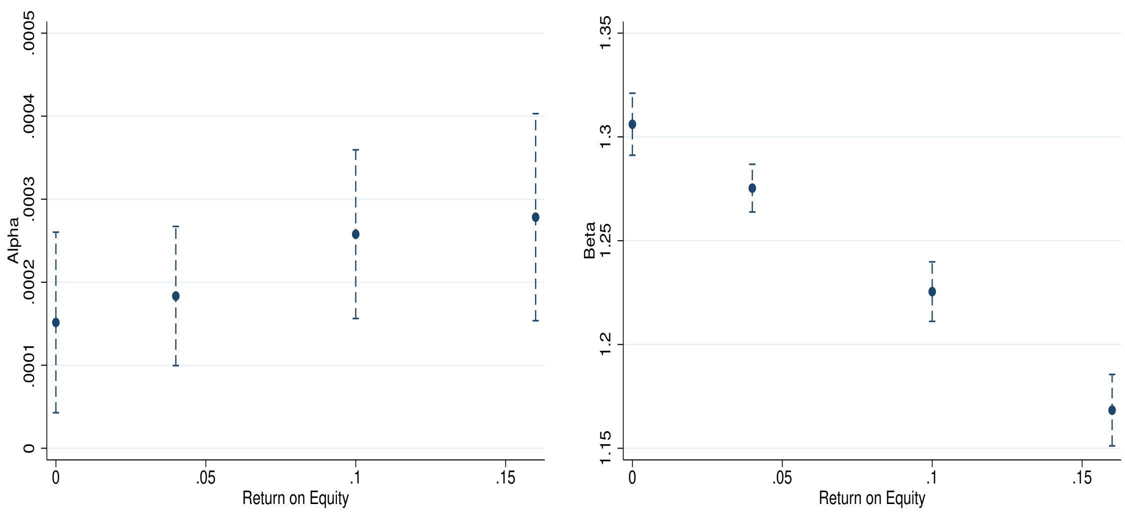


(a) Alpha

(b) Beta

Figure 3.8. Alpha and Beta: Pre Earnings Announcement

Note: we estimate pre-annual earnings announcement alpha and beta using event days $(-100, -1)$ respectively. Results similar if controlling for industry and year.

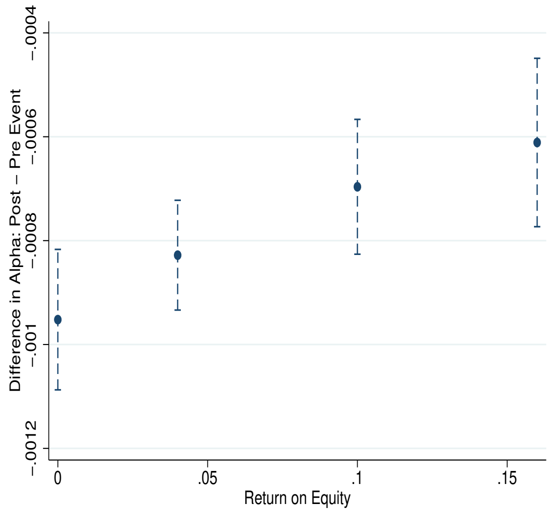


(a) Alpha

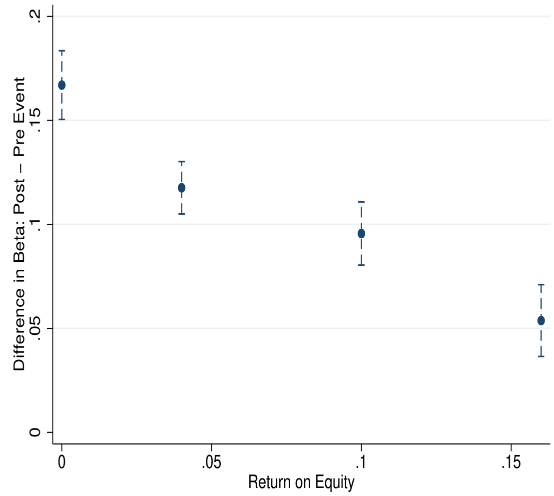
(b) Beta

Figure 3.9. Alpha and Beta: Post Earnings Announcement

Note: we estimate post-annual earnings announcement alpha and beta using event days $(+1, +100)$ respectively. Results similar if controlling for industry and year.



(a) Alpha



(b) Beta

Figure 3.10. Alpha and Beta: Post - Pre Earnings Announcement

Note: we estimate pre and post-annual earnings announcement alpha and beta using event days $(-100, -1)$ and $(+1, +100)$ respectively. Results similar if controlling for industry and year.

are not more mispriced than other firms. However, when we look at β, we observe that β is significantly higher for firms with low transparency whose ROE is from 0 to 4%. This finding suggests that firms with low-transparency are more exposed to systematic market risk. Higher β may be caused by low-quality firm level information and high uncertainties on firm's performance.

In an efficient market, investors are compensated with expected return commensurate to the risk in an individual stock. The higher the risk, the higher the expected return that investors would demand. As a consequence of a higher β, investors are taking more market risk by buying stocks of firms with ROE from 0 to 4% and they will only be doing so if they are compensated with a higher expected return. A higher expected return is equivalent to a lower current stock price. Depressed stock prices have adverse effects on firm's additional capital raising from stock market. In a seasoned equity offering (SEO) in which firms sell new shares to shareholders, firms are only able to sell shares at the current price. A depressed stock price would hurt firms' ability of raising additional capital from stock market, which may result in a binding financing constraint and force firms to forego worthy investment projects.

3.7 Can Investors Distinguish Good and Bad Firms in the Low Information Segment?

We present two pieces of evidence that investors are not able to tell *good* and *bad* firms apart in the low information segment.

Firstly, we construct two measures of earnings management for all firms in the low information segment based on accounting literature. Namely, real earnings management and discretionary accrual. Afterwards, we divide all firms in the low information segment into 5 sub-samples using the level of discretionary accrual and real earnings management in ascending order. In Figure ??, we plot the abnormal return variance for firms with different levels of earnings management in the low information segment. The five dots are the average of abnormal return variance (residual) for each sub-sample. The residual is predicted after regressing abnormal return variance on firm size, leverage, absolute value of unexpected earnings, return on equity, industry, and year dummies. We see that abnormal return variance is similar across groups of firms with different levels of earnings management. Similarly, we plot on Figure 3.12 the abnormal trading volume and find similar results.

In summary, our results suggest that investors can not distinguish good and bad firms in the low information segment based on levels of earnings management.

Secondly, we show that investors react identically to annual reports by firms that accidentally/temporarily fall into the low information segment and firms systematically stay in the low information segment. More specifically, we want to show investors cannot distinguish between future stayers and escapers. Here, we define escapers as of those firms that moves from low information segment to high information segment in the next year, and stayers as of those firms that still stay in the low information segment in the next year.

$$Y_t = \alpha + \beta_1 Escaper_{t+1} + \beta_2 Controls_t + \varepsilon_t \tag{3.6}$$

As shown in equation (9), we define *Escaper* as a dummy variable that equals 1 if a

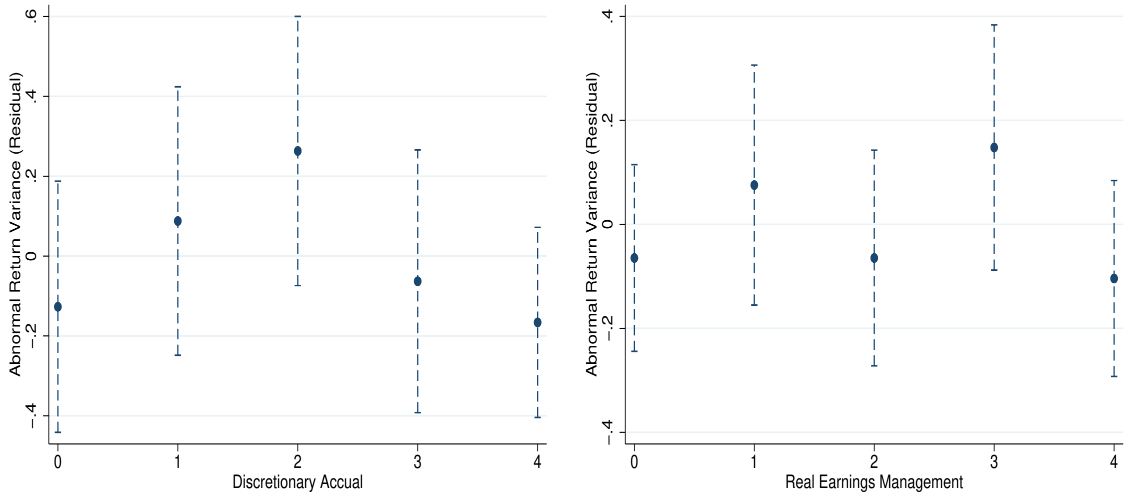


Figure 3.11. Abnormal Return Variance for Firms with Different Levels of Earnings Management in the Low Information Segment

Note: We divide all firms in the low information segment into 5 subsamples using the level of discretionary accrual and real earnings management in ascending order. The five dots are the average of abnormal return variance (residual) for each subsample. Residual is predicted after regressing abnormal return variance on firm size, leverage, absolute value of unexpected earnings, return on equity, industry, and year dummies. We find that investors can not distinguish good and bad firms in the low information segment based on levels of earnings management.

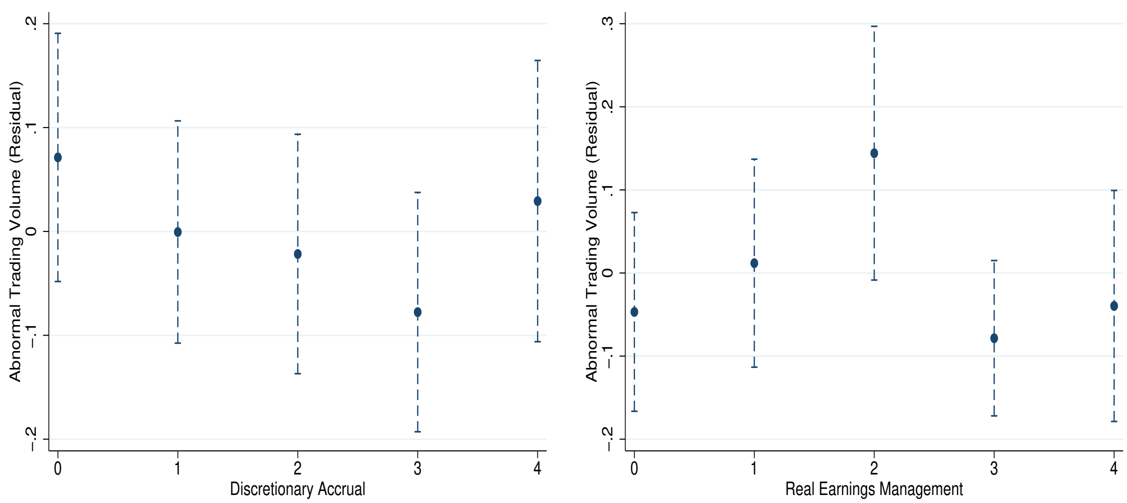


Figure 3.12. Abnormal Trading Volume for Firms with Different Levels of Earnings Management in the Low Information Segment

Note: We divide all firms in the low information segment into 5 subsamples using the level of discretionary accrual and real earnings management in ascending order. The five dots are the average of abnormal trading volume (residual) for each subsample. Residual is predicted after regressing abnormal trading volume on firm size, leverage, absolute value of unexpected earnings, return on equity, industry, and year dummies. We find that investors respond similarly to firms with different levels of earnings management in the low information segment in terms of abnormal trading volume around the dates of annual earnings announcement firm moves out of low information segment next year, and equals to 0 if a firm stays in the low information segment next year. We restrict our sample to only include these two types of firms. Y_t remains to be our short term financial measures and long term price informativeness measures. The significance of β_1 here indicates whether investors react differently to future escapers and stayers in the low information segment. Table 4 shows that none of our reaction measures differs significantly between future stayers and escapers. In other words, investors cannot accurately distinguish relatively good firms from bad firms in the low information segment.

Table 3.4. Investors Can Not Distinguish Escapers vs Stayers

	(1) ab_ret_var	(2) ab_trade_vol	(3) ΔR^2	(4) $\Delta\beta$	(5) $\Delta R^2/R^2$	(6) $\Delta\beta/\beta$
Escaper	-0.211	0.0136	-0.0103	-0.0000114	-0.0269	0.00273
	(0.183)	(0.0898)	(0.0120)	(0.0129)	(0.0421)	(0.0143)
ln_asset	-0.180**	-0.00757	0.0130***	0.0123**	0.0641***	0.00958*
	(0.0701)	(0.0334)	(0.00462)	(0.00493)	(0.0162)	(0.00549)
leverage	-0.250	0.0471	0.0477*	-0.0377	0.168*	-0.0402
	(0.400)	(0.194)	(0.0263)	(0.0281)	(0.0921)	(0.0313)
B/M	0.210**	-0.0824*	-0.0332***	-0.0153**	-0.137***	-0.0154**
	(0.0918)	(0.0434)	(0.00605)	(0.00646)	(0.0212)	(0.00719)
ROE	8.706	-1.505	-0.0843	-0.319	-1.255	-0.320
	(6.015)	(2.903)	(0.396)	(0.423)	(1.386)	(0.471)
Constant	5.658***	1.792***	-0.287***	-0.227**	-1.166***	-0.142
	(1.445)	(0.687)	(0.0953)	(0.102)	(0.333)	(0.113)
Observations	2544	1848	2544	2544	2544	2544
Adjusted R^2	0.002	0.001	0.010	0.003	0.015	0.002

Note: We have six dependent variables, respectively abnormal return variance (ab_ret_var), abnormal trading volume (ab_trade_vol), level change of β after firm's annual report ($\Delta\beta$), level change of R^2 after firm's annual report (ΔR^2), percent change of β after firm's annual report ($\Delta\beta/\beta$), percent change of R^2 after firm's annual report ($\Delta R^2/R^2$). Our sample is all the firms in the low information segment. *Escaper* is a dummy variable that equals 1 if a firm moves out of low information segment in the next year, and equals to 0 if a firm stays in the low information segment in the next year. Standard errors in parentheses. * (p<0.10), ** (p<0.05), *** (p<0.01).

3.8 Causal Evidence on Effects of Falling into Low Information Segment

Firms listed in China are divided into high and low information segment due to its delisting policy. In section 6, we show that firms in the low information segment suffer from adverse financial effects compared to firms in the high information segment. We attribute the adverse financial effects for firms in the low information segment to firms' massive earnings management and investors' distrust.

Obviously, it is natural to think that firms in the low information segment differ in many other dimensions from firms in the high information segment. We proceed in two steps to mitigate this endogeneity problem. Firstly, we control for as many firm observables as possible. In our regressions, we control for firm size, market to book ratio, unexpected earnings, leverage, industry and year fixed effects.

However, simply controlling for firm characteristics is not sufficient for causal inference. Our goal is to identify the impact of falling into low information segment on firms. We need to make sure that firm unobservables are not driving our results. The ideal experiment is to negatively shock some firms from the high information segment into the low information segment. For example as in figure 3.13, pick two firms from the high information segment at year t: firm A with 8% ROE and firm B with 13% ROE. We give both of them a negative 5% ROE shock in year t+1. In year t+1, firm A falls into the low information segment since now its ROE is below 4% whereas firm B stays in the high information segment. We can then compare investors' reaction to their announcement of ROE in year t+1 to determine the impact of falling into the low information segment. One might argue that firm A and B are different firms since they have different ROEs in year t which bias our results. Hence, we design a difference in difference estimation strategy to get rid of time-invariant firm fixed effects. More specifically, we first measure the change in investors' reaction from year t to t+1 for both firm A and B. We then take another difference between firm A's change and firm B's change. The difference in firm A and B's differences is the impact of falling into low information segment. Our identifying assumption is that time-varying firm effects do not impact firm A and B differently. We will manage to present evidence on that front.

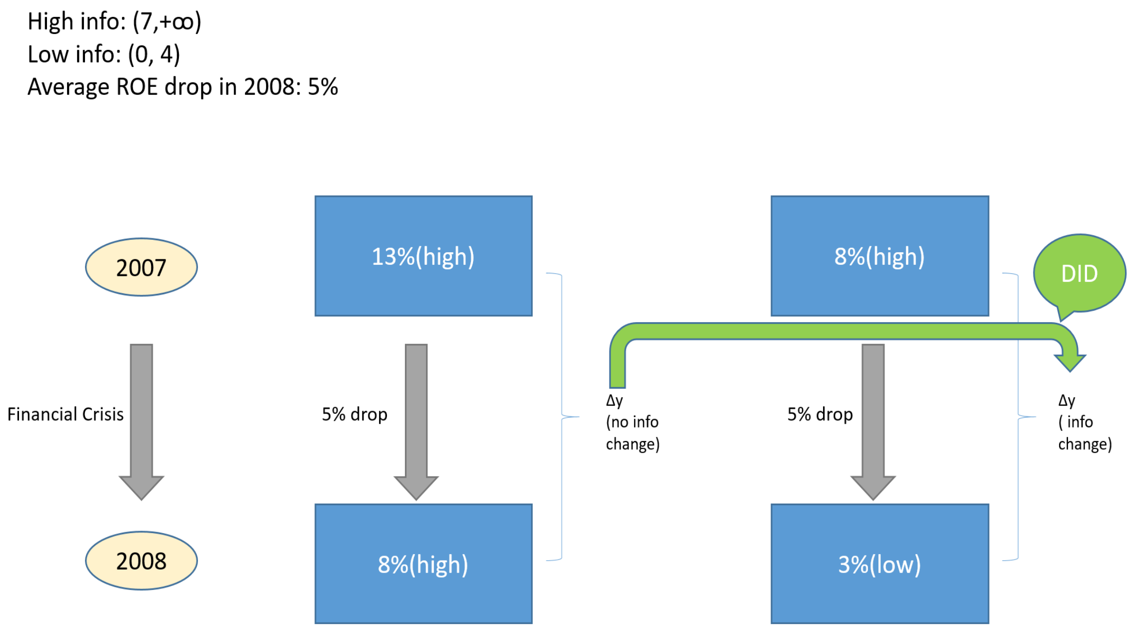


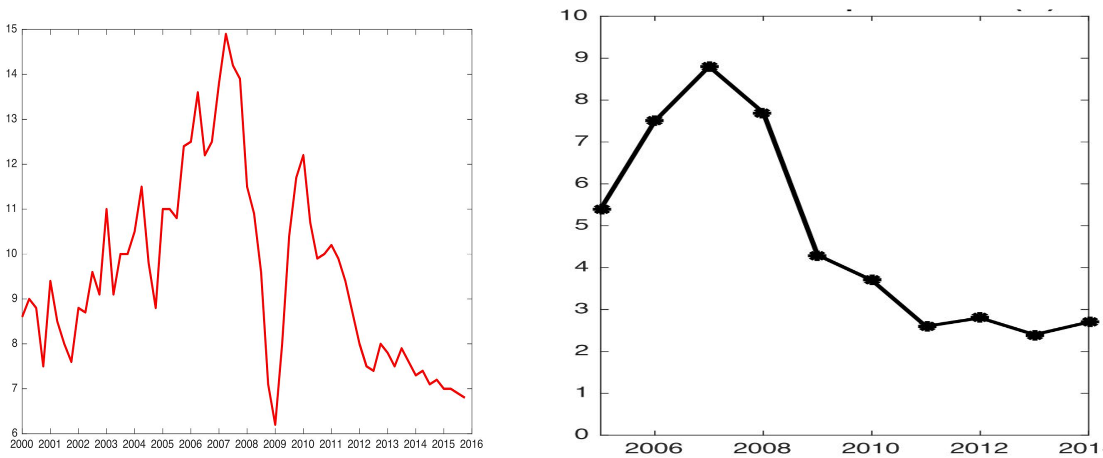
Figure 3.13. DID Design

The key ingredient of our identification strategy is a large exogenous negative shock to firms' ROE. We will first explain why the 2007-08 global financial crisis can be seen as an exogenous shock to firms listed in China. Afterwards, we implement the difference in differences estimation strategy to identify the effects of falling low information segment.

3.8.1 Why is the 2007-08 Global Financial Crisis an Exogenous Shock to Firms Listed in China?

China's booming export-driven economy took an unexpected hard hit in 2008 by the financial crisis (Chong-en et al., 2016). Figure 3.14 shows that China's average quarterly GDP growth rate from 2003 to 2007 had been over 10 %. However, China's quarterly GDP growth rate dropped from 13.9% in 2007Q4 to 7.1% in 2008Q4. In the meanwhile, export as a ratio of GDP also declined from 9% in 2007 to 8% in 2008. Hence, the financial crisis can be viewed as an major negative foreign demand shock to Chinese firms.

Moreover, it is reasonable to view financial crisis as an exogenous shock to listed firms in China since it was caused by sub-prime mortgage defaults in the U.S.. In addition, average ROE for all firms listed in China was 8.5% in 2007 and dived to 5.1% in 2008. The drop in ROE from 2007 to 2008 is even larger for firms in the tradable sector such as manufacturing. There were over 700 listed manufacturing firms in China in 2007. Their average ROE fell by over 4% due to the financial crisis, going from 10.7% in 2007 to 6.5% in 2008. In summary, the 2007/08 financial crisis is both an exogenous and sizeable negative shock to China's listed firms.



(a) China's Quarterly GDP Growth

(b) China's Annual Export/GDP Ratio

Figure 3.14. Impact of the 2007/08 Financial Crisis on China's Economy

Note: these two graphs are from Chong-en et al. (2016)

3.8.2 Estimation Strategy

Forecasting Model of ROE

Looking at each firm's ROE in year 2007 and 2008, we are able to identify a group of firms that were in the high information segment in year 2007 and dropped to low information segment in 2008. However, we can not say that every firm in this group plunges into the low information segment due to an exogenous shock. Some firms might switch from high to low information segment even without the financial crisis as a shock. Hence, we need to eliminate firms that switch information segment due to endogenous reasons unrelated to the financial crisis. More specifically, we define our treatment group as firms that fell into high information segment in 2007, forecasted to be staying in the high information segment in 2008, but actually fell into low information segment in 2008. In the contrary, we define our control group as firms that fell into high information segment in 2007, forecasted to be staying in the high information segment in 2008, and actually fell into high information segment in 2008.

We follow Fama and French (2000) in constructing our forecasting model of firm's ROE. We use the model to forecast each firm's ROE in 2008 based only on information available in 2007. The financial crisis came in as an unexpected shock to listed firms in China. If a firm that is forecasted to stay in the high information segment in 2008 but in reality dropped to low information segment, we are confident that this firm fell into the low information segment due to an exogenous reason that is not related to firm's fundamentals.

The forecasting model of firm's ROE has two stages. For the first stage, we regress $E(Y_t/BE_t)$ for the firms in our sample on variables meant to capture differences across firms in expected profitability for each year t. BE_t is a firm's total book equity at the end of year t; Y_t is earnings before interest and extraordinary items but after taxes. We then use the fitted values from this first-stage regression as the proxy for $E(Y_t/BE_t)$ for year t.

$$Y_t/BE_t = d_0 + d_1VE_t/BE_t + d_2VE_{t-1}/BE_{t-1} + d_3DD_t + d_4D_t/BE_t + \varepsilon_t \qquad (3.7)$$

We use three variables to explain expected profitability $E(Y_t/BE_t)$. (i) D_t/BE_t is the ratio of year t dividends to the book value of common equity at the end of the year. (ii) Fama and French (1999) find that firms that do not pay dividends tend to be much less profitable than dividend payers. Our second variable is a dummy, DD_t , that is 0 for dividend payers and 1 for nonpayers. (iii) We use the market-to-book equity ratio, VE_t/BE_t, to pick up variation in expected profitability missed by the dividend variables. Here VE_t is the firm's market equity value. We develop the model in two aspects: first, we add up the lagged term VE_{t-1}/BE_{t-1} to allow intertemporal effect of market-to-book equity ratio; second, we estimate the parameters d_0, d_1, d_2 and d_0 in a three year window to exclude short term noises. Also, we scale annual net income by book equity instead of book asset.

Table 5 shows the result for our first stage regression. We need $E(Y_t/BE_t)$ for both 2006 and 2007 to construct our second stage forecasting model. Similar to Fama and French (2000), we observe higher profitability associated with dividend payers and higher dividend payout ratio. Moreover, we get a positive contemporary and a negative lagged effect of market-to-book equity ratio.

Table 3.5. First stage regression for 2006 and 2007

	(1) 2006	(2) 2007
VE_t/BE_t	-0.0388***	-0.00917***
	(0.00138)	(0.000927)
$\mathrm{VE}_{t-1}/BE_{t-1}$	0.0279***	0.0193***
	(0.00148)	(0.00155)
DD_t	-0.0804***	-0.0751***
	(0.00905)	(0.00887)
D_t/BE_t	1.569***	1.280***
	(0.164)	(0.162)
Constant	0.0619***	0.0506***
	(0.00826)	(0.00810)
Observations	3847	3892
Adjusted R^2	0.270	0.136

Note: The independent variable is Y_t/BE_t. BE_t is a firm's total book equity at the end of year t. Y_t is earnings before interest and extraordinary items but after taxes. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

For the second stage, we use the following model based on the mean reversion in profitability.

$$CP_{t+1} = a + b_1 DFE_t + b_2 NDFE_t + b_3 SNDFE_t + b_4 SPDFE_t + c_1 CP_t + c_2 NCP_t + c_3 SNCP_t + c_4 SPCP_t + e_{t+1} \quad (3.8)$$

$CP_t = Y_t/BE_t - Y_{t-1}/BE_{t-1}$ is the change in profitability from $t-1$ to t; and $DFE_t =$

$Y_t/BE_t - E(Y_t/BE_t)$ is the deviation of profitability from its expected value; all other explanatory variables include negative deviations of profitability from its expected value ($NDFE_t$), squared negative deviations ($SNDFE_t$), squared positive deviations ($SPDFE_t$), negative changes in profitability (NCP_t), squared negative changes ($SNCP_t$), and squared positive changes ($SPCP_t$). Here, b_2, b_3, b_4 measure nonlinearity in the mean reversion of profitability, that is, in the speed of adjustment of profitability to its expected value. And c_2, c_3, and c_4 measure nonlinearity in the autocorrelation of changes in profitability.

For the financial crisis shock, we first estimate equation (11) using CP_{2007} as our independent variable and then forecast CP_{2008} with all explanatory variables in 2007. Using CP_{2008} as our forecast ROE change without financial crisis, we are able to classify firms that are exogenously shocked to fall into the low information segment. Table 6 shows the result for our second stage regression.

Difference in Differences Estimation

Here is our estimation equation:

$$Y_{it} = \alpha + \beta_1 * Post + \beta_2 * Treatment + \beta_3 * Post * Treatment + Controls_{it} + \varepsilon_i \tag{3.9}$$

where t= 2007 or 2008. i denotes firms listed in China. We only keep firms that have data in both year 2007 and 2008. Y_{it} is our outcome variable that can either be a financial effect or a real effect. *Post* is a dummy variable that equals 1 if year=2008 and 0 if year= 2007. We define our treatment group to be firms that were in the high information segment in 2007, forecasted to be in the high information segment in 2008, and actually fell into the low information segment in 2008. Respectively, our control group consists of firms that were also in the high information segment in 2007, forecasted to be in the high information segment in 2008, and actually stayed in the high information segment in 2008. More specifically, treatment =1 if ROE(07) $>$ 7%, forecasted ROE (08) $>$ 7%, and ROE(08) $\in$ (0, 4%). Respectively, treatment=0 if ROE(07) $>$ 7%, forecasted ROE (08) $>$ 7%, and ROE(08) $>$ 7%. We further restrict our control group

Table 3.6. Second stage regression

	(1) CP_{t+1}	(2) CP_{t+1}	(3) CP_{t+1}	(4) CP_{t+1}
DFE_t	-0.469***	0.0982	-0.312***	0.0904
	(0.0295)	(0.0675)	(0.0298)	(0.0702)
CP_t	-0.0867***	-0.0807***	-0.0454	-0.186**
	(0.0291)	(0.0266)	(0.0875)	(0.0877)
$NDFE_t$		-1.183***		-1.069***
		(0.140)		(0.165)
$SNDFE_t$		-0.239**		-0.252**
		(0.106)		(0.116)
$SPDFE_t$		-0.226***		-0.251***
		(0.0732)		(0.0763)
NCP_t			-0.0347	0.298*
			(0.163)	(0.169)
$SNCP_t$			0.845***	0.522***
			(0.181)	(0.189)
$SPCP_t$			0.0948	0.157
			(0.0964)	(0.0977)
Constant	0.0559***	-0.00394	0.0303***	0.00189
	(0.00481)	(0.00658)	(0.00587)	(0.00688)
Observations	1211	1211	1211	1211
Adjusted R^2	0.317	0.438	0.406	0.442

Note: The independent variable is CP_{2007}. We then use the parameters obtained in 2007 to forecast CP_{2008}. $CP_t = Y_t/BE_t - Y_{t-1}/BE_{t-1}$ is the change in profitability from $t-1$ to t; and $DFE_t = Y_t/BE_t - E(Y_t/BE_t)$ is the deviation of profitability from its expected value; all other explanatory variables include negative deviations of profitability from its expected value ($NDFE_t$), squared negative deviations ($SNDFE_t$), squared positive deviations ($SPDFE_t$), negative changes in profitability (NCP_t), squared negative changes ($SNCP_t$), and squared positive changes ($SPCP_t$). Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

(treatment=0) to be firms whose ROE in 2008 is lower than their ROE in 2007. Basically, we remove from our control group all the firms that had an increase in ROE from 2007 to 2008 so that our control group is more comparable to treatment group. We include controls such as market to book ratio, firm size, leverage, difference in ROE from 2007 to 2008, etc. All continuous variables are winsorized at 1% and 99% before all regressions. 2 The key coefficient of interest is β_3 which measures the difference in treatment and control's differences from 2007 to 2008. In other words, β_3 measures the impact of falling into low information segment due to an exogenous shock of ROE.

3.8.3 Regression Results

Table 7 shows the result for our difference-in-differences regression around 2008 financial crisis. For short-term reaction measures, both abnormal return variance and abnormal trading volume correspond to a significant negative coefficient. The mean of abnormal return variance and abnormal trading volume in 2007 are respectively 2.11 and 1.70, which indicates a 46.4% drop of abnormal return variance and a 22.2% drop of abnormal trading volume when a firm moves from the high information segment to the low information segment. On the contrary, for long-term financial measures, both $\Delta\beta/\beta$ and $\Delta R^2/R^2$ correspond to a significant positive coefficient. There is an 8% increase of β and a 16.4% increase of R^2 when a firm moves from the high information segment to the low information segment. As we expected, firms will co-move more with market index after falling into low information segment. Investors tend not to trust these firm's disclosure, which leads to a smaller proportion of firm level reliable information compared with the market overall influence. Moreover, firms falling into low information segment bear a higher systematic risk β and also a higher realized cost of equity.

Table 3.7. DID for financial crisis shock

	(1) ab_ret_var	(2) ab_trade_vol	(3) $\Delta\beta$	(4) ΔR^2	(5) $\Delta\beta/\beta$	(6) $\Delta R^2/R^2$
post	-1.042***	-0.795***	-0.0243	-0.195***	-0.0199	-0.395***
	(0.250)	(0.109)	(0.0196)	(0.0184)	(0.0215)	(0.0434)
treatment	0.472*	0.153	0.00270	-0.00936	-0.00237	-0.0368
	(0.286)	(0.129)	(0.0225)	(0.0210)	(0.0245)	(0.0496)
post× treatment	-0.978**	-0.377**	0.0728**	0.0648**	0.0805**	0.164**
	(0.398)	(0.176)	(0.0312)	(0.0292)	(0.0341)	(0.0690)
ln_asset	0.00173	-0.108***	0.0146**	0.0219***	0.0174**	0.0443***
	(0.0825)	(0.0354)	(0.00648)	(0.00607)	(0.00708)	(0.0143)
Firm Leverage	-0.939*	0.0137	-0.0259	0.0757*	-0.0109	0.226**
	(0.560)	(0.246)	(0.0439)	(0.0411)	(0.0480)	(0.0970)
B/M	0.464***	0.134*	-0.00254	-0.0368***	-0.0108	-0.0853***
	(0.168)	(0.0733)	(0.0132)	(0.0124)	(0.0144)	(0.0291)
ROE	0.0677	0.276	0.109	0.0369	0.0819	0.329
	(1.206)	(0.526)	(0.0947)	(0.0886)	(0.103)	(0.209)
ΔROE	-0.335	-0.840*	0.00545	0.185**	0.0174	0.238
	(1.090)	(0.479)	(0.0856)	(0.0802)	(0.0935)	(0.189)
Constant	2.168	4.027***	-0.306**	-0.419***	-0.352**	-0.838***
	(1.747)	(0.748)	(0.137)	(0.128)	(0.150)	(0.303)
Observations	683	516	683	683	683	683
Adjusted R^2	0.041	0.162	0.016	0.408	0.014	0.349

Note: We have six dependent variables, respectively abnormal return variance (ab_ret_var), abnormal trading volume (ab_trade_vol), level change of β after firm's annual report ($\Delta\beta$), level change of R^2 after firm's annual report (ΔR^2), percent change of β after firm's annual report ($\Delta\beta/\beta$), percent change of R^2 after firm's annual report ($\Delta R^2/R^2$). *Post* is a dummy variable that equals 1 if year=2008 and 0 if year= 2007. We define our treatment group to be firms that were in the high information segment in 2007, forecasted to be in the high information segment in 2008, and actually fell into the low information segment in 2008. Respectively, our control group consists of firms that were also in the high information segment in 2007, forecasted to be in the high information segment in 2008, and actually stayed in the high information segment in 2008. Standard errors in parentheses. * (p<0.10), ** (p<0.05), *** (p<0.01).

3.9 Conclusion

China's stock market is critical in allocating capital and aggregating firm level information efficiently. However, the efficiency of its stock market is severely held back by government policies and regulations. In this paper, we focus on the delisting policy in China's stock market, which is based on firms' reported earnings and hence incentivizes firms to engage in massive earnings management to stay listed.

We propose and test the spillover effect of earnings management by a set of firms on market reaction to other similar firms. More specifically, we show that the delisting policy endogenously divides China's stock market into high and low information segments. We document significant adverse consequences of firms falling into the low information segment including lower stock market investors' reaction, lower cumulative abnormal return around earnings announcements, insignificant earnings response coefficient, lower trading liquidity, higher systematic risk, and higher synchronicity. Our results can be supported by causal evidence using the 2007-08 financial crisis in the U.S. as an exogenous shock to listed firms in China.

Acknowledgement

This chapter, in part, is currently being prepared for submission for publication of the material. Chen, Jun. Liu, Yibin. The dissertation author was one of the primary investigators and authors of this paper.

We thank seminar participants at 2019 AAA Annual Meeting, 2019 American Finance Association Annual Meeting (Poster Session), 2018 the 13^{th} International Conference on Asia-Pacific Financial Markets (**Best Doctoral Paper**), 2018 UC San Diego Economics, 2018 UC San Diego Rady Finance Brownbag.

Appendix A

A.1 Constructing proxies for accrual and real earnings management

A.1.1 Proxies for accrual earnings management

Following Dechow et al. (1995), I use modified cross-sectional Jones model (Jones (1991)) to estimate discretionary accruals within each fiscal year and Fama-French 48 industries:

$$\frac{Total\ \ Accruals_{i,t}}{Assets_{i,t-1}} = \alpha_0 + \alpha_1 \frac{1}{Assets_{i,t-1}} + \alpha_2 \frac{\Delta REV_{i,t}}{Assets_{i,t-1}} + \alpha_3 \frac{PPE_{i,t}}{Assets_{i,t-1}} + \varepsilon_{i,t},$$

where i indexes firm and t indexes fiscal year. Total accruals ($Total\ \ Accruals_{i,t}$) are defined as earnings before extraordinary items and discontinued operations minus operating cash flows (taken from the statement of cash flows). $Assets_{i,t-1}$ is total assets at the end of year $t-1$. $\Delta REV_{i,t}$ is change in revenues from year $t-1$ to t. $PPE_{i,t}$ is the gross value of property, plant, and equipment. Each industry-year regression requires at least ten observations.

The estimated coefficients are used to compute the fitted normal accruals ($NA_{i,t}$) :

$$NA_{i,t} = \hat{\alpha}_0 + \hat{\alpha}_1 \frac{1}{Assets_{i,t-1}} + \hat{\alpha}_2 \frac{\Delta REV_{i,t} - \Delta AR_{i,t}}{Assets_{i,t-1}} + \hat{\alpha}_3 \frac{PPE_{i,t}}{Assets_{i,t-1}} + \varepsilon_{i,t}.$$

As in Dechow et al. (1995), $\Delta AR_{i,t}$ (change in accounts receivable from year $t-1$ to t is subtracted from the changes in revenues since credit sales might also provide a potential opportunity for accounting manipulation. The firm-year-specific discretionary accruals are calculated as $DA_{i,t} = Total\ \ Accruals_{i,t}/Assets_{i,t-1} - NA_{i,t}$. Following Cohen et al. (2008a), I compute the absolute value of discretionary accruals to proxy for earnings management, as well as positive and negative discretionary accruals.

For robustness checks, I use two alternative measures of discretionary accruals. Firstly, I replace $\Delta REV_{i,t}/Assets_{i,t-1}$ with $(\Delta REV_{i,t} - \Delta AR_{i,t})/Assets_{i,t-1}$ in the first stage regression. Secondly, I follow Kothari et al. (2005) in adjusting the estimated discretionary accruals for performance. I match each sample firm with another firm which is from the same fiscal year industry and has the closest return on assets as the given firm. The performance-matched discretionary accruals are then computed as each sample firm's discretionary accruals minus the discretionary accruals of the matched firm. My results are robust to using these two alternative measures.

A.1.2 Proxies for real earnings management

Real earnings management refers to management actions that deviate from normal operational practices and are undertaken with the primary objective of meeting certain earnings thresholds (Roychowdhury (2006), Zang (2012)).

Following Roychowdhury (2006), I examine three major components of real earnings management: abnormal cash flow from operations (*CFO*),[1] abnormal production costs (*Prod.*),[2] and abnormal discretionary expenses (*Disc.*).[3] I estimate the normal levels of CFO, production

[1]Firms can accelerate the timing of sales through price discounts or more lenient credit terms which temporarily increase earnings in the current periods. However, both price discounts and more lenient credit terms will lower cash flows in the current period after controlling for change in sales (Roychowdhury, 2006).

[2]Firms can overproduce to inflate earnings. Overproduction spreads the fixed overhead costs over a larger amount of units and thus lowering fixed costs per unit. Under the assumption that reduction in fixed costs per unit is more substantial than potential increases in the marginal cost per unit, overproduction reduces the cost of goods sold and hence increases earnings (Roychowdhury, 2006).

[3]Firms can boost current period earnings by cutting back on or slowing the growth of discretionary expenditures including R&D, advertising, and selling, general, and administrative (SG&A) expenditures (Roychowdhury, 2006).

costs, and discretionary expenses using the models developed by Dechow et al. (1998) as implemented in Roychowdhury (2006). More specifically, I run the following three cross-sectional regressions for each industry and year with at least ten observations to estimate normal level of CFO, production costs, and discretionary expenses, respectively. For CFO:

$$\frac{CFO_t}{A_{t-1}} = \alpha_0 + \alpha_1\left(\frac{1}{A_{t-1}}\right) + \alpha_2\left(\frac{S_t}{A_{t-1}}\right) + \alpha_3\left(\frac{\Delta S_t}{A_{t-1}}\right) + \varepsilon_t$$

where CFO_t is cash flow from operations in period t. A_{t-1} is the total assets in year $t-1$. S_t is sales in year t. ΔS_t is the change in sales from year $t-1$ to t.

Secondly, I estimate the normal level of production costs using the following regression:

$$\frac{PROD_t}{A_{t-1}} = \alpha_0 + \alpha_1\left(\frac{1}{A_{t-1}}\right) + \alpha_2\left(\frac{S_t}{A_{t-1}}\right) + \alpha_3\left(\frac{\Delta S_t}{A_{t-1}}\right) + \alpha_4\left(\frac{\Delta S_{t-1}}{A_{t-1}}\right) + \varepsilon_t$$

where $PROD_t$ is the sum of the cost of goods sold in year t and the change in inventory from $t-1$ to t. The residual is then used as a proxy for abnormal production costs. The higher the residual is, the larger is the amount of inventory overproduction, and the more significant the increase in reported earnings through reduction of the cost of goods sold.

Lastly, I estimate the normal level of discretionary expenditures using the following regression:

$$\frac{DISX_t}{A_{t-1}} = \alpha_0 + \alpha_1\left(\frac{1}{A_{t-1}}\right) + \alpha_2\left(\frac{S_{t-1}}{A_{t-1}}\right) + \varepsilon_t$$

where $DISX_t$ is the discretionary expenditures (i.e., the sum of R&D, adverting, and SG&A expenditures) in year t.[4] The abnormal level of discretionary expenditures is measured as the estimated residuals from the regression.

[4]As long as SG&A is available, advertising expense and R&D are set as zero if missing.

Appendix B

B.1 Proofs

Lemma 1. *For any equilibrium Γ, there exists an equivalent equilibrium Γ' such that $a_1' = t'$.*

Lemma 1, which is proved in steps below, demonstrates that we can restrict attention to equilibria in which $a_1 = t$ where the lowest element of the partition exactly coincides with no-disclosure.

Lemma 5. *The following statements hold:*

(i) Let $\tilde{x}$ be a continuous random variable on an open interval Y of $[0,1]$. Let $h(\cdot)$ be the density of $\tilde{x}$. Then for any $b \in Y$,

$$\int_Y (x - \mathbb{E}(\tilde{x}|\tilde{x} \in Y))^2 h(x)dx > \int_{Y \cap [0,b]} (x - \mathbb{E}(\tilde{x}|\tilde{x} \in Y \cap [0,b]))^2 h(x)dx + \int_{Y \cap (b,1]} (x - \mathbb{E}(\tilde{x}|\tilde{x} \in Y \cap (b,1]))^2 h(x)dx. \quad \text{(B.1)}$$

(ii) In any maximal equilibrium Γ, (ii.a) $t \in [a_1, a_2)$ and (ii.b) there exists an equivalent equilibrium Γ' such that $a_1' = t'$ and $a_i' = a_i$ for any $i \geqslant 2$.

Part (i) of Lemma 5 shows that it is always strictly better for the investors to be able to choose a partition with more elements (no matter where the cutoff is).

Proof of Part (i). Let $Y_1 = Y \cap [0,b]$ and $Y_2 = Y \cap (b,1]$. Then $Y = Y_1 \cup Y_2$ and $Y_1 \cap Y_2 = \emptyset$. Let $E_1 = \mathbb{E}(x|x \in Y_1)$ and $E_2 = \mathbb{E}(x|x \in Y_2)$. Let $m = \int_{Y_1} h(x)dx = (\int_{Y_1} h(x)dx)/(\int_Y h(x)dx) \in (0,1)$ and $1-m = \int_{Y_2} h(x)dx = (\int_{Y_2} h(x)dx)/(\int_Y h(x)dx) \in (0,1)$. Then $\mathbb{E}(x|x \in Y) = E_1 m + E_2(1-m)$. So the LHS of Eq (B.1) can be written as

$$
\begin{aligned}
&\int_Y (x - \mathbb{E}(x|x \in Y))^2 h(x)dx \\
=&\int_Y (x - E_1 m - E_2(1-m))^2 h(x)dx \\
=&\int_Y x^2 h(x)dx - 2(E_1 m + E_2(1-m))\int_Y xh(x)dx + (E_1 m + E_2(1-m))^2 \\
=&\int_Y x^2 h(x)dx - (E_1 m + E_2(1-m))^2.
\end{aligned}
$$

The RHS of Eq (B.1) can be written as

$$
\begin{aligned}
&\int_{Y\cap[0,b]} (x - \mathbb{E}(x|x \in Y \cap [0,b]))^2 h(x)dx + \int_{Y\cap(b,1]} (x - \mathbb{E}(x|x \in Y \cap (b,1]))^2 h(x)dx \\
=&\int_{Y_1} (x-E_1)^2 h(x)dx + \int_{Y_2} (x-E_2)^2 h(x)dx \\
=&[\int_{Y_1} x^2 h(x)dx + \int_{Y_2} x^2 h(x)dx] - 2[E_1 \int_{Y_1} xh(x)dx + E_2 \int_{Y_2} xh(x)dx] + [E_1^2 m + E_2^2(1-m)] \\
=&\int_Y x^2 h(x)dx - 2[E_1^2 m + E_2^2(1-m)] + [E_1^2 m + E_2^2(1-m)] \\
=&\int_Y x^2 h(x)dx - (E_1^2 m + E_2^2(1-m)).
\end{aligned}
$$

Because the quadratic function is strictly convex, $(mE_1 + (1-m)E_2)^2 < mE_1^2 + (1-m)E_2^2$. Hence the LHS of Eq (B.1) is strictly greater than the RHS of Eq (B.1), which completes the proof. □

We then prove (ii.a) by the following two results. First, we show that the first equilibrium cutoff a_1 is less than or equal to t.

Claim 2. *In any equilibrium, $a_1 \leqslant t$.*

Proof. Given an equilibrium information set $\{a_i\}_{i=1}^{I-1}$, suppose that $t < a_1$. Then the investors receive message ND if the firm gets no signal or a signal less than t; message v for v greater than or equal to t. Let

$$q \equiv (1-p) + p\int_{a_0}^{t} f(v)dv$$

and

$$r \equiv (1-p) + p\int_{a_0}^{a_1} f(v)dv.$$

By Eq. (2.1),

$$\begin{aligned}
P(1) &= [q/r][\mu\frac{1-p}{q} + \int_{a_0}^{t} v\frac{f(v)}{(q-(1-p))/p}dv\frac{q-(1-p)}{q}] + [(r-q)/r][\int_{t}^{a_1} v\frac{f(v)}{(r-q)/p}dv] \\
&= \mu\frac{1-p}{r} + \frac{p}{r}\int_{a_0}^{t} vf(v)dv + \frac{p}{r}\int_{t}^{a_1} vf(v)dv \\
&= \mu\frac{1-p}{r} + \frac{p}{r}\int_{a_0}^{a_1} vf(v)dv.
\end{aligned}$$

If the firm changes the disclosure policy to $t' = a_1$, then the firm saves cost $c[p\int_t^{a_1} f(v)dv] > 0$ for $p > 0$. But the first element $P(1)'$ is

$$\begin{aligned}
P(1)' &= \mu\frac{1-p}{r} + \int_{a_0}^{a_1} v\frac{f(v)}{(r-(1-p))/p}dv\frac{r-(1-p)}{r} \\
&= \mu\frac{1-p}{r} + \frac{p}{r}\int_{a_0}^{a_1} vf(v)dv \\
&= P(1).
\end{aligned}$$

It is clear that the market price does not change in other partition elements under these two disclosure policies, because the firm always reveals the signal if it exists. So by Eq. (2.2), the policy t' gives the firm an expected off higher than the policy t by the amount $c[p\int_t^{a_1} f(v)dv]$. Hence it must be that $t \geqslant a_1$, i.e., there is no value in disclosing below a_1. □

Next, we show that the second cutoff a_2 is greater than t.

Claim 3. *In any equilibrium, $a_2 > t$.*

Proof. Fix an equilibrium information set $\{a_i\}_{i=1}^{I-1}$. Suppose that $a_2 \leqslant t$. Consider another information set $\{a_i'\}_{i=1}^{I-1}$ such that $a_i' = a_{i+1}$ for $i = 1, \ldots, I-2$ and $a_{I-1}' > a_{I-2}' = a_{I-1}$ for some $a_{I-1} < a_{I-1}' < \infty$. We claim that the investors will get a strictly higher expected payoff from $\{a_i'\}_{i=1}^{I-1}$ than $\{a_i\}_{i=1}^{I-1}$ given the firm's disclosure cutoff t. Let $a_j \leqslant t < a_{j+1}$ for $j \geqslant 2$.

It is clear that the price in the elements $k = j+1, \ldots, I-2$ of $\{a_i'\}_{i=1}^{I-1}$ or $k+1 = j+2, \ldots, I-1$ of $\{a_i\}_{i=1}^{I-1}$ satisfy $\hat{P}(k)' = (\int_{a_{k-1}'}^{a_k'} v f(v) dv)/(\int_{a_{k-1}'}^{a_k'} f(v) dv) = (\int_{a_k}^{a_{k+1}} v f(v) dv)/(\int_{a_k}^{a_{k+1}} f(v)$ $\hat{P}(k+1)$. So the market will respond with the same price for values between $a_j' = a_{j+1} > t$ and $a_{I-2}' = a_{I-1}$ when the firm gets a signal. Then by Eq (2.2), the investors' expected payoff will be the same in this case.

Furthermore, when the firm gets a signal, the value will be revealed only if $v \geqslant t$. Since $a_j \leqslant t$, the investors are able to distinguish the firm's signal (if revealed) from *ND* under $\{a_i\}_{i=1}^{I-1}$. So the price $\hat{P}(1)$ for the disclosure *ND* (under $\{a_i\}_{i=1}^{I-1}$) is determined by

$$
\begin{aligned}
\hat{P}(1) &= \mu \frac{1-p}{(1-p) + p\int_0^t f(v) dv} + \int_0^t v \frac{f(v)}{\int_0^t f(v) dv} dv \frac{p \int_0^t f(v) dv}{(1-p) + p\int_0^t f(v) dv} \\
&= \frac{1-p}{(1-p) + p\int_0^t f(v) dv} \mu + \frac{p}{(1-p) + p\int_0^t f(v) dv} \int_0^t v f(v) dv.
\end{aligned}
$$

Since $j \geqslant 2$, $a_1' \leqslant a_{j-1}' = a_j \leqslant t$. Then the investors are able to distinguish the firm's signal from *ND* under $\{a_i'\}_{i=1}^{I-1}$ as well and the price $\hat{P}(1)'$ for the disclosure *ND* is exactly the same as $\hat{P}(1)$ because of the same expression. Hence the market will respond with price $\hat{P}(1)' = \hat{P}(1)$ if the firm's signal value is below t or if there is no signal. Then the investors' expected payoff will be the same as well.

Let us now consider the cases where $t \leqslant v < a_j' = a_{j+1}$ and the firm gets a signal. Since $a_{j-1}' = a_j \leqslant t < a_{j+1} = a_j'$, the price $\hat{P}(j+1)$ and $\hat{P}(j)'$ are determined by

$$
\hat{P}(j+1) = \int_t^{a_{j+1}} v \frac{f(v)}{\int_t^{a_{j+1}} f(v) dv} dv = \int_t^{a_j'} v \frac{f(v)}{\int_t^{a_j'} f(v) dv} dv = \hat{P}(j)'.
$$

So the market will respond with price $\hat{P}(j)' = \hat{P}(j+1)$ when the firm's signal is between t and $a'_j = a_{j+1}$. Then the investors' expected payoff is still the same in this case.

Finally, investors will get a strictly higher expected payoff for signals greater than $a'_{I-2} = a_{I-1}$ by Lemma 5 (i). Because the distribution f has a positive measure in this region, the investors can do strictly better from $\{a'_i\}_{i=1}^{I-1}$ than $\{a_i\}_{i=1}^{I-1}$, which contradicts the equilibrium assumption. Therefore we conclude that $a_2 > t$. □

Part (ii.b) of Lemma 5 implies that every equilibrium outcome can be supported by a strategy profile involving $t = a_1$. Hence it is without loss of generality to restrict our attention to monotonic equilibria in which $t = a_1$.

Proof of Part (ii.b). By Claims 2 and 3, the equilibrium disclosure policy satisfies $a_1 \leqslant t < a_2$. We show that if there is an equilibrium in which $a_1 < t < a_2$, there is another equilibrium in which $a'_1 = t$ with everything else the same. Consider the proposed strategy of investors $\{a'_i\}_{i=1}^{I-1}$ such that $a'_1 = t$ and $a'_j = a_j$ for $j = 2, \ldots, I-1$.

First, we show that the firm has no incentive to deviate from t given $\{a'_i\}_{i=1}^{I-1}$. Because $a'_1 = t$, the firm does not want to disclose more, i.e. choosing a lower cutoff, by the similar argument as the proof of Lemma 2. If the firm gains by choosing a larger cutoff, it is then greater than a'_1 and also a_1. Note that the firm gets the same expected payoff from t given the two information sets, because $a_1 < t < a_2$ and $a'_1 = t < a'_2 = a_2$. Moreover, the firm still gets the same expected payoff from any cutoff t' greater than t, because $a_1 < a'_1 < t'$ (*ND* is sent if there is no signal or the value is less than t' and the two information sets only differ in the first cutoff). So if there is a profitable deviation to a larger disclosure cutoff under $\{a'_i\}_{i=1}^{I-1}$, there must be a profitable deviation to a larger disclosure cutoff under $\{a_i\}_{i=1}^{I-1}$, which contradicts the equilibrium assumption. Hence the firm has no incentive to deviate from t.

Next, we show that the investors have no incentive to deviate from $\{a'_i\}_{i=1}^{I-1}$ given t. It is clear that the investors get the same expected payoff from $\{a'_i\}_{i=1}^{I-1}$ and from $\{a_i\}_{i=1}^{I-1}$ given t. So if investors have a profitable deviation from $\{a'_i\}_{i=1}^{I-1}$, they must also have a profitable deviation from $\{a_i\}_{i=1}^{I-1}$, which contradicts the equilibrium assumption. Hence the investors have no incentive to deviate from $\{a_i'\}_{i=1}^{I-1}$. Therefore, we have shown that $(t, \{a_i'\}_{i=1}^{I-1})$ is an equilibrium strategy profile. $\square$

Lemma 2. *A solution $\{a_i^\dagger\}$ to program (K_0) satisfies*

$$a_i^\dagger = \frac{\mathbb{E}[\tilde{v}|a_i^\dagger \leqslant \tilde{v} < a_{i+1}^\dagger] + \mathbb{E}[\tilde{v}|a_{i-1}^\dagger \leqslant \tilde{v} < a_i^\dagger]}{2} \tag{2.4}$$

for $i = 2, \ldots, I-1$.

Proof. Assume that the firm will fully disclose the signal when being perfectly informed. The investors minimize the ex-ante loss function given by Eq (2.3). The objective function is rewritten below:

$$p\textstyle\sum_{i=2}^{I}\int_{\hat{a}_{i-1}}^{\hat{a}_i}(v-\mathbb{E}[\tilde{v}|\hat{a}_{i-1}\leqslant\tilde{v}<\hat{a}_i])^2 f(v)dv+(1-p)\int_{\hat{a}_1}^{1}(v-P(1))^2 f(v)dv+\int_{\hat{a}_0}^{\hat{a}_1}(v-P(1))^2 f(v)dv,$$

where $P(1) = \frac{pF(\hat{a}_1)\mathbb{E}(\tilde{v}|\hat{a}_0\leqslant\tilde{v}<\hat{a}_1)+(1-p)\mathbb{E}(\tilde{v})}{pF(\hat{a}_1)+(1-p)}$, $\hat{a}_0 = 0$, and $\hat{a}_I = 1$. The conditional expectation is given by $\mathbb{E}[\tilde{v}|a_{i-1} \leqslant \tilde{v} < a_i] = \int_{a_{i-1}}^{a_i} vf(v)dv / \int_{a_{i-1}}^{a_i} f(v)dv$. Then the first term of Equation (2.3) can be written as

$$\begin{aligned}
&p\sum_{i=2}^{I}\int_{\hat{a}_{i-1}}^{\hat{a}_i}(v-\mathbb{E}[\tilde{v}|\hat{a}_{i-1}\leqslant\tilde{v}<\hat{a}_i])^2 f(v)dv\\
=&p\sum_{i=2}^{I}[\int_{\hat{a}_{i-1}}^{\hat{a}_i}v^2 f(v)dv-2\,\mathbb{E}[\tilde{v}|\hat{a}_{i-1}\leqslant\tilde{v}<\hat{a}_i]\int_{\hat{a}_{i-1}}^{\hat{a}_i}vf(v)dv+\mathbb{E}[\tilde{v}|\hat{a}_{i-1}\leqslant\tilde{v}<\hat{a}_i]^2\int_{\hat{a}_{i-1}}^{\hat{a}_i}f(v)dv]\\
=&p\sum_{i=2}^{I}[\int_{\hat{a}_{i-1}}^{\hat{a}_i}v^2 f(v)dv-2(\int_{\hat{a}_{i-1}}^{\hat{a}_i}vf(v)dv)^2/(\int_{\hat{a}_{i-1}}^{\hat{a}_i}f(v)dv)+(\int_{\hat{a}_{i-1}}^{\hat{a}_i}vf(v)dv)^2/(\int_{\hat{a}_{i-1}}^{\hat{a}_i}f(v)dv)]\\
=&p\sum_{i=2}^{I}[\int_{\hat{a}_{i-1}}^{\hat{a}_i}v^2 f(v)dv-(\int_{\hat{a}_{i-1}}^{\hat{a}_i}vf(v)dv)^2/(\int_{\hat{a}_{i-1}}^{\hat{a}_i}f(v)dv)]
\end{aligned}$$

The cutoff $\hat{a}_i$ for $i = 2, \ldots, I-1$ only appears in the first term of Equation (2.3). Given the other cutoffs, each $\hat{a}_i$ ($i = 2, \ldots, I-1$) minimizes

$$\textstyle\int_{\hat{a}_{i-1}}^{\hat{a}_i}v^2 f(v)dv-(\int_{\hat{a}_{i-1}}^{\hat{a}_i}vf(v)dv)^2/(\int_{\hat{a}_{i-1}}^{\hat{a}_i}f(v)dv)+\int_{\hat{a}_i}^{\hat{a}_{i+1}}v^2 f(v)dv-(\int_{\hat{a}_i}^{\hat{a}_{i+1}}vf(v)dv)^2/(\int_{\hat{a}_i}^{\hat{a}_{i+1}}f(v)dv).$$

Assume that all terms are differentiable and the conditions for Dominated Convergence Theorem are satisfied. So we can take first order condition with respect to $\hat{a}_i$ ($i = 2, \ldots, I-1$) and interchange derivatives and integrals. By Leibniz integral rule,

$$
\begin{aligned}
&\hat{a}_i^2 f(\hat{a}_i) - [2(\int_{\hat{a}_{i-1}}^{\hat{a}_i} vf(v)dv)(\hat{a}_i f(\hat{a}_i))/(\int_{\hat{a}_{i-1}}^{\hat{a}_i} f(v)dv) \\
&-(\int_{\hat{a}_{i-1}}^{\hat{a}_i} vf(v)dv)^2 (f(\hat{a}_i))/(\int_{\hat{a}_{i-1}}^{\hat{a}_i} f(v)dv)^2] \\
&+(-\hat{a}_i^2 f(\hat{a}_i)) - [-2(\int_{\hat{a}_i}^{\hat{a}_{i+1}} vf(v)dv)(\hat{a}_i f(\hat{a}_i))/(\int_{\hat{a}_i}^{\hat{a}_{i+1}} f(v)dv) \\
&+(\int_{\hat{a}_i}^{\hat{a}_{i+1}} vf(v)dv)^2 (f(\hat{a}_i))/(\int_{\hat{a}_i}^{\hat{a}_{i+1}} f(v)dv)^2] \\
=&2\hat{a}_i f(\hat{a}_i)[(\int_{\hat{a}_i}^{\hat{a}_{i+1}} vf(v)dv)/(\int_{\hat{a}_i}^{\hat{a}_{i+1}} f(v)dv) - (\int_{\hat{a}_{i-1}}^{\hat{a}_i} vf(v)dv)/(\int_{\hat{a}_{i-1}}^{\hat{a}_i} f(v)dv)] \\
&+f(\hat{a}_i)[((\int_{\hat{a}_{i-1}}^{\hat{a}_i} vf(v)dv)/(\int_{\hat{a}_{i-1}}^{\hat{a}_i} f(v)dv))^2 - ((\int_{\hat{a}_i}^{\hat{a}_{i+1}} vf(v)dv)/(\int_{\hat{a}_i}^{\hat{a}_{i+1}} f(v)dv))^2] \\
=&f(\hat{a}_i)[(\int_{\hat{a}_i}^{\hat{a}_{i+1}} vf(v)dv)/(\int_{\hat{a}_i}^{\hat{a}_{i+1}} f(v)dv) - (\int_{\hat{a}_{i-1}}^{\hat{a}_i} vf(v)dv)/(\int_{\hat{a}_{i-1}}^{\hat{a}_i} f(v)dv)] \\
&[2\hat{a}_i - (\int_{\hat{a}_i}^{\hat{a}_{i+1}} vf(v)dv)/(\int_{\hat{a}_i}^{\hat{a}_{i+1}} f(v)dv) - (\int_{\hat{a}_{i-1}}^{\hat{a}_i} vf(v)dv)/(\int_{\hat{a}_{i-1}}^{\hat{a}_i} f(v)dv)] = 0.
\end{aligned}
$$

Because $f(a_i) > 0$ for all $a_i \in [0,1]$ and $(\int_{a_i}^{a_{i+1}} vf(v)dv)/(\int_{a_i}^{a_{i+1}} f(v)dv) = \mathbb{E}[\tilde{v}|a_i \leqslant \tilde{v} < a_{i+1}] > \mathbb{E}[\tilde{v}|a_{i-1} \leqslant \tilde{v} < a_i] = (\int_{a_{i-1}}^{a_i} vf(v)dv)/(\int_{a_{i-1}}^{a_i} f(v)dv)$, the optimal $\hat{a}_i$ ($i = 2, \ldots, I-1$) satisfies

$$
\begin{aligned}
\hat{a}_i &= [(\int_{\hat{a}_i}^{\hat{a}_{i+1}} vf(v)dv)/(\int_{\hat{a}_i}^{\hat{a}_{i+1}} f(v)dv) + (\int_{\hat{a}_{i-1}}^{\hat{a}_i} vf(v)dv)/(\int_{\hat{a}_{i-1}}^{\hat{a}_i} f(v)dv)]/2 \\
&= (\mathbb{E}[\tilde{v}|\hat{a}_i \leqslant \tilde{v} < \hat{a}_{i+1}] + \mathbb{E}[\tilde{v}|\hat{a}_{i-1} \leqslant \tilde{v} < \hat{a}_i])/2.
\end{aligned}
$$

for $i = 2, \ldots, I-1$. Next, we derive the solution to $\hat{a}_1$ that are involved in both the endpoints of integration and $P(1)$. The terms that involve $\hat{a}_1$ are

$$
\begin{aligned}
& p\int_{\hat{a}_1}^{\hat{a}_2}(v-\mathbb{E}[\tilde{v}|\hat{a}_1 \leqslant \tilde{v} < \hat{a}_2])^2 f(v)dv + (1-p)\int_{\hat{a}_1}^{1}(v-P(1))^2 f(v)dv + \int_{\hat{a}_0}^{\hat{a}_1}(v-P(1))^2 f(v)dv \\
=& p[\int_{\hat{a}_1}^{\hat{a}_2} v^2 f(v)dv - (\int_{\hat{a}_1}^{\hat{a}_2} v f(v)dv)^2/(\int_{\hat{a}_1}^{\hat{a}_2} f(v)dv)] + (1-p)\int_{\hat{a}_1}^{1}(v-P(1))^2 f(v)dv \\
+& \int_{\hat{a}_0}^{\hat{a}_1}(v-P(1))^2 f(v)dv.
\end{aligned}
$$

We take first order condition with respect to $\hat{a}_1$:

$$
\begin{aligned}
& p(-\hat{a}_1^2 f(\hat{a}_1)) - p[-2(\int_{\hat{a}_1}^{\hat{a}_2} v f(v)dv)(\hat{a}_1 f(\hat{a}_1))/(\int_{\hat{a}_1}^{\hat{a}_2} f(v)dv) \\
+& (\int_{\hat{a}_1}^{\hat{a}_2} v f(v)dv)^2 (f(\hat{a}_1))/(\int_{\hat{a}_1}^{\hat{a}_2} f(v)dv)^2] \\
+& (1-p)[-(\hat{a}_1-P(1))^2 f(\hat{a}_1) + \int_{\hat{a}_1}^{1} 2f(v)(v-P(1))(-\frac{\partial}{\partial \hat{a}_1}P(1))dv] \\
+& (\hat{a}_1-P(1))^2 f(\hat{a}_1) + \int_{\hat{a}_0}^{\hat{a}_1} 2f(v)(v-P(1))(-\frac{\partial}{\partial \hat{a}_1}P(1))dv \\
=& -p\hat{a}_1^2 f(\hat{a}_1) + p(\int_{\hat{a}_1}^{\hat{a}_2} v f(v)dv f(\hat{a}_1)/\int_{\hat{a}_1}^{\hat{a}_2} f(v)dv)[2\hat{a}_1 - \int_{\hat{a}_1}^{\hat{a}_2} v f(v)dv/\int_{\hat{a}_1}^{\hat{a}_2} f(v)dv] \\
+& p(\hat{a}_1-P(1))^2 f(\hat{a}_1) - \frac{\partial}{\partial \hat{a}_1}P(1)[(1-p)\int_{\hat{a}_1}^{1} 2(v-P(1))f(v)dv + \int_{0}^{\hat{a}_1} 2(v-P(1))f(v)dv] = 0,
\end{aligned}
$$

where $P(1) = \frac{pF(\hat{a}_1)\mathbb{E}(\tilde{v}|0\leqslant\tilde{v}<\hat{a}_1)+(1-p)\mathbb{E}(\tilde{v})}{pF(\hat{a}_1)+(1-p)}$ and $\frac{\partial}{\partial \hat{a}_1}P(1) = \frac{p^2 f(\hat{a}_1)F(\hat{a}_1)(\hat{a}_1-\mathbb{E}(\tilde{v}|0\leqslant\tilde{v}<\hat{a}_1))+(1-p)pf(\hat{a}_1)(\hat{a}_1-\mathbb{E}(\tilde{v}))}{(pF(\hat{a}_1)+(1-p))^2}$. Hence the interior solution is characterized by Equation (2.5). By continuity of the loss function, the minimum either attains at the interior solution where the first order condition holds or at 0, because $\hat{a}_1 = 1$ is clearly dominated by $\hat{a}_1 = 0$ given Equation (2.4). Therefore, for the ideal information set of investors, Equation (2.4) is satisfied, and either Equation (2.5) or the corner solution $\hat{a}_1 = 0$ holds. □

Proposition 1. *When $\tilde{v}$ is uniformly distributed, all equilibrium information structures induce monotone partitions on the state space.*

Proof. We show that in the optimal information set, the set that induces any price must be connected with Lebesgue measure one. Suppose that price p is induced by disclosures in intervals $(k-n,k)$ and $(k+x,k+x+m)$, where $m,n>0$, and $x \geqslant 0$. The price formed by rational expectation is $p=(k-n+k+k+x+k+x+m)/4 = k+(m-n)/4+x/2$. We show that x must be zero in the optimal information set. The expected pricing error from these intervals is given by

$$
\begin{aligned}
&\int_{k-n}^{k}(v-p)^2dv+\int_{k+x}^{k+x+m}(v-p)^2dv\\
=&\frac{1}{3}[v-(k+(m-n)/4+x/2)]^3|_{k-n}^{k}\\
+&\frac{1}{3}[v-(k+(m-n)/4+x/2)]^3|_{k+x}^{k+x+m}\\
=&\frac{1}{3}[-((m-n)/4+x/2)^3+(n+(m-n)/4+x/2)^3\\
&+(x+m-((m-n)/4+x/2))^3)-(x-((m-n)/4+x/2))^3)]\\
=&\frac{1}{3}[3((m-n)/4+x/2)n((m-n)/4+x/2+n)\\
&+3(x-((m-n)/4+x/2))m(x-((m-n)/4+x/2)+m)+m^3+n^3]\\
=&(\frac{x}{2}+\frac{m-n}{4})(\frac{x}{2}+\frac{m+3n}{4})n\\
+&(\frac{x}{2}-\frac{m-n}{4})(\frac{x}{2}+\frac{3m+n}{4})m+\frac{1}{3}(m^3+n^3)\\
=&\frac{m+n}{4}x^2+\frac{(m+n)^2}{4}x-\frac{3(m-n)^2(m+n)}{16}+\frac{1}{3}(m^3+n^3).
\end{aligned}
$$

It is then clear that x should be minimized at zero. Furthermore, the pricing error from these intervals that induce the same price is strictly increasing in x. Hence if there is a set that is not connected in the state space, we can always permute the intervals so that each set is connected, which reduces the expected pricing error. $\square$

Lemma 6 demonstrates that types withholding their signals in a monotonic equilibrium are the lowest ones, which is useful in proving Proposition 3.

Lemma 6. *If a type v induces the nondisclosure price in a monotonic equilibrium, then the nondisclosure price is induced by all types below v as well.*

Proof. Suppose that a type $v' < v$ induces a price p' different than the nondisclosure price in a monotonic equilibrium. The nondisclosure price cannot be lower than p' in a monotonic equilibrium, because the nondisclosure price is induced by higher informed types and all uninformed types. If the nondisclosure price is higher than p', then the type v' will have a strict incentive to deviate to nondisclosure, which contradicts the equilibrium definition. So the nondisclosure price is induced by all types below v in equilibrium. □

Proposition 3. *The equilibrium with the lowest disclosure cutoff gives investors the highest expected payoff over all equilibria that induce interval partitions.*

Proof. We focus on equilibria with interval structures in which the induced price is the same in each interval. Let t be the highest type in the maximal equilibrium that induces the nondisclosure price.[1] Suppose that investors gain a strictly higher payoff in another equilibrium with a higher cutoff type $t' > t$. Because any type above t will disclose the signal if being informed in the maximal equilibrium by Lemma 6 and $t' > t$, the investor can raise a_1 and induce the same information set as the "better" equilibrium, which would generate a strictly higher payoff than the maximal equilibrium by the hypothesis. The profitable deviation implies that the maximal "equilibrium" strategy profile is actually not an equilibrium. Hence the maximal equilibrium (equilibrium with the lowest cutoff) is optimal among all equilibria from the perspective of investors. □

Lemma 7. *If Inequality (2.11) does not hold, the equilibrium disclosure cutoff t is greater than $a_1^{\dagger}$.*

Proof. By Lemma 1, we focus on the equilibrium in which $t = a_1$ without loss of generality.

[1] If there is no disclosure cost (but some probability that the firm is not informed), some types below t might disclose *but would still induce the nondisclosure price*. They are indifferent between disclosure or not because the same price is induced. If there is a positive disclosure cost, no type below t will disclose her signal.

Suppose by contradiction that Inequality (2.11) does not hold in equilibrium, i.e.,

$$\mathbb{E}(\tilde{v}|\tilde{v} \in [a_1^\dagger, a_2^\dagger)) - c < \frac{pF(a_1^\dagger)\mathbb{E}(\tilde{v}|\tilde{v} \leqslant a_1^\dagger) + (1-p)\mathbb{E}(\tilde{v})}{pF(a_1^\dagger) + (1-p))},$$

and the disclosure cutoff t is less than or equal to $a_1^\dagger$. By Lemma 2, the investors must best respond to a cutoff no greater than $a_1^\dagger$ by choosing the information set $\{a_i^\dagger\}_{i=1}^{I-1}$. The firm will then be strictly worse off if revealing the signal in $[a_1^\dagger, a_2^\dagger)$ than concealing it by the hypothesis. The profitable deviation for the firm shows that the disclosure cutoff t must be greater than $a_1^\dagger$. □

Lemma 8. *In the maximal equilibrium, the firm is either indifferent between disclosing the signal in A_2 or withholding it, or $t = a_1^\dagger$.*

Proof. Because we restrict our attention to equilibrium with $t = a_1$, it must be that

$$\mathbb{E}(\tilde{v}|\tilde{v} \in [a_1, a_2)) - c \geqslant \frac{pF(a_1)\mathbb{E}(\tilde{v}|\tilde{v} \leqslant a_1) + (1-p)\mathbb{E}(\tilde{v})}{pF(a_1) + (1-p))},$$

because firm with signal in A_2 is willing to disclose. If Inequality (2.11) holds, then $t = a_1^\dagger$ by the definition of $a^\dagger$. If Inequality (2.11) does not hold, we claim that the weak inequality above must hold with equality in the *maximal* equilibrium. Suppose not. We show that there is another equilibrium with a strictly lower cutoff.

Consider an equilibrium with cutoff t and information set $\{a_i\}_{i=1}^{I-1}$ that is given by Equation (2.9). In other words, $a_2, \ldots, a_{I-1}$ are given by Equation (2.10). Because it is assumed that $\mathbb{E}(\tilde{v}|\tilde{v} \in [a_1, a_2)) - c > \frac{pF(a_1)\mathbb{E}(\tilde{v}|\tilde{v} \leqslant a_1) + (1-p)\mathbb{E}(\tilde{v})}{pF(a_1) + (1-p))}$, by continuity there are $t' < t$ arbitrarily close to t and an information set $\{a_i'\}_{i=1}^{I-1}$ satisfying $a_1' = t'$ and Equation (2.10) such that $\mathbb{E}(\tilde{v}|\tilde{v} \in [a_1', a_2')) - c \geqslant \frac{pF(a_1')\mathbb{E}(\tilde{v}|\tilde{v} \leqslant a_1') + (1-p)\mathbb{E}(\tilde{v})}{pF(a_1') + (1-p))}$. It is clear that the firm is best responding to $\{a_i'\}_{i=1}^{I-1}$. Further, we claim that the investors are best responding as well. By Lemma 7, the disclosure cutoff t is greater than $a_1^\dagger$ when Inequality (2.11) does not hold. Recall that $a_1^\dagger$ minimizes the investors' expected pricing error (given that $a_2, \ldots, a_{I-1}$ are determined by Equation (2.10)). By convexity, the expected pricing error is increasing in a_1 for a_1 greater than $a_1^\dagger$. Hence the investors minimize pricing error by choosing a_1' equal to t' and, together with Lemma 3, are best responding to the firm's disclosure strategy given by t'. Therefore, we construct an equilibrium with a strictly lower cutoff, implying that the original one is not the maximal equilibrium, which completes the proof. $\square$

Lemma 9. *Suppose that Assumption 2 holds. For two sequences $a \equiv \{a_i\}_{i=0}^{I}$ and $a' \equiv \{a_i'\}_{i=0}^{I}$ ($I \geqslant 2$) that solve Equations (2.10) and (2.12), $a_0 = a_0' < a_1' < a_1$ implies that $a_i < a_i'$ for all $2 \leqslant i \leqslant I$.*

Proof. First, we show that $a_0 = a_0' < a_1' < a_1$ implies $a_2 < a_2'$. Choose $\{a_i''\}_{i=0}^{2}$ with $a_1' = a_1'' < a_2'' = a_2$ so that Equation (2.12) holds. Then $a_1' = a_1'' < a_1$. By Assumption 2, $a_0'' < a_0$. So $a_0 = a_0' > a_0''$. It follows that

$$
\begin{aligned}
&\frac{1-p}{(1-p)+pF(a_1')}\mu + \frac{pF(a_1')}{(1-p)+pF(a_1')}\mathbb{E}[\tilde{v}|a_0' \leqslant \tilde{v} < a_1'] \\
>&\frac{1-p}{(1-p)+pF(a_1'')}\mu + \frac{pF(a_1'')}{(1-p)+pF(a_1'')}\mathbb{E}[\tilde{v}|a_0'' \leqslant \tilde{v} < a_1''].
\end{aligned}
$$

Then $\mathbb{E}[\tilde{v}|a_1' \leqslant \tilde{v} < a_2'] - c > \mathbb{E}[\tilde{v}|a_1'' \leqslant \tilde{v} < a_2''] - c$ by Equation (2.12). Hence $a_2' > a_2'' = a_2$.

Next, we show that for $I \geqslant 3$, $a_i < a_i'$ for all $2 \leqslant i \leqslant I$. Suppose by way of contradiction that $a_j \geqslant a_j'$ for some $3 \leqslant j \leqslant I$; suppose further that j is the smallest index greater than 2 such that this inequality is satisfied, so that $a_i < a_i'$ for all i such that $2 \leqslant i < j$. Because $a_2' > a_2$, there must be at least one index $2 \leqslant i < j$ such that $a_i < a_i'$. Choose $\{\hat{a}_i''\}_{i=0}^{j}$ with $a_{j-1}' = \hat{a}_{j-1}'' < \hat{a}_j'' = a_j$ so that Equation (2.10) and (2.12) hold. Then $a_{j-1} < a_{j-1}' = \hat{a}_{j-1}''$ by the definition of j. By Assumption 2, $a_i < \hat{a}_i''$ for all $0 \leqslant i \leqslant j-1$. Furthermore, $\hat{a}_j'' = a_j \geqslant a_j'$ and $a_{j-1}' = \hat{a}_{j-1}''$ by the assumption and definition. Then by Equation (2.10),

$$
\begin{aligned}
\mathbb{E}[v|a_{j-2}' \leqslant v < a_{j-1}'] =& 2a_{j-1}' - \mathbb{E}[v|a_{j-1}' \leqslant v < a_j'] \\
\geqslant& 2\hat{a}_{j-1}'' - \mathbb{E}[v|\hat{a}_{j-1}'' \leqslant v < \hat{a}_j''] \\
=& \mathbb{E}[v|\hat{a}_{j-2}'' \leqslant v < \hat{a}_{j-1}''],
\end{aligned}
$$

which implies that $a'_{j-2} \geqslant \hat{a}''_{j-2}$. So $a'_i \geqslant \hat{a}''_i$ for all $0 \leqslant i \leqslant j-2$ by Assumption 2.[2] Hence $a'_i \geqslant \hat{a}''_i > a_i$ for all $0 \leqslant i \leqslant j-2$, particularly $a'_1 > a_1$, which leads to a contradiction. $\square$

Let $\{\bar{a}_i\}_{i=0}^{I}$ be the sequence that satisfies Equation (2.10) and (2.12). When the cost c is small or probability of being informed p is large, the investors choose the information structure $\{a_i^{\dagger}\}_{i=1}^{I-1}$; When c is large or p is small, the equilibrium partition is given by $\{\bar{a}_i\}_{i=1}^{I-1}$. The following lemmas are useful to prove Proposition 4. Lemma 10 shows that $\bar{a}_i$ increases in c for $i = 1, \ldots, I-1$. We consider two sequences with cost $c > c'$ and show that $a_i^c > a_i^{c'}$ for all $i = 1, \ldots, I-1$ by contradiction.

Lemma 10. *Given I and p, $\bar{a}_1$ of the sequence $\{\bar{a}_i\}_{i=0}^{I}$ is strictly increasing in c.*

Proof. Suppose not. Then there exists $c > c'$ such that $\bar{a}_{I-1}(c) < \bar{a}_{I-1}(c')$. This is because otherwise $\bar{a}_1(c) > \bar{a}_1(c')$ by Assumption 2. Let $\{\bar{a}'_i(c')\}_{i=0}^{I}$ be another sequence that satisfies $\mathbb{E}[\tilde{v}|\bar{a}'_1(c') \leqslant \tilde{v} < \bar{a}'_2(c')] - c' = \frac{1-p}{(1-p)+pF(\bar{a}'_1(c'))}\mu + \frac{pF(\bar{a}'_1(c'))}{(1-p)+pF(\bar{a}'_1(c'))}\mathbb{E}[\tilde{v}|\bar{a}'_0(c') \leqslant \tilde{v} < \bar{a}'_1(c')]$ and $\bar{a}'_i(c') = \bar{a}_i(c)$ for $i = 1, \ldots, I$. Then $\bar{a}'_i(c') < \bar{a}_i(c')$ for all $0 \leqslant i \leqslant I-1$ by Assumption 2, particularly $\bar{a}'_0(c') < \bar{a}_0(c') = \bar{a}_0(c)$. But observe that

$$
\begin{aligned}
\mathbb{E}[\tilde{v}|\bar{a}'_0(c') \leqslant \tilde{v} < \bar{a}_1(c)] &= \mathbb{E}[\tilde{v}|\bar{a}'_0(c') \leqslant \tilde{v} < \bar{a}'_1(c')] \\
&= \left(\mathbb{E}[\tilde{v}|\bar{a}'_1(c') \leqslant \tilde{v} < \bar{a}'_2(c')] - c' - \frac{1-p}{(1-p)+pF(\bar{a}'_1(c'))}\mu\right) \\
&\quad / \frac{pF(\bar{a}'_1(c'))}{(1-p)+pF(\bar{a}'_1(c'))} \\
&> \left(\mathbb{E}[\tilde{v}|\bar{a}'_1(c') \leqslant \tilde{v} < \bar{a}'_2(c')] - c - \frac{1-p}{(1-p)+pF(\bar{a}'_1(c'))}\mu\right) \\
&\quad / \frac{pF(\bar{a}'_1(c'))}{(1-p)+pF(\bar{a}'_1(c'))} \\
&= \left(\mathbb{E}[\tilde{v}|\bar{a}_1(c) \leqslant \tilde{v} < \bar{a}_2(c)] - c - \frac{1-p}{(1-p)+pF(\bar{a}_1(c))}\mu\right) \\
&\quad / \frac{pF(\bar{a}_1(c))}{(1-p)+pF(\bar{a}_1(c))} \\
&= \mathbb{E}[\tilde{v}|\bar{a}_0(c) \leqslant \tilde{v} < \bar{a}_1(c)],
\end{aligned}
$$

[2]This follows immediately by Assumption 2 if $a'_{j-2} > \hat{a}''_{j-2}$. If $a'_{j-2} = \hat{a}''_{j-2}$, $a'_i = \hat{a}''_i$ for all $i \leqslant j-1$ by a straightforward induction argument on Equation (2.10) and the continuity assumption about prior density.

(B.3)

where the first and the third equalities follow from the construction $\bar{a}'_i(c') = \bar{a}_i(c)$ for $i = 1, \ldots, I-1$, the second and the fourth equalities follow from the definition of the sequences, and the inequality follows from $c > c'$. Then $\bar{a}'_0(c') > \bar{a}_0(c)$, which implies a contradiction. Hence $\bar{a}_i(c) > \bar{a}_i(c')$ for $1 \leqslant i \leqslant I-1$ by Assumption 2. □

The next two lemmas show that $\bar{a}_i$ decreases in p for $i = 1, \ldots, I-1$. When the probability of getting a signal is small, the firm withholds the bad signal as if no signal was received, similar to the intuition in Dye (1985) and Jung and Kwon (1988). When the probability of getting a signal is large, the firm knows that no disclosure will be interpreted as an extremely bad signal and hence would like to disclose more. But anticipating that investors will not choose the first cutoff to be lower than $a_1^\dagger$, the firm will set the disclosure cutoff t exactly to be $a_1^\dagger$ to save the cost.

Lemma 11. *For any sequence $\{\bar{a}_i\}_{i=0}^{I}$ with $a_0 = 0$ and $a_I = 1$ such that Equations (2.10) and (2.12) hold,* $\mathbb{E}[\tilde{v}|\bar{a}_1 \leqslant \tilde{v} < \bar{a}_2] - c < \mu$.

Proof. This lemma follows directly from Equation (2.12). Observe that

$$\begin{aligned}\mu &= \mathbb{E}[\tilde{v}|\bar{a}_0 \leqslant \tilde{v} < \bar{a}_1]F(\bar{a}_1) + \mathbb{E}[\tilde{v}|\bar{a}_1 \leqslant \tilde{v} \leqslant \bar{a}_I](1 - F(\bar{a}_1)) \\ &> \mathbb{E}[\tilde{v}|\bar{a}_0 \leqslant \tilde{v} < \bar{a}_1]F(\bar{a}_1) + \mathbb{E}[\tilde{v}|\bar{a}_0 \leqslant \tilde{v} < \bar{a}_1](1 - F(\bar{a}_1)) = \mathbb{E}[\tilde{v}|\bar{a}_0 \leqslant \tilde{v} < \bar{a}_1]\end{aligned}$$

under the continuous distribution F with strictly positive density everywhere. Hence we must have

$$\begin{aligned}\mathbb{E}[\tilde{v}|\bar{a}_1 \leqslant \tilde{v} < \bar{a}_2] - c &= \frac{1-p}{(1-p) + pF(\bar{a}_1)}\mu + \frac{pF(\bar{a}_1)}{(1-p) + pF(\bar{a}_1)}\mathbb{E}[\tilde{v}|\bar{a}_0 \leqslant \tilde{v} < \bar{a}_1] \\ &< \frac{1-p}{(1-p) + pF(\bar{a}_1)}\mu + \frac{pF(\bar{a}_1)}{(1-p) + pF(\bar{a}_1)}\mu = \mu.\end{aligned}$$

$\square$

Lemma 12. *Given I and c, $\bar{a}_1$ of the sequence $\{\bar{a}_i\}_{i=0}^{I}$ is strictly decreasing in p.*

Proof. Suppose not. Then there exists $p > p'$ such that $\bar{a}_{I-1}(p) > \bar{a}_{I-1}(p')$. This is because otherwise $\bar{a}_1(p) < \bar{a}_1(p')$ by Assumption 2. Let $\{\bar{a}'_i(p')\}_{i=0}^{I}$ be another sequence that satisfies $\mathbb{E}[\tilde{v}|\bar{a}'_1(p') \leqslant \tilde{v} < \bar{a}'_2(p')] - c = \frac{1-p'}{(1-p')+p'F(\bar{a}'_1(p'))}\mu + \frac{p'F(\bar{a}'_1(p'))}{(1-p')+p'F(\bar{a}'_1(p'))}\mathbb{E}[\tilde{v}|\bar{a}'_0(p') \leqslant \tilde{v} < \bar{a}'_1(p')]$ and $\bar{a}'_i(p') = \bar{a}_i(p)$ for $i = 1, \ldots, I$. Then $\bar{a}'_i(p') > \bar{a}_i(p')$ for all $0 \leqslant i \leqslant I-1$ by Assumption 2, particularly $\bar{a}'_0(p') > \bar{a}_0(p') = \bar{a}_0(p)$. But observe that

$$
\begin{aligned}
\mathbb{E}[\tilde{v}|\bar{a}'_0(p') \leqslant \tilde{v} < \bar{a}_1(p)] &= \mathbb{E}[\tilde{v}|\bar{a}'_0(p') \leqslant \tilde{v} < \bar{a}'_1(p')] \\
&= (\mathbb{E}[\tilde{v}|\bar{a}'_1(p') \leqslant \tilde{v} < \bar{a}'_2(p')] - c - \mu)/(\frac{p'F(\bar{a}'_1(p'))}{(1-p')+p'F(\bar{a}'_1(p'))}) + \mu \\
&< (\mathbb{E}[\tilde{v}|\bar{a}'_1(p') \leqslant \tilde{v} < \bar{a}'_2(p')] - c - \mu)/(\frac{pF(\bar{a}'_1(p'))}{(1-p)+pF(\bar{a}'_1(p'))}) + \mu \\
&= (\mathbb{E}[\tilde{v}|\bar{a}_1(p) \leqslant \tilde{v} < \bar{a}_2(p)] - c - \mu)/(\frac{pF(\bar{a}_1(p))}{(1-p)+pF(\bar{a}_1(p))}) + \mu
\end{aligned}
$$

$= (\mathbb{E}[\tilde{v}|\bar{a}_1(p) \leqslant \tilde{v} < \bar{a}_2(p)] - c - \frac{1-p}{(1-p)+pF(\bar{a}_1(p))}\mu)/\frac{pF(\bar{a}_1(p))}{(1-p)+pF(\bar{a}_1(p))}$

$= \mathbb{E}[\tilde{v}|\bar{a}_0(p) \leqslant \tilde{v} < \bar{a}_1(p)]$,where the first and the fifth equalities follow from the construction $\bar{a}'_i(c') = \bar{a}_i(c)$ for $i = 1, \ldots, I-1$, the second and the seventh equalities follow from the definition of the sequences, and the inequality follows from $p > p'$ and Lemma 11. Then $\bar{a}'_0(p') < \bar{a}_0(p)$, which implies a contradiction. Hence $\bar{a}_i(p) < \bar{a}_i(p')$ for $1 \leqslant i \leqslant I-1$ by Assumption 2. $\square$

Proposition 4. *The voluntary disclosure cutoff t increases in the disclosure cost c and decreases in the probability of being informed p.*

Proof. By Lemma 1, we can restrict attention to equilibria in which $t = a_1$ without loss of generality. If the cost is small enough so that Inequality (2.11) holds, $t = a_1^\dagger$ which is constant in

c. If the cost is large enough such that Inequality (2.11) no longer holds, the equilibrium cutoffs are given by $\{\bar{a}_i\}_{i=0}^{I}$. By Lemma 10, $\bar{a}_1$ is strictly increasing in the cost c. Hence the disclosure cutoff t is (weakly) increasing in c overall and strictly increasing when c is large.

Similarly, if the probability is large enough so that Inequality (2.11) holds, then $t = a_1^\dagger$ which is decreasing in the probability p. If the probability is small enough such that Inequality (2.11) no longer holds, the equilibrium cutoffs are given by $\{\bar{a}_i\}_{i=0}^{I}$. By Lemma 12, $\bar{a}_1$ is strictly decreasing in p as well. Hence the disclosure cutoff t is decreasing in p. □

Proposition 5. *The expected pricing error is increasing in the disclosure cost c and decreasing in the probability of being informed p. As an example, in the special case of uniform cash flows $\tilde{v}$,*

(i) The pricing error conditional on disclosure is decreasing in the disclosure cost c and increasing in the probability of being informed p;

(ii) For sufficiently large cost, the expected pricing error is first strictly decreasing and then strictly increasing in attention capacity I. The pricing error conditional on disclosure is strictly decreasing in attention capacity I for I sufficiently large, i.e., when Inequality (2.11) does not hold.

Proof of general distributions. When Inequality (2.11) is satisfied, the pricing error is determined by $\{a_i^\dagger\}_{i=0}^{I}$ and not affected by the cost c or probability p.

When Inequality (2.11) is not satisfied, the cutoffs are given by Equations (2.10) and (2.12). In the maximal equilibrium, all types below t induce the nondisclosure price and types above t disclose their signals. By Proposition 4, the cutoff t increases in the cost c and decreases in the probability p of being informed. So when cost increases or probability decreases, the disclosure threshold becomes higher and the investors cannot do better, because fewer types are providing information. We show further that the expected pricing error is *strictly* increasing in the extent of frictions.

The expected pricing error when Inequality (2.11) does not hold is given by

$$\mathbb{E}L \equiv p\sum_{j=2}^{I}\int_{a_{j-1}}^{a_j}(\mathbb{E}[\tilde{v}|a_{j-1}\leqslant\tilde{v}<a_j]-\tilde{v})^2 f(\tilde{v})d\tilde{v}+(1-p)\int_{a_1}^{1}(\tilde{v}-P(1))^2 f(\tilde{v})d\tilde{v}+\int_{a_0}^{a_1}(\tilde{v}-P(1))^2 f(\tilde{v})d\tilde{v},$$

where $P(1) = \frac{pF(a_1)\mathbb{E}(\tilde{v}|a_0\leqslant\tilde{v}<a_1)+(1-p)\mathbb{E}(\tilde{v})}{pF(a_1)+(1-p)}$. Note that $a_1,\ldots,a_{I-1}$ are all functions of c and p. The derivative of $\mathbb{E}L$ with respect to c is given by the chain rule,

$$\begin{aligned}
&\frac{d\,\mathbb{E}L}{dc}\\
=&p\sum_{j=2}^{I-1}\frac{da_j}{dc}\frac{d}{da_j}\Big(\int_{a_{j-1}}^{a_j}(\mathbb{E}[\tilde{v}|a_{j-1}\leqslant\tilde{v}<a_j]-\tilde{v})^2 f(\tilde{v})d\tilde{v}\\
&+\int_{a_j}^{a_{j+1}}(\mathbb{E}[\tilde{v}|a_j\leqslant\tilde{v}<a_{j+1}]-\tilde{v})^2 f(\tilde{v})d\tilde{v}\Big)\\
&+\frac{da_1}{dc}\frac{d}{da_1}\Big(p\int_{a_1}^{a_2}(\mathbb{E}[\tilde{v}|a_1\leqslant\tilde{v}<a_2]-\tilde{v})^2 f(\tilde{v})d\tilde{v}\\
&+(1-p)\int_{a_1}^{1}(\tilde{v}-P(1))^2 f(\tilde{v})d\tilde{v}+\int_{a_0}^{a_1}(\tilde{v}-P(1))^2 f(\tilde{v})d\tilde{v}\Big)
\end{aligned}\tag{B.3}$$

Since $\mathbb{E}[\tilde{v}|a_{j-1}^t\leqslant\tilde{v}<a_j^t]$ and $P(1)$ are the investor's rational pricing to a signal that would minimize the pricing error, and since $a_0\equiv 0, a_I\equiv 1$, it follows by the Envelope Theorem that for $j=2,\ldots,I-1$,

$$\begin{aligned}
&\frac{d}{da_j}\Big(\int_{a_{j-1}}^{a_j}(\mathbb{E}[\tilde{v}|a_{j-1}\leqslant\tilde{v}<a_j]-\tilde{v})^2 f(\tilde{v})d\tilde{v}+\int_{a_j}^{a_{j+1}}(\mathbb{E}[\tilde{v}|a_j\leqslant\tilde{v}<a_{j+1}]-\tilde{v})^2 f(\tilde{v})d\tilde{v}\Big)\\
=&f(a_j)[(\mathbb{E}[\tilde{v}|a_{j-1}\leqslant\tilde{v}<a_j]-a_j)^2-(\mathbb{E}[\tilde{v}|a_j\leqslant\tilde{v}<a_{j+1}]-a_j)^2];
\end{aligned}$$

and

$$\begin{aligned}
&\frac{d}{da_1}\Big(p\int_{a_1}^{a_2}(\mathbb{E}[\tilde{v}|a_1\leqslant\tilde{v}<a_2]-\tilde{v})^2 f(\tilde{v})d\tilde{v}+(1-p)\int_{a_1}^{1}(\tilde{v}-P(1))^2 f(\tilde{v})d\tilde{v}\\
&+\int_{a_0}^{a_1}(\tilde{v}-P(1))^2 f(\tilde{v})d\tilde{v}\Big)\\
=&f(a_1)[(a_1-P(1))^2-p(\mathbb{E}[\tilde{v}|a_1\leqslant\tilde{v}<a_2]-a_1)^2-(1-p)(a_1-P(1))^2].
\end{aligned}$$

Because $(\mathbb{E}[\tilde{v}|a_{j-1}\leqslant\tilde{v}<a_j]-a_j)^2=(\mathbb{E}[\tilde{v}|a_j\leqslant\tilde{v}<a_{j+1}]-a_j)^2$ for all $j=2,\ldots,I-1$ by

(2.10), Equation (B.3) is simplified to

$$\frac{d\mathbb{E}L}{dc} = \frac{da_1}{dc} f(a_1)p[(P(1)-a_1)^2 - (\mathbb{E}[\tilde{v}|a_1 \leqslant \tilde{v} < a_2] - a_1)^2]. \tag{B.4}$$

By Proposition 4, $da_1/dc > 0$. Furthermore, $(P(1)-a_1^\dagger)^2 = (\mathbb{E}[\tilde{v}|a_1^\dagger \leqslant \tilde{v} < a_2^\dagger] - a_1^\dagger)^2$ in the unconstrained problem. As c increases, a_1 will increase but a_2 will decrease by Proposition 4 and Lemma 9. It follows from the continuity of $(P(1)-a_1)^2 - (\mathbb{E}[\tilde{v}|a_1 \leqslant \tilde{v} < a_2] - a_1)^2$ with respect to a_1 that $(P(1)-a_1)^2 > (\mathbb{E}[\tilde{v}|a_1 \leqslant \tilde{v} < a_2] - a_1)^2$ for $a_1 > a_1^\dagger$.[3] So $(P(1)-a_1)^2 - (\mathbb{E}[\tilde{v}|a_1 \leqslant \tilde{v} < a_2] - a_1)^2 > 0$ and $d\mathbb{E}L/dc > 0$ by Equation (B.4).

The comparative static analysis with respect to p uses the similar argument except that $da_1/dp < 0$. The extra term that is the partial derivative of $\mathbb{E}L$ with respect to p is

$$\sum_{j=2}^{I} \int_{a_{j-1}}^{a_j} (\mathbb{E}[\tilde{v}|a_{j-1} \leqslant \tilde{v} < a_j] - \tilde{v})^2 f(\tilde{v}) d\tilde{v} - \int_{a_1}^{1} (\tilde{v} - P(1))^2 f(\tilde{v}) d\tilde{v} < 0.$$

It is hence clear that $d\mathbb{E}L/dp < 0$. □

Proof of Part (i) of the uniform case. The perceived quality of disclosure, however, increases in c and decreases in p. Let $\mathbb{E}(L|v \geqslant t)$ denote the pricing error conditional on disclosure which is given by

$$\mathbb{E}(L|v \geqslant t) \equiv \frac{1}{1-F(a_1)} \sum_{j=2}^{I} \int_{a_{j-1}}^{a_j} (\mathbb{E}[\tilde{v}|a_{j-1} \leqslant \tilde{v} < a_j] - \tilde{v})^2 f(\tilde{v}) d\tilde{v}.$$

Similarly because $(\mathbb{E}[\tilde{v}|a_{j-1} \leqslant \tilde{v} < a_j] - a_j)^2 = (\mathbb{E}[\tilde{v}|a_j \leqslant \tilde{v} < a_{j+1}] - a_j)^2$ for $j = 2, \ldots, I-1$, the derivative of $\mathbb{E}(L|v \geqslant t)$ with respect to c can be simplified to

$$\begin{aligned} \frac{d\mathbb{E}(L|v \geqslant t)}{dc} &= \frac{da_1}{dc} [-\frac{1}{1-F(a_1)} f(a_1)(\mathbb{E}[\tilde{v}|a_1 \leqslant \tilde{v} < a_2] - a_1)^2 \\ &+ \frac{f(a_1)}{(1-F(a_1))^2} \sum_{j=2}^{I} \int_{a_{j-1}}^{a_j} (\mathbb{E}[\tilde{v}|a_{j-1} \leqslant \tilde{v} < a_j] - \tilde{v})^2 f(\tilde{v}) d\tilde{v}] \end{aligned}$$

[3] Note that given a_1 and $a_0 = 0$, a_2 is determined by Equation (2.12). As $a_2, \ldots, a_{I-1}$ declines, there is no a_1 such that $(P(1)-a_1)^2 = (\mathbb{E}[\tilde{v}|a_1 \leqslant \tilde{v} < a_2] - a_1)^2$ with $a_0 = 0$ by Assumption 1.

$$= \text{d a}_1 \frac{}{dc \frac{f(a_1)}{1-F(a_1)} [\mathbb{E}(L|v \geqslant t) - (\mathbb{E}[\tilde{v}|a_1 \leqslant \tilde{v} < a_2] - a_1)^2].}$$

It is clear that the sign of the derivative depends on the average pricing error in the disclosure region and the pricing error at the point a_1. Hence it is in general ambiguous, which depends on the probability density function f. But if the distribution is uniform, $\mathbb{E}(L|v \geqslant t) < (\mathbb{E}[\tilde{v}|a_1 \leqslant \tilde{v} < a_2] - a_1)^2$. The result follows because a_1 is a boundary point and the pricing error in the interior of the partition elements is smaller than the pricing error at the boundary. So the average pricing error is smaller as well. This implies that $d\,\mathbb{E}(L|v \geqslant t)/dc < 0$ from $da_1/dc > 0$ by Proposition 4. Hence it is clear in the case of uniform distribution that the pricing error conditional on disclosure declines in the disclosure cost c, which supports our intuition.

Likewise, we can perform exactly the same analysis for the comparative statics with respect to p. The result is ambiguous in general as well, but the pricing error conditional on disclosure strictly increases in p in the uniform case when Equation (2.11) is not satisfied. □

The pricing error *conditional* on disclosure is not part of the *expected* pricing error from the disclosure region. In the latter case, the pricing error from disclosure strictly decreases (increases) in cost (probability) for any general distribution (that satisfies Assumption 2).

Proof of Part (ii) of the uniform case: (a) Expected pricing error. Finally, let us consider the comparative statics with respect to attention capacity I. We show next that for given cost and probability of being informed, it is not necessary that the investors would strictly prefer equilibrium partitions with more steps (larger I's).[4] The comparative statics of expected pricing error with respect to I are ambiguous, but we find that it is U-shaped if we assume uniform distribution. In general, it is ambiguous (and clearly not monotonic). When investors have more attention, they would not incorrectly classify the marginal discloser as better firms. The firm

[4]Still, the investors base their pricing choice on rational expectations and the prior distribution is fixed. Nevertheless, the equilibria with more steps are not, ceteris paribus, more informative.

then has less incentive to disclose because of less price in response, which is detrimental to the quality of investors' information and can become a dominant force for sufficiently large capacity.

It is clear that the expected pricing error is decreasing in the partition size if the information set is given by $\{a_i^\dagger\}_{i=1}^{I-1}$, i.e., if Inequality (2.11) is satisfied, because the partition is the *unique* optimal information set with more attention capacity and more disclosure. We will examine below the change in expected pricing error as I increases when Inequality (2.11) does not hold.

Fix p and c, and let $\bar{a}(I)$ be the maximal equilibrium of size I. We shall argue that $\bar{a}(I)$ can be continuously deformed to the (maximal) equilibrium of size $I+1$ and express how the expected pricing error changes throughout the deformation.

Let $a^t \equiv (a_0^t, a_1^t, \ldots, a_{I+1}^t)$ be the partition that satisfies

$$\mathbb{E}[\tilde{v}|a_1 \leqslant \tilde{v} < a_2] - c = \frac{pF(a_1)\,\mathbb{E}[\tilde{v}|a_0 \leqslant \tilde{v} < a_1] + (1-p)\mathbb{E}(\tilde{v})}{(1-p) + pF(a_1)} \tag{B.5}$$

for $i = 1$ and

$$a_i = \frac{\mathbb{E}[\tilde{v}|a_i \leqslant \tilde{v} < a_{i+1}] + \mathbb{E}[\tilde{v}|a_{i-1} \leqslant \tilde{v} < a_i]}{2}. \tag{B.6}$$

for $i = 2, \ldots, I-1$ with $a_0^t = 0$, $a_1^t = t$, and $a_{I+1}^t = 1$. If $t = \bar{a}_1(I)$ then $a_I^t = 1$, and if $t = \bar{a}_1(I+1)$ then $a^t = \bar{a}(I+1)$ and (2.10) is satisfied for all $i = 2, \ldots, I$. We will next write down the partial derivative of the expected pricing error $\mathbb{E}L(t)$ with respect to t when $t \in [\bar{a}_1(I), \bar{a}_1(I+1)]$, which is a non-degenerate interval by Lemma 14.

By definition, $\mathbb{E}L(t)$ is given by

$\mathbb{E}L(t) \equiv p\sum_{j=2}^{I+1}\int_{a_{j-1}^t}^{a_j^t}(\mathbb{E}[\tilde{v}|a_{j-1}^t \leqslant \tilde{v} < a_j^t] - \tilde{v})^2 f(\tilde{v})d\tilde{v}$
$+ (1-p)\int_{a_1^t}^{1}(\tilde{v} - P(1))^2 f(\tilde{v})d\tilde{v} + \int_{a_0}^{a_1^t}(\tilde{v} - P(1))^2 f(\tilde{v})d\tilde{v},$
where $P(1) = \frac{pF(a_1^t)\mathbb{E}(\tilde{v}|a_0 \leqslant \tilde{v} < a_1^t) + (1-p)\mathbb{E}(\tilde{v})}{pF(a_1^t) + (1-p)}$. The Envelope Theorem yields

$$\begin{aligned}\frac{d\,\mathbb{E}L(t)}{dt} =& p\sum_{j=2}^{I} f(a_j^t)\frac{da_j^t}{dt}[(\mathbb{E}[\tilde{v}|a_{j-1}^t \leqslant \tilde{v} < a_j^t] - a_j^t)^2 - (\mathbb{E}[\tilde{v}|a_j^t \leqslant \tilde{v} < a_{j+1}^t] - a_j^t)^2] \\ &+ f(a_1^t)\frac{da_1^t}{dt}[(a_1^t - P(1))^2 - p(\mathbb{E}[\tilde{v}|a_1^t \leqslant \tilde{v} < a_2^t] - a_1^t)^2 - (1-p)(a_1^t - P(1))^2].\end{aligned}$$

Note that $(\mathbb{E}[\tilde{v}|a^t_{j-1} \leqslant \tilde{v} < a^t_j] - a^t_j)^2 = (\mathbb{E}[\tilde{v}|a^t_j \leqslant \tilde{v} < a^t_{j+1}] - a^t_j)^2$ for $j = 2, \ldots, I-1$ by (2.10). Then we can simplify the expression to

$$\begin{aligned}\frac{d\,\mathbb{E}L(t)}{dt} =& pf(a^t_1)\frac{da^t_1}{dt}[(a^t_1 - P(1))^2 - (\mathbb{E}[\tilde{v}|a^t_1 \leqslant \tilde{v} < a^t_2] - a^t_1)^2] \\ &+ pf(a^t_I)\frac{da^t_I}{dt}[(\mathbb{E}[\tilde{v}|a^t_{I-1} \leqslant \tilde{v} < a^t_I] - a^t_I)^2 - (\mathbb{E}[\tilde{v}|a^t_I \leqslant \tilde{v} < a^t_{I+1}] - a^t_I)^2].\end{aligned}$$

So the change in the expected pricing error when I increases to $I+1$ is given by

$$\begin{aligned}\Delta(\mathbb{E}L) =& \int_{\bar{a}_1(I)}^{\bar{a}_1(I+1)} \frac{d\,\mathbb{E}L(t)}{dt}dt \\ =& p\int_{\bar{a}_1(I)}^{\bar{a}_1(I+1)} f(a^t_1)\frac{da^t_1}{dt}[(a^t_1 - P(1))^2 - (\mathbb{E}[\tilde{v}|a^t_1 \leqslant \tilde{v} < a^t_2] - a^t_1)^2]dt \\ &+ p\int_{\bar{a}_1(I)}^{\bar{a}_1(I+1)} f(a^t_I)\frac{da^t_I}{dt}[(\mathbb{E}[\tilde{v}|a^t_{I-1} \leqslant \tilde{v} < a^t_I] - a^t_I)^2 - (\mathbb{E}[\tilde{v}|a^t_I \leqslant \tilde{v} < a^t_{I+1}] - a^t_I)^2]dt \\ =& p\int_{\bar{a}_1(I)}^{\bar{a}_1(I+1)} [(a_1 - P(1))^2 - (\mathbb{E}[\tilde{v}|a_1 \leqslant \tilde{v} < a_2] - a_1)^2]f(a_1)da_1 \\ &- p\int_{\bar{a}_I(I+1)}^{1} [(\mathbb{E}[\tilde{v}|a_{I-1} \leqslant \tilde{v} < a_I] - a_I)^2 - (\mathbb{E}[\tilde{v}|a_I \leqslant \tilde{v} < a_{I+1}] - a_I)^2]f(a_I)da_I.\end{aligned} \tag{B.7}$$

Let us take a closer look at the two terms in (B.7). First,

$$(a_1 - P(1))^2 - (\mathbb{E}[\tilde{v}|a_1 \leqslant \tilde{v} < a_2] - a_1)^2 > 0$$

for all $a_1 \in (\bar{a}_1(I), \bar{a}_1(I+1)]$ when c is sufficiently large by Equation (2.11).[5] Further,

$$(\mathbb{E}[\tilde{v}|a_{I-1} \leqslant \tilde{v} < a_I] - a_I)^2 - (\mathbb{E}[\tilde{v}|a_I \leqslant \tilde{v} < a_{I+1}] - a_I)^2 > 0$$

for all $a_I \in [\bar{a}_I(I+1), \bar{a}_I(I))$.

When I is sufficiently large,
$|\bar{a}_1(I+1) - \bar{a}_1(I)| > |\bar{a}_I(I) - \bar{a}_I(I+1)|$ and $(\bar{a}_1(I) - P(1))^2 - (\mathbb{E}[\tilde{v}|\bar{a}_1(I) \leqslant \tilde{v} < \bar{a}_2(I)] - \bar{a}_1(I))^2$
$= \min_{a_1 \in (\bar{a}_1(I), \bar{a}_1(I+1)]} (a_1 - P(1))^2 - (\mathbb{E}[\tilde{v}|a_1 \leqslant \tilde{v} < a_2] - a_1)^2 > (\mathbb{E}[\tilde{v}|\bar{a}_{I-1}(I) \leqslant \tilde{v} < \bar{a}_I(I)] - \bar{a}_I(I))^2 -$

[5] When $c = 0$, $P(1) = \mathbb{E}[\tilde{v}|a_1 \leqslant \tilde{v} < a_2]$. Then the expected pricing error is strictly decreasing in I for all I.

$(\mathbb{E}[\tilde{v}|\bar{a}_I(I) \leqslant \tilde{v} < 1] - \bar{a}_I(I))^2 = \max_{a_I \in (\bar{a}_I(I+1), \bar{a}_I(I)]} (\mathbb{E}[\tilde{v}|a_{I-1} \leqslant \tilde{v} < a_I] - a_I)^2 - (\mathbb{E}[\tilde{v}|a_I \leqslant \tilde{v} < a_{I+1}] - a_I)^2$

. Because $f(a_1) = f(a_I)$, $\Delta \mathbb{E} L$ is positive. □

Proof of Part (ii) of the uniform case: (b) Pricing error conditional on disclosure. The pricing error conditional on disclosure is given by

$$\mathbb{E}(L|v \geqslant t) \equiv \frac{1}{1 - F(a_1^t)} \sum_{j=2}^{I+1} \int_{a_{j-1}^t}^{a_j^t} (\mathbb{E}[\tilde{v}|a_{j-1}^t \leqslant \tilde{v} < a_j^t] - \tilde{v})^2 f(\tilde{v}) d\tilde{v}.$$

The Envelope Theorem yields

$$\begin{aligned}
&\frac{d\,\mathbb{E}(L|v \geqslant t)}{dt} \\
=&\frac{da_1^t}{dt}[-\frac{1}{1-F(a_1^t)} f(a_1^t)(\mathbb{E}[\tilde{v}|a_1^t \leqslant \tilde{v} < a_2^t] - a_1^t)^2 \\
+&\frac{f(a_1^t)}{(1-F(a_1^t))^2} \sum_{j=2}^{I+1} \int_{a_{j-1}^t}^{a_j^t} (\mathbb{E}[\tilde{v}|a_{j-1}^t \leqslant \tilde{v} < a_j^t] - \tilde{v})^2 f(\tilde{v}) d\tilde{v}] \\
&+ \frac{da_I^t}{dt} \frac{f(a_I^t)}{1-F(a_1^t)} [(\mathbb{E}[\tilde{v}|a_{I-1}^t \leqslant \tilde{v} < a_I^t] - a_I^t)^2 - (\mathbb{E}[\tilde{v}|a_I^t \leqslant \tilde{v} < a_{I+1}^t] - a_I^t)^2] \\
=&\frac{da_1^t}{dt} \frac{f(a_1^t)}{1-F(a_1^t)} [\mathbb{E}(L|v \geqslant t) - (\mathbb{E}[\tilde{v}|a_1^t \leqslant \tilde{v} < a_2^t] - a_1^t)^2] \\
+&\frac{da_I^t}{dt} \frac{f(a_I^t)}{1-F(a_1^t)} [(\mathbb{E}[\tilde{v}|a_{I-1}^t \leqslant \tilde{v} < a_I^t] - a_I^t)^2 - (\mathbb{E}[\tilde{v}|a_I^t \leqslant \tilde{v} < a_{I+1}^t] - a_I^t)^2].
\end{aligned}$$

Note that $a_1^t = t$. For the uniform distribution, $\mathbb{E}(L|v \geqslant t) < (\mathbb{E}[\tilde{v}|a_1^t \leqslant \tilde{v} < a_2^t] - a_1^t)^2$ and $(\mathbb{E}[\tilde{v}|a_{I-1}^t \leqslant \tilde{v} < a_I^t] - a_I^t)^2 > (\mathbb{E}[\tilde{v}|a_I^t \leqslant \tilde{v} < a_{I+1}^t] - a_I^t)^2$. So the first term is negative. When Inequality (2.11) holds, both t and a_I^t decreases in I by Lemma 13. So $da_I^t/dt > 0$ and the second term is positive. Hence the sign of $d\,\mathbb{E}(L|v \geqslant t)/dt$ is ambiguous and the effect of change in attention capacity on pricing error conditional on disclosure is indeterminate overall. When Inequality (2.11) does not hold, $da_I^t/dt < 0$. In this case, both terms are negative. So $d\,\mathbb{E}(L|v \geqslant t)/dt < 0$ and the pricing error conditional on disclosure is decreasing in I by Proposition 6. □

The next two lemmas will be used to show Proposition 6.

Lemma 13. *Let $a_0 = 0$ and $a_I = 1$. For unconstrained information sets $a^{\dagger}(I)$ with size I and $a^{\dagger}(I+1)$ with size $I+1$, $a^{\dagger}_{i-1}(I) < a^{\dagger}_{i}(I+1) < a^{\dagger}_{i}(I)$ for all $i = 1,\ldots,I$.*

Proof. That $a^{\dagger}_{i-1}(I) < a^{\dagger}_{i}(I+1)$ follows from Assumption 1. If $a^{\dagger}_{i-1}(I) \geqslant a^{\dagger}_{i}(I+1)$ for some $i = 1,\ldots,I$, then $a^{\dagger}_{I-1}(I) \geqslant a^{\dagger}_{I}(I+1)$ by Assumption 1. This leads to a contradiction of $a^{\dagger}_{0}(I) = a^{\dagger}_{0}(I+1) = 0$.

That $a^{\dagger}_{i}(I+1) < a^{\dagger}_{i}(I)$ for $i = 1,\ldots,I$ is by induction on I. For $I = 1$, the lemma is vacuously true. Suppose that $I > 1$ and that the conclusion of the Lemma is true for all $i = 1,\ldots,I-1$. Let $a^{\dagger}(I+1)$ and $a^{\dagger}(I)$ be as in the statement of the Lemma. Suppose by way of contradiction that $a^{\dagger}_{j}(I+1) \geqslant a^{\dagger}_{j}(I)$ for some j such that $0 < j < I$; suppose further that j is the smallest index greater than 0 such that this inequality is satisfied, so that $a^{\dagger}_{i}(I+1) < a^{\dagger}_{i}(I)$ for all i such that $0 < i < j$. Let ${}^{x}a \equiv ({}^{x}a_j, {}^{x}a_{j+1}, \ldots, {}^{x}a_I)$ be the partial partition that satisfies (2.4) for $i = j+1,\ldots,I-1$ with ${}^{x}a_I = a^{\dagger}_{I}(I) = 1$ and ${}^{x}a_{I-1} = x$. It follows from Assumption 1 and continuity of ${}^{x}a$ in x that there is an $\tilde{x} < a^{\dagger}_{I}(I+1)$ such that $a^{\dagger}_{j}(I+1) = {}^{\tilde{x}}a_j$. Let ${}^{\tilde{x}}a \equiv \tilde{a}$. We assumed that $a^{\dagger}_{j}(I+1) \geqslant a^{\dagger}_{j}(I)$. So $\tilde{a}_i \geqslant a^{\dagger}_{i}(I)$ for $j \leqslant i \leqslant I-1$ by Assumption 1. This implies that there is a unique $\tilde{a}_{j-1} \in [0, \tilde{a}_j)$ such that $\mathbb{E}[\tilde{v}|\tilde{a}_j \leqslant \tilde{v} < \tilde{a}_{j+1}] = 2\tilde{a}_j - \mathbb{E}[\tilde{v}|\tilde{a}_{j-1} \leqslant \tilde{v} < \tilde{a}_j]$ and $\mathbb{E}[\tilde{v}|\tilde{a}_j \leqslant \tilde{v} < \tilde{a}_{j+1}] \leqslant 2\tilde{a}_j - \mathbb{E}[\tilde{v}|a \leqslant \tilde{v} < \tilde{a}_j]$ for $a \leqslant \tilde{a}_{j-1}$. Then $\mathbb{E}[\tilde{v}|\tilde{a}_j \leqslant \tilde{v} < \tilde{a}_{j+1}] \leqslant 2\tilde{a}_j - \mathbb{E}[\tilde{v}|a^{\dagger}_{j-1}(I) \leqslant \tilde{v} < \tilde{a}_j]$ by Assumption 1. Further because $a^{\dagger}_{j}(I+1) = \tilde{a}_j$,

$$
\begin{aligned}
\mathbb{E}[\tilde{v}|\tilde{a}_j \leqslant \tilde{v} < a^{\dagger}_{j+1}(I+1)] =& \mathbb{E}[\tilde{v}|a^{\dagger}_{j}(I+1) \leqslant \tilde{v} < a^{\dagger}_{j+1}(I+1)] \\
=& 2a^{\dagger}_{j}(I+1) - \mathbb{E}[\tilde{v}|a^{\dagger}_{j-1}(I+1) \leqslant \tilde{v} < a^{\dagger}_{j}(I+1)] \\
=& 2\tilde{a}_j - \mathbb{E}[\tilde{v}|a^{\dagger}_{j-1}(I+1) \leqslant \tilde{v} < \tilde{a}_j] \\
>& 2\tilde{a}_j - \mathbb{E}[\tilde{v}|a^{\dagger}_{j-1}(I) \leqslant \tilde{v} < \tilde{a}_j] \\
\geqslant& \mathbb{E}[\tilde{v}|\tilde{a}_j \leqslant \tilde{v} < \tilde{a}_{j+1}],
\end{aligned}
$$

where the second equality follows from the definition of the sequence $\{a^{\dagger}_{i}(I+1)\}_{i=0}^{I+1}$ and the first inequality follows from $a^{\dagger}_{j-1}(I+1) < a^{\dagger}_{j-1}(I)$. So $a^{\dagger}_{j+1}(I+1) > \tilde{a}_{j+1}$. But $a^{\dagger}_{j+1}(I+1) < \tilde{a}_{j+1}$ by the induction hypothesis, because $a^{\dagger}_{j}(I+1) = \tilde{a}_j$ and $a^{\dagger}_{I+1}(I+1) = \tilde{a}_I$. Hence the contradiction establishes the desired conclusion. $\square$

In particular, we show that $a_1^\dagger(I+1) < a_1^\dagger(I)$, which is equal to the disclosure cutoffs under these two attention levels.

Lemma 14. *Let $a_0 = 0$ and $a_I = 1$. For partitions $\bar{a}(I)$ with size I and $\bar{a}(I+1)$ with size $I+1$ that satisfy Equations (2.10) and (2.12), $\bar{a}_1(I) < \bar{a}_1(I+1)$.*

Proof. This lemma follows directly from Lemma 9. Consider two partitions $\{\bar{a}_i(I)\}_{i=0}^{I}$ with size I and $\{\bar{a}_i(I+1)\}_{i=0}^{I+1}$ with size $I+1$ such that Equations (2.10) and (2.12) are satisfied. Suppose by way of contradiction that $\bar{a}_1(I) \geqslant \bar{a}_1(I+1)$. Then $\bar{a}_i(I) \leqslant \bar{a}_i(I+1)$ for all $2 \leqslant i \leqslant I$ by Lemma 9, which contradicts $\bar{a}_I(I) = \bar{a}_{I+1}(I+1) = 1$. $\square$

The only way to have more partition elements is to increase the first cutoff $\bar{a}_1$ (and all subsequent cutoffs will decline). Intuitively, if the partition is finer, the firm's gain from inducing a slightly higher price $\mathbb{E}[\tilde{v}|\bar{a}_1 \leqslant \tilde{v} < \bar{a}_2]$ becomes smaller, which is outweighed by the cost of disclosure.

Proposition 6. *The disclosure cutoff t is first strictly decreasing and then strictly increasing in the partition size I.*

Proof. By Lemma 1, we restrict attention to equilibria in which $t = a_1$ without loss of generality. By Lemma 13, the first disclosure cutoff $a_1^\dagger(I)$ of the optimal information set is strictly decreasing in I. By Lemma 14, the first cutoff $\bar{a}_1(I)$ of the sequence $\{\bar{a}_i(I)\}_{i=0}^{I}$ is strictly increasing in I. So if the investors have very limited attention such that Inequality (2.11) holds, then $t = a_1^\dagger$ is strictly decreasing in I. If the investors are able to pay a lot of attentions to the signal such that Inequality (2.11) does not hold, the equilibrium cutoffs are given by $\{\bar{a}_i\}_{i=0}^{I}$ and $\bar{a}_1$ is strictly increasing in I. Hence the disclosure cutoff t is first strictly decreasing and then strictly increasing in the partition size I. $\square$

Our comparative statics results with respect to the disclosure cost c and probability of being informed p would still hold for cash flow v following a general distribution on an unbounded support. In this case, we focus on the equilibrium with the lowest disclosure threshold. Without imposing the Monotonicity Conditions, it is possible that there are multiple solutions to the optimal information set. Hence multiple equilibria could arise.

Proposition 10. *The voluntary disclosure threshold t increases in the disclosure cost c and decreases in the probability of being informed p.*

Proof. When the constraints are slack, the disclosure threshold t does not depend on the cost c or the probability p. When the constraints bind, the equilibrium information set is given by Equations (2.10) and (2.12). So the equilibrium threshold is a function of c and p. We define L as follows:

$$L = \mathbb{E}[\tilde{v}|a_1 \leqslant \tilde{v} < a_2] - c - [\frac{pF(a_1)\,\mathbb{E}[\tilde{v}|a_0 \leqslant \tilde{v} < a_1] + (1-p)\mathbb{E}(\tilde{v})}{(1-p) + pF(a_1)}]. \tag{B.8}$$

When a_1 and a_2 are part of the (constrained) equilibrium, $L = 0$ by Eq (2.12). We will first sign the derivative $\partial L/\partial a_1$, $\partial L/\partial p$, and $\partial L/\partial c$ in order to determine the sign of $\partial a_1/\partial p$ and $\partial a_1/\partial c$ by the Implicit Function Theorem.

First, let us find the sign of $\partial L/\partial a_1$. If a_1 is close to zero, $\mathbb{E}[\tilde{v}|a_1 \leqslant \tilde{v} < a_2] - c$ is no greater than $\mathbb{E}[\tilde{v}|0 \leqslant \tilde{v} < a_2]$, while $\frac{pF(a_1)\,\mathbb{E}[\tilde{v}|a_0 \leqslant \tilde{v} < a_1] + (1-p)\mathbb{E}(\tilde{v})}{(1-p)+pF(a_1)} \geqslant \mathbb{E}(\tilde{v}) - \varepsilon$ for $\varepsilon > 0$ arbitrarily small by continuity. When $I \geqslant 3$, a_2 is less than 1, which implies that $\mathbb{E}[\tilde{v}|0 \leqslant \tilde{v} < a_2] < \mathbb{E}(\tilde{v}) - \varepsilon$ for $\varepsilon > 0$ small enough. Hence, if a_1 is close to zero, L must be less than zero. If a_1 is close to the upper bound 1, a_2 must be close to the upper bound as well and $\mathbb{E}[\tilde{v}|a_1 \leqslant \tilde{v} < a_2] - c \geqslant 1 - \varepsilon' - c$ for $\varepsilon' > 0$ arbitrarily small. Moreover, $\frac{pF(a_1)\,\mathbb{E}[\tilde{v}|a_0 \leqslant \tilde{v} < a_1] + (1-p)\mathbb{E}(\tilde{v})}{(1-p)+pF(a_1)}$ is no greater than $\mathbb{E}(\tilde{v})$. Whenever there is some type v' who would like to disclose the type, we can find ε' small enough such that $1 - \varepsilon' > v'$ and all types above $1 - \varepsilon'$ (including $1 - \varepsilon'$) would all (strictly) prefer to disclose. Hence $1 - \varepsilon' - c > \mathbb{E}(\tilde{v})$. It follows that $L = \mathbb{E}[\tilde{v}|a_1 \leqslant \tilde{v} < a_2] - c - [\frac{pF(a_1)\,\mathbb{E}[\tilde{v}|a_0 \leqslant \tilde{v} < a_1] + (1-p)\mathbb{E}(\tilde{v})}{(1-p)+pF(a_1)}] > 0$ if a_1 is close to 1. This shows that there is at least one value of a_1 such that $L = 0$ holds. Furthermore, because $L < 0$ when a_1 is close to zero, the function L must cross zero from below at the first solution of a_1 to $L = 0$. So $\partial L/\partial a_1 > 0$.

The derivative of L with respect to p is given by

$$\partial L/\partial p = \quad \frac{F(a_1)(\mathbb{E}(\tilde{v}) - \mathbb{E}[\tilde{v}|a_0 \leqslant \tilde{v} < a_1])}{((1-p)+pF(a_1))^2} > 0.$$

It is then clear that $\partial a_1/\partial p = -(\partial L/\partial p)/(\partial L/\partial a_1) < 0$ and the disclosure threshold is strictly decreasing in p in the constrained case.

The derivative of L with respect to c is given by

$$\partial L/\partial c = -1 < 0.$$

Hence $\partial a_1/\partial c = -(\partial L/\partial c)/(\partial L/\partial a_1) > 0$ and the disclosure threshold is strictly increasing in c in the constrained case. $\square$

Lemma 4. *The value of information to firms V exceeds the social value of information V^*.*

Proof. Let $\mu =: \mathbb{E}(\tilde{v})$ be the mean value of future cash flows. The social value of information V^* is given by

$$\begin{aligned}
V^* &= \int_0^1 [(vx^*(v) - \psi(x^*(v))) - (vx^*(\mu) - \psi(x^*(\mu)))] f(v)dv \\
&= \int_0^1 (vx^*(v) - \psi(x^*(v)))f(v)dv - [x^*(\mu)\int_0^1 vf(v)dv - \psi(x^*(\mu))] \\
&= \int_0^1 (vx^*(v) - \psi(x^*(v)))f(v)dv - (\mu x^*(\mu) - \psi(x^*(\mu))).
\end{aligned}$$

The firm's private value of information is greater than the social value, i.e., $V > V^*$, as shown below:

$$
\begin{aligned}
V &= \int_t^1 [(vx^*(v) - \psi(x^*(v))) - (\mathbb{E}(\tilde{v}|ND)x^*(\mathbb{E}(\tilde{v}|ND)) - \psi(x^*(\mathbb{E}(\tilde{v}|ND))))]f(v)dv \\
&> \int_0^1 [(vx^*(v) - \psi(x^*(v))) - (\mathbb{E}(\tilde{v}|ND)x^*(\mathbb{E}(\tilde{v}|ND)) - \psi(x^*(\mathbb{E}(\tilde{v}|ND))))]f(v)dv \\
&= \int_0^1 (vx^*(v) - \psi(x^*(v)))f(v)dv - (\mathbb{E}(\tilde{v}|ND)x^*(\mathbb{E}(\tilde{v}|ND)) - \psi(x^*(\mathbb{E}(\tilde{v}|ND)))) \\
&> \int_0^1 (vx^*(v) - \psi(x^*(v)))f(v)dv - (\mu x^*(\mu) - \psi(x^*(\mu))) = V^*,
\end{aligned}
$$

where the first inequality follows because $vx^*(v) - \psi(x^*(v)) > (<)\, \mathbb{E}(\tilde{v}|ND)x^*(\mathbb{E}(\tilde{v}|ND)) - \psi(x^*(\mathbb{E}(\tilde{v}|ND)))$ for $v > (<)\ t$, and the second inequality follows because $\mu > \mathbb{E}(\tilde{v}|ND)$ and $vx^*(v) - \psi(x^*(v))$ is strictly increasing in v by the Envelope Theorem. □

Proposition 8. *Suppose that there is only acquisition cost and no disclosure cost, i.e., $c = 0$. For **any** finite information capacity, the value of information to firms is less than the full-information case.*

Proof. Consider the firm's private value of information V_f when investors only have finite information capacity I. Let t be the equilibrium disclosure threshold in the rational (full attention) model. Let t^f be the equilibrium disclosure threshold when the attention capacity is I. Because the firm will disclose any signal above $t^f = a_1$ (and conceal otherwise), the value at capacity I is given by

$$
\begin{aligned}
V_f = &\int_{t^f}^1 (P(i|D(v) \in A_i)x^*(P(i|D(v) \in A_i)) - \psi(x^*(P(i|D(v) \in A_i))))f(v)dv \\
&- \int_{t^f}^1 (P(1)x^*(P(1)) - \psi(x^*(P(1))))f(v)dv
\end{aligned}
$$

where the third equality follows by the formation of market price. Because $vx^*(v) - \psi(x^*(v))$ is strictly convex in v by the convexity of $\psi(\cdot)$,

$$\begin{aligned}
&\mathbb{E}(\tilde{v}|\tilde{v}\in(a_{i-1},a_i])x^*(\mathbb{E}(\tilde{v}|\tilde{v}\in(a_{i-1},a_i]))-\psi(x^*(\mathbb{E}(\tilde{v}|\tilde{v}\in(a_{i-1},a_i])))\\
<&\int_{a_{i-1}}^{a_i}(vx^*(v)-\psi(x^*(v)))\frac{f(v)}{F(a_i)-F(a_{i-1})}dv
\end{aligned}$$

for $i=2,\ldots,I$ by Jensen's Inequality. It is follows that

$$\begin{aligned}
&\sum_{i=2}^{I}(\mathbb{E}(\tilde{v}|\tilde{v}\in(a_{i-1},a_i])x^*(\mathbb{E}(\tilde{v}|\tilde{v}\in(a_{i-1},a_i]))-\psi(x^*(\mathbb{E}(\tilde{v}|\tilde{v}\in(a_{i-1},a_i]))))(F(a_i)-F(a_{i-1}))\\
<&\int_{a_1}^{1}(vx^*(v)-\psi(x^*(v)))f(v)dv\\
=&\int_{t^f}^{1}(vx^*(v)-\psi(x^*(v)))f(v)dv,
\end{aligned}$$

where the equality is implied from the equilibrium condition. Hence the firm's value with finite capacity I satisfies

$$\begin{aligned}
V_f&<\int_{t^f}^{1}[(vx^*(v)-\psi(x^*(v)))-(P(1)x^*(P(1))-\psi(x^*(P(1))))]f(v)dv\\
&\leqslant\int_{t^f}^{1}[(vx^*(v)-\psi(x^*(v)))-(E(\tilde{v}|ND)x^*(\mathbb{E}(\tilde{v}|ND))-\psi(x^*(\mathbb{E}(\tilde{v}|ND))))]f(v)dv\\
&\leqslant\int_{t}^{1}[(vx^*(v)-\psi(x^*(v)))-(E(\tilde{v}|ND)x^*(\mathbb{E}(\tilde{v}|ND))-\psi(x^*(\mathbb{E}(\tilde{v}|ND))))]f(v)dv=V,
\end{aligned}$$

where V is the firm's private value of information in the full attention case. The second inequality above holds because $P(1)\geqslant\mathbb{E}(\tilde{v}|ND)$ (the non-disclosure price in the case of full attention) by the minimum principle and the function $vx^*(v)-\psi(x^*(v))$ is strictly increasing in v. The third inequality holds because

(1) If $t^f \geqslant t$,

$$\begin{aligned}
&\int_{t^f}^{1}[(vx^*(v)-\psi(x^*(v)))-(E(\tilde{v}|ND)x^*(\mathbb{E}(\tilde{v}|ND))-\psi(x^*(\mathbb{E}(\tilde{v}|ND))))]f(v)dv \\
=&\int_{t}^{1}[(vx^*(v)-\psi(x^*(v)))-(E(\tilde{v}|ND)x^*(\mathbb{E}(\tilde{v}|ND))-\psi(x^*(\mathbb{E}(\tilde{v}|ND))))]f(v)dv- \\
&\int_{t}^{t^f}[(vx^*(v)-\psi(x^*(v)))-(E(\tilde{v}|ND)x^*(\mathbb{E}(\tilde{v}|ND))-\psi(x^*(\mathbb{E}(\tilde{v}|ND))))]f(v)dv.
\end{aligned} \tag{B.9}$$

Since $vx^*(v)-\psi(x^*(v))$ is strictly increasing in v, $vx^*(v)-\psi(x^*(v)) > \mathbb{E}(\tilde{v}|ND)x^*(\mathbb{E}(\tilde{v}|ND))-\psi(x^*(\mathbb{E}(\tilde{v}|ND)))$ for all $v > \mathbb{E}(\tilde{v}|ND) = t$, where $\mathbb{E}(\tilde{v}|ND) = t$ follows from the equilibrium condition in the case of full attention. Hence the second term of Equation (B.9) is nonnegative and the result follows.

(2) If $t^f < t$,

$$\begin{aligned}
&\int_{t^f}^{1}[(vx^*(v)-\psi(x^*(v)))-(E(\tilde{v}|ND)x^*(\mathbb{E}(\tilde{v}|ND))-\psi(x^*(\mathbb{E}(\tilde{v}|ND))))]f(v)dv \\
=&\int_{t}^{1}[(vx^*(v)-\psi(x^*(v)))-(E(\tilde{v}|ND)x^*(\mathbb{E}(\tilde{v}|ND))-\psi(x^*(\mathbb{E}(\tilde{v}|ND))))]f(v)dv+ \\
&\int_{t^f}^{t}[(vx^*(v)-\psi(x^*(v)))-(E(\tilde{v}|ND)x^*(\mathbb{E}(\tilde{v}|ND))-\psi(x^*(\mathbb{E}(\tilde{v}|ND))))]f(v)dv.
\end{aligned} \tag{B.10}$$

Since $vx^*(v)-\psi(x^*(v))$ is strictly increasing in v, $vx^*(v)-\psi(x^*(v)) < \mathbb{E}(\tilde{v}|ND)x^*(\mathbb{E}(\tilde{v}|ND))-\psi(x^*(\mathbb{E}(\tilde{v}|ND)))$ for all $v < \mathbb{E}(\tilde{v}|ND) = t$. Hence the second term of Equation (B.10) is negative and the result follows.

Therefore, the firm gains less from acquiring information, which implies the role of inattention in reducing excessive information acquisition. □

Proposition 9. *Let A_1^1 be the first element of the investor's information set in period 1. Let a_1^2 be the first cutoff that the investor selects in period two. Let t_1 be the disclosure threshold in period one. The cutoff a_1^2 will be lower if the investor does not observe A_1^1 in period one or if the realized cash flow in period one falls below the disclosure threshold t_1 (when the observation about the signal is A_1^1).*

Proof. We prove the first part of the proposition. Suppose that the investor observes something other than A_1^1 in the first period. Then they know that the firm does obtain a signal and disclose. So the probability p_2 that the firm receives a signal in period 2 is λ_1. Because $\lambda_1 > \lambda_0$ and the probability $\phi(v_1)$ if A_1^1 is instead observed is a weighted average of λ_0 and λ_1, the probability p_2 attains the highest possible value when the observation is not A_1^1. By Proposition 4, with a higher belief about the signal endowment, the investor will choose a lower cutoff a_1^2 in period 2 in equilibrium.

We next show that if v_1 in period one is lower than t_1, the cutoff a_1^2 in period 2 will be lower than the case in which $v_1 \geqslant t_1$. First, the realization of v_1 (given the *observed disclosure*) does not affect the investor's choice of information set and firm's disclosure in period two if the investor observes something other than A_1^1 in period one, because the cash flows between the two periods are independent and p_1 is known to be λ_1. If the investor observes A_1^1 in period one, the firm either receives no signal or withdraws a low signal. The probability p_2 that the firm gets a signal in period two is given by $\frac{1-\lambda}{M}\lambda_0 + (1-\frac{1-\lambda}{M})\lambda_1$, where $M = (1-\lambda) + \lambda \Pr(s_1 < t_1 | v_1)$. The conditional probability is equal to

$$\Pr(s_1 < t_1 | v_1) = \begin{cases} q(v_1) + (1-q(v_1)) \int_0^{t_1} dG & \text{if } v_1 < t_1 \\ (1-q(v_1)) \int_0^{t_1} dG & \text{if } v_1 \geqslant t_1. \end{cases}$$

If the cash flow is less than t_1, $\Pr(s_1 < t_1 | v_1)$ will be higher and M will be larger, because $q(v_1) + (1-q(v_1)) \int_0^{t_1} dG \geqslant \int_0^{t_1} dG > (1-q(v_1')) \int_0^{t_1} dG$ for any $v_1' \geqslant t_1$ by $q > 0$ and $\int_0^{t_1} dG \leqslant 1$. It follows that the weight $\frac{1-\lambda}{M}$ will be smaller and the complement $1 - \frac{1-\lambda}{M}$ will be larger. Further because $\lambda_1 > \lambda_0$, $\phi(v_1)$ must be larger. Hence the investor believes that the firm has a higher chance to get a signal in period 2, i.e., p_2 is larger. By Proposition 4, with a higher belief about the signal endowment, the investor will choose a lower cutoff a_1^2 in period 2 in equilibrium, which completes the proof. $\square$

B.2 Tables and Figures in Section 2.6

Table B.1. Variable Definitions

Variables	Definitions	Sources
Dependent Variables		
$Management\ Forecast\ (MF_{i,t})$	Indicator equals to 1 if a firm i makes a forecast in year t, and zero otherwise.	I/B/E/S
Variables of Interest		
$Capacity\ (percent)_{i,t-1}$	Percentage of institutional ownership in year $t-1$	Thomson Reuters
$Capacity\ (percent)^2_{i,t-1}$	Squared term of $Capacity\ (percent)$ in year $t-1$	Thomson Reuters
$Capacity\ (ratio)_{i,t-1}$	$(Ins - Ins(LT))/(1 - Ins(LT))$ where Ins is institutional ownership and $Ins(LT)$ ownership by by institutions with > 5 % of shares	Thomson Reuters
$Capacity\ (ratio)^2_{i,t-1}$	Squared term of $Capacity\ (ratio)$ in year $t-1$	Thomson Reuters
Control Variables		
$EPS\ Increase_{i,t-1}$	Indicator equals to one if firm i reports an increase in earnings per share from year $t-2$ to $t-1$	Compustat
$Abs.\ EPS\ Change_{i,t-1}$	Absolute value of the change in earnings per share from year $t-2$ to $t-1$	Compustat
$Book\ to\ Market_{i,t-1}$	Book value of equity / market value of equity	Compustat/CRSP
$Size_{i,t-1}$	Natural log of market capitalization	CRSP
$Loss_{i,t-1}$	Indicator equals to 1 if earnings < 0 in year $t-1$	Compustat
$Leverage\ Ratio_{i,t-1}$	Total liabilities divided by total assets measured for firm i in year $t-1$	Compustat

Table B.2. Sample Selection

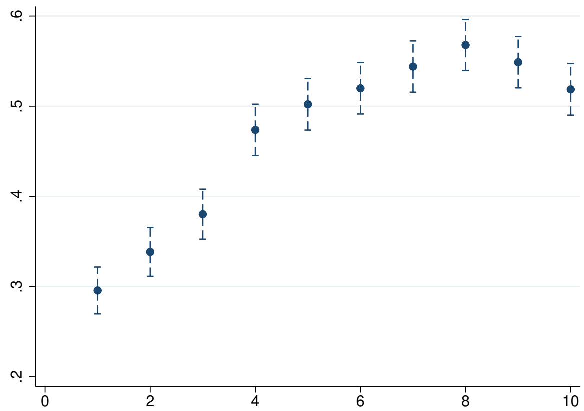
	Details	# Firm-Year	# Firms	# MF
Step 1	I/B/E/S EA sample (US firms) 1/1/2004 - 12/31/2016	67,239	10,945	
	a) : Non-missing current or prior EA date	62,359	10,035	
Step 2	I/B/E/S CIG sample 1/1/2004 - 12/31/2016			
	a) : Matched to I/B/E/S EA	62,359	10,035	70,198
	b) : Keep management forecasts (MF) after prior EA date and at least 6 months before current period end	62,359	10,035	28,787
	c) : Keep only earliest MF	62,359	10,035	12,769
Step 3	Keep obs. with data from CRSP, Compustat, Thompson Reuters	50,703	7,864	11,451
Step 4	Other Sample Selections:			
	a) : Drop firms that always forecast	46,748	7,339	7,392
	b) : Drop firms that never forecast	16,508	2,583	7,392
	Total	**16,508**	**2,583**	**7,392**

Note: this table summarizes our sample selection procedures. Annual earnings announcements (EA) and management forecasts (MF) are obtained from I/B/E/S. Firms in our sample must also have information on prices from CRSP, fundamentals from Compustat, and institutional ownership from Thompson Reuters.

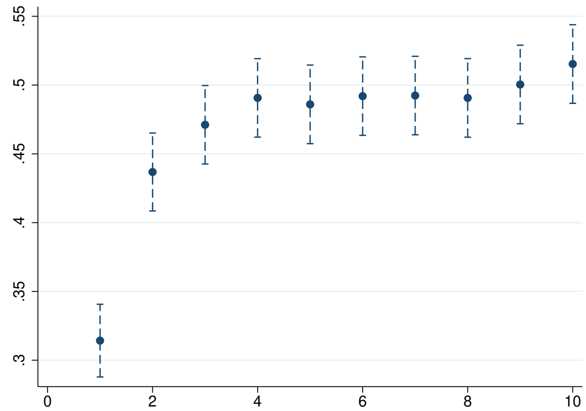
Table B.3. Summary Statistics

	N	mean	sd	p10	p25	p50	p75	p90
Management Forecast	16,508	0.45	0.50	0.00	0.00	0.00	1.00	1.00
Ins Holding Ratio	15,901	0.64	0.30	0.17	0.42	0.68	0.86	0.98
Ins Holding Percent	15,901	0.69	0.28	0.26	0.53	0.75	0.90	0.99
Quasi-indexer	16,508	0.45	0.24	0.03	0.29	0.49	0.63	0.74
Transient	16,508	0.14	0.10	0.00	0.06	0.13	0.21	0.28
Dedicated	16,508	0.04	0.06	0.00	0.00	0.00	0.06	0.12
Earnings Per Share	16,508	1.50	3.04	-0.31	0.32	1.15	2.27	3.72
Total Asset	16,508	8,867.31	27,606.47	110.86	326.13	1,178.02	4,585.02	16,931.30
Leverage Ratio	16,459	0.53	0.23	0.22	0.36	0.53	0.70	0.86
Book to Market	15,323	0.67	5.71	0.17	0.28	0.47	0.73	1.08
Market Cap	15,372	6,145.79	20,694.78	106.27	316.50	1,013.20	3,449.15	12,566.46
Return on Assets	16,508	0.01	0.16	-0.09	0.01	0.04	0.08	0.12
Loss	16,508	0.15	0.36	0.00	0.00	0.00	0.00	1.00
EPS Increase	16,508	0.68	0.47	0.00	0.00	1.00	1.00	1.00

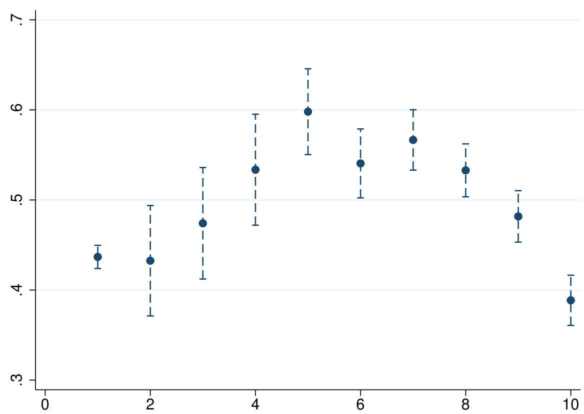
Note: this table reports summary statistics on our sample. Management Forecast is an indicator variable which equals to one when a firm makes a forecast. 45% of all firm-years have management forecasts. Information on institutional ownership is obtained from Thomson Reuters. Firm fundamentals including earnings per share, total asset, leverage ratio, return on assets are from Compustat. Market capitalization is calculated as the product of number of shares times closing price obtained from CRSP.



(a) *Quasi-indexer* $(t-1)$



(b) *Transient* $(t-1)$



(c) *Dedicated* $(t-1)$

Figure B.1. Likelihood of Management Forecast Across Deciles of Three Types of Institutional Ownership

Note: Figure B.1 plots percentage of firms with management forecasts in year t across deciles of ownership by *Quasi-indexers*(sub-figure a), *Transient* investors (sub-figure b), and *Dedicated* investors in year $t-1$. All sub-figures plot the 95% confidence interval around the mean values for each decile.

Table B.4. Investor Attention and Management Forecast

This table presents results from estimating the relation between investor attention and the likelihood of management forecast using the following specification:

$$MF_{i,t} = \alpha_t + \alpha_j + \beta Capacity_{i,t-1} + \gamma Capacity^2_{i,t-1} + Controls_{i,t-1} + \varepsilon_{i,t}$$

where α_t is year fixed effect and α_j industry fixed effect. The dependent variable $MF_{i,t}$ equals to one if a firm i makes a forecast on future earnings in year t. The variable of interest is $Capacity^2_{i,t-1}$, which is the squared term of lagged either $Capacity\ (percent)$ and $Capacity\ (ratio)$ at firm level. All independent variables are lagged one period relative to management forecasts. t statistics are in parentheses and standard errors are clustered at firm level. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Panel A: Ordinary Least Squares Regressions

Dep. Var.	Management Forecast ($MF_{i,t}$)					
$Capacity\ (ratio)_{i,t-1}$	0.610***		0.603***		0.109***	
	(11.33)		(11.66)		(3.81)	
$Capacity\ (ratio)^2_{i,t-1}$	−0.271***		−0.256***		−0.012***	
	(-5.91)		(-6.04)		(-4.22)	
$Capacity\ (percent)_{i,t-1}$		0.536***		0.470***		0.126***
		(6.67)		(6.11)		(3.96)
$Capacity\ (percent)^2_{i,t-1}$		−0.217***		−0.149**		−0.00046***
		(-3.03)		(-2.20)		(-3.89)
$Loss_{i,t-1}$					−0.180***	−0.179***
					(-9.22)	(-9.25)
$EPS\ Increase_{i,t-1}$					0.0198**	0.0191**
					(2.16)	(2.08)
$Abs.\ EPS\ Change_{i,t-1}$					−0.022***	−0.0225***
					(-3.17)	(-3.21)
$Leverage\ Ratio_{i,t-1}$					0.0620	0.0599
					(1.43)	(1.38)
$Size_{i,t-1}$					0.0342***	0.0338***
					(5.21)	(5.12)
$Book\ to\ Market_{i,t-1}$					−0.0440**	−0.0475**
					(-2.34)	(-2.51)
$Constant$	0.203***	0.208***	-0.0126	-0.0180	-0.0291	-0.0380
	(13.92)	(10.84)	(-0.11)	(-0.15)	(-0.19)	(-0.25)
Year FE			✓	✓	✓	✓
Industry FE			✓	✓	✓	✓
Controls					✓	✓
Observations	16,508	16,508	16,508	16,508	16,508	16,508
Adjusted R^2	0.035	0.026	0.146	0.138	0.193	0.192

Table B.4. Investor Attention and Management Forecast (Continued)
Panel B Logit Regressions

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.	Management Forecast ($MF_{i,t}$)					
$Capacity\ (ratio)_{i,t-1}$	2.865***		3.241***		0.787***	
	(10.50)		(10.80)		(4.42)	
$Capacity\ (ratio)^2_{i,t-1}$	−1.368***		−1.502***		−0.139***	
	(−6.23)		(−6.44)		(−2.82)	
$Capacity\ (percent)_{i,t-1}$		2.448***		2.466***		1.384***
		(6.66)		(6.22)		(12.29)
$Capacity\ (percent)^2_{i,t-1}$		−1.066***		−0.894***		−0.00499***
		(−3.41)		(−2.68)		(−12.04)
$Constant$	−1.334***	−1.279***	−2.641***	−2.640***	−2.343***	−2.443***
	(−16.87)	(−13.10)	(−3.23)	(−3.21)	(−2.64)	(−3.10)
Year FE			✓	✓	✓	✓
Industry FE			✓	✓	✓	✓
Controls					✓	✓
Observations	16,508	16,508	16,508	16,508	16,508	16,508
Adjusted R^2	0.035	0.026	0.146	0.138	0.193	0.192

Table B.5. Results from Spline Regressions

This table presents our results from estimating a spline regression that treats the relation between the likelihood of management forecast and investor attention as piecewise linear. We estimate a separate slope for each side of a threshold τ of investor attention as follows:

$$MF_{i,t} = \alpha_t + \alpha_j + \beta_1(Capacity_{i,t-1} - \tau < 0) + \beta_2(Capacity_{i,t-1} - \tau \geqslant 0) + Controls_{i,t-1} + \varepsilon_{i,t}.$$

If our theoretical prediction holds, we expect to see that $\beta_1 > 0$ and $\beta_2 < 0$. By eyeballing Figure 4, we conjecture that the threshold is around the 80^{th} percentile of both $Capacity(percent)$ and $Capacity(ratio)$. For robustness, we set $\tau = 70^{th}, 75^{th}, 80^{th}, 85^{th}$ percentile of both $Capacity(ratio)$ (Panel A) and $Capacity(percent)$ (Panel B). Panel C presents our results estimated from the Multivariate Adaptive Regression Spline (MARS) method.

Panel A: $Capacity(ratio)$ **with pre-specified** τ

	(1) $\tau = 70^{th} pctile$	(2) $\tau = 75^{th} pctile$	(3) $\tau = 80^{th} pctile$	(4) $\tau = 85^{th} pctile$
$Capacity(ratio) - \tau < 0$	0.413***	0.415***	0.412***	0.395***
	(19.91)	(14.93)	(15.51)	(15.58)
$Capacity(ratio) - \tau \geqslant 0$	-0.212***	-0.226***	-0.245***	-0.274***
	(-3.80)	(-2.90)	(-2.70)	(-2.59)
Year FE	✓	✓	✓	✓
4-digit SIC FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Observations	16,508	16,508	16,508	16,508
Adjusted R^2	0.149	0.154	0.155	0.152

Panel B: $Capacity(percent)$ **with pre-specified** τ

	(1) $\tau = 70^{th} pctile$	(2) $\tau = 75^{th} pctile$	(3) $\tau = 80^{th} pctile$	(4) $\tau = 85^{th} pctile$
$Capacity(ratio) - \tau < 0$	0.371***	0.370***	0.371***	0.368***
	(12.56)	(13.08)	(13.69)	(14.09)
$Capacity(ratio) - \tau \geqslant 0$	-0.285**	-0.334**	-0.439***	-0.409**
	(-2.42)	(-2.51)	(-2.74)	(-2.11)
Year FE	✓	✓	✓	✓
4-digit SIC FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Observations	16,508	16,508	16,508	16,508
Adjusted R^2	0.145	0.146	0.148	0.147

Table B.5. Results from Spline Regressions (Continued)

Panel C: Multivariate Adaptive Regression Spline (MARS) with optimal threshold τ^*

	(1) $\tau^* = 79^{th}$ *pctile of Capacity(ratio)*	(2) $\tau^* = 81^{th}$ *pctile of Capacity(percent)*
$Capacity\ (ratio)$ - $\tau < 0$	0.397***	
	(15.24)	
$Capacity\ (ratio)$ - $\tau \geqslant 0$	-0.208**	
	(-2.49)	
$Capacity\ (percent)$ - $\tau < 0$		0.366***
		(13.87)
$Capacity\ (percent)$ - $\tau \geqslant 0$		-0.303**
		(-2.01)
Year FE	✓	✓
4-digit SIC FE	✓	✓
Controls	✓	✓
Observations	16,508	16,508
Adjusted R^2	0.153	0.159

Table B.6. Three Types of Institutional Investor Ownership and Management Forecast

This table presents results from estimating the relation between each of the three types of institutional investor ownership and the likelihood of management forecast using the following specification:

$$MF_{i,t} = \alpha_t + \alpha_j + \beta Capacity_{i,t-1} + \gamma Capacity^2_{i,t-1} + Controls_{i,t-1} + \varepsilon_{i,t}$$

where α_t is year fixed effect and α_j industry fixed effect. The dependent variable $MF_{i,t}$ equals to one if a firm i makes a forecast in year t. We replace $Capacity_{i,t-1}$ and $Capacity^2_{i,t-1}$ with either ownership by *Quasi-indexer*, *Transient*, or *Dedicated* investors at $t-1$ and their squared terms, respectively. Standard errors are clustered at firm-level and t statistics are in parentheses. * indicates statistical significance at the 10% level, ** at the 5% level, and *** at the 1% level.

	Management Forecast		
	(1)	(2)	(3)
Quasi-indexer	0.485***		
	(4.30)		
Quasi-indexer^2	-0.102		
	(-0.85)		
Transient		0.879***	
		(4.78)	
Transient^2		-1.387***	
		(-3.11)	
Dedicated			0.385
			(1.28)
Dedicated^2			-3.792***
			(-2.80)
Constant	0.0140	0.148	0.194
	(0.07)	(0.79)	(1.06)
Year FE	✓	✓	✓
4-digit SIC FE	✓	✓	✓
Controls	✓	✓	✓
Observations	16,508	16,508	16,508
Adjusted R^2	0.189	0.172	0.170

Appendix C

C.1 Appendix

C.1.1 Additional Market Transparency Literature

The value of market efficiency is one of the most important questions in the finance literature. First, it is the essential assumption for most of the modern asset pricing models. Second, in spite of many findings about return anomalies, Fama (1998) supports the market efficiency and shows most long-term return anomalies tend to disappear with reasonable changes in technique.

Among all the factors that contribute to the efficiency, market transparency has been mostly used and well documented. Using laboratory experiments, Bloomfield and O'Hara (1999) shows that higher transparency increases the informational efficiency of transaction prices. Recent papers also shed light on the effects of the corporate bond transparency. Using a complete record of all US OTC secondary trades in corporate bonds (TRACE), Edwards et al. (2004) finds that transaction costs of corporate bonds are higher than in equities and decrease significantly with trade size. Moreover, later Bessembinder et al. (2006) further shows the trade execution costs significantly dropped after an increase of the transaction reporting transparency.

Instead of directly studying market transparency, most papers use disclosure level as a proxy. Various financial and real effects have been studied under a variation of disclosure level. Using the 1990 annual reports of 122 manufacturing firm, Botosan (1997) finds that for firms that attract a low analyst following, greater disclosure is associated with a lower cost of equity capital. Similarly, Sengupta (1998) provides evidence that firms with high disclosure quality ratings from financial analysts enjoy a lower effective interest cost of issuing debt. Healy et al. (1999) shows that the disclosure rating increases are accompanied by increases in sample firms' stock returns, institutional ownership, analyst following, and stock liquidity. Recent papers also tried to distinguish various information sources inside regular disclosures. Easley and O'Hara (2004) find that investors demand a higher return to hold stocks disclosing a greater percentage of private information.

C.1.2 Details on Discretionary Accrual and Real Earnings Management

We review four parts of the earnings management literature: discretionary accrual, real earnings management, and market reaction to earnings management.

Firstly, we briefly explain what accruals are and why they are important. The total accruals are managers' estimates about future cash flows. By recording accruals, a company can measure what it owes and also what cash revenue it expects to receive in the future. Annual accounting earnings is the sum of accruals and current cash flows. Adding accruals to accounting earnings gives a more complete picture of a firm's fundamental performance than just current cash flows.

The non-discretionary component of accruals reflects business conditions that naturally affect accruals, which is largely out of manager's control. However, managers can adjust their estimates of firms' future cash flows, within the flexibility of accounting regulations. The component of accruals at managers' discretion is called the discretionary accruals. According to Dechow (1994), discretionary accruals often provide managers with opportunities to manipulate earnings.

Managers can also manage earnings through real earning management. Roychowdhury (2006) define real earnings management as management actions that deviate from normal operational practices, undertaken with the primary objective of meeting certain earnings thresholds. The accounting literature captures real earnings management by checking whether firms use price discounts to generate unsustainable sales, overproduce and put additional output to inventory to report a lower cost of goods sold, cut discretionary expenses such as R&D, advertising, and selling, general, and administrative (SG&A) expenditures to inflate current year's earnings.

There is also a strand of literature that examine non-operational real earnings management. Bartov (1993) and Herrmann et al. (2003) document that firms in the U.S. and other developed countries manipulate the timing and magnitude of transactions inducing sales of fixed asset and financial securities. HAW et al. (2005) and Chen and Yuan (2004) study the non-operational real earnings management in the context of China. They found that Chinese firms manage their earnings by selling financial securities and real estate properties, restructuring debt, and obtaining government subsidies.

Lastly, we review the literature on market reaction to firms' earnings management. Hayn (1995), Burgstahler and Dichev (1997), and Degeorge et al. (1999) found that a significantly large number of firms has an annual earnings that is either slightly greater than zero or just beats analyst forecasts. Bartov et al. (2002) and Bhojraj et al. (2009) reported that firms manage accruals and cut discretionary expenses to just beat analyst forecasts. Their stocks' performance improves in the short term. However, HAW et al. (2005) concluded that investors are able to differentiate the quality of earnings and discount the earnings suspected of a greater degree of management.

Discretionary Accruals (DA)

We measure each firm's DA using modified Jones model:

$$\frac{Accruals_t}{A_{t-1}} = \alpha_0 + \alpha_1 \frac{1}{A_{t-1}} + \alpha_2 \frac{\Delta S_t - \Delta AR_t}{A_{t-1}} + \alpha_3 \frac{PPE_t}{A_{t-1}} + \varepsilon_t \tag{C.1}$$

where $Accruals_t$ is calculated by subtracting a firm's operating cash flow from its operating income in year t. PPE_t is the gross property, plant, and equipment and A_{t-1} is a firm's total asset in year $t-1$. ΔS_t is the change in sales from year $t-1$ to t and ΔAR_t is the change in account receivables from year $t-1$ to t. We estimate the above cross-sectional regression for each industry-year group with at least 20 observations. The estimated residuals, capturing the abnormal part of accruals, proxy for firms' accrual-based earnings management.

Real Earnings Management

Real earnings management refers to management actions that deviate from normal operational practices, undertaken with the primary objective of meeting certain earnings thresholds (Roychowdhury (2006), Zang (2011)).

Following Roychowdhury (2006), we examine two major components of real earnings management: production costs and discretionary expenses. Facing enormous pressure to report a positive earnings, firm could increase earnings by reducing the cost of goods sold by overproducing inventory and cutting discretionary expenditures, including R&D, advertising, and selling, general, and administrative (SG&A) expenditures. The former is measured by the abnormal level of production costs, the latter by the abnormal level of discretionary expenditures. Subsequent studies using the same methods provide further evidence that these measures capture real activities manipulation (Cohen et al., 2008b; Cohen and Zarowin, 2010b).

We estimate the normal level of production costs using the following regression:

$$PROD_t/A_{t-1} = \alpha_0 + \alpha_1(1/A_{t-1}) + \alpha_2(S_t/A_{t-1}) + \alpha_3(\Delta S_t/A_{t-1}) + \alpha_4(\Delta S_{t-1}/A_{t-1}) + \varepsilon_t \quad \text{(C.2)}$$

where $PROD_t$ is the sum of the cost of goods sold in year t and the change in inventory from $t-1$ to t. A_{t-1} is the total assets in year $t-1$. S_t is sales in year t. ΔS_t is the change in sales from year $t-1$ to t. We estimate the above equation cross-sectionally for each industry-year with at least 20 observations. The abnormal level of production cost (RM_{PROD}) is measured as the estimated residual. The higher the residual, the larger is the amount of inventory overproduction, and the greater is the increase in reported earnings through reducing the costs of goods sold.

Furthermore, we estimate the normal level of discretionary expenditures using the following regression:

$$DISX_t/A_{t-1} = \alpha_0 + \alpha_1(1/A_{t-1}) + \alpha_2(S_{t-1}/A_{t-1}) + \varepsilon_t \tag{C.3}$$

where $DISX_t$ is the discretionary expenditures (i.e., the sum of R&D, adverting, and SG&A expenditures) in year t. We estimate the above cross-sectional regression for industry-year groups with at least 20 observations. The abnormal level of discretionary expenditures is measured as the estimated residual from the regression. We multiply the residuals by -1 to get RM_{DISX} so that higher values of RM_{DISX} imply greater amounts of cut on discretionary expenditures by firms to inflate reported earnings. We construct an aggregate measure of firm level real earnings management (RM) by taking the sum of RM_{PROD} and RM_{DISX}.

C.1.3 Conceptual Framework

There are two types of firms in China: T (truth) and L (lie). T firms report truthfully about their earnings with a low level of noise. The noise mainly comes from the pressure to beat last years' earnings. These firms can be considered as firms that would normally report a positive earning and hence do not face a delisting risk. That's why they don't have a strong incentive to manage their earnings. Moreover, the payoff of investing in T firms is normally distributed with a mean R and a low variance.

On the other hand, L type firms are the ones that manage a great amount of its earnings. It is costly for them to manage earnings. They would usually cut back on R&D, investment, advertising to do so. The major incentive comes from China's delisting policy. Think about these L-type firms as those who would normally report a negative earnings and do face a delisting risk. As a result, these firms sacrifice future growth to manage their earnings from negative to positive. They would not want to report a high ROE due to convex cost in managing earnings and also tax. Due to earnings management, the payoff for investors in investing in L-type firms is normally distribute with a mean R and a higher variance.

On the investor side, there is a mass 1 of investors who maximize mean-variance utility with the same risk aversion level. To generate trading, we have a fraction of informed investors and the rest are uninformed investors. Informed investors pay attention to earnings report and update their belief of firms' payoff after observing firms' signals whereas uninformed investors do not. All the firms publish the same signal which equals the true type plus a Gaussian noise. The informed traders also know that high EM segment has much more L-type firms than T-type firms whereas it is the other way around in the low EM segment. However, investors do not know the true type of each firm.

C.1.4 Real Effects

In progress. Right now we mainly want to focus on short-term firm performance related real effects, such as CEO turnover, CEO compensation change. In addition, we plan to explore real effects on firms' investment, R&D, patents, etc. There are two empirical difficulties in investigating firms' real effects. First, these real effect measures are not fully collected back in 2007. We do not have enough data comparable with the financial measures we studied before. We may need to change to years after 2012 with better data. Second, investment, RD and patents normally come into play over a really long period. Firms in the low information segment may prefer other real effects with faster payoffs.

C.1.5 Delisting Threat vs. Non-delisting Threat

Table 8 and 9 show that investors react less to firms with delisting threat (negative earnings last year) compared with firms without delisting threat (positive earnings last year).

Table C.1. Delisting Threat vs Non-Delisting Threat 2009-2016 China

	Abnormal Return Variance					
	ROE∈ (0,0.06)		FROE ∈ (-0.1,0.06)		FROA ∈ (-0.1,0.03)	
Delisting Threat	-0.228**	-0.200**	-0.212**	-0.221**	-0.224***	-0.200**
	(0.0967)	(0.101)	(0.0959)	(0.104)	(0.0867)	(0.0929)
Firm Size		-0.0272		0.0134		-0.0169
		(0.0385)		(0.0418)		(0.0344)
Firm Leverage		-0.0386		-0.0692		0.0278
		(0.192)		(0.218)		(0.223)
Industry effect	No	Yes	No	Yes	No	Yes
Year effect	No	Yes	No	Yes	No	Yes
Observations	2126	2126	1686	1686	1897	1897
Adjusted R^2	0.002	0.018	0.002	0.017	0.002	0.032

Note: In the parentheses below coefficient estimates are standard errors adjusted for heteroskedasticity and firm-level clustering. All continuous variables are winsorized at the 1st and 99th percentile. *, **, and *** denote significance at the 0.1, 0.05, and 0.01 level, respectively.

Table C.2. Delisting Threat vs Non-Delisting Threat 2009-2016 China

	Abnormal Trading Volume					
	ROE∈ (0,0.06)		FROE ∈ (-0.1,0.06)		FROA ∈ (-0.1,0.03)	
Delisting Threat	-0.141***	-0.0845*	-0.166***	-0.127***	-0.158***	-0.123***
	(0.0481)	(0.0456)	(0.0487)	(0.0473)	(0.0431)	(0.0424)
Firm Size		0.00119		-0.00289		-0.0273**
		(0.0144)		(0.0163)		(0.0134)
Firm Leverage		0.144*		0.164*		0.219**
		(0.0841)		(0.0911)		(0.105)
Industry effect	No	Yes	No	Yes	No	Yes
Year effect	No	Yes	No	Yes	No	Yes
Observations	2191	2171	1468	1468	1691	1691
Adjusted R^2	0.003	0.179	0.005	0.201	0.005	0.193

Note: In the parentheses below coefficient estimates are standard errors adjusted for heteroskedasticity and firm-level clustering. All continuous variables are winsorized at the 1st and 99th percentile. *, **, and *** denote significance at the 0.1, 0.05, and 0.01 level, respectively.

C.1.6 Regression Result for Cross-Segment Investor Reaction

Table C.3. Stock Market Reactions Surrounding Earnings Announcements: 2009-2016

	Abnormal Return Variance		Abnormal Trading Volume	
	(1)	(2)	(3)	(4)
Return on Equity	1.228***	1.476***	0.722***	0.676***
	(0.255)	(0.256)	(0.110)	(0.110)
Firm Size		-0.0152		-0.00548
		(0.0192)		(0.00707)
Firm Leverage		-0.231**		-0.00176
		(0.112)		(0.0408)
|Unexpected Earnings|		0.672		-0.0316
		(1.537)		(0.173)
Year effect	No	Yes	No	Yes
Industry effect	No	Yes	No	Yes
Observations	9243	8393	8610	8567
R^2	0.00223	0.0239	0.00486	0.162

Note: In the parentheses below coefficient estimates are robust t-statistics based on standard errors adjusted for heteroskedasticity and firm-level clustering. All continuous variables are winsorized at the 1st and 99th percentile. *, **, and *** denote significance at the 0.1, 0.05, and 0.01 level, respectively.

C.1.7 Bunching Estimator

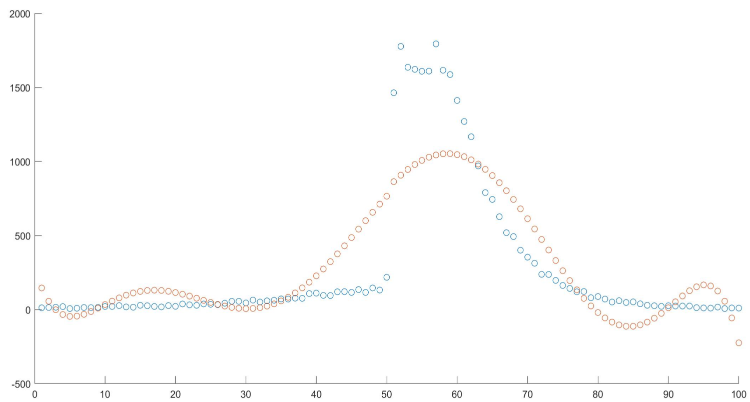


Figure C.1. Counterfactual ROE Distribution

Note: In order to estimate the counterfactual ROE distribution without any manipulation, we use bunching estimator following Chetty et al. (2011) with a polynomial approximation that ignores the effects of data around the threshold. Blue dotted line shows the actual ROE distribution in China. Red dotted line is the estimation of counterfactual ROE distribution without any earnings manipulation.

Bibliography

Abramova, I., Core, J.E., Sutherland, A., 2020. Institutional investor attention and firm disclosure. The Accounting Review .

Acharya, V.V., DeMarzo, P., Kremer, I., 2011. Endogenous information flows and the clustering of announcements. American Economic Review 101, 2955–79.

Aghamolla, C., Corona, C., Zheng, R., 2019. No reliance on guidance: Counter-signaling in management forecasts. Rand Journal of Economics, forth .

Ali, A., Klasa, S., Li, O.Z., 2008. Institutional stakeholdings and better-informed traders at earnings announcements. Journal of Accounting and Economics 46, 47–61.

Asthana, S., Balsam, S., 2001. The effect of edgar on the market reaction to 10-k filings. Journal of Accounting and Public Policy 20, 349–372.

Asthana, S., Balsam, S., Sankaraguruswamy, S., 2004. Differential response of small versus large investors to 10-k filings on edgar. The Accounting Review 79, 571–589.

Banerjee, S., 2011. Learning from prices and the dispersion in beliefs. The Review of Financial Studies 24, 3025–3068.

Banerjee, S., Kremer, I., 2010. Disagreement and learning: Dynamic patterns of trade. The Journal of Finance 65, 1269–1302.

Barberis, N., Thaler, R., 2003. A survey of behavioral finance. Handbook of the Economics of Finance 1, 1053–1128.

Bartov, E., 1993. The timing of asset sales and earnings manipulation. Accounting Review , 840–855.

Bartov, E., Givoly, D., Hayn, C., 2002. The rewards to meeting or beating earnings expectations. Journal of accounting and economics 33, 173–204.

Basu, R.S., Pierce, S., Stephan, A., 2020. The effect of investor inattention on voluntary disclosure. Available at SSRN 3071399 .

Beatty, A., Liao, S., Yu, J.J., 2013. The spillover effect of fraudulent financial reporting on peer firms' investments. Journal of Accounting and Economics 55, 183–205.

Becker, C.L., DeFond, M.L., Jiambalvo, J., Subramanyam, K., 1998. The effect of audit quality on earnings management. Contemporary Accounting Research 15, 1–24.

Bergemann, D., Shen, J., Xu, Y., Yeh, E., 2012. Multi-dimensional mechanism design with limited information, in: Proceedings of the 13th ACM Conference on Electronic Commerce, pp. 162–178.

Bertomeu, J., Ma, P., Marinovic, I., 2019. How often do managers withhold information? The Accounting Review, forth. .

Bertomeu, J., Marinovic, I., 2016. A theory of hard and soft information. The Accounting Review 91, 1–20.

Bessembinder, H., Maxwell, W., Venkataraman, K., 2006. Market transparency, liquidity externalities, and institutional trading costs in corporate bonds. Journal of Financial Economics 82, 251–288.

Beyer, A., Cohen, D.A., Lys, T.Z., Walther, B.R., 2010. The financial reporting environment: Review of the recent literature. Journal of accounting and economics 50, 296–343.

Beyer, A., Dye, R.A., 2020. Debt and voluntary disclosuredebt and voluntary disclosure. The Accounting Review .

Bhojraj, S., Hribar, P., Picconi, M., McInnis, J., 2009. Making sense of cents: An examination of firms that marginally miss or beat analyst forecasts. The Journal of Finance 64, 2361–2388.

Bird, A., Karolyi, S.A., 2016. Do institutional investors demand public disclosure? The Review of Financial Studies 29, 3245–3277.

Blankespoor, E., 2019. The impact of information processing costs on firm disclosure choice: Evidence from the xbrl mandate. Journal of Accounting Research 57, 919–967.

Blankespoor, E., deHaan, E., Marinovic, I., 2020a. Disclosure processing costs, investors' information choice, and equity market outcomes: A review. Journal of Accounting and Economics, forthcoming .

Blankespoor, E., deHaan, E., Marinovic, I., 2020b. Disclosure processing costs, investors information choice, and equity market outcomes: A review. Journal of Accounting & Economics (JAE), Forthcoming .

Blankespoor, E., Miller, G.S., White, H.D., 2014. The role of dissemination in market liquidity: Evidence from firms' use of twitter. The Accounting Review 89, 79–112.

Bloomfield, R., Fischer, P.E., 2011. Disagreement and the cost of capital. Journal of Accounting Research 49, 41–68.

Bloomfield, R., O'Hara, M., 1999. Market transparency: Who wins and who loses? The Review of Financial Studies 12, 5–35.

Boone, A.L., White, J.T., 2015. The effect of institutional ownership on firm transparency and information production. Journal of Financial Economics 117, 508–533.

Botosan, C.A., 1997. Disclosure level and the cost of equity capital. The Accounting Review 72, 323–349.

Budanova, S., Cianciaruso, D., Marinovic, I., 2020. The ambiguity of earnings announcements. Management Science .

Burgstahler, D., Dichev, I., 1997. Earnings management to avoid earnings decreases and losses. Journal of accounting and economics 24, 99–126.

Burgstahler, D., Eames, M., 2006. Management of earnings and analysts' forecasts to achieve zero and small positive earnings surprises. Journal of Business Finance & Accounting 33, 633–652.

Bushee, B.J., 1998. The influence of institutional investors on myopic r&d investment behavior. The Accounting review , 305–333.

Bushee, B.J., Noe, C.F., 2000. Corporate disclosure practices, institutional investors, and stock return volatility. Journal of Accounting Research , 171–202.

Caskey, J.A., 2009. Information in equity markets with ambiguity-averse investors. The Review of Financial Studies 22, 3595–3627.

Chakraborty, A., Harbaugh, R., 2010. Persuasion by cheap talk. American Economic Review 100, 2361–82.

Chang, Y.C., Hsiao, P.J., Ljungqvist, A., Tseng, K., 2020a. Testing disagreement models. Working paper. .

Chang, Y.C., Ljungqvist, A., Tseng, K., 2020b. Do corporate disclosures constrain strategic analyst behavior? Working paper. .

Chen, K.C., Yuan, H., 2004. Earnings management and capital resource allocation: Evidence from china's accounting-based regulation of rights issues. The Accounting Review 79, 645–665.

Chen, Q., Corona, C., Zhang, Y., 2017. Communication with endogenous investor attention .

Chen, Q., Goldstein, I., Jiang, W., 2006. Price informativeness and investment sensitivity to stock price. The Review of Financial Studies 20, 619–650.

Chen, Q., Hemmer, T., Zhang, Y., 2007. On the relation between conservatism in accounting standards and incentives for earnings management. Journal of Accounting Research 45, 541–565.

Cheng, Q., Luo, T., Yue, H., 2013. Managerial incentives and management forecast precision. The Accounting Review 88, 1575–1602.

Cheong, F.S., Thomas, J., 2011. Why do eps forecast error and dispersion not vary with scale? implications for analyst and managerial behavior. Journal of Accounting Research 49, 359–401.

Cheong, F.S., Thomas, J., 2018. Management of reported and forecast eps, investor responses, and research implications. Management Science 64, 4277–4301.

Chetty, R., Friedman, J., Olsen, T., Pistaferri, L., 2011. Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from danish tax records. The Quarterly Journal of Economics 126, 749–804.

Cheynel, E., Levine, C.B., 2020. Public disclosures and information asymmetry: A theory of the mosaic. The Accounting Review 95, 79–99.

Chiba, S., Leong, K., 2013. Cheap talk with outside options. Department of Management, Università Ca'Foscari Venezia Working Paper .

Chong-en, B., Changtai, H., Michael, S.Z., 2016. The long shadow of a fiscal expansion. Brookings Papers on Economic Activity , 15–16.

Christensen, H.B., Floyd, E., Liu, L.Y., Maffett, M., 2017. The real effects of mandated information on social responsibility in financial reports: evidence from mine-safety records. Journal of Accounting and Economics 64, 284–304.

Christensen, H.B., Hail, L., Leuz, C., 2016. Capital-market effects of securities regulation: Prior conditions, implementation, and enforcement. The Review of Financial Studies 29, 2885–2924.

Cohen, D.A., Dey, A., Lys, T.Z., 2008a. Real and accrual-based earnings management in the pre-and post-sarbanes-oxley periods. The Accounting Review 83, 757–787.

Cohen, D.A., Dey, A., Lys, T.Z., 2008b. Real and accrual-based earnings management in the pre-and post-sarbanes-oxley periods. The accounting review 83, 757–787.

Cohen, D.A., Zarowin, P., 2010a. Accrual-based and real earnings management activities around seasoned equity offerings. Journal of Accounting and Economics 50, 2–19.

Cohen, D.A., Zarowin, P., 2010b. Accrual-based and real earnings management activities around seasoned equity offerings. Journal of accounting and Economics 50, 2–19.

Collins, D.W., Kothari, S., 1989a. An analysis of intertemporal and cross-sectional determinants of earnings response coefficients. Journal of Accounting and Economics 11, 143–181.

Collins, D.W., Kothari, S., 1989b. An analysis of intertemporal and cross-sectional determinants of earnings response coefficients. Journal of accounting and economics 11, 143–181.

Crawford, V.P., Sobel, J., 1982. Strategic information transmission. Econometrica: Journal of the Econometric Society , 1431–1451.

Da, Z., Engelberg, J., Gao, P., 2011. In search of attention. The Journal of Finance 66, 1461–1499.

Daniel, K., Hirshleifer, D., Teoh, S.H., 2002. Investor psychology in capital markets: Evidence and policy implications. Journal of Monetary Economics 49, 139–209.

Dechow, P., Ge, W., Schrand, C., 2010. Understanding earnings quality: A review of the proxies, their determinants and their consequences. Journal of Accounting and Economics 50, 344–401.

Dechow, P.M., 1994. Accounting earnings and cash flows as measures of firm performance: The role of accounting accruals. Journal of Accounting and Economics 18, 3–42.

Dechow, P.M., Dichev, I.D., 2002. The quality of accruals and earnings: The role of accrual estimation errors. The Accounting Review 77, 35–59.

Dechow, P.M., Kothari, S.P., Watts, R.L., 1998. The relation between earnings and cash flows. Journal of Accounting and Economics 25, 133–168.

Dechow, P.M., Sloan, R.G., Sweeney, A.P., 1995. Detecting earnings management. The Accounting Review , 193–225.

DeFond, M.L., Jiambalvo, J., 1994. Debt covenant violation and manipulation of accruals. Journal of Accounting and Economics 17, 145–176.

Degeorge, F., Patel, J., Zeckhauser, R., 1999. Earnings management to exceed thresholds. The Journal of Business 72, 1–33.

DellaVigna, S., Pollet, J.M., 2009. Investor inattention and friday earnings announcements. The Journal of Finance 64, 709–749.

Dranove, D., Jin, G.Z., 2010. Quality disclosure and certification: Theory and practice. Journal of Economic Literature 48, 935–63.

Duguay, R., Rauter, T., Samuels, D., 2019. The impact of open data on public procurement. Working paper. .

Durnev, A., Morck, R., Yeung, B., Zarowin, P., 2003. Does greater firm-specific return variation mean more or less informed stock pricing? Journal of Accounting Research 41, 797–836.

Dworczak, P., Martini, G., 2019. The simple economics of optimal persuasion. Journal of Political Economy forthcoming.

Dye, R.A., 1985. Disclosure of nonproprietary information. Journal of Accounting Research , 123–145.

Dye, R.A., Hughes, J.S., 2018. Equilibrium voluntary disclosures, asset pricing, and information transfers. Journal of Accounting and Economics 66, 1–24.

Easley, D., O'Hara, M., 2004. Information and the cost of capital. The Journal of Finance 59, 1553–1583.

Easton, P.D., Zmijewski, M.E., 1989. Cross-sectional variation in the stock market response to accounting earnings announcements. Journal of Accounting and Economics 11, 117–141.

Edwards, A.K., Harris, L., Piwowar, M.S., 2004. Corporate bond market transparency and transaction costs. Fifteenth Annual Utah Winter Finance Conference .

Einhorn, E., 2005. The nature of the interaction between mandatory and voluntary disclosures. Journal of Accounting Research 43, 593–621.

Einhorn, E., Ziv, A., 2008. Intertemporal dynamics of corporate voluntary disclosures. Journal of Accounting Research 46, 567–589.

Emery, L.P., Gulen, H., 2019. Expanding horizons: The effect of information access on geographically biased investing. Working paper. .

Ewert, R., Wagenhofer, A., 2005. Economic effects of tightening accounting standards to restrict earnings management. The Accounting Review 80, 1101–1124.

Fama, E.F., 1970. Efficient capital markets: A review of theory and empirical work. The journal of Finance 25, 383–417.

Fama, E.F., 1998. Market efficiency, long-term returns, and behavioral finance. Journal of Financial Economics 49, 283–306.

Fama, E.F., French, K.R., 2000. Forecasting profitability and earnings. The Journal of Business 73, 161–175.

Fang, V.W., Huang, A.H., Wang, W., 2017. Imperfect accounting and reporting bias. Journal of Accounting Research 55, 919–962.

Farrell, J., Rabin, M., 1996. Cheap talk. Journal of Economic Perspectives 10, 103–118.

Ferri, F., Zheng, R., Zou, Y., 2018. Uncertainty about managers' reporting objectives and investors' response to earnings reports: Evidence from the 2006 executive compensation disclosures. Journal of Accounting and Economics 66, 339–365.

Fischer, P.E., Heinle, M.S., Verrecchia, R.E., 2016. Beliefs-driven price association. Journal of Accounting and Economics 61, 563–583.

Fischer, P.E., Stocken, P.C., 2001. Imperfect information and credible communication. Journal of Accounting Research 39, 119–134.

Fischer, P.E., Verrecchia, R.E., 2000. Reporting bias. The Accounting Review 75, 229–245.

Friedman, H.L., Heinle, M.S., 2016. Taste, information, and asset prices: Implications for the valuation of csr. Review of Accounting Studies 21, 740–767.

Friedman, J.H., 1991. Multivariate adaptive regression splines. The Annals of Statistics , 1–67.

Gabaix, X., 2019. Behavioral inattention, in: Handbook of Behavioral Economics: Applications and Foundations 1. Elsevier. volume 2, pp. 261–343.

Gao, M., Huang, J., 2020. Informing the market: The effect of modern information technologies on information production. The Review of Financial Studies 33, 1367–1411.

Gipper, B., Leuz, C., Maffett, M., 2019. Public Oversight and Reporting Credibility: Evidence from the PCAOB Audit Inspection Regime. The Review of Financial Studies .

Goldstein, I., Yang, S., Zuo, L., 2020. The real effects of modern information technologies. Working paper. .

Gomez, E., 2020. The effect of mandatory disclosure dissemination on information asymmetry: Evidence from the implementation of the edgar system. Working paper. .

Goodman, T.H., Neamtiu, M., Shroff, N., White, H.D., 2014. Management forecast quality and capital investment decisions. The Accounting Review 89, 331–365.

Gray, R.M., Neuhoff, D.L., 1998. Quantization. IEEE transactions on information theory 44, 2325–2383.

Griffin, P.A., 2003. Got information? investor response to form 10-k and form 10-q edgar filings.

Review of Accounting Studies 8, 433–460.

Guan, Y., Lobo, G.J., Tsang, A., Xin, X., 2020. Societal trust and management earnings forecasts. The Accounting Review .

Guttman, I., Kremer, I., Skrzypacz, A., 2014. Not only what but also when: A theory of dynamic voluntary disclosure. American Economic Review 104, 2400–2420.

Hart, S., Kremer, I., Perry, M., 2017. Evidence games: Truth and commitment. American Economic Review 107, 690–713.

HAW, I.M., Qi, D., Wu, D., Wu, W., 2005. Market consequences of earnings management in response to security regulations in china. Contemporary Accounting Research 22, 95–140.

Hayn, C., 1995. The information content of losses. Journal of accounting and economics 20, 125–153.

Healy, P.M., Hutton, A.P., Palepu, K.G., 1999. Stock performance and intermediation changes surrounding sustained increases in disclosure. Contemporary Accounting Research 16, 485–520.

Healy, P.M., Wahlen, J.M., 1999. A review of the earnings management literature and its implications for standard setting. Accounting Horizons 13, 365–383.

Heater, J.C., Liu, Y., Matthies, B., 2017. Managerial response to non-fundamental price shocks. Working paper. .

Heinle, M.S., Verrecchia, R.E., 2016. Bias and the commitment to disclosure. Management Science 62, 2859–2870.

Herrmann, D., Inoue, T., Thomas, W.B., 2003. The sale of assets to manage earnings in japan. Journal of Accounting Research 41, 89–108.

Hirshleifer, D., Lim, S.S., Teoh, S.H., 2009. Driven to distraction: Extraneous events and underreaction to earnings news. The Journal of Finance 64, 2289–2325.

Hirshleifer, D., Teoh, S.H., 2003. Limited attention, information disclosure, and financial reporting. Journal of accounting and economics 36, 337–386.

Hirshleifer, D.A., Lim, S.S., Teoh, S.H., 2004. Disclosure to an audience with limited attention. Available at SSRN 604142 .

Ivanov, M., 2010a. Communication via a strategic mediator. Journal of Economic Theory 145, 869–884.

Ivanov, M., 2010b. Informational control and organizational design. Journal of Economic Theory 145, 721–751.

Jiang, X., Yang, M., 2017. Properties of optimal accounting rules in a signaling game. Journal of Accounting and Economics 63, 499–512.

Jin, L., Myers, S.C., 2006. R2 around the world: New theory and new tests. Journal of financial Economics 79, 257–292.

Jones, J.J., 1991. Earnings management during import relief investigations. Journal of Accounting Research , 193–228.

Jung, M.J., Naughton, J.P., Tahoun, A., Wang, C., 2017. Do firms strategically disseminate? evidence from corporate use of social media. The Accounting Review 93, 225–252.

Jung, W.O., Kwon, Y.K., 1988. Disclosure when the market is unsure of information endowment of managers. Journal of Accounting Research 26, 146–153.

Kempf, E., Manconi, A., Spalt, O., 2017. Distracted shareholders and corporate actions. The Review of Financial Studies 30, 1660–1695.

Kim, J.B., Kim, J.W., Lim, J.H., 2019. Does xbrl adoption constrain earnings management? early evidence from mandated us filers. Contemporary Accounting Research 36, 2610–2634.

Kim, J.M., Taylor, D.J., Verrecchia, R.E., 2020. Voluntary disclosure when private information and proprietary costs are jointly determined. Available at SSRN 3595320 .

Kolotilin, A., Li, H., 2019. Relational communication. arXiv preprint arXiv:1901.05645 .

Kolotilin, A., Zapechelnyuk, A., 2019. Persuasion meets delegation. arXiv preprint arXiv:1902.02628 .

Kos, N., 2012. Communication and efficiency in auctions. Games and Economic Behavior 75, 233–249.

Kothari, S., 2001. Capital markets research in accounting. Journal of Accounting and Economics 31, 105–231.

Kothari, S.P., Leone, A.J., Wasley, C.E., 2005. Performance matched discretionary accrual measures. Journal of Accounting and Economics 39, 163–197.

Krishna, V., Morgan, J., 2001. Asymmetric information and legislative rules: Some amendments. American Political science review 95, 435–452.

Kumar, P., Langberg, N., Sivaramakrishnan, K., 2012. Voluntary disclosures, corporate control, and investment. Journal of Accounting Research 50, 1041–1076.

Levine, R., Zervos, S., 1996. Stock market development and long-run growth. The World Bank Economic Review 10, 323–339.

Li, E.X., Ramesh, K., 2009. Market reaction surrounding the filing of periodic sec reports. The Accounting Review 84, 1171–1208.

Li, V., 2015. Do false financial statements distort peer firms' decisions? The Accounting Review 91, 251–278.

Li, X., Yang, H.I., 2016. Mandatory financial reporting and voluntary disclosure: The effect of mandatory ifrs adoption on management forecasts. The Accounting Review 91, 933–953.

Liang, Y., Marinovic, I., Varas, F., 2018. The credibility of financial reporting: A reputation-based approach. The Accounting Review 93, 317–333.

Liu, Y., 2020. Information acquisition cost and earnings management: Evidence from the implementation of EDGAR .

Lo, K., 2008. Earnings management and earnings quality. Journal of Accounting and Economics 45, 350–357.

Lo, K., Ramos, F., Rogo, R., 2017. Earnings management and annual report readability. Journal of Accounting and Economics 63, 1–25.

Lu, J., 2019. Limited attention: Implications for financial reporting. Available at SSRN 3278995 .

Maćkowiak, B., Matějka, F., Wiederholt, M., 2018. Survey: Rational inattention, a disciplined behavioral model .

Miao, B., Teoh, S.H., Zhu, Z., 2016. Limited attention, statement of cash flow disclosure, and the valuation of accruals. Review of Accounting Studies 21, 473–515.

Milgrom, P.R., 1981. Good news and bad news: Representation theorems and applications. The Bell Journal of Economics , 380–391.

Morck, R., Yeung, B., Yu, W., 2000a. The information content of stock markets: why do emerging markets have synchronous stock price movements? Journal of financial economics 58, 215–260.

Morck, R., Yeung, B., Yu, W., 2000b. The information content of stock markets: why do emerging markets have synchronous stock price movements? Journal of Financial Economics 58, 215–260.

Morgan, J., Stocken, P.C., 2003. An analysis of stock recommendations. RAND Journal of economics , 183–203.

Niessner, M., 2015. Strategic disclosure timing and insider trading. Working paper. .

Pevzner, M., Xie, F., Xin, X., 2015a. When firms talk, do investors listen? the role of trust in stock market reactions to corporate earnings. Journal of Financial Economics 117, 190–223.

Pevzner, M., Xie, F., Xin, X., 2015b. When firms talk, do investors listen? the role of trust in stock market reactions to corporate earnings announcements. Journal of Financial Economics 117, 190–223.

Qi, D., Wu, W., Haw, I.M., 2000. The incremental information content of sec 10-k reports filed under the edgar system. Journal of Accounting, Auditing & Finance 15, 25–46.

Rappoport, D., 2020. Evidence and skepticism in verifiable disclosure games. Available at SSRN 2978288 .

Roll, R., 1998. R2. The Journal of Finance 43, 541–566.

Roychowdhury, S., 2006. Earnings management through real activities manipulation. Journal of Accounting and Economics 42, 335–370.

Rubinstein, A., 1998. Modeling bounded rationality. MIT press.

Samuels, D., 2020. Government procurement and changes in firm transparency. The Accounting Review, forthcoming .

Samuels, D., Taylor, D.J., Verrecchia, R.E., 2020a. The economics of misreporting and the role of public scrutiny. Journal of Accounting and Economics .

Samuels, D., Taylor, D.J., Verrecchia, R.E., 2020b. The economics of misreporting and the role of public scrutiny. Journal of Accounting and Economics, forth. .

Schantl, S.F., Wagenhofer, A., 2020. Deterrence of financial misreporting when public and private enforcement strategically interact. Journal of Accounting and Economics , 101311.

Schoenfeld, J., 2017. The effect of voluntary disclosure on stock liquidity: New evidence from index funds. Journal of Accounting and Economics 63, 51–74.

Sengupta, P., 1998. Corporate disclosure quality and the cost of debt. The Accounting Review

73, 459–474.

Shavell, S., 1994. Acquisition and disclosure of information prior to sale. The RAND Journal of Economics , 20–36.

Shon, J., Yan, M., 2015. R&d cuts and subsequent reversals: Meeting or beating quarterly analyst forecasts. European Accounting Review 24, 147–166.

Sims, C.A., 2003. Implications of rational inattention. Journal of monetary Economics 50, 665–690.

Sims, C.A., 2010. Rational inattention and monetary economics, in: Handbook of Monetary Economics. Elsevier. volume 3, pp. 155–181.

Stocken, P., 2000. Credibility of voluntary disclosure. Rand Journal of Economics 31, 359–374.

Stocken, P., 2013. Strategic accounting disclosure. Foundations and Trends® in Accounting 7, 197–291.

Suijs, J., Wielhouwer, J.L., 2019. Disclosure policy choices under regulatory threat. The RAND Journal of Economics 50, 3–28.

New York Times, 1982. S.E.C. data: Difficult hunt. May 19.

New York Times, 1993. Plan opens more data to public. Oct 22.

New York Times, 1995. S.E.C. seeks to keep free internet service. Aug 16.

The SEC, 1987. 1987 annual report.

The SEC, 1996. 1996 annual report.

The Washington Post, 1993. What cost public information? SEC's new data system prompts debate on commercial control of government records. Jun 21.

Wall Street Journal, 1992. SEC's electronic filing plan draws criticism on price from Nader group. Jun 30.

Tsang, A., Xie, F., Xin, X., 2019. Foreign institutional investors and corporate voluntary disclosure around the world. The Accounting Review 94, 319–348.

Van Nieuwerburgh, S., Veldkamp, L., 2010. Information acquisition and under-diversification. The Review of Economic Studies 77, 779–805.

Veldkamp, L.L., 2011. Information choice in macroeconomics and finance. Princeton University Press.

Verrecchia, R.E., 1983. Discretionary disclosure. Journal of Accounting and Economics 5, 179–194.

Zang, A.Y., 2011. Evidence on the trade-off between real activities manipulation and accrual-based earnings management. The Accounting Review 87, 675–703.

Zang, A.Y., 2012. Evidence on the trade-off between real activities manipulation and accrual-based earnings management. The Accounting Review 87, 675–703.