Title

Essays in Financial Accounting and Corporate Governance

Author

Chen, Jun

Publication Date

2022

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

Essays in Financial Accounting and Corporate Governance

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

in

Management

by

Jun Chen

Committee in charge:

Professor Joseph Engelberg, Chair
Professor Rossen Valkanov, Co-Chair
Professor Roger Gordan
Professor Jun Liu
Professor William Mullins

2022

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

University of California San Diego

2022

DEDICATION

I dedicate this dissertation to my loving and supportive family—to my wife who always encourages me to follow my dreams, to Tucker who stays with me day and night, and to my parents and sister who have supported me every step of the way.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I am indebted to my advisors Joseph Engelberg and Rossen Vakalnov for their unlimited help throughout my Ph.D. journey. I would like to express my deepest appreciation to Jun Liu for all the time and support he gave to prepare me for my academic career. I am extremely grateful to Roger Gordan, John Hughes, and William Mullins for their detailed guidance and invaluable insights on my thesis. I want to thank Yibin Liu who is not only a talented coauthor but also a great friend. I also wish to thank Rady faculty members for their helpful suggestions on both research and teaching, especially Snehal Banerjee, Jeremy Bertomeu, Krisztina Buti, Eric Floyd, Mario Milone, Allan Timmermann, and Richard Townsend. Last but definitely not least, my family and friends are always there for me along the way.

Chapter 1 is currently being prepared for submission of publication of the material. It is solely authored by the dissertation author.

Chapter 2 is currently being prepared for submission of publication of the material, and is coauthored with John Hughes, Jun Liu, and Dan Yang. The dissertation author is a primary investigator of this material.

Chapter 3 is currently being prepared for submission of publication of the material, and is coauthored with Yibin Liu. The dissertation author is a primary investigator of this material.

VITA

2014	B.S. in Mathematics, Peking University, China
2015	Master of Finance, University of California San Diego
2022	Ph.D. in Management, University of California San Diego

ABSTRACT OF THE DISSERTATION

Essays in Financial Accounting and Corporate Governance

by

Jun Chen

Doctor of Philosophy in Management

University of California San Diego, 2022

Professor Joseph Engelberg, Chair
Professor Rossen Valkanov, Co-Chair

This dissertation comprises three papers examining several questions in finance and accounting. A common thread is investigating the strategic interactions between public firms and stock market investors. Chapter 1 studies how investors with short-term horizons can impact firms' behaviors. Chapters 2 and 3 examine the impact of corporate disclosure and the market pricing of information.

In Chapter 1, I use the unique features of the margin trading system in China to identify the causal impact of transient investors on managerial myopia. Specifically, I employ a regression discontinuity design that exploits the ranking procedure that determines a stock's margin trading

eligibility. I find that margin traders are extremely short-term oriented and cause a sharp increase in stock share turnover. Moreover, marginable firms cater to these transient investors by manipulating current earnings and reducing long-term investments. Consistent with managerial myopia, these firms experience a short-term price increase but a long-term decline in operating performance.

Chapter 2 is joint work with John Hughes, Jun Liu, and Dan Yang. We reexamine the relation between disclosure indices and cost of equity capital employing an empirical specification similar to that of Botosan (1997) for a substantially larger sample over an extended time frame made possible by textual analysis. Our results provide no support for a hypothesis of a negative relation between disclosure indices and implied cost of equity capital. Rather, consistent with a bias of implied cost of equity capital as a proxy for expected return depicted by Hughes et al. (2009), we find strong evidence of a positive relation.

Chapter 3 is joint work with Yibin Liu. We exploit an earnings-based delisting policy and examine its adverse effect on investor trust in earnings news. Besides providing prominent visual evidence of large-scale earnings management at the required earnings threshold, we find that firms close to this threshold are trusted less by investors, regardless of whether they have manipulated earnings. Moreover, we provide causal evidence by studying firms that approach this threshold due to a plausibly exogenous profitability shock. Our results suggest that earnings-based regulations with harsh punishment may lead to a decline in investor trust.

Chapter 1

Do Transient Investors Cause Managerial Myopia? Evidence from Margin Traders in China

1.1 Introduction

Managerial myopia, the phenomenon that managers pursue short-term performance at the expense of firms' long-term growth, has raised substantial concerns in the past three decades. A long tradition of economic research blames this on transient investors because they tend to overweight current earnings in the stock price (Porter, 1992; Froot et al., 1992). Consistent with this view, survey evidence by Graham et al. (2005) suggests that 80% of executives are willing to reduce discretionary spending and 55.3% would delay a profitable project, just in pursuit of an earnings target possibly desired by short-term speculators. Beyer et al. (2014) find from another survey that two-thirds of companies would deviate from strategic decisions and focus on current performance under the pressure of transient investors. Over time, this perspective has been supported by not only executives but industry professionals and policymakers.[1]

Yet causal evidence of this widespread short-termism remains scarce, as it requires identifying transient investors and then isolating the variation in their trading activities that is not driven by the same economic forces that affect corporate behaviors.[2] The latter is easily violated if, for example, firms with a recent structure optimization may simultaneously reduce discretionary expenses (a form of real earnings management) and attract active investors. Using a rank-based margin trading deregulation in China, I address these difficulties by comparing corporate reporting, operating, and investment decisions for firms that barely passed the experiment cutoff to those of firms that just missed it. A key identification assumption, that margin traders are more short-term oriented than non-leveraged investors, has long been recognized

[1] For example, Hillary Clinton, the presidential candidate, proposed a sharp increase in the tax rate on short-term investments in 2015 and claimed that it would reduce firms' pressure to show near-term gains (https://www.wsj.com/articles/clinton-capital-gains-taxes-on-short-term-investments). BlackRock consecutively wrote annual letters (2012–2019) to all companies it invested in and urged them to fight against speculators' short-term demand and focus on long-term growth (https://www.blackrock.com/corporate/investor-relations/2016-larry-fink-ceo-letter).

[2] To my knowledge, there is no paper that identifies the causal effect of speculators (or short-term investors) on myopic corporate behaviors. The closest one is Cremers et al. (2020), who find a negative correlation between transient institutional ownership and long-term investment around firms' addition to the Russell 2000 index. However, unlike the plausibly exogenous index inclusion, the investor ownership in their setting is endogenously decided.

in the literature due to the leverage risk and borrowing cost of margin trading (Hardouvelis, 1990; Brunnermeier and Pedersen, 2009; Heimer and Simsek, 2019). In a recent rule to restrict leveraged derivatives, the SEC even listed excessive borrowing as one of the most important driving factors for short-term speculation.[3]

The theoretical link between transient investors and managerial myopia has been well established. Most closely related to this paper is the catering theory proposed by Baker et al. (2006), where managers rationally cater to short-term investors' demand for earnings target or free cash flows and make corporate decisions that can encourage overvaluation. Short-term speculators are willing to pay a premium for these catering stocks either due to rational reasons, such as the option to sell the stock to other optimistic investors (Bolton et al., 2006), or behavioral bias like underestimating the extent of myopia (Goldman and Slezak, 2006). This mispricing can persist, at least in the short run, due to limits of arbitrage widely accepted in the literature (Miller, 1977). Meanwhile, managers can benefit directly from stock-based compensation and implicitly from the reputation for maintaining stock price (Narayanan, 1985). Earlier theories also attribute managerial myopia to investors' trading sensitivity to earnings news (Stein, 1989; von Thadden, 1995), a feature that particularly accords with leveraged investors due to overconfidence and margin requirements (Garleanu and Pedersen, 2011; Barber et al., 2020). Unlike the catering theory, stock prices in these models are unaffected because markets see through managers' incentives to manipulate earnings.

China equity markets provide an ideal laboratory to examine the role of short-term speculators. Although having the world's second-largest stock market, China has long banned margin trading in fear of causing irrational speculation among retail investors. From 2013 to 2015, Chinese regulators ranked all stocks based on their market value and trading volume and made those top-ranked stocks eligible for margin trading. This natural experiment lends itself to my research question for two reasons. First, margin traders in China are short-term retail

[3]See details in the SEC final rules: 17 CFR Parts 239, 249, 270, and 274 released on November 2, 2020 (https://www.sec.gov/rules/final/2020/ic-34084.pdf).

investors who, on average, only hold stocks for two weeks compared to a four-month holding period of an average investor (Section 1.2). Second, firms close to the experiment cutoff can be seen as randomly chosen for the margin trading eligibility. Firms are not likely to manipulate their rankings close to the cutoff because both the event day and the number of treated firms are unknown ex-ante.

To identify the causal effect of margin traders, I employ a fuzzy regression discontinuity (RD) design where I focus on stocks close to the experiment cutoff. The experiment is implemented over three rounds with approximately one year in between, which mitigates the concern that margin trading eligibility may coincide with other firm-level shocks that affect corporate decisions. Moreover, since RD's treatment is as good as randomly assigned local to the cutoff (Lee and Lemieux, 2010), my setting efficiently alleviates the anticipation effect, in which firms may predict future treatment and adjust corporate decisions beforehand (Hansman et al., 2021).

I begin by examining the effect of margin trading eligibility on stock market reactions. Relative to non-treated peers, marginable stocks experience a 46–56% increase in share turnover, which corresponds to a 32–36% decline in investor average holding period. This frequent trading pattern clearly demonstrates margin traders' short-term orientation, a key feature that leads to managerial myopia under the catering theory in Baker et al. (2006). Moreover, I find positive returns for marginable stocks that rise from 2% during event days to 11% over three months. Since firms around the cutoff can be seen as randomly treated, the abnormal returns are unlikely to reflect any fundamental changes. On the other hand, this is consistent with short-term overvaluation driven by transient investors, which can pressure the manager to increase current performance to justify the inflated price (Jensen, 2005; Chi and Gupta, 2009).

Given the effect on share turnover and stock prices, I test whether firms cater to transient investors and undertake myopic actions to encourage, or at least maintain, the overvaluation. Specifically, I examine the top two shortsighted behaviors chosen by executives in the survey of Graham et al. (2005), namely earnings management and underinvestment. Relative to non-treated peers, marginable firms increase absolute discretionary accruals by 5.4% and reduce discretionary expenses by 1.7%. As for investment decisions, marginable firms reduce their capital expenditures by 4.6% relative to the control group (all scaled by beginning-of-year total assets). A similar decline happens to the asset growth rate of marginable firms as well.

I consider several alternative explanations for my results. One potential interpretation is that margin traders may monitor firms' operating activities and prevent the long-existing overinvestment and overspending. In contradiction to this monitoring hypothesis, I find that marginable firms experience a continuous decline in operating profit and equity valuation starting from the next year of the deregulation. Another explanation is that discretionary accruals may reflect changes in firms' investment and growth opportunities (Fairfield et al., 2003). This is less of a problem in my setting since accruals are measured one year before the investment variables due to the implementation timeline (more details in Section 1.5.5). Nonetheless, I consider several ways to control for firm growth and investment and find that my results hold with multiple specifications.

Managerial myopia has long been recognized in the literature. Theoretical evidence attributes this myopia to various reasons, such as CEO reputation (Narayanan, 1985), takeover threats (Stein, 1988), and earnings inflation (Stein, 1989). In the wake of the research on the real effect of financial markets, recent models such as Bolton et al. (2006) and Goldman and Slezak (2006) show that transient investors may encourage managerial myopia by weighing the executive compensation more heavily on the current stock price. However, empirical tests have generally fallen behind due to the difficulty of identifying transient investors and isolating their causal impact.

A growing number of empirical studies have attempted to link managerial myopia to investor short-termism. In his early works, Bushee shows that transient institutional investors tend to overweight current earnings and are associated with lower R&D expenditures (Bushee, 1998, 2001). Using Bushee's classification of institutions, other studies find similar results in the setting of acquisition monitoring (Chen et al., 2007) and around firms' addition to the Russell index (Cremers et al., 2020). However, since institutional ownership is inevitably correlated with many firm characteristics, existing literature still calls for causal evidence of this horizon alignment hypothesis. Moreover, the effect of retail investors, an active and important group of short-term traders, on managerial myopia is largely ignored.

This paper makes three contributions. First, my results confirm the survey evidence of widespread managerial myopia driven by transient investors (Graham et al., 2005; Beyer et al., 2014). In order to tackle the long debating endogeneity issue in this field, I exploit a natural experiment that shocks the participation of margin traders in China and provide some of the first causal evidence for this speculation-induced myopia. Notably, this paper relates to yet differs greatly in the investor composition with the literature that examines the effect of blockholder horizon on manager incentives. Besides the studies using Bushee's institution classifications, the existing literature tends to proxy for shareholder horizon using specific controlling blockholders with myopic incentives, say, venture capitalists (Cadman and Sunder, 2014), or option vesting conditions that implicitly indicate blockholder short-termism (Edmans et al., 2017). Therefore, one thing in common is that these studies all focus on large institutions and controlling blockholders and leave the role played by retail investors untouched. Others attempt to tackle this issue by examining investment decisions under general capital market pressure, such as public listing and increased reporting frequency (Asker et al., 2015; Kraft et al., 2018). Benefiting from the unique institutional setting in China, this paper is also among the first to examine the real effect of transient retail investors.

Second, this paper demonstrates at least one channel through which trading in secondary financial markets can affect firms' operating decisions (Bond et al., 2012). Rather than assuming investors play a role through implicit channels such as closed-door intervention and exit threat (McCahery et al., 2016), I show that unleashed retail speculators dramatically increase the share turnover and stock prices. Consistent with the catering theory, managers are pressured to cater to investor short-termism and join the notorious earnings game to maintain the inflated prices (Fuller and Jensen, 2010). My findings also echo the heated discussion about a speculative bubble driven by retail traders, especially with the rapidly growing Robinhood during COVID-19

(Welch, 2020). A clear example, undoubtedly, is the recent GameStop short squeeze that boosts its stock price by nearly 30 times in a month.

Third, my findings contribute to the policy debate on the benefits and costs of margin trading. Prior work suggests that margin trading can facilitate information flow (Seguin, 1990), provide liquidity (Kahraman and Tookes, 2017), yet, on the other hand, may also cause destabilizing speculation (Hardouvelis, 1990), generate liquidity spirals (Brunnermeier and Pedersen, 2009), and encourage risk-taking (Ben-David et al., 2018). As one of the first papers to study the real effect of margin trading, I find that margin traders can exert short-term pressure on firms and lead to myopic corporate behaviors. Consistent with Heimer and Simsek (2019), my results suggest that margin trading should be cautiously regulated, especially in a retail investor concentrated market.

The remainder of the paper is organized as follows. Section 1.2 provides a description of the margin trading system in China. Section 1.3 describes the data and the variable definitions. Section 1.4 explains the RD design. The empirical analysis of the impact of margin trading eligibility on stock performance and corporate decisions is in Section 1.5. Section 1.6 concludes.

1.2 Institutional Setting

Ever since the re-establishment of the stock market in the early 1990s, China has long banned margin trading in fear of causing irrational speculation among retail investors. Aiming to further develop its financial market, China launched a pilot scheme in March 2010 to lift the margin trading ban for designated stocks. As the counterpart of the SEC, the China Securities Regulatory Commission (CSRC) then announced that this pilot scheme would become a routine practice for both the Shanghai Stock Exchange and Shenzhen Stock Exchange. From early 2010 to late 2014, over 950 stocks were approved for margin trading throughout five major rounds.[4]

[4]The CSRC implemented another round in late 2019. However, compared to the original 50% initial margin requirement (same as the U.S.), this newest round requires a much higher minimum initial margin at 100%, meaning a margin trader has to pay all the purchase price with a leverage ratio of zero. Although securities may be discounted to provide initial collateral, this adjustment substantially lowered margin traders' borrowing ability.

The pilot scheme was implemented in two phases. In the first phase, Pilot A and B, firms in the corresponding market indexes were selected. On February 13, 2010, firms included in the Shanghai 50 Index and Shenzhen 40 Index got approved for margin trading. On November 25, 2011, the extended list further accommodated firms included in the Shanghai 180 Index and Shenzhen 100 Index.[5] However, this first phase has raised great concerns for identification. First, firms can easily predict future treatment and adjust operating decisions beforehand (Hansman et al., 2021). For example, suppose a firm in the Shanghai 180 Index reduces discretionary spending in preparation for the next round and adjusts back to normal afterward. In that case, we could even find an opposite effect of margin trading eligibility on firm operating decisions using the standard difference-in-differences approach. Second, stocks inside and outside the market indexes inevitably differ in many crucial features such as institutional ownership, ETF tracking, and board structure, leading to poorer comparability between treated and non-treated firms (Appel et al., 2016).

Therefore, this paper focuses on the second phase, including the next three rounds, where China uses a ranking procedure to decide the experiment eligibility. Specifically, the CSRC first excluded stocks that were extremely small and volatile.[6] Then, within each exchange, either Shanghai (SH) or Shenzhen (SZ), the CSRC ranked the remaining stocks according to the formula in Equation (1.1) and chose the top-ranked stocks to add to the margin trading list. The formula essentially calculates a weighted average of a stock's equity value and trading volume scaled by the entire exchange. Using this procedure, the CSRC further implemented three rounds: (1) 276 firms in Round 1 on January 31, 2013; (2) 206 firms in Round 2 on September 16, 2013;

[5]These market indexes generally include firms with the highest market value within each exchange. The Shanghai 50 Index covers the largest 50 firms in the Shanghai Exchange. Likewise, the Shanghai 180 Index covers the top 180 firms, including the aforesaid 50 firms. The Shenzhen 40 Index (also known as the Shenzhen Component Index) and the Shenzhen 100 Index are established in the same manner.

[6] Requirements: (1) have been traded for more than three months; (2) have either more than 200 million tradable shares or a market value over 800 million; (3) have more than 4,000 shareholders; (4) have not experienced any of the followings in the previous three months: (a) daily share turnover less than 20 percent of the share turnover of the market index; (b) the average of the absolute value of the daily price change deviates from that of the market index by more than 4 percent; (c) the maximum price difference scale by the average price is higher than that of the market index by five times; (5) have completed the share reform; (6) are not specially treated stocks; and (7) other unspecified conditions.

(3) 218 firms in Round 3 on September 22, 2014. The event day and the number of treated firms are both unknown ex-ante, making it unlikely, at least around the cutoff, to predict future margin trading eligibility.

$$\text{Ranking Index}_{\text{i}} = 2 \times \frac{\text{Average Tradable Market Value of Stock i}}{\text{Average Tradable Market Value of All Stocks in SH/SZ}} + \frac{\text{Average Trading Volume of Stock i}}{\text{Average Trading Volume of All Stocks in SH/SZ}} \tag{1.1}$$

Table 1.1 summarizes the eligibility decision rule, event day, and the number of newly marginable stocks for both stock exchanges across five rounds. My sample from Round 1 to Round 3 mainly covers mid-cap and small-cap firms that take up about 20% of the market value.

Margin trading rules in China are generally similar to those applied in the United States. For eligible stocks, minimum initial margins are set at 50% (that is, a margin trader may borrow up to 50% of the purchase price), and the minimum maintenance margins are set at 37.5% (that is, after purchase, a margin call will occur once the loan takes up more than 62.5% of the value of the stock held by the trader). Securities can be used to provide initial collateral with a pre-determined discount rate based on asset class and risk level (for example, 65% for stocks and 90% for ETFs), another feature shared with the United States. To prevent margin traders from using some of the borrowed money on non-marginable stocks, the CSRC set up a much higher requirement to extract cash from the margin account. I will discuss this feature and how margin accounts work in practice in more detail in the Appendix.

Alternative ways to obtain leverage in China are associated with extra costs and strict restrictions. Throughout the sample period of this paper, stock futures are only available for the market indexes, which means that investors can only increase their leverage by investing in the entire market.[7] Besides borrowing directly from the brokerage firms, investors may also take leveraged positions through informal channels, such as private financing companies, peer-to-peer

[7] As of December 2016, only three stock index futures are available in China, including Shanghai/Shenzhen 300 Index, China Securities 500 Index, and Shanghai 50 Index.

lending, and wealth management products from commercial banks. However, there are some crucial drawbacks to these informal channels. First, interest rates provided by these shadow financing lenders are around 25%, much higher than the 8.5% interest rates of brokerage firms (Bian et al., 2021). Second, these shadow financing contracts are not regulated or protected by the CSRC, which means investors have to bear an additional risk upon the already pressured leverage position. Importantly, any alternative channel would go against finding significant results in my empirical analysis.

Margin trading takes up a significant portion of the market value in China. Figure 1.1 plots the average outstanding margin debt scaled by total tradable market capitalization for newly marginable stocks in each round. All three rounds display a steep jump of the outstanding margin debt after the event day, then reach a stable level at 5%–7%. Compared with the 2% margin debt in the United States, margin trading plays a much more important role in China.

Due to the leverage risk and borrowing cost, margin trading has long been recognized as a short-term trading strategy. Moreover, almost all margin accounts are owned by retail investors.[8] To ensure market stability, the CSRC requires a qualified investor to have a trading account for at least six months, with a total account value of more than 500,000 RMB (USD 80,000). Unlike U.S. equity markets, China provides stock-level daily margin positions and trading data. In Figure 1.2, I estimate the average holding period for margin traders and that of average investors for all marginable stocks in rounds 1, 2, and 3. The holding period of margin traders (average investors) is calculated as the reciprocal of the stock-level/market-level share turnover of margin (total) trading activities from 2013 to 2016. Margin traders, on average, only hold stocks for two weeks, while an average investor in China holds stocks for 13 to 16 weeks. This difference is even more dramatic when we include firms in the pilot rounds or exclude margin trading activities when calculating average investors' holding period. Overall, margin traders are extremely short-term orientated in China.

A policy-related concern is that margin trading and short selling are simultaneously

[8] Institutional investors are banned from conducting margin trades through brokers in China.

approved for designated stocks. However, short selling in China has long been criticized for its minimal functionality. Previous studies such as Chang et al. (2014) attribute this deficiency to several reasons, including the high transaction cost, the up-tick rule, and the fact that most retail investors in China steer clear of short selling due to insufficient understanding of its mechanism. But most importantly, Chinese retail investors face a highly limited supply of security lending, especially for those mid-cap and small-cap firms examined in this paper.[9] As a result, the total outstanding shares of short selling take up less than 1% of those of margin trading; from a cross-nation perspective, daily shorting volume accounts for merely 0.49% of total trading volume during the sample period, which is negligible as compared to that of 20% in the United States (Boehmer and Wu, 2013). Moreover, my findings of positive *event-time* returns after the experiment (Section 1.5.2) further help to distinguish between margin trading and short selling. While short selling can better incorporate negative information and lead to lower stock prices (Grullon et al., 2015), margin trading, on the contrary, often stirs up speculation and boosts the stock price above its fundamental value (Hardouvelis, 1990). Benefiting from the RD design, I am able to identify positive *event-time* returns for treated stocks, which is consistent with the above consensus that margin trading dominates short selling in China. Importantly, however, since short-sellers have been shown to monitor and discipline firms' reporting and operating decisions (Massa et al., 2015; Fang et al., 2016), any synchronous short-selling activities would prevent me from identifying the myopic corporate actions driven by margin traders.

1.3 Sample, Data, and Variable Definitions

This paper investigates public firms that trade on the Shanghai Stock Exchange and the Shenzhen Stock Exchange in China. I collect information on the implementation details and daily margin trading data from the official exchange websites. Daily stock market data and

[9] Many media and professionals raised concern for the limited supply of security lending in China. For more details, please refer to the following websites: (1) http://finance.sina.com.cn/stock/quanshang/qsyj/20140710 (in Chinese); (2) https://www.globalinvestorgroup.com/articles/3431778/china-a-shares-lending-slow-to-progress (in English).

annual accounting information come from CSMAR (analogous to a combination of Compustat, CSRP, and I/B/E/S). My sample period is from January 2011, two years before Round 1 began, through December 2016. I exclude firms that have been treated more than once and firms in the financial industry.[10]

After excluding stocks not eligible for the ranking procedure, the full sample from Round 1 to Round 3 includes 672 newly-treated stocks and 4,083 non-treated stock-round observations (1,868 unique stocks). By design, the RD sample includes stocks close to the experiment cutoff in each round. As suggested by Calonico, Cattaneo, and Titiunik (2014), denoted CCT hereafter, I choose the bandwidth using a data-driven MSE-optimal selector (more details in Section 1.4.3). On average, the local sample covers 40% of the treated firms, more or less, depending on the examined outcome variable, and 57% have entered the local sample for one or more tests.

The two myopic behaviors in the paper are earnings management and long-term investment cuts. To measure accrual-based earnings management, I use the absolute value of discretionary accruals calculated from the modified Jones model and refer to it as *ADA* throughout the analysis.[11] Discretionary accruals are estimated from cross-sectional regressions of total accruals on changes in sales and PP&E within each industry. Therefore, higher *ADA* indicates that a firm reports accruals far from its industry standard. For real earnings management activities, I use abnormal discretionary expenses measured as the change of SG&A expenses and denote it as $\Delta DISEXP$ hereafter. Reducing such spending will boost current period earnings and lead to higher current period cash flows if firms generally pay such expenses in cash.

I use the abnormal capital expenditures $\Delta CAPEX$ and changes in total assets $\Delta ASSETS$

[10]Marginable firms with negative earnings for two consecutive years (ST firms) may be removed from the margin trading list and rejoin later when their earnings turn positive. These firms take up less than 1% of my sample and do not affect my results. Besides the convention of removing the financial industry in the literature, financial firms also directly or indirectly serve as the leverage provider for margin traders.

[11]Since margin trading eligibility is almost permanent in China, firms with excessively high profits are incentivized to save "extra" earnings through accruals in a repeated game. Moreover, because earnings smoothing is even more important than hitting earnings target in the eyes of executives (Graham et al., 2005), *ADA* is the appropriate measure to use to determine whether accruals management occurs (e.g., Bergstresser and Philippon (2006) or Cornett et al. (2008)). In contrast, real earnings management is always directional because firms face huge costs adjusting operating decisions and cannot accurately control the manipulation results like in the "accruals game".

as measures of corporate investment.[12] $\Delta CAPEX$ is equal to the annual change in capital expenditures scaled by beginning-of-year total assets. $\Delta ASSETS$ is equal to the percent change in total assets. Throughout this paper, I mainly examine the annual change of those non-stationary variables to alleviate the impact of any omitted time-invariant firm characteristics. A detailed variable construction is provided in the Appendix.

Table 1.2 reports descriptive statistics for firm-round observations local to the eligibility cutoff. As these statistics depend on the bandwidth, I provide results for multiple bandwidth choices. The mean book value of assets of local firms is around 800 million RMB (about 125 million US dollars). As an important validity test for the RD design, I verify that none of the fundamental variables displays a significant discontinuity at the cutoff in Section 1.5.4.

1.4 Empirical Methodology

This section explains the replication of the ranking procedure and the fuzzy regression discontinuity design. For each round, non-marginable firms are first ranked based on Equation (1.1) and will then be added to the margin trading list if their ranking order is higher than the number of treated firms. Therefore, my identification strategy comes from the discontinuity in margin trading eligibility at the experiment cutoff.

1.4.1 Defining the Standardized Forcing Variable

Following CSRC's rules, for each non-marginable stock eligible for the ranking procedure, I calculate its index using Equation (1.1) and denote stock i's index for round k as $Index_i^k$, where $k = 1, 2, 3$.[13] Next, as suggested by Hansman et al. (2021), I determine the experiment cutoff as

[12] I omit R&D in the main analysis because: (1) R&D data in China is of relatively low quality and reduces my sample by about 30%; (2) R&D spending can be attributed to both real earnings management and long-term investment cuts, making it difficult to differentiate between these two myopic behaviors. My results do hold when including R&D in either of these two activities (Section 1.5.4).

[13] The time window used to calculate the ranking index is not officially confirmed by the CSRC. Following Hansman et al. (2021), I use three calendar months before the announcement date because: (1) the 3-month window aligns with the calculation window listed in the screening criteria; (2) industry sources also suggest a 3-month pre-event window.

follows. First, for each round and each exchange, I sort all stocks based on their ranking indexes. Second, I take the exact number X of treated stocks in that round and set the cutoff to be the ranking index of the *Xth* highest-ranked stock, regardless of whether that stock is treated or not. This procedure ensures that the ranking index solely decides whether a firm is placed above or below the experiment cutoff. The cutoff for round k and exchange E is denoted as C_E^k.

Lastly, I normalize the cutoff to zero and generate a standardized forcing variable for stocks in each round and each exchange. This procedure makes the forcing variable comparable across different rounds and identifies the average treatment effect on the observed distribution of individuals close to the cutoff.[14] For example, a standardized forcing variable V_i^k of 1 (-1) indicates that stock i has a ranking index one standard deviation higher (lower) than the experiment cutoff in round k.

A requirement for the regression discontinuity design is that firms cannot accurately manipulate their ranking to be treated or not *close* to the experiment cutoff (McCrary, 2008). In my setting, firms are not likely to do so because: (1) firms do not know the event day and the number of treated firms ex-ante; (2) the inputs for the ranking formula are determined by investors' daily trading activities with high-frequency variations. Figure 1.3 plots the frequency density of firm-round observations within each bin of the forcing variable around the cutoff. I also plot local polynomial density estimates (solid blue and red) and robust bias-corrected confidence intervals (shaded blue and red). There is no evidence of bunching at the cutoff. Following Cattaneo et al. (2020), I further run a manipulation test at the cutoff. The final test result is $T = 0.3326$, with a p-value of 0.7394, suggesting no statistical evidence of manipulation either.

[14] Another advantage of round standardization is that it makes the final sample equally distributed among three rounds. Since the ranking formula generates a score scaled by the entire exchange, firms in Round 3 (slightly smaller) are naturally more densely distributed around the threshold than firms in Round 1 (slightly larger). Without round normalization, the final sample would include a disproportionate number (45%) of firms in Round 3.

1.4.2 Fuzzy RD Regression Specifications

I use a fuzzy regression discontinuity design to study the causal effect of margin trading eligibility. The treatment discontinuity is not perfectly sharp at the cutoff because: (1) the specific window used to calculate the ranking index is not officially confirmed by the CSRC; (2) the CSRC occasionally exercises its discretion to decide a firm's inclusion. Therefore, I use a standard two-stage least squares (2SLS) estimation for this fuzzy RD setting.

The first stage is given by Equation (1.2), where I use an indicator of being above the cutoff as an instrument variable for margin trading eligibility.

$$Marginable_i^k = \beta_1 \cdot \mathbf{1}[V_i^k > 0] + R(V_i^k) + L(V_i^k) + \eta^k + \varepsilon_i^k \tag{1.2}$$

Here, $Marginable_i^k$ is a dummy variable that equals one if stock i is added to the margin trading list in round k, and zero otherwise. $\mathbf{1}[V_i^k > 0]$ is an indicator that equals to one if stock i is ranked above the cutoff in round k, and zero otherwise. Gelman and Imbens (2019) suggest that using high-order polynomials in RD analysis is a flawed approach due to its noisy estimates, order-sensitive results, and poor coverage of confidence intervals. Therefore, I include separate linear controls of the forcing variable V_i^k that allow for different slopes on the right-hand side of the cutoff $R(V_i^k)$ and on the left-hand side of the cutoff $L(V_i^k)$. η^k represents the round fixed effect. There is no need for firm fixed effects since non-marginable firms are only treated once in the experiment. The resulting estimates of β_1, representing the additional probability of being marginable when ranked above the cutoff, are reported in Section 1.5.1.

For the second stage, I estimate a similar relationship between margin trading eligibility and the outcome variable. For each outcome variable, I estimate the following equation:

$$Y_i^k = \beta_2 \cdot \widehat{Marginable}_i^k + R(V_i^k) + L(V_i^k) + \eta^k + \varepsilon_i^k \tag{1.3}$$

$\widehat{Marginable}_i^k$ is the predicted value of margin trading eligibility from the first stage. The resulting estimate of β_2 represents the effect of margin trading eligibility on the outcome variable Y_i^k. Different control variables are included depending on the outcome variable examined.

1.4.3 Bandwidth Selection

RD design restores the treatment randomness by focusing on firms close to the experiment cutoff. Therefore, choosing a proper bandwidth is important in my setting. The main objective is to maintain the exogenous treatment variation (smaller bandwidth) and provide sufficient statistical power in estimation (larger bandwidth). In the analysis that follows, I choose a fully data-driven bandwidth selector suggested by Calonico, Cattaneo, and Titiunik (2014). As the current state of the art, CCT stand out in the RD literature because they provide mean squared error optimal bandwidths and reduce the inference bias caused by overestimated bandwidths in alternative methods.[15]

Since there is no perfect bandwidth, I examine alternative bandwidth choices in the robustness check. Specifically, I increase/decrease the CCT bandwidth by 10% and 25% (which expand/shrink my sample size by about 15% and 35%) and redo the RD analysis. Because CCT choose the optimal bandwidth based on the distribution of the dependent variable, the bandwidth and sample size slightly varies with the outcome variable we are examining.

1.5 Results

1.5.1 First-Stage Regressions

In this section, I show that there is indeed a discontinuity in the margin trading eligibility at the cutoff. Figure 1.4 plots the margin trading eligibility against the forcing variable. I include all firms within the CCT bandwidth, and each point denotes the average probability of being marginable for firms within each evenly-spaced bin. Separate regression lines, along with

[15] Alternative methods include rule-of-thumb, cross-validation, and other MSE-optimal bandwidth selectors (Imbens and Kalyanaraman, 2012). MSE-optimal selectors gradually take the place of the former two methods because they provide solid theoretical support and a data-driven algorithm. Of the last category, Calonico et al. (2014) improve on prior studies by offering a more conservative bandwidth with smaller inference bias. See a more detailed review of these methods in Kahraman and Tookes (2017).

90% confidence intervals based on heteroscedasticity-consistent standard errors, are shown on both sides of the eligibility cutoff. As expected, the RD plot shows a sharp discontinuity in the treatment probability at the experiment cutoff. Table 1.3 reports the first-stage regression, directly running margin trading eligibility on the indicator of being above the cutoff. Consistent with Figure 1.4, regression results show a dramatic jump of the treatment probability of 54% at the cutoff. The estimates are almost identical with the triangular kernel, and I use the uniform (rectangular) kernel for my main analysis hereafter.[16]

Both the RD plot and regression results confirm an evident discontinuity in margin trading eligibility at the cutoff. Besides the evidence of no bunching at the cutoff (Section 1.4.1), I further confirm there is no significant discontinuity of any firm characteristics in Section 1.5.4. Overall, the margin trading deregulation in China is a great fit for the regression discontinuity design.

1.5.2 Stock Market Reactions

In this section, I examine the impact of margin trading eligibility on share turnover and stock prices. Under the catering theory (Baker et al., 2006), managers cater to short-term traders' demand for earnings target and free cash flows, and take myopic actions that can encourage overvaluation. In equilibrium, investors expect managers' catering behaviors and trade against each other to profit in the speculative bubble (Bolton et al., 2006). Therefore, we expect to see an increase in share turnover, indicating the short-term horizon of margin traders, and an increase in stock prices, representing the overvaluation driven by the competing speculators.

[16] The triangular kernel gives a higher (lower) weight for firms close to (far from) the cutoff, while the uniform kernel equally weights all firms within the CCT bandwidth. The uniform kernel better suits my setting because: (1) the coefficient of estimates is easier to interpret; (2) firms closer to the cutoff in the fuzzy RD design are more likely to be non-compliers, that is, treated (untreated) when ranked below (above) the cutoff. My results do hold with different kernel choices.

Share Turnover

In this section, I test whether marginable stocks experience an increase in abnormal share turnover, measured as the change of share turnover before and after the implementation date using event windows of 4 weeks, 8 weeks, and 12 weeks.[17] The share turnover itself is defined as the daily share volume over tradable shares outstanding (Lo and Wang, 2000). Besides the benefit of controlling time-invariant omitted variables, using abnormal share turnover also alleviates the concern that trading volume is included in the ranking formula. Any pre-event volume difference has been excluded in this procedure. Figure 1.5 plots the abnormal share turnover against the forcing variable. To control for time-series variation, I demean the outcome variable using the average values of all stocks within the CCT bandwidth for the round. Each point denotes the average abnormal share turnover within each bin of width 0.05. Separate regression lines along with 90% confidence intervals are shown on both sides of the eligibility cutoff. All figures display an evident discontinuity in abnormal share turnover at the threshold.

Panel A of Table 1.4 reports the reduced form regression in which I run the abnormal share turnover on the indicator of being above the cutoff. Therefore, the estimated coefficient of 0.569 matches closely with the cutoff jump of abnormal share turnover in the 8-week case of Figure 1.5. Panel B reports the fuzzy RD results using 2SLS estimation. The outcome variable is the abnormal share turnover, and the main explanatory variable of interest is an indicator variable of margin trading eligibility. Following Chordia et al. (2007), I control for the previous month's absolute stock return, stock price, analyst following, forecast dispersion, and other annual frequency variables including earnings surprise, leverage, book-to-market ratio, market capitalization, and beta. Although these controls are not exhaustive, one of the biggest advantages of the RD design is that it compares similar firms, which renders it fairly robust to omitted variables. Depending on different event windows, marginable stocks experience a 0.8–1% (18–22%) increase in daily (monthly) share turnover relative to non-marginable peers. For a

[17] The pre-event window is fixed at 12 weeks before the event day, the same length as the ranking procedure.

representative firm with a median monthly share turnover of 39% in my sample, margin trading eligibility leads to a 46–56% increase in share turnover, corresponding to a 32–36% decline of investor average holding period. Both the economic magnitude and the statistical significance of these results are large, suggesting that margin trading eligibility leads to a significant decline in the average shareholder horizon.

Returns

This section examines the impact of margin trading eligibility on stock prices, measured with the cumulative returns over event days and in a longer post-event window of three months. Figure 1.6 plots the cumulative returns against the forcing variable. All specifications are the same as before. Both figures display an evident jump in stock returns at the cutoff. Table 1.5 reports the reduced form regression and the fuzzy RD regression using 2SLS estimation. The outcome variable is the stock returns, measured as the raw/market-adjusted cumulative abnormal returns (CAR) right after the implementation day (event window [0,1]) and raw/DGTW-adjusted cumulative returns in a longer post-event window of three months.[18] Relative to non-treated peers, marginable stocks experience positive returns of 2.3–2.4% during event days and 9–11% over three months.[19]

Overall, I find that margin trading eligibility leads to higher stock prices. Since firms around the cutoff can be seen as randomly treated, this abnormal return is unlikely to reflect any fundamental changes. Instead, the gradually increasing prices are consistent with the equilibrium investor-manager dynamics in the catering theory, where investors expect managers' catering behaviors and trade against each other to profit in the speculative bubble.

[18] I follow the same procedure as in Daniel et al. (1997) and sort all stocks into 125 ($5 \times 5 \times 5$) portfolios based on size, book-to-market ratio, and momentum. The equal-weighted returns for each portfolio serve as the benchmark return for all stocks in the same portfolio.

[19] My results are consistent with Hansman et al. (2021), who find positive returns for marginable stocks over three months and twelve months. Although I find similar results for a longer period, I remain conservative about the post-event window since the shortest time gap between rounds is eight months.

1.5.3 Managerial Myopia

In this section, I test whether managers cater to investor short-termism and undertake myopic behaviors to encourage, or at least maintain, the mispricing. Following survey evidence of Graham et al. (2005), I examine the top two shortsighted behaviors chosen by executives, namely earnings management and long-term investment cut.

Earnings Management

I use absolute discretionary accruals *ADA* to measure accrual-based earnings management and abnormal discretionary expenses Δ*DISEXP* to proxy for real earnings management. Higher *ADA* means a firm adjusts accruals far from its industry-standard. On the other hand, lower Δ*DISEXP* suggests a decline in firms' discretionary spending, a common approach to boost current period earnings and increase cash flows. Figure 1.7 plots *ADA* and Δ*DISEXP* against the forcing variable. All specifications are the same as before. RD plots show an evident jump of *ADA* and a clear drop of Δ*DISEXP* at the cutoff.

Table 1.6 reports the reduced form regression and the fuzzy RD regression using 2SLS estimation. The outcome variables are *ADA* and Δ*DISEXP*, and the main explanatory variable of interest is an indicator variable of margin trading eligibility. Following Hribar and Nichols (2007) and Yu (2008), I control for market capitalization, market-to-book ratio, leverage, operating ROA, operating cash flow, cash flow volatility, the growth rate of assets, and lagged discretionary accruals. I also report pre-event (one year before the experiment) results as a placebo test. Relative to non-treated peers, marginable firms increase discretionary accruals by 5.4% and reduce discretionary expenses by 1.7% (scaled by total assets), suggesting an increase in both accrual-based and real earnings management. In contrast, there is no discontinuity in the pre-event period, indicating my results are not driven by any mechanical difference between treatment and control groups. In Section 1.5.4, I examine alternative measures, including abnormal discretionary accruals Δ*ADA* and a slightly modified version of Δ*DISEXP* that includes R&D expenses, and find similar results.

The magnitude of these effects is generally in line with magnitudes documented in other studies on the impact of short-termism on earnings management. For example, Cornett et al. (2008) find a one-sigma increase in the option compensation variable is associated with an increase in absolute discretionary accruals that takes up 4.5% of total assets. Ali and Zhang (2015) find that CEOs in their earlier years are associated with 0.95% lower abnormal discretionary expenses as scaled by assets, which further goes up to 1.94–2.67% after controlling for different monitoring parties such as institutional investors, analysts, and independent boards. Although these studies focus directly on the manager horizon itself, they provide a valuable reference to analyze the impact of investor short-termism in my setting.

Due to the specific timeline of the experiment, *ADA* is measured one period before $\Delta DISEXP$. This is because, unlike the accruals earnings management that takes as little time as a firm would spend on preparing the annual report, operational decision making requires multiple rounds of board meetings and even more time in the actual implementation (Badertscher, 2011). Therefore, I calculate discretionary accruals using annual reports published after the deregulation (Iliev, 2010) and require real earnings management (as well as investment variables in the next section) to overlap at least six months to be counted in a fiscal year (Grullon et al., 2015). Since all three rounds are implemented at the end of the year, *ADA* is measured one year before $\Delta DISEXP$.[20] The insignificant pre-event results provide additional support that the estimation period for each outcome variable is valid. Notably, my results may not suggest a substitutive relation between accrual-based and real earnings management since the lag of $\Delta DISEXP$ may well result from insufficient time for operational adjustment in my setting (Zang, 2012).

Long-Term Investment

I use abnormal capital expenditures $\Delta CAPEX$ and changes in total assets $\Delta ASSETS$ to measure long-term investment. Figure 1.8 plots $\Delta CAPEX$ and $\Delta ASSETS$ against the forcing

[20] All firm fiscal years end in December in China. Although Round 1 was implemented in January, firms still have enough time for accruals management. This is because less than 1% of firms publish annual reports in January, and more than 95% do so after March. My results do hold with only firms that publish annual reports after February or March.

variable. All specifications are the same as before. RD plots show a noticeable drop of both $\Delta CAPEX$ and $\Delta ASSETS$ at the experiment cutoff. Table 1.7 reports the reduced form regression and the fuzzy RD regression using 2SLS estimation. As mentioned in the last section, investment variables are required to overlap at least six months to be counted in a fiscal year. Following Grullon et al. (2015), I control for variables that may affect firm investment in event studies, including operating cash flow, lagged total assets, and past profitability (operating ROA).

Relative to non-treated peers, marginable firms reduce capital expenditures by 4.6% and lower their asset growth rates by 28%. Although the effects on asset growth rates appear to be large, they should be considered together with the average asset growth rate of 19% during the sample period. In other words, the actual decline in asset growth for a marginable firm is around 9%. Once again, there is no discontinuity in the pre-event period. In the robustness check, I include R&D expenses in the long-term investment and find similar results. The magnitude of these effects is generally in line with other studies on the impact of capital market pressure on long-term investment. For example, Asker et al. (2015) find that public firms are associated with 2.4–4.9% lower gross investment and 9.1% lower asset growth rates compared to private firms. Kraft et al. (2018) find a slightly smaller decline in capital expenditures that takes up 1.8–1.9% of total assets along with the transition of firms from annual reporting to quarterly reporting. Considering the dramatic change in share turnover and stock prices in China, my results that are about the same level as the public listing status and twice that of the increased reporting frequency are fairly reasonable.

My results suggest that margin trading eligibility leads to a decline in firms' long-term investment. Survey evidence of Graham et al. (2005) shows that 56% of the executives are willing to delay profitable projects and capital expenditures in pursuit of investors' demand for current period earnings and cash flows. My findings confirm this consensus and provide one of the first pieces of evidence on a causal link between short-term speculators and managerial myopia.

1.5.4 Robustness Check

Alternative Bandwidths

Regression discontinuity design stands out in field studies since it restores the treatment randomness by focusing on firms close to the experiment cutoff. Hence a proper bandwidth is important in the RD setting. Throughout the main analysis, I let the CCT data-driven selector choose the optimal bandwidth for each outcome variable. In this section, I check whether my results are sensitive to alternative bandwidth choices. Table 1.8 reports the fuzzy RD results with alternative bandwidths. Specifically, I increase/decrease the CCT bandwidth by 10% and 25% (which expand/shrink my sample size by about 15% and 35%) and redo the RD analysis for all outcome variables, including the abnormal share turnover, returns, *ADA*, $\Delta DISEXP$, $\Delta CAPEX$, and $\Delta ASSETS$. Overall, my findings are robust to alternative bandwidths.

Placebo Tests

Like any other RD design study, one alternative interpretation is that the ranking index itself can predict future corporate behaviors and market reactions. To ensure that my results are not driven by variation in the ranking procedure, I repeat the RD analysis around multiple placebo cutoffs. Specifically, I set placebo cutoffs above and below the actual cutoff 0, using a fixed distance of 0.5 and a flexible distance based on the CCT bandwidth for each outcome variable. There is no need for a 2SLS estimation since, for placebo tests, the first stage regression is no longer predictive. Table 1.9 reports the reduced form RD regression results for all outcome variables at placebo cutoffs. Unlike the main analysis, I do not observe a significant cutoff discontinuity in any outcome variables. Therefore, the ranking index itself is unlikely to explain my results.

Validity Tests

I have confirmed no pre-event discontinuity of any dependent variable in the main analysis. In this section, I check the extent to which covariates exhibit discontinuity at the cutoff.

Validity tests are important in the RD design because discontinuity in pre-event covariates may indicate that firms above and below the cutoff are systematically different. Table 1.10 reports the fuzzy RD results for all control variables used in this study, measured one year before the margin trading deregulation. All specifications are the same as the main analysis. Overall, there is no significant discontinuity in any control variables at the cutoff.

Alternative Measures

In this section, I examine whether my results are sensitive to alternative measures. For accrual-based earnings management, I use *ADA* rather than Δ*ADA* in the main analysis because the latter may not be suitable for event studies. Since discretionary accruals are defined as the residuals of cross-sectional regressions within each industry year, a change of *ADA* does not necessarily imply a firm engage in accruals management. In other words, Δ*ADA* not only depends on firms themselves but also on their industry peers. However, like other outcome variables, Δ*ADA* has the advantage of alleviating ex-ante accruals difference between firms above and below the cutoff.

R&D is another widely used variable in managerial myopia literature. However, R&D data is of relatively low quality in China. Throughout my sample period, only two-thirds of firms in the local sample report R&D information. Even for those disclosing firms, R&D expenses merely take up 30% of capital expenditures and 20% of SG&A expenses. Compared with an almost 1:1 ratio between R&D and CAPEX in the United States, R&D plays a much less important role in Chinese firms' investment policy. Moreover, reducing R&D expenditures has been attributed to both real earnings management and long-term investment cuts in the literature, making it difficult to differentiate between these two myopic behaviors. Nonetheless, I respectively include R&D in abnormal discretionary expenses (Δ*DISEXPRD*) and in long-term investment (Δ*CAPEXRD*) and redo the analysis.

Table 1.11 reports the regression results for all variables discussed above. For a better comparison with previous results, I use the same bandwidth as that of the main analysis for each outcome variable. Overall, my results are robust to these alternative measures.

1.5.5 Alternative Explanations

So far, my results suggest that margin traders, as short-term speculators, pressure the manager to focus on current earnings and sacrifice long-term growth. In this section, I evaluate multiple alternatives to this explanation.

Managerial Myopia or External Monitoring?

One explanation for my results is that margin traders may monitor firms' operating activities and prevent long-existing overinvestment and overspending. This monitoring hypothesis is questionable in the first place because margin traders in China are retail investors with limited information sources and little involvement in the board meeting. Nonetheless, firms may still choose to reduce overinvestment simply because they fear getting caught for suboptimal operating decisions when put in the spotlight. Since firms' optimal operating strategy is unobservable, it remains an empirical question whether reducing investment and discretionary spending is beneficial in the long run.

I examine the long-term performance of marginable firms to test whether they adjust operating decisions in a myopic way. Myopic behaviors, by definition, would lead to a decline in long-term operating profitability, as well as lower equity valuation when markets find out firms' real sacrifice. Notably, the standard RD design no longer applies because non-treated firms in an earlier round may enter the treatment group in the next one. To tackle this issue, I adopt a dynamic version of the regression discontinuity design and estimate the following equation:

$$Y_{i,t+\tau} = \beta^{\tau} \cdot Marginable_{it} + R^{\tau}(V_{it}) + L^{\tau}(V_{it}) + \alpha_{\tau} + \eta_{c} + \lambda_{it} + \varepsilon_{it\tau} \tag{1.4}$$

This specification follows expression (7) in Cellini et al. (2010). $Marginable_{it}$ is a dummy variable that equals one if stock i is approved for margin trading at time t, and zero otherwise. The term β^{τ} estimates the effect of margin trading eligibility at time t on outcome variables τ

periods later. For example, $\beta^{\tau=2}$ estimates the treatment effect two years after the deregulation. For each firm-round (i,t), I include observations for firm i from period $t-2$ (two years before) to $t+4$ (four years after) and exclude the event year t. Therefore, the effect on future outcome variables is estimated based on pre-event firm characteristics. The coefficient β^{τ} and separate linear controls of the forcing variable $R^{\tau}(V_{it})$ and $L^{\tau}(V_{it})$ are allowed to vary for $\tau > 0$, and constrained to zero for $\tau < 0$, and standard errors are clustered at the firm level. The parameters α_{τ} and η_{c} are fixed effects for time periods relative to the event and calendar years. Additionally, pooling multiperiod data allows me to include firm-round fixed effect λ_{it} and control for time-invariant firm characteristics. However, the CCT bandwidth selector cannot be applied to the dynamic RD design. For robustness, I report results for multiple bandwidth choices.

Table 1.12 presents the effect of margin trading eligibility on firm long-term performance. Marginable firms experience a decline in operating ROA starting from the next year of the deregulation. The estimated coefficient -0.02 corresponds to a ROA drop of 36%, representing 27% of the standard deviation (mean and standard deviation of ROA are 0.056 and 0.074, respectively). This declining profitability persists in years $t+3$ and $t+4$, although only marginally significant. Stock returns show similar results, where marginable firms experience a 17% price decline in the next year and a slightly downward trend after that. Overall, my results are against the interpretation that margin traders serve as external monitors and prevent suboptimal operating decisions.

Accruals, Growth, and Investment

Prior literature shows that a firm's accruals correlate with its growth opportunity and investment level (Fairfield et al., 2003; Zhang, 2007; Wu et al., 2010). Since I use discretionary accruals to proxy for accrual-based earnings management, an alternative explanation is that the accruals discontinuity may result from an investment change.

Before diving into a more detailed analysis, I would like to explain why this alternative explanation is unlikely in the first place. In my setting, discretionary accruals are measured one year before the investment variables (more details in Section 1.5.3). This unique feature benefits from the fact that accruals management takes as little time as a firm would spend on preparing the annual report, while, in contrast, operating decisions require multiple rounds of board meetings and even more time in the actual implementation. This sequential relation between accruals and investment makes it almost impossible for firm investment to reversely affect accruals in the preceding year.

Nonetheless, I adopt several controls for firms' investment and growth to thoroughly investigate this concern. Besides the asset growth rate $\Delta ASSETS$ that I have already controlled in the main analysis, I further add $\Delta CAPEX$ to control for firms' growth opportunities. The results are reported in Table 1.13. I include $\Delta CAPEX$ in column (1) and the squared terms of both $\Delta ASSETS$ and $\Delta CAPEX$ to account for a nonlinear effect of investment on accruals. The coefficient on the margin trading eligibility, $Marginable_i^k$, are barely affected by the inclusion of these controls. In columns (3) and (4), I construct a slightly modified version of $\Delta CAPEX$ by including R&D expenses and denote it $\Delta CAPEXRD$. Again, the results are similar to those reported in Table 1.6.

1.5.6 Discussion

Before concluding, I would like to discuss the extent to which my finding can be generalized outside of the institutional setting in this paper. Although recent models of managerial myopia have been written with developed markets in mind, the core assumption of these models—information asymmetry and heterogeneous beliefs—also applies, if not more, to emerging markets. Moreover, besides the common underlying mechanism, China and the U.S. are comparable in market size, regulation rules, and margin trading patterns. As of the end of 2015, China's stock market capitalization stood at \$8.2 trillion, the second-largest in the world, ranking behind only the U.S. equity markets (\$25 trillion). China also set up almost identical minimum requirements of initial margins (50%) and maintenance margins (37.5%) as those of the United States. Not surprisingly, margin traders in China and the U.S. display similar trading patterns in which they increase margin positions when the stock market performs better and decrease them when the market gets worse.[21] All the similarities listed above, coupled with the overwhelming blame on speculators from the U.S. entrepreneurs, alleviate the external validity concern on the causal link between investor short-termism and managerial myopia.

The magnitude of these myopic effects, however, may be larger in China than in the United States. An outstanding feature of China is that retail investors dominate the stock markets and account for 85% of the total trading volume (25% in the U.S.). As a result, Chinese stocks generally have higher share turnover and a shorter average holding period. Notably, this higher share turnover itself should not affect my findings since I essentially compare the abnormal share turnover for firms on either side of the experiment cutoff. It is the possibility—overconfident or irrational retail investors freed from financial constraints may generate higher speculative pressure than institutional investors do—that might contribute to a larger impact of margin trading eligibility in China than in the United States. Nevertheless, in the wake of retail investor participation along with the rapidly growing Robinhood (Welch, 2020), this concern should not be central to the generalization of my results. The recent short squeeze of GameStop best demonstrates that even in a developed financial market like the United States, speculative retail investors can well push the stock price far above its fundamental value. In a word, although the magnitude of the effects in China may not be identical to developed markets, my paper still provides a valuable reference to future regulations around the world.

Finally, an RD-specific external validity concern is whether firms close to the experiment cutoff are representative of the China economy. This is less of a problem in my setting since, out of the 672 treated firms in my sample period, 383 or 57% of firms have entered the local sample for one or more tests. Except for those extremely large and tiny firms, my results can apply to a large group of firms in China.

[21] Since NYSE only discloses monthly market-level margin positions, I compare the correlation between monthly market returns and monthly changes in outstanding margin positions in the two countries. This correlation is 0.64 in China and is 0.60 in the United States, both positive and significant.

1.6 Conclusions

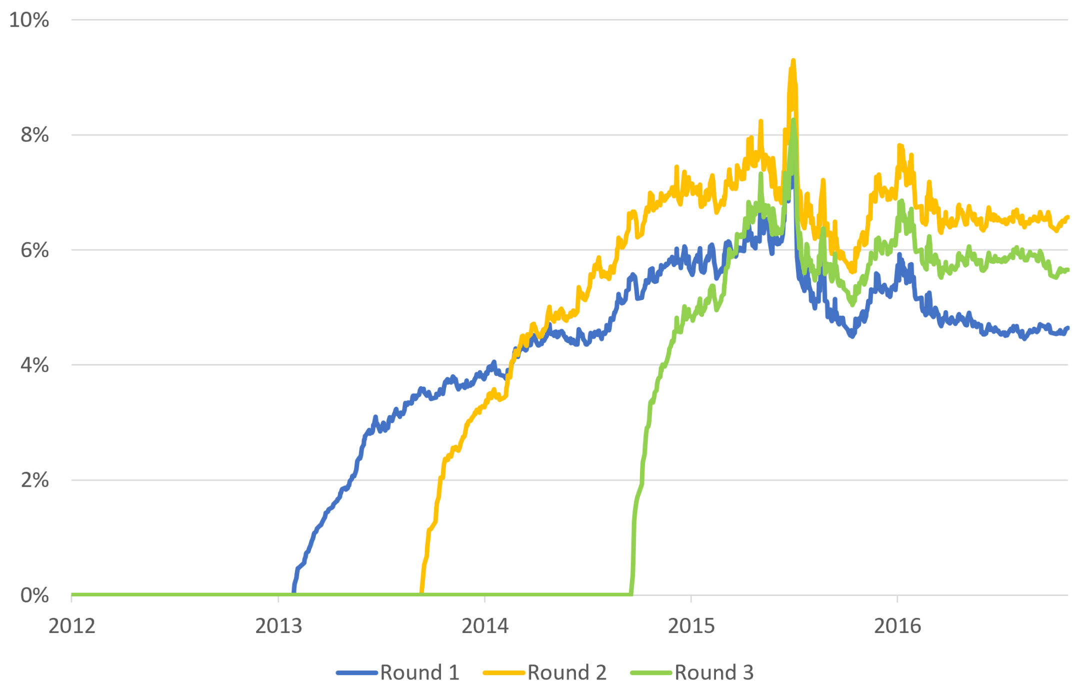
This paper investigates whether speculative retail investors impact stock prices and distort corporate decisions. To provide causal evidence, I exploit a rank-based margin trading deregulation in China and compare corporate reporting, operating, and investment decisions for firms that barely passed the experiment cutoff to those of firms that just missed it. Building on prior literature, my key identification assumption is that, on average, margin traders are more speculative and short-term orientated than non-leveraged investors. The data supports this assumption.

My results show that relative to non-treated peers, marginable stocks experience a dramatic increase in share turnover and hence a decline in investor average holding period. This frequent trading pattern clearly demonstrates margin traders' short-term orientation, a key feature that leads to managerial myopia under the catering theory. Moreover, I find gradually increasing returns for marginable stocks, consistent with the notion that in equilibrium, investors expect managers' catering behaviors and trade against each other to profit in the speculative bubble. On the other side of the coin, managers cater to investor short-termism and undertake myopic behaviors such as earnings management and long-term investment cuts to encourage or maintain this overvaluation. Lastly, consistent with managerial myopia, marginable firms experience a continuous decline in operating performance and equity valuation starting from the next year of the deregulation.

My findings highlight investor short-termism as a key driving factor for managerial myopia. A number of authors have emphasized the adverse effect of short-term institutions and controlling blockholders. However, the role played by retail investors, an active and important group of short-term traders, is largely ignored. Benefiting from the unique institutional setting in China, my paper not only confirms the myopic effect of speculative retail investors but also contributes to the broader literature that studies the real effect of secondary financial markets. My results give support to a worldwide policy change that aims for a longer investment horizon as well as a tightening of the leverage constraint in both China and the United States.

1.7 Acknowledgements

Chapters 1, in full, is currently being prepared for submission for publication of the material. Jun Chen, the dissertation author, is the primary investigator and author of this paper.

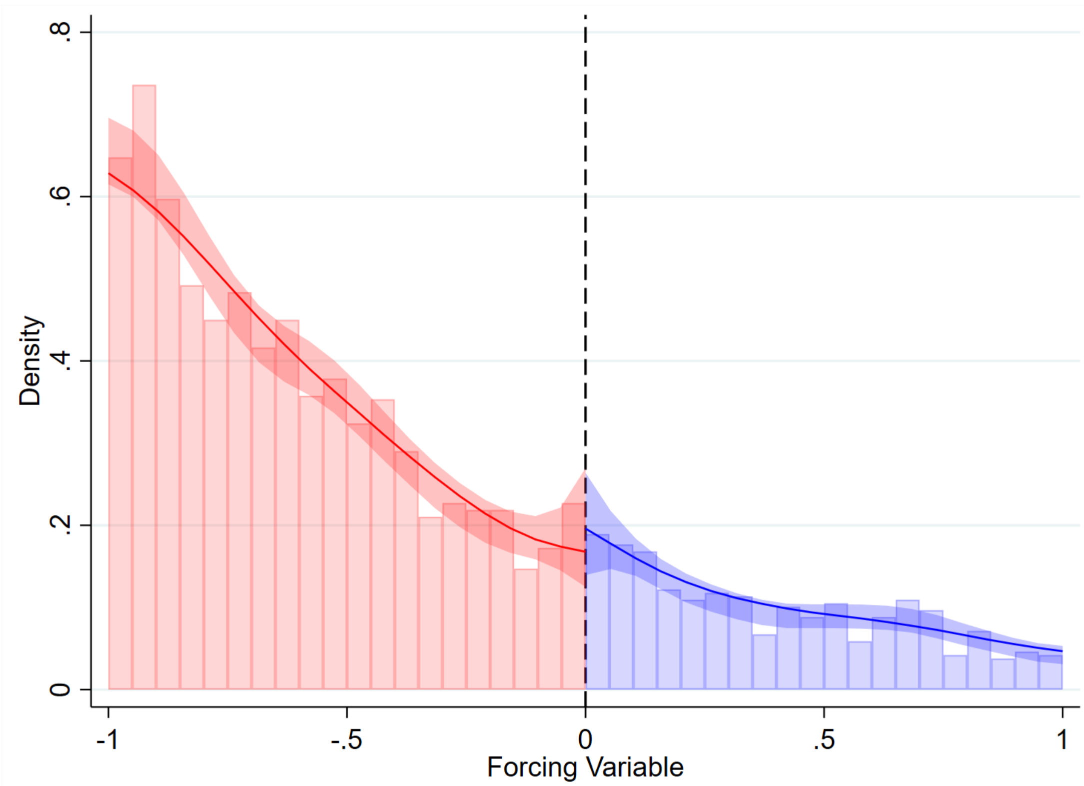


This figure plots the average stock-level outstanding margin debt scaled by total tradable market capitalization for newly marginable stocks in rounds 1, 2, and 3.

Figure 1.1. Outstanding Margin Debt by Round

This figure plots the average holding period of margin traders and average investors for all newly marginable stocks in rounds 1, 2, and 3. The left column calculates the average holding period of margin traders (average investors) as the reciprocal of the *stock-level* share turnover of margin (total) trading activities from 2013 to 2016. The right column instead uses the *market-level* share turnover of the margin (total) trading activities. For margin trading, the share turnover is calculated as the daily repay amount scaled by outstanding margin debt; for total trading activities, the share turnover is calculated as the daily share volume over tradable shares outstanding.

Figure 1.2. Average Holding Period: Margin Trader vs. Average Investor



This figure shows the frequency density of firm-round observations in each bin of the forcing variable around the eligibility cutoff. Vertical bars represent the histogram estimate of the forcing variable. Following Cattaneo et al. (2020), I plot local polynomial density estimates (solid blue and red) and robust bias-corrected confidence intervals (shaded blue and red) on both sides of the experiment cutoff.

Figure 1.3. Distribution of Stocks around the Eligibility Cutoff

Figure 1.4. Discontinuity in Margin Trading Eligibility at the Cutoff

This figure plots the margin trading eligibility against the forcing variable. I include all firms within the CCT bandwidth, and each point denotes the average probability of being marginable for firms within each evenly-spaced bin. Separate regression lines, along with 90% confidence intervals based on heteroscedasticity-consistent standard errors, are shown on both sides of the eligibility cutoff.

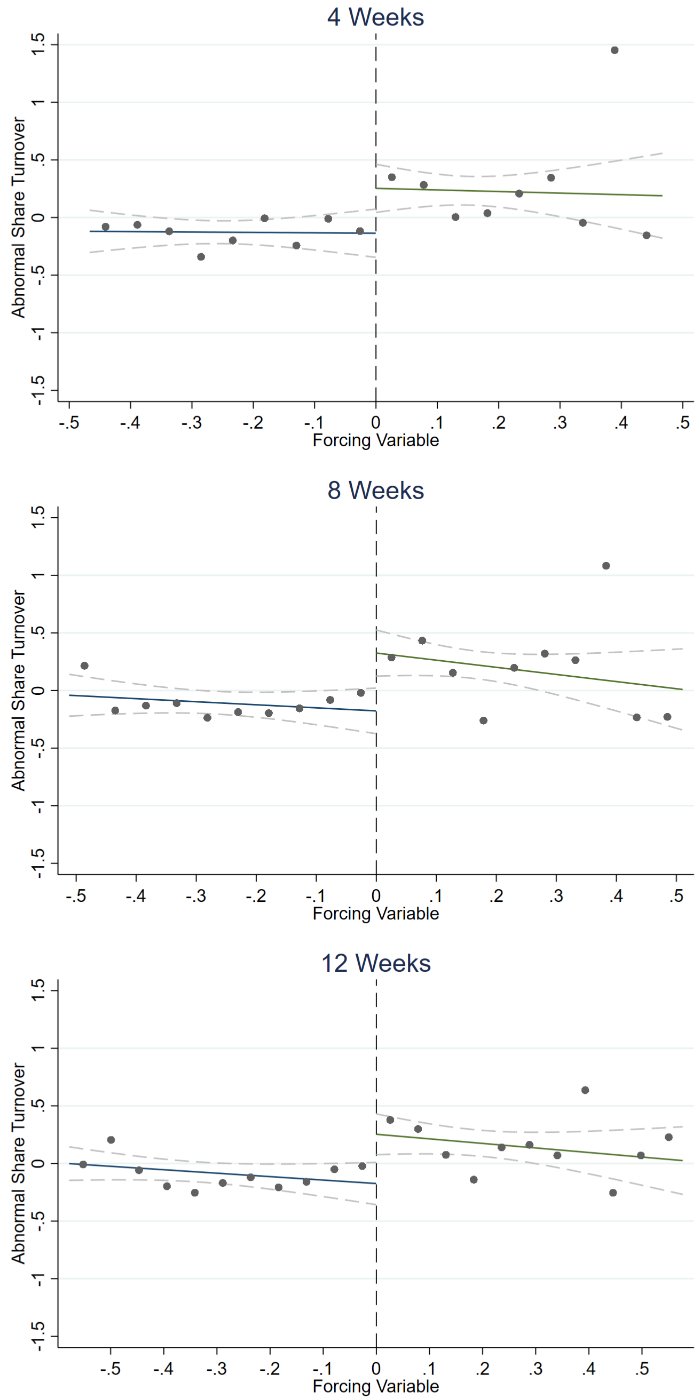


Figure 1.5. Discontinuity in Abnormal Share Turnover at the Cutoff

This figure plots the abnormal share turnover against the forcing variable. I include all firms within the CCT bandwidth and measure the abnormal share turnover as the change of daily share turnover before and after the implementation date using event windows of 4 weeks, 8 weeks, and 12 weeks. To control for time-series variation in abnormal share turnover, I demean each firm-round observation using the average values of all stocks in that round. Each point denotes the average abnormal share turnover within each bin of width 0.05. Separate regression lines, along with 90% confidence intervals based on heteroscedasticity-consistent standard errors, are shown on both sides of the eligibility cutoff.

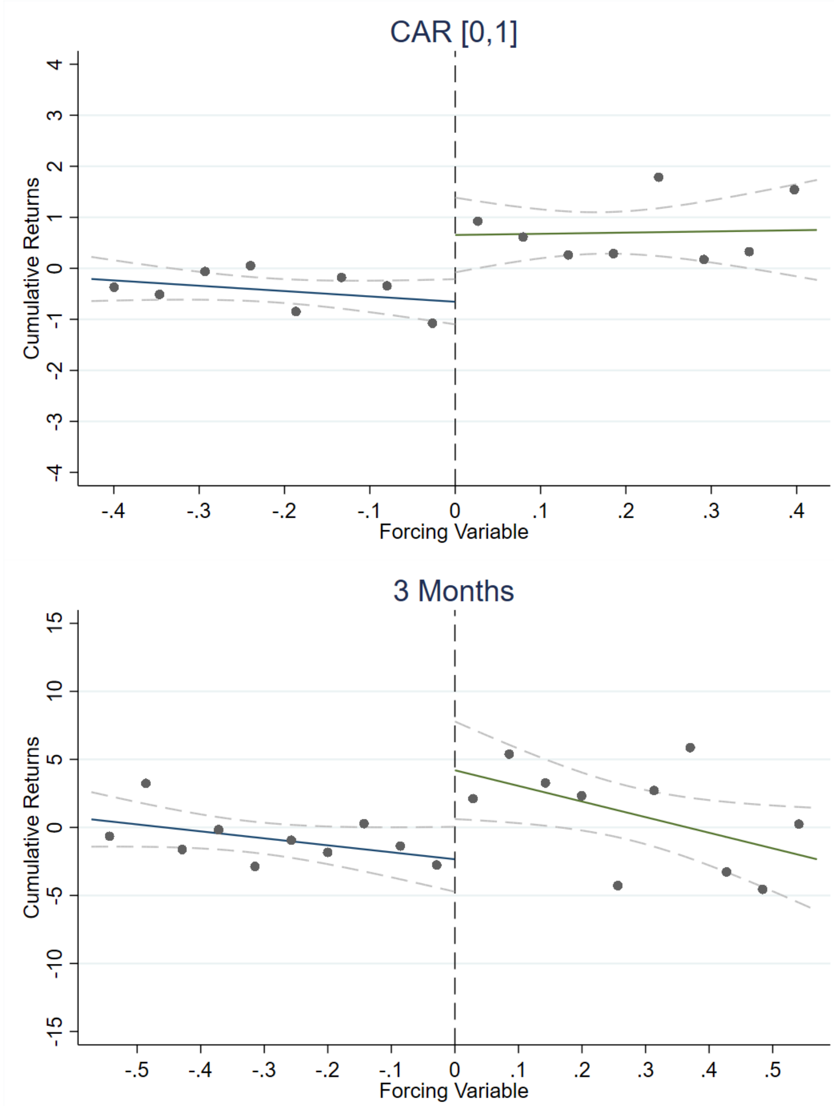


Figure 1.6. Discontinuity in Returns at the Cutoff

This figure plots the cumulative returns against the forcing variable. I include all firms within the CCT bandwidth and measure the outcome variable as the cumulative returns right after the implementation day (event window [0,1]) and cumulative returns in a longer post-event window of three months. To control for time-series variation in returns, I demean each firm-round observation using the average values of all stocks in that round. Each point denotes the average cumulative returns within each bin of width 0.05. Separate regression lines, along with 90% confidence intervals based on heteroscedasticity-consistent standard errors, are shown on both sides of the eligibility cutoff.

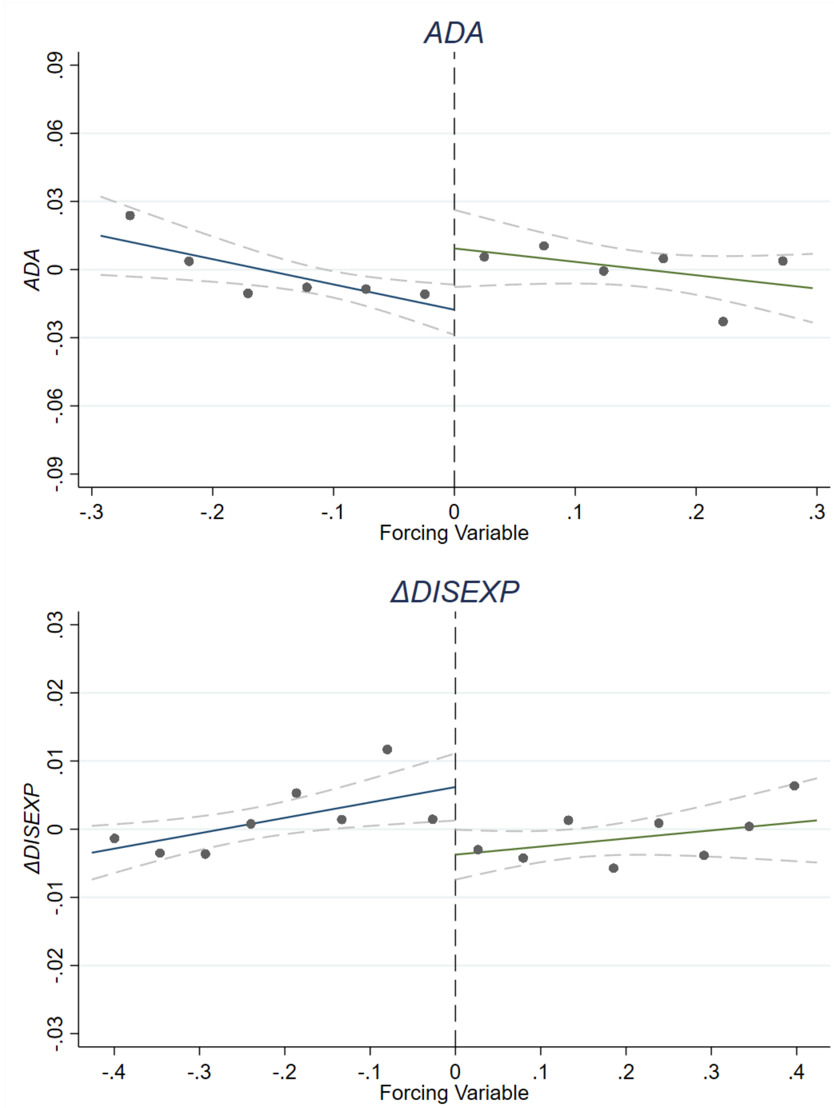


Figure 1.7. Discontinuity in Earnings Management at the Cutoff

This figure plots *ADA* and Δ*DISEXP* against the forcing variable for all firms within the CCT bandwidth. To control for time-series variation in accrual-based and real earnings management, I demean each firm-round observation using the average values of all stocks in that round. Each point denotes the average of the outcome variable within each bin of width 0.05. Separate regression lines, along with 90% confidence intervals based on heteroscedasticity-consistent standard errors, are shown on both sides of the eligibility cutoff.

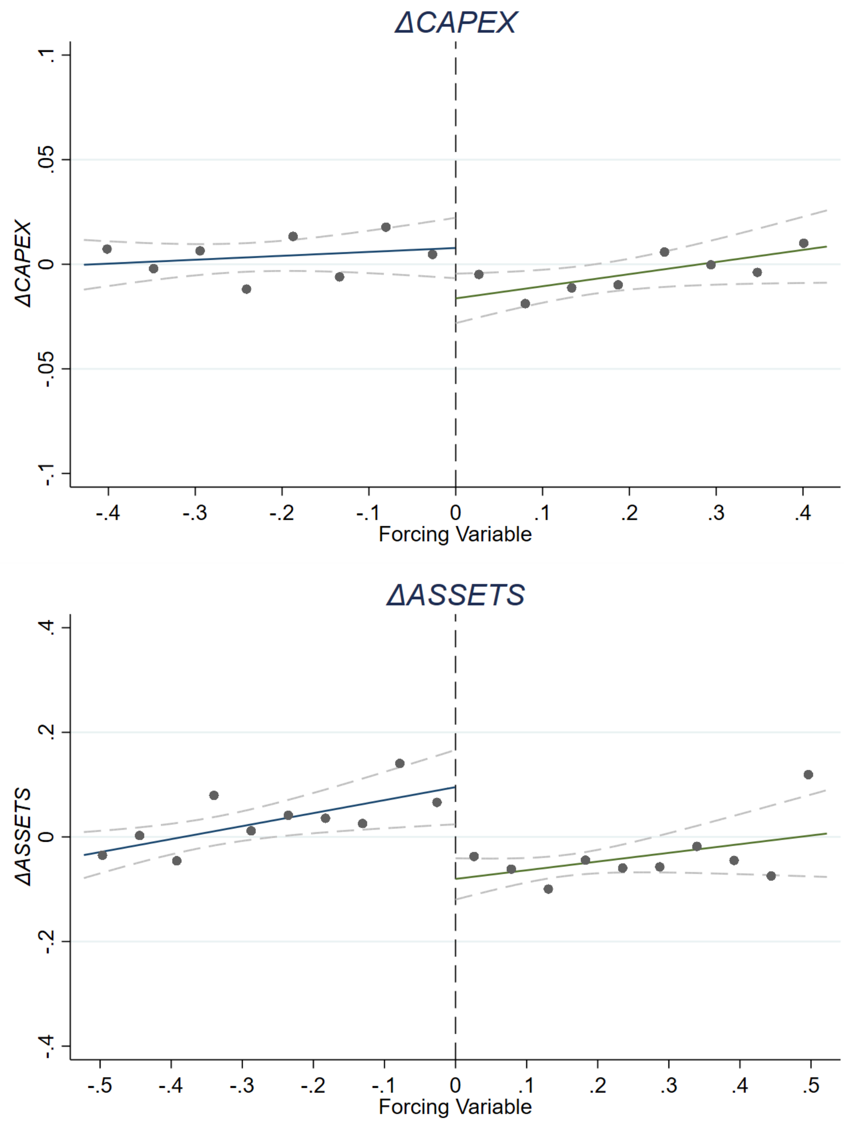


Figure 1.8. Discontinuity in Long-Term Investment at the Cutoff

This figure plots $\Delta CAPEX$ and $\Delta ASSETS$ against the forcing variable for all firms within the CCT bandwidth. To control for time-series variation in corporate investment, I demean each firm-round observation using the average values of all stocks in that round. Each point denotes the average of the outcome variable within each bin of width 0.05. Separate regression lines, along with 90% confidence intervals based on heteroscedasticity-consistent standard errors, are shown on both sides of the eligibility cutoff.

Table 1.1. Implementation Timeline

			# of Treated Stocks by Round		
Round #	Eligibility Decision	Event Day	Shanghai	Shenzhen	Both
Pilot A	Market index	March 31, 2010	50	40	90
Pilot B	Market index	December 5, 2011	130	59	189
1	Ranking formula	January 31, 2013	163	113	276
2	Ranking formula	September 16, 2013	104	102	206
3	Ranking formula	September 22, 2014	104	114	218

Table 1.2. Summary Statistics

	Bandwidth = 0.4			Bandwidth = 0.5			Bandwidth = 0.6		
Variable	Mean	Median	S.D.	Mean	Median	S.D.	Mean	Median	S.D.
Total Assets	8.42	4.37	12.10	8.08	4.22	11.48	7.75	4.09	10.94
Market Cap.	7.39	6.12	6.20	7.22	5.92	5.94	6.99	5.85	5.55
B/M	0.45	0.37	0.32	0.44	0.37	0.31	0.45	0.38	0.31
ROA	5.61	4.69	7.37	5.57	4.69	7.42	5.52	4.46	7.27
Leverage	0.48	0.49	0.21	0.47	0.48	0.21	0.47	0.48	0.21
Cash Flow	5.74	5.27	8.86	5.63	5.14	8.92	5.36	5.10	8.91
Cash Flow Vol.	11.18	7.57	14.22	11.30	7.56	14.92	11.50	7.61	14.89
ΔNI	3.33	1.61	6.03	3.45	1.68	6.31	3.57	1.72	6.48
Analyst No.	7.32	3.00	12.02	6.92	3.00	11.62	6.67	3.00	11.21
Dispersion	11.15	5.82	15.14	10.94	4.96	15.13	10.88	4.36	15.42
Beta	1.12	1.11	0.32	1.13	1.11	0.32	1.13	1.12	0.31
Share Turnover	2.50	1.79	2.18	2.51	1.83	2.20	2.52	1.84	2.21
ADA	6.83	4.52	7.53	6.93	4.43	7.90	6.93	4.48	7.81
$\Delta DISEXP$	1.66	0.83	2.70	1.64	0.82	2.73	1.58	0.81	2.69
$\Delta CAPEX$	1.18	0.14	6.87	1.22	0.17	7.00	1.16	0.15	6.98
$\Delta ASSETS$	19.26	12.37	35.39	18.88	12.09	34.61	19.19	12.00	35.37

This table reports descriptive statistics for observations local to the eligibility cutoff within the bandwidth of 0.4, 0.5, and 0.6. All firm characteristics are measured in the fiscal year closest to the implementation date. Variable definitions are provided in the Appendix. All variables are winsorized at the 1% and 99% levels. *Total Assets* and *Market Cap.* are in 100 millions RMB. *ROA*, *Cash Flow*, *Cash Flow Vol.*, ΔNI, *Dispersion*, *Share Turnover*, *ADA*, $\Delta DISEXP$, $\Delta CAPEX$, and $\Delta ASSETS$ are in percentage points.

Table 1.3. First Stage: Discontinuity in Margin Trading Eligibility at the Cutoff

	Marginable	
Above the Cutoff	0.539***	0.524***
	(0.064)	(0.070)
Firms Below	332	409
Firms Above	234	253
CCT Bandwidth	0.342	0.401
Kernel	Uniform	Triangular

This table reports the first-stage regression results estimated by the following equation:

$$Marginable_i^k = \beta_1 \cdot \mathbf{1}[V_i^k > 0] + R(V_i^k) + L(V_i^k) + \eta^k + \varepsilon_i^k$$

$Marginable_i^k$ is a dummy variable that equals one if stock i is approved for margin trading in round k. $\mathbf{1}[V_i^k > 0]$ is a dummy variable that equals one if stock i's ranking index is above the experiment cutoff in round k. $R(V_i^k)$ and $L(V_i^k)$ are linear controls of the forcing variable on the right-hand and left-hand side of the cutoff. η^k represents the round fixed effect. The bandwidth is automatically chosen by a data-driven algorithm as described in Calonico et al. (2014). All standard errors are calculated using the heteroskedasticity-robust nearest-neighbor estimator (Calonico et al., 2019). I show coefficient estimates of β_1 and report standard errors in parentheses. All estimates are similar among different kernel choices, and I use the uniform (rectangular) kernel for the main analysis hereafter. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 1.4. Abnormal Share Turnover

Panel A: Reduced Form

	4 Weeks	8 Weeks	12 Weeks
Above the cutoff	0.453***	0.569***	0.455***
	(0.164)	(0.169)	(0.154)

Panel B: Fuzzy RD

	4 Weeks	8 Weeks	12 Weeks
Marginable	0.796***	0.972***	0.748***
	(0.300)	(0.303)	(0.260)
Firms Below	468	528	636
Firms Above	264	277	305
CCT Bandwidth	0.468	0.512	0.578
Controls	Yes	Yes	Yes

Panel A reports the reduced form regression in which I run the outcome variable on the indicator of being above the cutoff. Panel B reports the fuzzy RD regression results using 2SLS estimation.

$$Y_i^k = \beta_2 \cdot Marginable_i^k + R(V_i^k) + L(V_i^k) + X_i^k + \eta^k + \varepsilon_i^k$$

The outcome variable Y_i^k is the abnormal share turnover (in percentage) for firm i in round k, measured as the change of daily share turnover before and after the implementation date using event windows of 4 weeks, 8 weeks, and 12 weeks. The main explanatory variable of interest $Marginable_i^k$ is an indicator variable that equals one if firm i is approved for margin trading in round k. An indicator of whether stock i's ranking index is above the cutoff, $\mathbf{1}[V_i^k > 0]$, is used as an instrument for $Marginable_i^k$. Control variables X_i^k are as follows: previous month's absolute stock return and log stock price; log(1+*Analyst No.*), where *Analyst No.* is the previous month's number of analysts who follow a firm and report forecasts; previous month's forecast dispersion, defined as the standard deviations of earnings per share forecasts reported by analysts; earnings surprise defined as the absolute value of the current ROA minus last year's ROA; other annual frequency controls including leverage, book-to-market ratio, market capitalization, and beta. $R(V_i^k)$ and $L(V_i^k)$ are separate linear controls of the forcing variable V_i^k on the right-hand and left-hand side of the cutoff. η^k represents the round fixed effect. The bandwidth is automatically chosen by the CCT data-driven algorithm as described in Calonico et al. (2014). All standard errors are calculated using the heteroskedasticity-robust nearest-neighbor estimator (Calonico et al., 2019). I show coefficient estimates of β_2 and report standard errors in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 1.5. Returns

Panel A: Reduced Form

	CAR [0,1]		3 Months	
Adjustment	Raw	Market	Raw	DGTW
Above the cutoff	1.308**	1.345**	6.540**	5.365**
	(0.539)	(0.559)	(2.710)	(2.397)

Panel B: Fuzzy RD

	CAR [0,1]		3 Months	
Adjustment	Raw	Market	Raw	DGTW
Marginable	2.335**	2.427**	10.597**	9.023**
	(0.977)	(1.026)	(4.443)	(4.088)
Firms Below	430	387	684	618
Firms Above	254	248	328	311
CCT Bandwidth	0.427	0.403	0.573	0.535

Panel A reports the reduced form regression in which I run the outcome variable on the indicator of being above the cutoff. Panel B reports the fuzzy RD regression results using 2SLS estimation.

$$Y_i^k = \beta_2 \cdot Marginable_i^k + R(V_i^k) + L(V_i^k) + \eta^k + \varepsilon_i^k$$

The outcome variable Y_i^k is the stock returns (in percentage) for firm i in round k. I measure raw/market-adjusted cumulative abnormal returns (CAR) right after the implementation day (event window [0,1]) and raw/DGTW-adjusted cumulative returns in a longer post-event window of three months. The main explanatory variable of interest $Marginable_i^k$ is an indicator variable that equals one if firm i is approved for margin trading in round k. An indicator of whether stock i's ranking index is above the cutoff, $\mathbf{1}[V_i^k > 0]$, is used as an instrument for $Marginable_i^k$. $R(V_i^k)$ and $L(V_i^k)$ are separate linear controls of the forcing variable V_i^k on the right-hand and left-hand side of the cutoff. η^k represents the round fixed effect. The bandwidth is automatically chosen by the CCT data-driven algorithm as described in Calonico et al. (2014). All standard errors are calculated using the heteroskedasticity-robust nearest-neighbor estimator (Calonico et al., 2019). I show coefficient estimates of β_2 and report standard errors in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 1.6. Earnings Management

Panel A: Reduced Form

	ADA		*ΔDISEXP*	
	Post	Pre	Post	Pre
Above the Cutoff	0.028***	−0.006	−0.009***	−0.005
	(0.011)	(0.010)	(0.003)	(0.004)

Panel B: Fuzzy RD

	ADA		*ΔDISEXP*	
	Post	Pre	Post	Pre
Marginable	0.054***	−0.011	−0.017***	−0.009
	(0.021)	(0.019)	(0.006)	(0.007)
Firms Below	253	335	447	417
Firms Above	179	232	255	247
CCT Bandwidth	0.297	0.362	0.427	0.417
Controls	Yes	Yes	Yes	Yes

Panel A reports the reduced form regression in which I run the outcome variable on the indicator of being above the cutoff. Panel B reports the fuzzy RD regression results using 2SLS estimation.

$$Y_i^k = \beta_2 \cdot Marginable_i^k + R(V_i^k) + L(V_i^k) + X_i^k + \eta^k + \varepsilon_i^k$$

The outcome variables Y_i^k are respectively *ADA* and Δ*DISEXP* for firm i in round k. The main explanatory variable of interest $Marginable_i^k$ is an indicator variable that equals one if firm i is approved for margin trading in round k. An indicator of whether stock i's ranking index is above the cutoff, $\mathbf{1}[V_i^k > 0]$, is used as an instrument for $Marginable_i^k$. Control variables X_i^k include market capitalization, market-to-book ratio, leverage, operating return on assets, cash flow volatility, asset growth rate, operating cash flow, and lagged discretionary accruals. I also report pre-event results (one year before the deregulation) as a placebo test. $R(V_i^k)$ and $L(V_i^k)$ are separate linear controls of the forcing variable V_i^k on the right-hand and left-hand side of the cutoff. η^k represents the round fixed effect. The bandwidth is automatically chosen by the CCT data-driven algorithm as described in Calonico et al. (2014). All standard errors are calculated using the heteroskedasticity-robust nearest-neighbor estimator (Calonico et al., 2019). I show coefficient estimates of β_2 and report standard errors in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 1.7. Long-Term Investment

Panel A: Reduced Form

	$\Delta CAPEX$		$\Delta ASSETS$	
	Post	Pre	Post	Pre
Above the Cutoff	-0.025^{**}	0.012	-0.171^{***}	0.022
	(0.011)	(0.009)	(0.049)	(0.053)

Panel B: Fuzzy RD

	$\Delta CAPEX$		$\Delta ASSETS$	
	Post	Pre	Post	Pre
Marginable	-0.046^{**}	0.019	-0.286^{***}	0.042
	(0.021)	(0.015)	(0.085)	(0.100)
Firms Below	450	567	605	330
Firms Above	258	287	305	229
CCT Bandwidth	0.428	0.521	0.523	0.343
Controls	Yes	Yes	Yes	Yes

Panel A reports the reduced form regression in which I run the outcome variable on the indicator of being above the cutoff. Panel B reports the fuzzy RD regression results using 2SLS estimation.

$$Y_i^k = \beta_2 \cdot Marginable_i^k + R(V_i^k) + L(V_i^k) + X_i^k + \eta^k + \varepsilon_i^k$$

The outcome variables Y_i^k are respectively $\Delta CAPEX$ and $\Delta ASSETS$ for firm i in round k. The main explanatory variable of interest $Marginable_i^k$ is an indicator variable that equals one if firm i is approved for margin trading in round k. An indicator of whether stock i's ranking index is above the cutoff, $\mathbf{1}[V_i^k > 0]$, is used as an instrument for $Marginable_i^k$. Control variables X_i^k include past profitability (operating ROA), operating cash flow, and lagged total assets. I also report pre-event results (one year before the deregulation) as a placebo test. $R(V_i^k)$ and $L(V_i^k)$ are separate linear controls of the forcing variable V_i^k on the right-hand and left-hand side of the cutoff. η^k represents the round fixed effect. The bandwidth is automatically chosen by the CCT data-driven algorithm as described in Calonico et al. (2014). All standard errors are calculated using the heteroskedasticity-robust nearest-neighbor estimator (Calonico et al., 2019). I show coefficient estimates of β_2 and report standard errors in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 1.8. Alternative Bandwidths

Panel A: Dependent Variable = Abnormal Share Turnover				
Bandwidth	+25%	+10%	−10%	−25%
Marginable	0.772***	0.740***	0.800***	0.665**
	(0.213)	(0.237)	(0.290)	(0.328)
Observations	1277	1074	830	671
Panel B: Dependent Variable = Returns				
Bandwidth	+25%	+10%	−10%	−25%
Marginable	1.807**	2.220**	2.934***	3.197***
	(0.837)	(0.946)	(1.053)	(1.135)
Observations	833	724	567	481
Panel C: Dependent Variable = *ADA*				
Bandwidth	+25%	+10%	−10%	−25%
Marginable	0.047**	0.052***	0.051**	0.044*
	(0.018)	(0.019)	(0.022)	(0.025)
Observations	532	478	388	325
Panel D: Dependent Variable = Δ*DISEXP*				
Bandwidth	+25%	+10%	−10%	−25%
Marginable	−0.013***	−0.014**	−0.018***	−0.016**
	(0.005)	(0.006)	(0.007)	(0.007)
Observations	928	785	626	529
Panel E: Dependent Variable = Δ*CAPEX*				
Bandwidth	+25%	+10%	−10%	−25%
Marginable	−0.037**	−0.036*	−0.046**	−0.049**
	(0.017)	(0.018)	(0.023)	(0.025)
Observations	937	793	629	533
Panel F: Dependent Variable = Δ*ASSETS*				
Bandwidth	+25%	+10%	−10%	−25%
Marginable	−0.266***	−0.282***	−0.256***	−0.296***
	(0.064)	(0.076)	(0.089)	(0.109)
Observations	1201	1012	793	648

This table reports the fuzzy RD results with alternative bandwidths. Specifically, I increase/decrease the CCT optimal bandwidth by 10% and 25% and redo the RD analysis for all outcome variables, including the 12-week abnormal share turnover, event-day returns, *ADA*, Δ*DISEXP*, Δ*CAPEX*, and Δ*ASSETS*. All specifications and control variables are the same as the main analysis. I show coefficient estimates of β_2 and report standard errors in parentheses, along with the firm-round observations within each alternative bandwidth. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 1.9. Placebo Cutoffs

Panel A: Dependent Variable = Abnormal Share Turnover				
Cutoff	+0.5	−0.5	+CCT	−CCT
Marginable	−0.092	−0.094	−0.172	0.244
	(0.223)	(0.168)	(0.235)	(0.151)
Observations	636	859	441	980
Panel B: Dependent Variable = Returns				
Cutoff	+0.5	−0.5	+CCT	−CCT
Marginable	−0.740	0.114	−0.142	0.276
	(0.753)	(0.391)	(0.663)	(0.370)
Observations	543	1105	744	1345
Panel C: Dependent Variable = *ADA*				
Cutoff	+0.5	−0.5	+CCT	−CCT
Marginable	−0.002	−0.007	−0.003	0.011
	(0.010)	(0.007)	(0.008)	(0.008)
Observations	543	1103	966	1061
Panel D: Dependent Variable = Δ*DISEXP*				
Cutoff	+0.5	−0.5	+CCT	−CCT
Marginable	0.004	−0.004	0.002	−0.001
	(0.004)	(0.003)	(0.004)	(0.003)
Observations	625	1054	651	1403
Panel E: Dependent Variable = Δ*CAPEX*				
Cutoff	+0.5	−0.5	+CCT	−CCT
Marginable	−0.006	0.005	−0.002	0.010
	(0.013)	(0.008)	(0.012)	(0.008)
Observations	437	1295	615	1713
Panel F: Dependent Variable = Δ*ASSETS*				
Cutoff	+0.5	−0.5	+CCT	−CCT
Marginable	0.034	0.014	0.046	0.035
	(0.043)	(0.040)	(0.063)	(0.038)
Observations	1001	986	615	1129

This table reports results of placebo tests, in which I replicate the RD analysis around placebo cutoffs set at 0.5, −0.5, and one CCT bandwidth above and below the actual cutoff 0. Outcome variables include the 12-week abnormal share turnover, event-day returns, *ADA*, Δ*DISEXP*, Δ*CAPEX*, and Δ*ASSETS*. All specifications and control variables are the same as the main analysis. I show coefficient estimates of β_1 from the reduced form regression and report standard errors in parentheses, along with firm-round observations within the CCT bandwidth auto-selected at each placebo cutoff. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 1.10. Validity Tests

Variables	*Total Assets*	*Market Cap.*	*B/M*	*ROA*	*Leverage*	*Asset Growth*
Marginable	2.549	0.920	0.132	0.713	−0.021	0.065
	(3.508)	(1.657)	(0.098)	(2.045)	(0.058)	(0.103)
Firms	566	629	565	539	568	559

Variables	*Cash Flow*	*Cash Flow Vol.*	*ΔNI*	*Analyst No.*	*Dispersion*	*Beta*
Marginable	−0.396	0.556	−1.654	−4.593	1.782	0.068
	(2.862)	(3.646)	(1.739)	(4.079)	(4.689)	(0.080)
Firms	554	728	536	794	473	600

This table reports the fuzzy RD regression results using 2SLS estimation.

$$Y_i^k = \beta_2 \cdot Marginable_i^k + R(V_i^k) + L(V_i^k) + \eta^k + \varepsilon_i^k$$

The outcome variables Y_i^k include all control variables used in the main analysis for firm i in round k, measured one year before the margin trading deregulation. $Marginable_i^k$ is an indicator variable that equals one if firm i is approved for margin trading in round k. An indicator of whether stock i's ranking index is above the cutoff, $\mathbf{1}[V_i^k > 0]$, is used as an instrument for $Marginable_i^k$. $R(V_i^k)$ and $L(V_i^k)$ are separate linear controls of the forcing variable V_i^k on the right-hand and left-hand side of the cutoff. η^k represents the round fixed effect. The bandwidth is automatically chosen by the CCT data-driven algorithm as described in Calonico et al. (2014). All standard errors are calculated using the heteroskedasticity-robust nearest-neighbor estimator (Calonico et al., 2019). I show coefficient estimates of β_2 and report standard errors in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 1.11. Alternative Measures

Panel A: Reduced Form

	ΔADA	$\Delta DISEXPRD$	$\Delta CAPEXRD$
Above the Cutoff	0.027^{***}	-0.011^{***}	-0.027^{**}
	(0.010)	(0.004)	(0.011)

Panel B: Fuzzy RD

	ΔADA	$\Delta DISEXPRD$	$\Delta CAPEXRD$
Marginable	0.053^{**}	-0.020^{***}	-0.049^{**}
	(0.021)	(0.007)	(0.021)
Firms Below	244	447	450
Firms Above	173	255	258
CCT Bandwidth	0.297	0.427	0.428
Controls	Yes	Yes	Yes

Panel A reports the reduced form regression in which I run the outcome variable on the indicator of being above the cutoff. Panel B reports the fuzzy RD regression results using 2SLS estimation.

$$Y_i^k = \beta_2 \cdot Marginable_i^k + R(V_i^k) + L(V_i^k) + X_i^k + \eta^k + \varepsilon_i^k$$

The outcome variables Y_i^k are defined as follows: ΔADA is the annual change of ADA, representing the abnormal portion of accrual-based earnings management; $\Delta DISEXPRD$ is a slightly modified version of $\Delta DISEXP$, defined as the annual change of SG&A and R&D expenses scaled by beginning-of-year total assets; $\Delta CAPEXRD$ is a slightly modified version of $\Delta CAPEX$, defined as the annual change of capital expenditures and R&D expenses scaled by beginning-of-year total assets. All specifications and control variables are the same as before. The same bandwidth as that of the main analysis is used for each outcome variable. All standard errors are calculated using the heteroskedasticity-robust nearest-neighbor estimator (Calonico et al., 2019). I show coefficient estimates of β_2 and report standard errors in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 1.12. Long-Term Performance

	ROA			*Returns*		
Bandwidth	0.3	0.4	0.5	0.3	0.4	0.5
One year later, $t+1$	−0.017**	−0.018**	−0.016**	−0.176**	−0.163**	−0.167***
	(0.008)	(0.008)	(0.008)	(0.068)	(0.067)	(0.061)
Two years later, $t+2$	−0.019**	−0.021**	−0.021**	−0.077	−0.089	−0.080
	(0.010)	(0.009)	(0.009)	(0.060)	(0.059)	(0.054)
Three years later, $t+3$	−0.018*	−0.017	−0.014	−0.058	−0.066*	−0.063*
	(0.011)	(0.011)	(0.010)	(0.040)	(0.038)	(0.035)
Four years later, $t+4$	−0.018	−0.018*	−0.011	−0.041	−0.045	−0.052
	(0.011)	(0.011)	(0.010)	(0.043)	(0.041)	(0.039)
R^2	0.521	0.513	0.505	0.483	0.483	0.480
Observations	2912	3870	5060	2894	3848	5034
No. of firm-rounds	493	655	856	493	655	856

This table presents the effect of margin trading eligibility on firm long-term performance using a dynamic regression discontinuity model.

$$Y_{i,t+\tau} = \beta^\tau \cdot Marginable_{it} + R^\tau(V_{it}) + L^\tau(V_{it}) + \alpha_\tau + \eta_c + \lambda_{it} + \varepsilon_{it\tau}$$

The outcome variables are operating ROA (return on assets) and stock returns. $Marginable_{it}$ is an indicator of margin trading eligibility, and β^τ estimates the effect of margin trading eligibility at time t on firm outcomes at $t+\tau$. β^τ and separate linear controls of the forcing variable $R^\tau(V_{it})$ and $L^\tau(V_{it})$ are allowed to vary for $\tau > 0$, and constrained to zero for $\tau < 0$, and standard errors are clustered at the firm level. The parameters α_τ, η_c, and λ_{it} are fixed effects for time periods relative to the event, calendar years, and firm-rounds. The bandwidth is respectively set at 0.3, 0.4, and 0.5. I show coefficient estimates of β^τ ($\tau = 1,2,3,4$) and report standard errors in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 1.13. The Effect of Margin Trading Eligibility on *ADA* Controlling for Investment Growth

	Absolute Discretionary Accruals (*ADA*)			
	(1)	(2)	(3)	(4)
$Marginable_i^k$	0.051**	0.050**	0.052**	0.051**
	(0.021)	(0.021)	(0.021)	(0.021)
V_i^k	−0.125**	−0.125**	−0.126**	−0.125**
	(0.054)	(0.053)	(0.054)	(0.053)
$R(V_i^k)$	0.038	0.038	0.035	0.034
	(0.070)	(0.069)	(0.070)	(0.069)
Market Cap.	−0.001**	−0.001**	−0.001**	−0.001**
	(0.001)	(0.001)	(0.001)	(0.001)
Leverage	0.023	0.029*	0.023	0.030*
	(0.017)	(0.018)	(0.018)	(0.018)
MB	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Cash Flow	−0.056	−0.060	−0.056	−0.062
	(0.057)	(0.055)	(0.057)	(0.055)
Cash Flow Vol.	0.053	0.052	0.054	0.051
	(0.051)	(0.050)	(0.051)	(0.050)
ROA	0.019	0.072	0.018	0.073
	(0.095)	(0.091)	(0.095)	(0.091)
Lagged ADA	0.204***	0.193***	0.207***	0.196***
	(0.064)	(0.064)	(0.064)	(0.064)
$\Delta ASSETS$	0.040***	0.003	0.042***	0.001
	(0.016)	(0.021)	(0.016)	(0.022)
$\Delta CAPEX$	0.048	0.014		
	(0.054)	(0.062)		
$\Delta ASSETS^2$		0.000		0.000
		(0.000)		(0.000)
$\Delta CAPEX^2$		0.002		
		(0.003)		
$\Delta CAPEXRD$			0.030	0.010
			(0.047)	(0.060)
$\Delta CAPEXRD^2$				0.002
				(0.003)
Constant	0.005	0.006	0.005	0.005
	(0.015)	(0.015)	(0.015)	(0.015)
R^2	0.264	0.277	0.263	0.276
Observations	432	432	432	432

This table presents the effect of margin trading eligibility on *ADA* after controlling for firms' investment and growth. All regression specifications are the same as before. Besides the control variables used in the main analysis, I further include $\Delta CAPEX$ in column (1) and the squared terms $\Delta ASSETS^2$ and $\Delta CAPEX^2$ in column (2). In columns (3) and (4), I use a slightly modified version of $\Delta CAPEX$ by including both capital expenditures and R&D expenditures and denote it $\Delta CAPEXRD$. Heteroscedasticity-consistent standard errors are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Chapter 2

Disclosure and Cost of Equity Capital Revisited

2.1 Introduction

A prominent issue in accounting research is the relation between disclosures of firm-specific information such as that contained in financial reports and the cost of equity capital (i.e., expected return). Neoclassical theory on asset pricing argues that expected return is entirely composed of a risk-free return and a systematic risk premium. Idiosyncratic risks are presumed to be eliminated through diversification, implying that information on such risks has no effect on expected return. However, in apparent contradiction to theory, empirical studies such as Botosan (1997) offer evidence of a negative association between information contained in firm disclosures and implied cost of equity capital after controlling for effects of systematic risk. Limitations of her study are the small sample due to reliance on hand-collected data in creating indices as measures of firm disclosures and the questionable efficacy of implied cost of equity capital as a proxy for expected return.

In this study, we exploit recently developed textual analysis in a replication of Botosan (1997)'s study applied to a greatly expanded sample in both cross-section and time. Our sample consists of 43,806 firm-year observations spanning years from 1995 to 2019, for which we have sufficient data to construct the disclosure indices that we employ in our analysis. Applying filters for implied cost of equity capital estimation and controls for systematic risk leaves 28,284 firm-year observations for data acquired from Value Line and 37,341 firm-year observations acquired from I/B/E/S. With Value Line, we compute implied cost of equity capital employing Brav et al. (2005)'s model parameterized by price targets, dividend forecasts, and dividend growth rates, while with I/B/E/S data, we take the average of implied cost of equity capital employing models by Claus and Thomas (2001), Gebhardt et al. (2001), Ohlson and Juettner-Nauroth (2005), and Easton (2004) as suggested by Hail and Leuz (2006).

For most of our analysis, we use a modestly refined version of Botosan's original construction of disclosure indices introduced by Francis et al. (2008). The advantage of their version is enabling a validity check on the construction of indices by comparing descriptive statistics of index values. Francis et al. (2008)'s sample consists of 677 firms in fiscal 2001. Our textual analysis yielded 627 observations for 2001, for which the means, standard deviations, and quartile distributions closely correspond to those statistics in their study.[1] We extended the comparison of descriptive statistics to the full sample for which disclosure indices are feasibly constructed and find remarkably similar results. We emphasize analyses using Value Line data consistent with Botosan's original study. However, similar results are obtained using more extensive I/B/E/S data.

Similar to Botosan (1997), we regress estimates of implied cost of equity capital on proxies for systematic risk factors and disclosure indices for each of the 25 years composing our sample. Briefly summarizing our principal results under Value Line, we find that the coefficient of the total disclosure index is negative in only four years and never significantly negative in explaining the implied cost of equity capital as hypothesized by Botosan (1997). The annual coefficients are positive in the remaining 21 years and significantly positive at at least the 10% level in 10 years. We simulate the data generating process a thousand times using estimated parameters producing a similar pattern of signs and significance for yearly coefficients of the total disclosure index. Following Fama and MacBeth (1973)'s procedure under the further assumption that coefficients are time-invariant, we find the average cross-sectional coefficient of the total disclosure index is significantly positive. A further panel regression with year and industry fixed effects for both industry and time also yields a similar result for the coefficient of the total disclosure index. Collectively, our results lead us to reject the hypothesis that firm-specific disclosures serve to reduce the cost of equity capital (expected return). Rather, we are left with strong evidence of a positive relation between such disclosures and the implied cost of equity capital.

A plausible explanation for a positive relation between implied cost of equity capital and

[1] The coefficients of the total disclosure index with controls for systematic risks in both our study for 2001 and that of Francis et al. (2008) are similarly insignificant, suggesting that the effects of disclosure in their study on implied cost of equity capital, absent controls for systematic risk, may not be attributable to firm disclosures that contain idiosyncratic risks.

measures of disclosure is offered by Hughes et al. (2009), who model the difference between the implied cost of equity capital and expected return under the assumption that expected returns are stochastic. Merton (1973)'s seminal study relates stochastic expected returns to random states of nature that induce changes in investors' opportunity set. In particular, Hughes et al. (2009) adopt Merton's illustration with stochastic expected returns in the form of stochastic betas in their analysis. The assumption of stochastic expected return is well supported by several empirical studies, including Campbell (1991), Jagannathan and Wang (1996), and Fama and French (1997). Hughes et al. (2009) identify factors including the volatility of cash flows that could account for a bias in implied cost of equity capital as a proxy for expected return. We extend the independent identically distributed expected return setting of Hughes et al. (2009) to a general mean-reverting expected return setting. We adopt a continuous-time model (in contrast to discrete-time) to parsimoniously depict a relation between cash flow volatility and bias in implied cost of equity capital.

Reflecting on the role of disclosure indices vis-à-vis the cost of equity capital, the information provided by financial data from EDGAR is, generally speaking, about past events accompanied by estimates of future events. Whether disclosures upon which disclosure indices are based are voluntary or mandatory is inextricably mixed, notwithstanding Botosan's and others' view of voluntary disclosure as a driver of cross-sectional variations by firms seeking to reduce their cost of equity capital.[2] We, alternatively, interpret disclosure indices drawn in this fashion as descriptive of the level of cash flow news, a significant portion of which pertains to idiosyncratic risks. In theory, under this interpretation, there should be no association of disclosure indices with the expected return after controlling for systematic risk. However, as evident from our model, this prediction does not apply to the implied cost of equity capital when the expected return is stochastic. In this case, the implied cost of equity capital is a biased proxy for expected return, a bias that could explain a positive association with disclosure.

[2]A history of voluntary disclosure may be viewed as an implicit commitment, further blurring the distinction with mandatory disclosure.

Empirically, consistent with this explanation, we find that idiosyncratic volatility estimated from residual returns, as a measure of idiosyncratic risks that manifest in cash flow news, is positively correlated with both disclosure indices and implied cost of equity capital.[3]

2.1.1 Literature

Botosan (1997) provided a foundation for subsequent studies seeking to relate firm-specific information to the cost of equity capital. Botosan's sample consists of 122 hand-collected observations for manufacturing firms drawn from annual reports in 1990. The objective of her study is to empirically link indices thought to capture voluntary disclosure present in annual reports to the implied cost of equity capital as a proxy for expected returns after controlling for the effects of systematic risks. Results are consistent with a negative association. The study is seminal in the sense of lending impetus to various later studies examining pricing effects of information and its properties vis-à-vis the cost of equity capital. Botosan and Plumlee (2002) reaffirm the negative association between disclosure of information contained in annual reports and the implied cost of equity capital. A further study by Botosan et al. (2004) considers the relations between precisions of public and private information and the implied cost of equity capital. They find a negative association for the former that is more than offset by a positive association for the latter. Francis et al. (2008) refine the empirical specification in an attempt to better distinguish disclosures that are voluntary and consider the potential effects of earnings quality. They report a significant negative association between disclosure and implied cost of equity capital, but only before removing effects of systematic risks. This effect is diminished by conditioning on earnings quality, a feature that we do not explore.

Other empirical studies less directly linked to our platform include Francis et al. (2004, 2005), who explore the effects of accruals quality and other properties of earnings on the cost of equity capital. The latter uses excess returns from a three-factor model to assess whether

[3] We acknowledge that empirical evidence discussed by Lambert (2009) as well as Hughes et al. (2009) is mixed with regard to conditions under which implied cost of capital as a measure of expected return is increasing in the volatility of cash flows, thereby recommending caution in interpreting these results.

accruals quality is a priced risk factor, an implication challenged by Core et al. (2008). Kothari et al. (2009) employ content analysis of reports by management, analysts, and the business press to assess the differential effects of favorable versus unfavorable disclosures on rolling forward estimates of a three-factor model from which they infer the cost of equity capital. Other dimensions of firm disclosures effects on the cost of equity capital include Dhaliwal et al. (2011)'s study of the impact of initiating corporate disclosures of social responsibility, and Li (2010)'s study of consequences mandated adoption of International Financial Reporting Standards further refined by Daske et al. (2013) to capture firm discretion in implementation.

Given a well-functioning market in which diversification eliminates pricing effects of idiosyncratic risks, we would not expect to find an impact of information pertaining to such risks to affect the cost of equity capital. In theory, the absence of pricing effects carries over to private information, as shown by Hughes et al. (2007) and Lambert et al. (2007). Although we do not address the distinction between voluntary and mandatory disclosure in assessing the effects of firm disclosures on implied cost of equity capital[4], Cheynel (2013) models the effects of voluntary disclosures on a firm's sensitivity to systematic risks and shows that betas, as a measure of that sensitivity, are lower conditional on disclosure than on non-disclosure. However, the inclusion of an interaction variable of betas and disclosure indices in order to capture the effects of disclosures on factor loadings does not alter our results in a qualitative sense.

Last, as noted, Hughes et al. (2009) model the effects of stochastic expected returns on implied cost of capital. In such a setting, they show that the implied cost of capital provides a biased estimate of expected returns. Their analysis depicts factors driving that bias, including the volatility of cash flows. In turn, we provide a continuous-time analog to their model that identifies conditions under which the bias in implied cost of equity capital is increasing in cash flow volatility and, hence, in expected disclosures as a reflection of that volatility.

[4]We note that the distinction between what is voluntary versus mandatory is unclear in the items that compose disclosure indices in both Botosan (1997) and Francis et al. (2008). We view the disclosure indices as a measure of the level of cash flow news that the firm is expected to provide but not limited to what is dictated by accounting standards.

Neoclassical theory argues that information on idiosyncratic risks should have no effect on expected returns. Whether theory holds in practice is an empirical question. Our study contributes to the literature in several ways. First, through textual analysis of firm disclosures, we substantially expand samples employed in earlier studies by Botosan (1997) and others employing similar sampling rules and tests. Second, we provide extensive evidence calling into question findings supporting the view that firms lower their cost of equity capital through greater disclosure of firm-specific information. Third, our findings of a positive relation between disclosure indices and implied cost of equity capital are traceable to a bias in implied cost of equity capital as a proxy for expected return characterized by Hughes et al. (2009).[5]

The remainder of this paper is organized as follows: Section 2.2 describes our disclosure indices, estimates of implied cost of equity capital, and sampling rules. Section 2.3 contains our empirical analysis of the relation between disclosure indices and the implied cost of equity capital. Section 2.4 sets forth our model on bias in implied cost of equity capital and correlations of predicted relations. Section 2.5 concludes.

2.2 Disclosure Indices, Estimates of Implied Cost of Equity, and Sampling Rules

2.2.1 Disclosure Indices

Our construction of disclosure indices is based on Francis et al. (2008)'s modified version of Botosan (1997)'s approach.[6] The total disclosure measure consists of four categories, of which the contents are listed in Table 2.1. We began our coding by downloading all 10-K reports from the EDGAR system. For each 10-K report, according to the coding scheme, we assign the

[5]Empirically, allowing for interaction between disclosure indices and betas as a measure of systematic risk has no qualitative effects.

[6]Given that EDGAR did not exist in 1990, the year of Botosan (1997)'s data, we cannot establish the validity of our replication of her coding scheme via textual analysis. Accordingly, we resort to employing the coding scheme of Francis et al. (2008). The latter excludes background information and MDA arguing that these measures are less reflective of voluntary disclosure as a result of SEC restrictions. These categories were insignificant in Botosan's study, suggesting little impact of their exclusion. As noted earlier, this shift to Francis et al. (2008)'s coding enables a comparison of summary statistics as a means of verifying the validity of our recourse to textual analysis.

binary element (e.g., whether the firm discloses a forecasted cash flow or not) a value of one, if existing, or zero otherwise. For non-binary elements (e.g., number of quarters that firm discloses sales and net income), we convert it to binary variable depending on whether it is above or below the median value reported by all firms in the same year. Firms above the median receive a value of one and otherwise zero. The above procedures generate a value of zero or one for each of the elements in the coding scheme. We then scale this raw score by the maximum score in that year to obtain a percentage-based score for each firm year. Our results are robust to other ways to aggregate the disclosure score. For example, we have the same results if we assign equal weight to each category instead of each element. Neither are the results sensitive to directly using raw scores. Appendix B.2 provides a detailed description of our textual analysis procedure for constructing disclosure indices.

In Table 2.2, we report descriptive statistics on our raw total/subcategory disclosure indices for 2001 (Panel B), the year for which their data were drawn, in comparison with those of Francis et al. (2008) (Panel A). Notwithstanding a small difference in sample sizes, the two distributions are nearly identical, lending confidence that our replication of disclosure indices captures essentially the same information. We further compare the above descriptive statistics with those for our entire sample (Panel C) and find the similarity distinctive. These comparisons establish a high level of validity for our textual-based determination of disclosure scores.

In Table 2.3, we report inter-correlations of raw disclosure indices in total and by category. The correlations are uniformly positive as one would expect given the likelihood that greater disclosure in one category would be accompanied by greater disclosure in another. These results are consistent with Botosan's. Furthermore, for each firm over time, our disclosure measure has an AR(1) coefficient of 0.9 and a coefficient of variation (standard deviation scaled by mean) of 0.2. Therefore, our disclosure measure is persistent over time for each firm.

2.2.2 Implied Cost of Equity Capital

We use an estimation model based on future price targets to construct our main proxy for the implied cost of equity capital. Specifically, we follow Brav et al. (2005) and Francis et al. (2008) to form the implied cost of equity capital measure ICC_{VL}, which is derived from Value Line data on analysts' four-year out price targets (TP), dividend forecasts (DIV), and dividend growth rates (g). Assuming dividends are reinvested at the firm cost of equity capital ICC_{VL}, Brav et al. (2005) suggest the following equation for the expected return:

$$(1+ICC_{VL})^4 = \frac{TP}{P} + \frac{DIV[\frac{(1+ICC_{VL})^4-(1+g)^4}{ICC_{VL}-g}]}{P} \tag{2.1}$$

where P = stock price nine days prior to the date of the Value Line report. For each firm year in our sample, we determine the value of ICC_{VL} that satisfies the above equation and use this as our estimate of the implied cost of equity capital. Following Francis et al. (2008), we use the average of the firm's four quarterly estimates of ICC_{VL} to form the annual estimation of the implied cost of equity capital for that year. We draw similar inferences if we use the quarterly ICC_{VL} estimation right after the annual report publication date.

We advance the implied cost of equity capital estimated from Value Line data as our principal measure for two reasons. First, it is essentially identical to the implied cost of equity capital measure used in Botosan (1997) and Francis et al. (2008) lending comparability to those previous studies on the relation between disclosure and cost of equity capital. Second, aside from the question of potential bias as a proxy for expected return, the Value Line based implied cost of equity capital has a higher construct validity than other measures regarding associations with firm risk attributes (Botosan and Plumlee, 2005) and a significant correlation with future realized returns (Francis et al., 2004). Moreover, due to the use of the four-year out price targets, Value Line makes fewer assumptions of the long-term growth rate than other estimation models.

To reinforce our findings using Value Line, we also consider alternative measures implementable from I/B/E/S data. An advantage of I/B/E/S data is a significant enlargement of our sample. Following Hail and Leuz (2006), we use four different models in estimating implied cost of equity capital, including those by Claus and Thomas (2001), Gebhardt et al. (2001), Ohlson and Juettner-Nauroth (2005) (as implemented by Gode and Mohanram (2003)), and Easton (2004), as well as the average of these four estimates as our measure of implied cost of equity capital. Each of the four models calculates the implied cost of equity capital as the internal rate of return that equates current stock price with the discounted future dividends or earnings. We describe each model in more detail below.

Claus and Thomas (2001):

$$P_t = bv_t + \sum_{\tau=1}^{T} \frac{\hat{x}_{t+\tau} - ICC_{CT} \cdot bv_{t+\tau-1}}{(1+ICC_{CT})^{\tau}} + \frac{(\hat{x}_{t+T} - ICC_{CT} \cdot bv_{t+T-1})(1+g)}{(ICC_{CT}-g)(1+ICC_{CT})^{T}} \tag{2.2}$$

Gebhardt et al. (2001):

$$P_t = bv_t + \sum_{\tau=1}^{T} \frac{\hat{x}_{t+\tau} - ICC_{GLS} \cdot bv_{t+\tau-1}}{(1+ICC_{GLS})^{\tau}} + \frac{(\hat{x}_{t+T+1} - ICC_{GLS} \cdot bv_{t+T})}{ICC_{GLS}(1+ICC_{GLS})^{T}} \tag{2.3}$$

Ohlson and Juettner-Nauroth (2005):

$$P_t = (\hat{x}_{t+1}/ICC_{OJ}) \cdot \left(g_{st} + ICC_{OJ} \cdot \hat{d}_{t+1}/\hat{x}_{t+1} - g_{lt}\right) / (ICC_{OJ} - g_{lt}) \tag{2.4}$$

Modified price-earnings growth (PEG) ratio model by Easton (2004):

$$P_t = \left(\hat{x}_{t+2} + ICC_{PEG} \cdot \hat{d}_{t+1} - \hat{x}_{t+1}\right) / ICC_{PEG}^{2} \tag{2.5}$$

The first two models are both special cases of the residual income valuation model. Specifically, $\hat{x}_{t+\tau}$ is the forecasted earnings per share of year $t+\tau$, bv_t is the book value per share at year t, and g is the annualized inflation rate. The major difference between these two is the assumption of the residual income growth rate. Claus and Thomas (2001) assume residual income grows at the expected inflation rate after five years ($T = 5$), while Gebhardt et al. (2001)

derive the residual incomes by linearly fading the forecasted return on equity to the industry median from $T = 3$ to $T = 12$, and assuming residual income to remain constant from $T = 12$ on.

The latter two models are based on the abnormal earnings growth valuation model developed by Ohlson and Juettner-Nauroth (2005). Following Gode and Mohanram (2003), the short-term growth rate g_{st} is estimated as the average between the forecasted earnings growth rate from year $t+1$ to $t+2$ and the five-year growth forecast provided by the I/B/E/S analysts. We use the annualized inflation rate to proxy for the long-term growth rate g_{lt}. As suggested by Gebhardt et al. (2001), we estimate the dividend payout ratio by dividing the actual dividends from the most recent fiscal year by earnings over the same time period and calculate the expected future net dividends per share $\hat{d}_{t+1}$ accordingly. Notably, both these two models require a positive change in forecasted earnings to yield a numerical solution. To be consistent with our Value Line based measure, we use stock price P_t nine days prior to the date of the I/B/E/S report in year t.

2.2.3 Sample and Variable Description

We collect accounting information from Compustat, stock price from CRSP, analyst forecasts from Value Line and I/B/E/S, and 10-K file and annual report information from SEC EDGAR. Our sample spans calendar years from 1995 (the year EDGAR was fully implemented) to 2019, which corresponds to 10-K files for fiscal years from 1994 to 2018. Applying a filter based on all firm-years followed by Value Line and removing firms for which we lack sufficient data for disclosure indices and control variables leaves a final sample of 28,284 firm-years, or 83.6% of all firm-years with sufficient data for the Value Line implied cost of equity capital estimation.[7]

Table 2.4 summarizes our sample selection procedures (Panel A) and reports the sample distribution across industries (Panel B). Overall, our sample is well diversified across different industries and years. None of the industries takes up more than 10% of our sample, of which

[7]Sample size under I/B/E/S is 37,341 firm-year observations.

business services, retail, utilities, and electronic equipment take up 5% or above. Our sample slightly grows over time because Value Line is following more and more firms and providing necessary forecast data to calculate the implied cost of equity capital. There are 2,791 individual firms in our sample, of which 286 firms appear only once, and 163 firms appear in the sample for all 25 years.

Table 2.5 reports the summary statistics for all variables used in this study, including the disclosure index, the implied cost of equity capital, and other control variables. We report both the raw disclosure score, *Raw Disc* $(Total)$, and the percentage-based disclosure score, *Disc* $(Total)$, and a more detailed subcategory disclosure summary can be found in Table 2.2. Consistent with the extant literature, Value Line provides a higher implied cost of equity capital ICC_{VL} than all types of I/B/E/S measures (ICC_{IBES}, ICC_{GLS}, ICC_{CT}, ICC_{OJ}, ICC_{PEG}). The control variables include beta (measured with the CAPM model using a minimum of 24 monthly returns over the 60 months prior to the annual report publication date), size (proxied by the log of firms' market value of equity), book-to-market ratio (proxied by the log of firms' book-to-market ratio), ROA (return on assets), idiosyncratic risk (measured as the residual volatility of the Fama-French 3-Factor model as in Ang et al. (2006)).

2.3 Empirical Analysis of Disclosure Effects on Implied Cost of Equity Capital

2.3.1 Yearly Cross-Sectional Regression (Value Line)

For each calendar year from 1995 to 2019, we run the following cross-sectional regression where size, Beta, and book-to-market ratio are presumed to capture the effects of systematic risk, *Disc* denotes the total percentage-based disclosure score, and *ICC* is the implied cost of equity capital using Value Line data and Brav et al. (2005)'s model:

$$\mathrm{ICC}_{it} = \alpha_t + \gamma_{1t}\ \mathrm{Disc}_{\ it} + \gamma_{2t}\beta_{it} + \gamma_{3t}\ \mathrm{Size}_{\ it} + \gamma_{4t} \ln \mathrm{B/M}_{it} + \varepsilon_{it} \tag{2.6}$$

Table 2.6 contains the yearly cross-sectional regression results.[8] Notably, the slope coefficient for Disc is negative (significantly) only in 4 years (0 years), but is positive (significantly) in 21 years (10 years). The high frequency of positive coefficients, nearly half of which are statistically significant, cannot be attributed to chance. As shown below, when pooling the data over all the years together, the slope coefficient for Disc is significantly positive.

These results are clearly inconsistent with the hypothesis that disclosures reduce the cost of equity capital.[9] Indeed, we are not aware of any rational expectations pricing theory that would explain these findings. As mentioned earlier, in section 2.4, we offer a plausible explanation based on the potential bias contained in the implied cost of equity capital as a proxy for expected return.

2.3.2 Pooling Cross-Sectional Regressions (Value Line)

Table 2.7 reports pooling regression results with Value Line based estimates of implied cost of equity capital. We regress the implied cost of equity capital on factors assumed to capture effects of systematic risk (size, beta, book-to-market ratio) and disclosure scores. As noted, our full sample includes calendar years from 1995 to 2019.[10] Column (1) reports the average coefficient estimates from yearly cross-sectional regressions, as Fama and MacBeth (1973). We report the standard Fama-MacBeth intertemporal t-statistics based on the Newey-West consistent standard error. The significant positive coefficients on total disclosure indices complement the results from our yearly regressions as reported in Table 2.6.

Panel regression results with year and industry fixed effects are presented in the last two columns of Table 2.7. Year fixed effects account for the time-series variation in the implied cost of equity capital across business cycles and control for any potential time trends of the disclosure

[8]As we discuss later, the negative signs on size are consistent with smaller firms tending to have more volatile cash flows, thereby contributing to a bias in implied cost of equity capital as a proxy for expected return.

[9]Botosan (1997) reports a negative coefficient on Disc from a similarly specified regression. Her data was hand collected from 1990 annual reports for 122 manufacturing firms and unavailable.

[10]We also ran our analysis excluding 2001 for the tech bubble and 2007-2008 for the financial crisis with qualitatively unchanged results.

index; the inclusion of industry fixed effects alleviate the concern that our finding is caused by industry-related disclosure norms and/or cross-industry variations of the implied cost of equity capital. Our results also persist after controlling for unobservable factors affecting a given industry in a given year, which are absorbed by year$\times$industry fixed effects. With or without these fixed effects and/or including other fixed effects, the regression results are similar. Column (2) and (3) reports the panel regression results with year fixed effects and industry fixed effects and year$\times$industry fixed effects to examine the extent to which time and/or industry variations contribute to our principal findings. Again, the slope coefficient on *Disc* is significantly positive.

2.3.3 Pooling Cross-Sectional Regressions (I/B/E/S)

In Table 2.8, we report similar pooling regression results to those in Table 2.7 using I/B/E/S estimates of implied cost of equity capital from 1995 to 2019. We calculate five different implied cost of equity capital measures from previous literature, including Claus and Thomas (2001), Gebhardt et al. (2001), Ohlson and Juettner-Nauroth (2005), and Easton (2004), and the average of these four measures. Column (1) reports Fama-Macbeth regression using the average. Column (2) to (5) reports panel regression results with five different estimates of implied cost of equity capital controlling for year$\times$industry fixed effects. The results are qualitatively in line with those reported using Value Line data. The difference in magnitudes of coefficients is traceable to the relative scale difference in the implied cost of equity capital measures.

Overall, the findings reported in the previous three tables are remarkably robust, leaving little doubt that results in Botosan (1997) are special to the year for which she gathered data and do not generalize to samples that extend over time.

2.4 Disclosure and Bias in Implied Cost of Equity Capital

2.4.1 Model of Bias in Implied Cost of Equity Capital as Proxy for Expected Return

We now show that the difference between the implied cost of equity capital and expected return due to the stochastic property of the latter produces a positive relation between disclosure measures and the implied cost of equity capital. Our model extends the identical independently distributed (in time-series) expected return setting of Hughes et al. (2009) to a more general mean-reverting expected return. We equate dividends in our model with cash flows available for distribution to shareholders in Hughes et al. (2009). We begin with a lemma that derives the implied cost of equity capital. We then show that the bias in implied cost of equity capital increases with dividend volatility, an analog to cash flow volatility in Hughes et al. (2009).

Lemma. *Suppose the dividend D_t of a stock follows a geometric Brownian motion*

$$dD_t = D_t \left(g dt + \sigma_d dB_t^d \right) \tag{2.7}$$

and the expected return μ_t follows an OU process

$$d\mu_t = -K\left(\mu_t -\right) dt + \sigma dB_t \tag{2.8}$$

where g, $\sigma_d > 0, K > 0, \bar{\mu}, and \sigma > 0$ are all constant, B_t^d and B_t are two standard Brownian motions with a constant correlation ρ under the physical measure. Then the price-dividend ratio Φ_t at time t is

$$\Phi_t \equiv \frac{P_t}{D_t} = \int_0^\infty e^{-(\hat{\mu}-g)x} e^{-\frac{\mu_t-\hat{\mu}}{K}\left(1-e^{-Kx}\right)+\frac{\sigma^2}{2K^2}\left(x-\frac{2}{K}\left(1-e^{-Kx}\right)+\frac{1}{2K}\left(1-e^{-2Kx}\right)\right)} dx \tag{2.9}$$

where

$$\hat{\mu} = \bar{\mu} + \frac{\rho\sigma\sigma_d}{K} \tag{2.10}$$

The implied cost of equity capital is then

$$\nu_t = g + \frac{1}{\Phi_t} \tag{2.11}$$

Proof. See Appendix.

We assume the time-varying expected return follows a mean-reverting process, which is used in most empirical studies where the process of the expected return is needed to be explicitly assumed. An example of stochastic expected returns is to assume $\mu_t = r_f + \lambda_t \beta$, where the riskless return r_f and β are constant and the market risk premium λ_t follows

$$d\lambda_t = -K\left(\lambda_t - \bar{\lambda}\right)dt + \sigma_\lambda dB_t \tag{2.12}$$

In this case,

$$\bar{\mu} = r_f + \bar{\lambda}\beta \tag{2.13}$$

and

$$\sigma = \beta\sigma_\lambda \tag{2.14}$$

The Lemma shows that, implied cost of equity capital ν_t does not equal the expected return μ_t. It is a non-linear function of expected return μ_t.[11] Some special cases will help to understand the relation between the cost of equity capital and expected return.

- When $K \to +\infty$ and σ is finite. In this case, $\frac{\sigma}{K} \to 0$, the expected return μ_t is effectively constant, $\mu_t = \bar{\mu}$. In this limit, one can show that $\Phi_t = 1/(g + \bar{\mu})$, and the cost of equity capital $v_t = \bar{\mu}$.

- When $K \to +\infty$ and $\sigma \to +\infty$, such that $\frac{\sigma}{K} \neq 0$ (K and σ goes to infinity proportionally). In this limit, one can show that $\Phi_t = 1/\left(g + \bar{\mu} + \frac{\rho\sigma\sigma_d}{K}\right)$, and the cost of equity capital

[11] Implied cost of equity capital depends on parameters ρ,K,σ, and g. The sign of ρ is qualitatively important as we discussed above. The dependence of implied cost of equity capital on K,σ, and g are non-monotonic, which is decided by the value of the expected return μ_t.

$v_t = \bar{\mu} + \frac{\rho\sigma\sigma_d}{K}$. Intuitively, when $\rho > 0$, the expected return μ_t is high when D_t is high, and vice versa. So high dividend states receive high discount, which leads to lower P/D value than when μ_t is constant, which in turn leads to a higher cost of equity capital. This intuition holds in general (also as the 2-state model described in Lambert (2009)). This case is the continuous-time analog of Hughes et al. (2009).

The risk of the dividend in general has two components, systematic and idiosyncratic. Thus the dividend volatility depends on beta as well as idiosyncratic volatility. For example, in a market model, $\sigma_d^2 = \beta_d^2\sigma_m^2 + \sigma_d^{i2}$,where β_d is the beta of the dividend, σ_m is the market volatility, and σ_d^i is the idiosyncratic volatility. In neoclassical asset pricing models, the expected return depends on beta, but not on idiosyncratic volatility.

Proposition. *As in the above lemma, suppose the dividend of a stock follows a geometric Brownian motion*

$$dD_t = D_t\left(gdt + \sigma_d dB_t^d\right) \tag{2.15}$$

and the expected return follows a OU process

$$d\mu_t = -K(\mu_t - \bar{\mu})\,dt + \sigma dB_t \tag{2.16}$$

where $g, \sigma_d > 0, K > 0, \bar{\mu}$, and $\sigma > 0$ are all constant, B_t^d and B_t are two standard Brownian motions with a constant correlation ρ under the physical measure. Then the implied cost of equity capital v_t increases with the idiosyncratic volatility σ_d^i of the dividend if $\rho > 0$.

Proof. Note that

$$\frac{\partial v_t}{\partial \sigma_d^i} = -\frac{1}{\Phi_t^2}\frac{\partial \Phi_t}{\partial \sigma_d}\frac{\partial \sigma_d}{\partial \sigma_d^i} \tag{2.17}$$

From $\hat{\mu} = \bar{\mu} + \frac{1}{K}\rho\sigma\sigma_d$, the derivative of Φ_t with respect to σ_d is

$$\frac{\partial \Phi_t}{\partial \sigma_d} = -\frac{\rho\sigma}{K}\int_0^{\infty}(x - \frac{1}{K}(1-e^{-Kx}))e^{-(\hat{\mu}-g)x}e^{-\frac{\mu_t-\hat{\mu}}{K}(1-e^{-Kx})+\frac{\sigma^2}{2K^2}(x-\frac{2}{K}(1-e^{-Kx})+\frac{1}{2K}(1-e^{-2Kx}))}dx \tag{2.18}$$

We used the assumption that the expected return does not depend on idiosyncratic volatility σ_d^i. Thus, the derivative of the expected return μ_t, as well as the parameter $\bar{\mu}$, with respect to the idiosyncratic volatility is zero. The result follows noting that $\frac{\partial \sigma_d}{\partial \sigma_d^i} > 0$ and $-\frac{\rho\sigma_d}{K} < 0$ because $\rho > 0$ and $\left(x - \frac{1}{K}\left(1-e^{-Kx}\right)\right) > 0$ for $x > 0$. ■

It is widely documented that the correlation (ρ) between dividends (cash flows) and expected returns is positive. Under this condition, the proposition establishes that the implied cost of equity capital increases with dividend volatility, given that the expected return is independent of the dividend volatility under neoclassical asset pricing theory. Plausibly assuming that firms with a higher disclosure index also have higher dividend volatility provides an explanation for our empirical finding of a positive association between implied cost of equity capital and disclosure indices as a measure of cash flow news.[12] Below, we take a closer look empirically at the links between disclosure indices, cash flow volatility, and potential bias in implied cost of equity capital.

2.4.2 Correlations between Disclosure, Idiosyncratic Risk, and Implied Cost of Equity Capital

In Table 2.9, we estimate the correlations between the Value Line implied cost of equity capital, cash flow volatility stemming from idiosyncratic risk, and disclosure indices. Idiosyncratic risk is measured as the idiosyncratic volatility of the Fama-French 3-factor model (Ang et al., 2006). Since the disclosure index slightly increases over time, we detrend our disclosure score at the year level. We get similar results if we control for both year and industry fixed effects.

[12] Prices move with the news. We view disclosure indices as a reflection of cash flow news, a large component of which pertains to idiosyncratic risk.

The correlations reported in Table 2.9 support our explanation for a positive relation between implied cost of equity capital and disclosure indices. Consistent with the Proposition that implied cost of equity capital is increasing with idiosyncratic volatility, we find significant positive correlations between the implied cost of equity capital, idiosyncratic volatility as a measure of cash flow volatility attributable to idiosyncratic risk, and disclosure indices as a measure of news that gives rise to idiosyncratic volatility.

In short, we believe we have made a strong case for the prospect that the positive relation between disclosure indices and implied cost of equity capital documented in section 2.3 is likely driven by a bias in the latter as a proxy for expected return.

2.5 Conclusion

Understanding the relation between information and the cost of equity capital is central to accounting research. Botosan (1997) was seminal in its examination of the impact that disclosures contained in financial reports might have on a firm's cost of equity capital, i.e., its expected return. Her findings were suggestive of an effect of greater disclosure serving to reduce expected return after controlling for systematic risks. This interpretation was controversial. Neoclassical theory maintains that information on idiosyncratic risks should have no such effect on expected return. The dissonance between the empirical findings and theory prompted us to reexamine the linkage between disclosure and expected return, given the availability of an empirical technology that enabled a greatly enlarged sample extending over a quarter century. Our results strongly reject the hypothesis of a negative relation between disclosure and implied cost of equity capital as a proxy for expected return over that time frame. On the contrary, we found compelling evidence of a positive relation.

A potential explanation for our results is the prospect of a systematic bias in implied cost of equity capital modeled by Hughes et al. (2009) and depicted in our parsimonious, continuous-time extension of their analysis. Specifically, in keeping with Merton (1973)'s insight of stochastic expected returns, we identify conditions under which a bias in implied cost of equity capital as a proxy for expected return is increasing in the volatility of cash flows. Viewing disclosure indices as depicting the level of cash flow news suggests a positive relation between disclosure and cash flow volatility. Moreover, given that firms are small relative to the market suggests a positive relation between disclosure and the volatility of residual returns from a factor model as a measure of idiosyncratic volatility. We confirm these relationships empirically.

As with most empirical studies, we are limited in identifying defensible proxies for variables that we cannot observe. Although we believe our study poses a strong case for a bias in implied cost of equity capital as a measure of expected return driving a positive association between disclosure indices and implied cost of equity capital, we acknowledge that the conditions for our conclusions in that regard to hold as discussed in Hughes et al. (2009) and Lambert (2009) are not immutable.

Finally, we note that the advance in analyzing the contents of financial reports through textual analysis opens a new chapter for empirical studies addressed to the role that information firms provide to investors affects what they require as an expected return. The expansion in sampling enabled in this study calls into question the generalizability of earlier work limited by the available technology at those times. It seems to us that there are many opportunities for accounting researchers to employ textual analysis in revisiting previous studies or initiating new inquiries seeking to connect information with asset pricing.

2.6 Acknowledgements

Chapter 2 is currently being prepared for submission for publication and is couathored with John Hughes, Jun Liu, and Dan Yang. Jun Chen, the dissertation author, is the primary investigator and author of this paper.

Table 2.1. The Coding Scheme to Analyze 10-K Filings

I. Summary of historical results
a. Return on assets or sufficient information to compute ROA (net income, tax rate, interest expense, and total assets)
b. Net profit margin or sufficient information to compute PM (net income, tax rate, interest expense, and sales)
c. Asset turnover or sufficient information to compute TAT (sales and total assets)
d. Return on equity or sufficient information to compute ROE (net income and total equity)
e. Number of quarters that firm discloses sales and net income
f. Trends in the industry
g. Discussion of corporate strategy
II. Other financial measures
a. Free cash flow (or cash flow other than those reported in SCF)
b. Economic profit, residual income type measure
c. Cost of capital (wacc, hurdle rate, EVA target rate)
III. Nonfinancial measures
a. Number of employees
b. Average compensation per employee
c. Percentage of sales in products designed in the past few (3–5) years
d. Market share
e. Units sold (or other output measure, e.g., production)
f. Unit selling price
g. Growth in units sold (or growth in other output measure, e.g., production)
h. Growth in investment (expansion plans, number of outlets, etc.)
IV. Projected information
a. Forecasted market share
b. Cash flow forecast
c. Capital expenditures, R&D expenditures, or general investment forecast
d. Profit forecast
e. Sales forecast
f. Other output forecast
g. Industry forecast

We use the same coding scheme as in Francis et al. (2008), including four major categories and 25 individual items.

Table 2.2. Summary Statistics of the Disclosure Index

Variable	N	Mean	Std. Dev.	10%	25%	Median	75%	90%
Panel A: FNO's Disclosure Index								
Raw Disc (Total)	677	7.98	2.97	4	6	8	10	12
Raw Disc Cat. I	677	2.30	1.50	1	1	2	3	5
Raw Disc Cat. II	677	0.15	0.38	0	0	0	0	1
Raw Disc Cat. III	677	3.18	1.52	1	2	3	4	5
Raw Disc Cat. IV	677	2.35	1.58	0	1	2	3	5
Panel B: Our Disclosure Index in FNO's Sample								
Variable	N	Mean	Std. Dev.	10%	25%	Median	75%	90%
Raw Disc (Total)	627	7.95	2.99	4	6	8	10	12
Raw Disc Cat. I	627	2.24	1.54	1	1	2	3	5
Raw Disc Cat. II	627	0.23	0.48	0	0	0	0	1
Raw Disc Cat. III	627	3.20	1.35	2	2	3	4	5
Raw Disc Cat. IV	627	2.28	1.61	0	1	2	3	5
Panel C: Our Disclosure Index in the Full Sample								
Variable	N	Mean	Std. Dev.	10%	25%	Median	75%	90%
Raw Disc (Total)	43,806	10.27	3.43	6	8	10	13	15
Raw Disc Cat. I	43,806	2.22	1.25	1	1	2	3	3
Raw Disc Cat. II	43,806	0.39	0.58	0	0	0	1	1
Raw Disc Cat. III	43,806	4.07	1.43	2	3	4	5	6
Raw Disc Cat. IV	43,806	3.59	1.78	1	2	4	5	6

This table summarizes the descriptive statistics of FNO's (Francis et al., 2008) disclosure index as well as our disclosure index in FNO's original sample (fiscal 2001) and the entire Value Line/IBES sample from 1995 to 2019. We report both the total raw disclosure index and individual disclosure indexes for each of the four subcategories. Detailed construction of these individual subcategories can be found in Tabel 1.

Table 2.3. Inter-correlations of raw disclosure indices

	Total	*Cat. I*	*Cat. II*	*Cat. III*	*Cat. IV*
Raw Disc (Total)	1.0000	0.4687	0.4624	0.7711	0.8428
		0.0000	0.0000	0.0000	0.0000
Raw Disc Category I	0.5036	1.0000	0.0910	0.1031	0.1125
	0.0000		0.0000	0.0000	0.0000
Raw Disc Category II	0.4636	0.1242	1.0000	0.2146	0.3364
	0.0000	0.0000		0.0000	0.0000
Raw Disc Category III	0.7690	0.1606	0.2186	1.0000	0.5554
	0.0000	0.0000	0.0000		0.0000
Raw Disc Category IV	0.8510	0.2074	0.3472	0.5632	1.0000
	0.0000	0.0000	0.0000	0.0000	

This tables presents the inter-correlations of raw disclosure indices in total and by category. The lower left shows the Spearman correlation, while the upper right shows the Pearson correlation.

Table 2.4. Sample Selection Procedures

Panel A: Sample Selection

	Value Line		I/B/E/S	
	N	%	N	%
Firm-years with sufficient data for the cost of capital estimation	33,837	100	51,464	100
No disclosure score	3,569	10.5	8,896	17.3
Insufficient data for controls	1,984	5.9	5,227	10.2
Total firm-year observations	28,284	83.6	37,341	72.6

Panel B: Sample divided by industry (top 20)

	Value Line		I/B/E/S	
	N	%	N	%
Business Services	2,543	8.99	4,130	11.06
Retail	2,064	7.30	2,258	6.05
Utilities	1,728	6.11	1,581	4.23
Electronic Equipment	1,649	5.83	2,180	5.84
Banking	1,302	4.60	4,182	11.20
Machinery	1,217	4.30	1,355	3.63
Insurance	1,141	4.03	1,758	4.71
Pharmaceutical Products	1,044	3.69	1,029	2.76
Wholesale	976	3.45	1,345	3.60
Petroleum and Natural Gas	971	3.43	980	2.62
Computers	917	3.24	1,316	3.52
Trading	847	2.99	1,456	3.90
Medical Equipment	832	2.94	1,079	2.89
Chemicals	831	2.94	814	2.18
Measuring and Control Equipment	714	2.52	802	2.15
Transportation	701	2.48	1,053	2.82
Construction Materials	692	2.45	683	1.83
Food Products	644	2.28	635	1.70
Automobiles and Trucks	621	2.20	589	1.58
Consumer Goods	532	1.88	592	1.59

This table shows our sample selection procedures and the distribution of firm-year observations by industry. Specifically, we use the Fama-French 48 industry classification and report the top 20 industries based on the Value Line sample.

Table 2.5. Summary Statistics

Variable	N	Mean	Std. Dev.	10%	25%	Median	75%	90%
Raw Disc (Total)	43,806	10.269	3.433	6	8	10	13	15
Disc (Total)	43,806	0.613	0.175	0.357	0.500	0.632	0.737	0.833
ICC_{VL}	28,284	0.142	0.083	0.048	0.086	0.132	0.188	0.255
ICC_{IBES}	37,341	0.098	0.027	0.069	0.080	0.094	0.110	0.132
ICC_{GLS}	41,820	0.063	0.030	0.027	0.043	0.062	0.081	0.098
ICC_{CT}	41,949	0.094	0.032	0.062	0.077	0.091	0.108	0.129
ICC_{OJ}	37,998	0.115	0.032	0.083	0.095	0.110	0.129	0.155
ICC_{PEG}	41,029	0.119	0.049	0.076	0.090	0.108	0.135	0.177
Realized Return	41,668	0.152	0.625	-0.403	-0.156	0.085	0.338	0.677
Beta	43,806	1.125	0.719	0.318	0.620	1.018	1.494	2.060
Idio. Risk	43,806	0.101	0.055	0.048	0.063	0.088	0.125	0.170
Size	43,806	7.148	1.710	4.986	5.927	7.047	8.243	9.467
B/M	43,806	0.535	0.375	0.166	0.282	0.460	0.692	0.977
ln B/M	43,806	-0.858	0.728	-1.796	-1.266	-0.776	-0.369	-0.023
ROA	43,804	0.040	0.081	-0.023	0.011	0.040	0.079	0.123
Analysts	41,585	9.238	7.276	2	4	7	13	20
Audit Fee	32,951	14.083	1.195	12.517	13.296	14.067	14.857	15.640

This table summarizes the descriptive statistics of all variables used in this study. *Raw Disc (Total)* = the raw disclosure score of the 10-K file; *Disc (Total)* = the percentage-based disclosure score scaled by the maximum disclosure index in that year; ICC_{VL} = the Value Line implied cost of capital as in Brav et al. (2005); ICC_{GLS}, ICC_{CT}, ICC_{OJ}, ICC_{PEG} are respectively the implied cost of capital as described in Gebhardt et al. (2001), Claus and Thomas (2001), Ohlson and Juettner-Nauroth (2005), and Easton (2004); ICC_{IBES} = the average of the above four measures of the implied cost of capital using the I/B/E/S data; *Beta*: beta coefficients estimated with the CAPM model using a minimum of 24 monthly returns over the 60 months prior to the annual report publication date ; *Idio. Risk*: idiosyncratic risk, measured as the residual volatility of the Fama-French 3-Factor model as in Ang et al. (2006); *Size*: the log of firms' market value of equity in millions of dollars; (*ln*) *B/M*: (the log of) the firm's book-to-market ratio; *ROA*: return on assets; *Analysts*: the number of analysts following the firm; *Audit Fee*: the log of the firms' auditing fees.

Table 2.6. Yearly Cross-Sectional Regressions

Year	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Size	-0.670***	-1.001***	-1.278***	-1.767***	-1.879***	-0.995***	-0.548***	-0.187	-0.214	-0.128	-0.261	-0.116	0.015
	(-2.74)	(-4.22)	(-6.03)	(-8.64)	(-9.66)	(-4.26)	(-2.99)	(-1.02)	(-1.16)	(-0.77)	(-1.54)	(-0.77)	(0.10)
Beta	1.929***	2.269***	1.845***	3.160***	1.905***	2.253***	3.694***	3.281***	0.349	1.150***	2.312***	1.204***	1.117***
	(2.95)	(3.82)	(3.81)	(6.33)	(3.63)	(3.72)	(9.02)	(10.38)	(1.19)	(4.46)	(8.80)	(4.94)	(4.26)
ln B/M	1.294**	0.462	-0.840*	0.181	1.812***	3.463***	1.421***	0.764*	1.112***	-0.100	0.004	-0.436	0.100
	(2.06)	(0.77)	(-1.71)	(0.38)	(4.54)	(7.78)	(4.11)	(1.96)	(2.73)	(-0.25)	(0.01)	(-1.22)	(0.26)
Disc	1.960	3.546**	4.771***	4.111***	1.275	-0.229	1.987	2.371*	2.008	0.163	0.337	1.814	-0.133
	(1.17)	(2.11)	(3.22)	(2.89)	(0.90)	(-0.14)	(1.59)	(1.80)	(1.45)	(0.13)	(0.26)	(1.42)	(-0.10)
R^2	0.035	0.043	0.055	0.115	0.176	0.147	0.118	0.122	0.020	0.023	0.082	0.034	0.017
Firms	782	858	968	976	1,086	975	1,059	1,116	1,149	1,139	1,112	1,105	1,147
Year	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	
Size	-0.297*	0.026	0.045	-0.214	-0.271*	-0.364***	-0.489***	-0.219*	-0.419***	-0.438***	-0.311**	-0.495***	
	(-1.92)	(0.17)	(0.33)	(-1.47)	(-1.89)	(-3.07)	(-4.22)	(-1.83)	(-3.19)	(-3.34)	(-2.16)	(-3.00)	
Beta	2.267***	1.360***	0.747***	2.705***	2.997***	1.747***	0.787***	3.037***	3.593***	1.769***	1.624***	4.013***	
	(7.49)	(3.54)	(2.62)	(8.51)	(9.06)	(6.38)	(3.03)	(9.45)	(10.04)	(5.14)	(4.47)	(9.00)	
ln B/M	1.220***	0.409	0.968***	1.064***	1.772***	0.566**	0.152	1.004***	0.555**	1.042***	2.145***	3.738***	
	(3.84)	(1.33)	(3.25)	(3.45)	(6.22)	(2.36)	(0.64)	(4.68)	(2.45)	(4.45)	(8.56)	(13.46)	
Disc	2.527*	2.260	2.346*	1.542	3.113**	-0.147	0.987	3.113**	1.107	-0.046	5.048***	4.388**	
	(1.80)	(1.60)	(1.70)	(1.12)	(2.00)	(-0.11)	(0.73)	(2.18)	(0.70)	(-0.03)	(2.88)	(2.22)	
R^2	0.087	0.021	0.020	0.083	0.133	0.061	0.030	0.114	0.112	0.059	0.091	0.243	
Firms	1,179	1,160	1,205	1,215	1,219	1,245	1,247	1,264	1,291	1,284	1,296	1,207	

This table reports the results of cross-sectional regressions of the Value Line implied cost of equity capital (in percentage) on known risk proxies (*Beta*, *Size*, and *ln B/M*), and the percentage-based disclosure indices *Disc*, for each calendar year from 1995 to 2019. The t-statistics are reported in the parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 2.7. Disclosure and the Implied Cost of Equity Capital (Value Line)

	Implied Cost of Capital		
	(1)	(2)	(3)
Size	-0.499***	-0.388***	-0.358***
	(-3.08)	(-5.76)	(-5.31)
Beta	2.125***	1.399***	1.492***
	(12.11)	(10.47)	(10.54)
ln B/M	0.955***	1.589***	1.450***
	(3.66)	(12.11)	(10.92)
Disc	2.009***	1.634***	1.770***
	(5.95)	(3.44)	(3.72)
R^2	0.082	0.204	0.270
Observations	28284	28284	28218
Year FE		Yes	No
Industry FE		Yes	No
Year*Ind FE		No	Yes
Method	Fama-Macbeth	Panel	Panel

This table reports results from regressing the implied cost of capital (in percentage) on known risk proxies (*Beta*, *Size*, and *ln B/M*) and the disclosure indices. *Disc* is the percentage-based total disclosure score. In columns (1), we follow Fama and MacBeth (1973) and report the mean of the annual coefficient estimates; t-statistics are calculated based on the Newey-West consistent standard error. In columns (2) and (3), we report panel regressions with year, industry, and year*industry fixed effects. Robust t-statistics adjusted for firm-level clustering are reported in the parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 2.8. Disclosure and the Implied Cost of Equity Capital (I/B/E/S)

	Implied Cost of Equity Capital					
	(1)	(2)	(3)	(4)	(5)	(6)
	ICC_{IBES}	ICC_{IBES}	ICC_{GLS}	ICC_{CT}	ICC_{OJ}	ICC_{PEG}
Size	-0.259***	-0.324***	-0.144***	-0.195***	-0.423***	-0.645***
	(-10.15)	(-18.77)	(-11.18)	(-9.68)	(-20.27)	(-22.40)
Beta	0.578***	0.480***	0.180***	0.172***	0.583***	1.178***
	(6.15)	(15.25)	(7.05)	(4.20)	(14.84)	(19.23)
ln B/M	1.049***	0.942***	1.451***	0.378***	0.739***	1.266***
	(12.60)	(24.66)	(37.48)	(8.07)	(16.48)	(20.14)
Disc	0.762***	0.660***	0.220**	0.396***	0.858***	1.799***
	(5.32)	(5.65)	(2.47)	(2.79)	(6.00)	(8.28)
R^2	0.195	0.382	0.693	0.224	0.294	0.288
Observations	37341	37287	41774	41902	37944	40989
Year*Ind FE		Yes	Yes	Yes	Yes	Yes

This table reports results from regressing the implied cost of capital (in percentage) on known risk proxies (*Beta*, *Size*, and *ln B/M*) and the disclosure indices. *Disc* is the percentage-based total disclosure score. In columns (1), we follow Fama and MacBeth (1973) and report the mean of the annual coefficient estimates; t-statistics are calculated based on the Newey-West consistent standard error. In columns (2)-(6), we report panel regressions with year*industry fixed effects. ICC_{GLS}, ICC_{CT}, ICC_{OJ}, ICC_{PEG} are respectively the implied cost of capital as described in Gebhardt et al. (2001), Claus and Thomas (2001), Ohlson and Juettner-Nauroth (2005), and Easton (2004); ICC_{IBES} = the average of the above four measures of the implied cost of capital using the I/B/E/S data. Robust t-statistics adjusted for firm-level clustering are reported in the parentheses *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 2.9. Relation between disclosure, idiosyncratic volatility, and implied cost of equity capital

	Disc	*Idiosyncratic VOL*	*Implied CoC*
Disc	1.0000	0.1879	0.0527
		0.0000	0.0000
Idiosyncratic VOL	0.1838	1.0000	0.2224
	0.0000		0.0000
Implied CoC	0.0501	0.2515	1.0000
	0.0000	0.0000	

This table presents the correlations between the disclosure index (*Disc*), idiosyncratic volatility (*Idiosyncratic VOL*) measured as the residual volatility from a Fama-French 3-factor model, and Value Line implied cost of equity capital (*Implied CoC*). The disclosure index is detrended at the year level. The lower left shows the Spearman correlation, while the upper right shows the Pearson correlation.

Chapter 3

Investor Distrust: The Unintended Effect of an Earnings-Based Delisting Policy

3.1 Introduction

The financial and real consequences of earnings management are central to accounting research. 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 marginally positive earnings compared to firms in the United States, as shown in Figure 3.1 below. Chinese investors are well aware of both the de-listing policy and what firms have been doing.

We categorize all firms with an annual ROE (return on equity) from 0 to 4% into the low information segment since many firms in this segment are suspects of earnings management. Correspondingly, firms with ROE in (7%,+∞) are categorized into the high information segment since these firms do not face an imminent pressure of delisting (as shown in figure 3.3). We provide evidence on why we divide China's stock market this way[1]. We also show that firms in

[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%.

the low information segment indeed have higher earnings management than other firms.

We 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 United States. 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 2000-2016.[2]

We provide empirical evidence that investors cannot distinguish which firm actually engages in earnings management. First, we divide all firms in the low information segment into five quantiles based on two widely used measures of earnings management, namely discretionary accruals 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' reactions 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

[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.

words, the 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 ROE drop but stayed in the 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 a 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 literature 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 endogeneity 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 have disengaged from unobservable country-level omitted variables. Our paper also provides a rare opportunity to study the real value of financial market development. Treating the low information segment as a developing, less supervised market, and the high information segment as a developed, more mature market, we offer 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 3.2 reviews the literature. Section 3.3 provides institutional background on China's delisting policy. Section 3.4 presents summary statistics on the data. We present evidence on firms' earnings management in Section 3.5. In section 3.6, we document the existence of two information segments and the financial effects of falling into the low information segment. Section 3.7 shows that investors treat good and bad firms indifferently in the low information segment. In section 3.8, we pin down a group of firms that exogenously fall into the low information segment due to the global financial crisis and present causal evidence on the effects of sliding into the low information segment. Section 3.9 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 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

[3] See Appendix for more details.

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 (1989), 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.

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

[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.

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.

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

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

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 3.4: 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 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.

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:

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

where DA_i is the discretionary accrual of firms i. $1_{\mathbf{ROE \in (0,0.04)}}$ 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:

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

where $1_{\mathbf{ROE \in (0,0.04)}}$ 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.

[6] See Appendix for details.

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

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%, +∞) 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 3.5 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.

Firstly, Figure 3.5 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.

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

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 3.7. 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 3.5.

From Figure 3.7, 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., 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 3.9. 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 3.9, 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 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.

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

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 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 Morck et al. (2000a); Jin and Myers (2006), 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 3.11 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 3.11 is consistent with our reasoning. We see from Figure 3.11 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%.

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 Morck et al. (2000a);

Jin and Myers (2006), 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 3.12 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 3.12 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 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 3.14, 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 3.16 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 3.6, we define *Escaper* as a dummy variable that equals 1 if a 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 3.4 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.

3.8 Causal Evidence on the Effect 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 3.6, 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 3.18, 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.

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 3.19 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.

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.

$$Y_t/BE_t = d_0 + d_1 VE_t/BE_t + d_2 VE_{t-1}/BE_{t-1} + d_3 DD_t + d_4 D_t/BE_t + \varepsilon_t \tag{3.7}$$

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 3.5 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.

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

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

$CP_t = Y_t/BE_t - Y_{t-1}/BE_{t-1}$ 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 3.6 shows the result for our second stage regression.

Difference in Differences Estimation

Here is our estimation equation:

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

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 (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.

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 3.7 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.

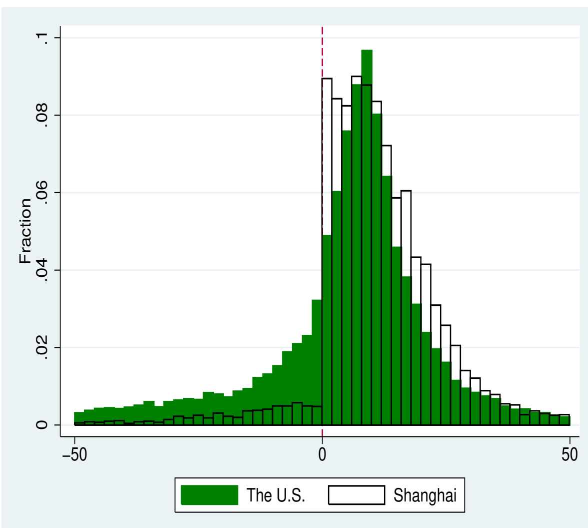
3.9 Conclusion

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.

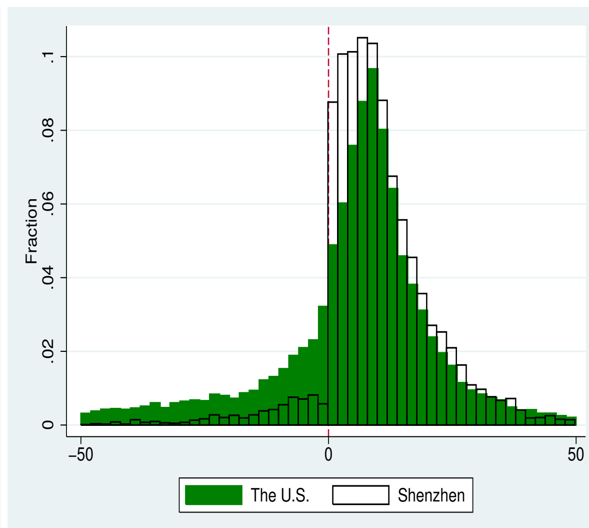
We propose and test the spillover effect of earnings management by a set of firms on market reaction to other similar firms. More specifically, we show that the delisting policy endogenously divides China's stock market into high and low information segments. We document significant adverse consequences 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.

3.10 Acknowledgements

Chapter 3 is currently being prepared for submission for publication and is couathored with Yibin Liu. Jun Chen, the dissertation author, is the primary investigator and author of this paper.



(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)

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 3.1(a) plots all the firms listed in the Shanghai Stock Exchange, whereas Figure 3.1(b) all the firms in the Shenzhen Stock Exchange.

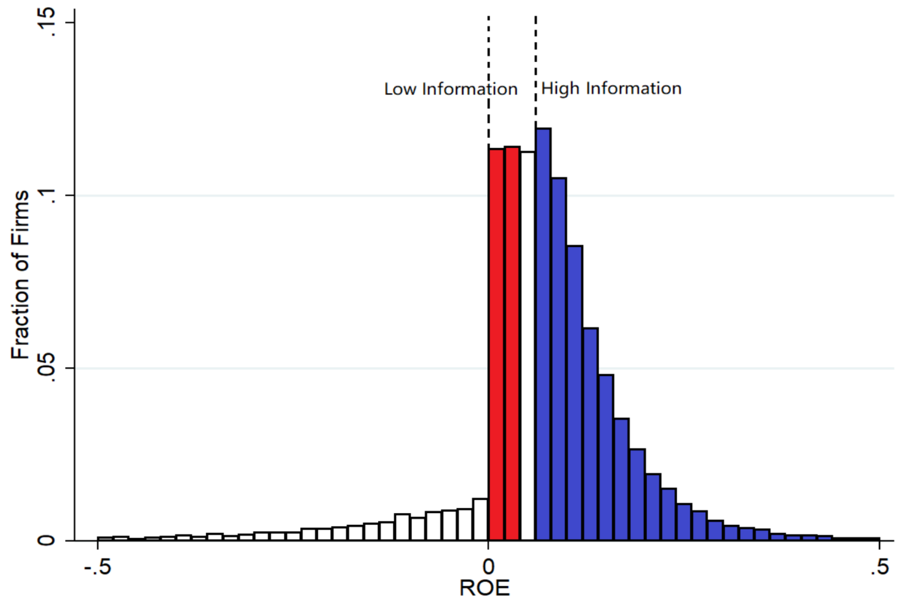


Figure 3.3. Two Information Segments in China

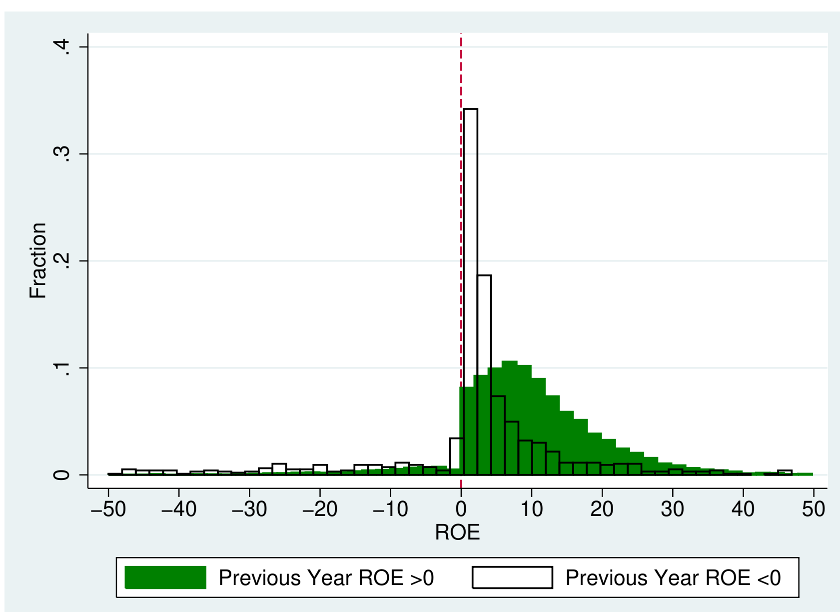
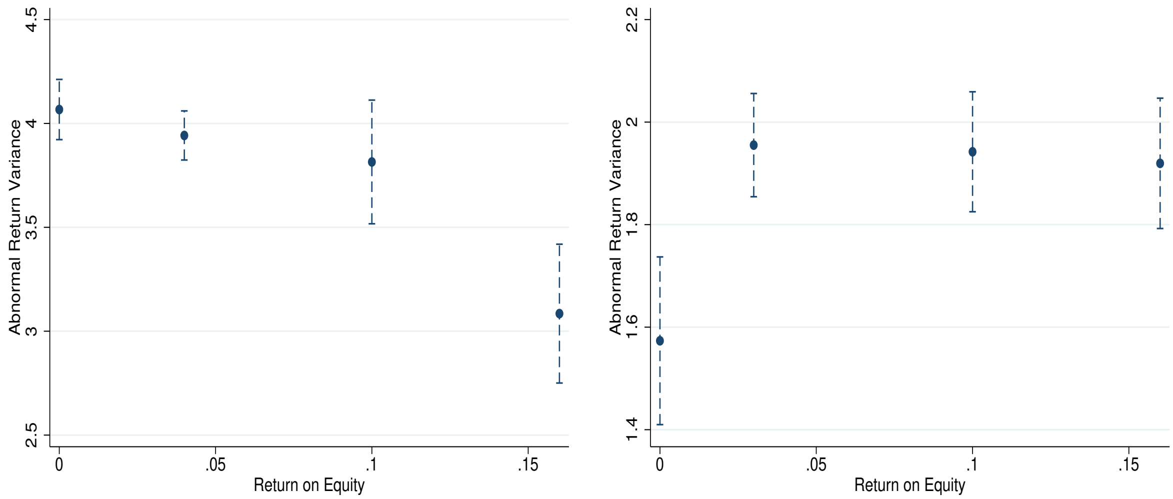
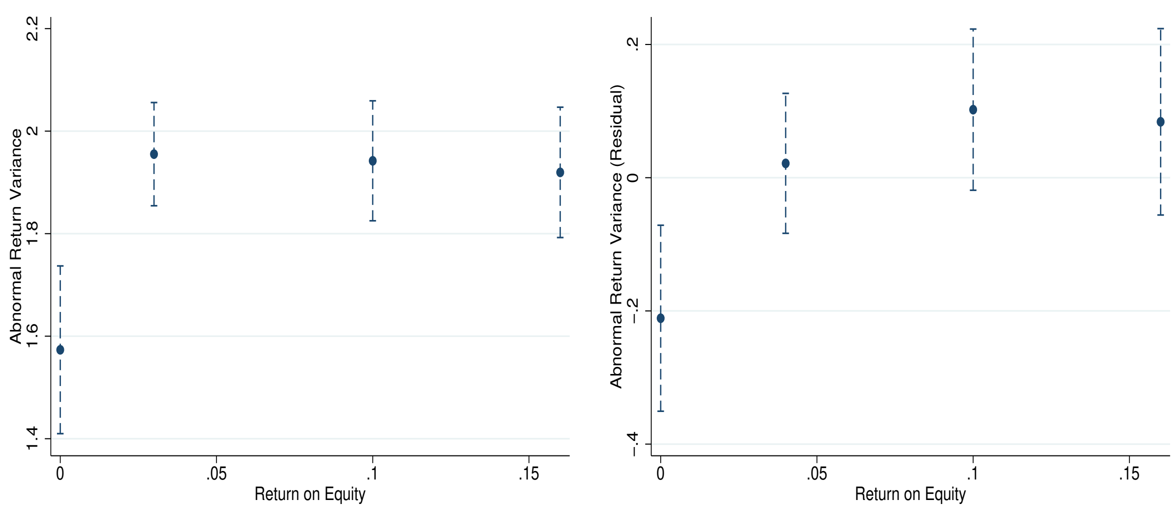


Figure 3.4. Firm ROE Distribution in China: Positive v.s. Negative ROE (Last Year)



(a) the U.S.　　**(b)** China

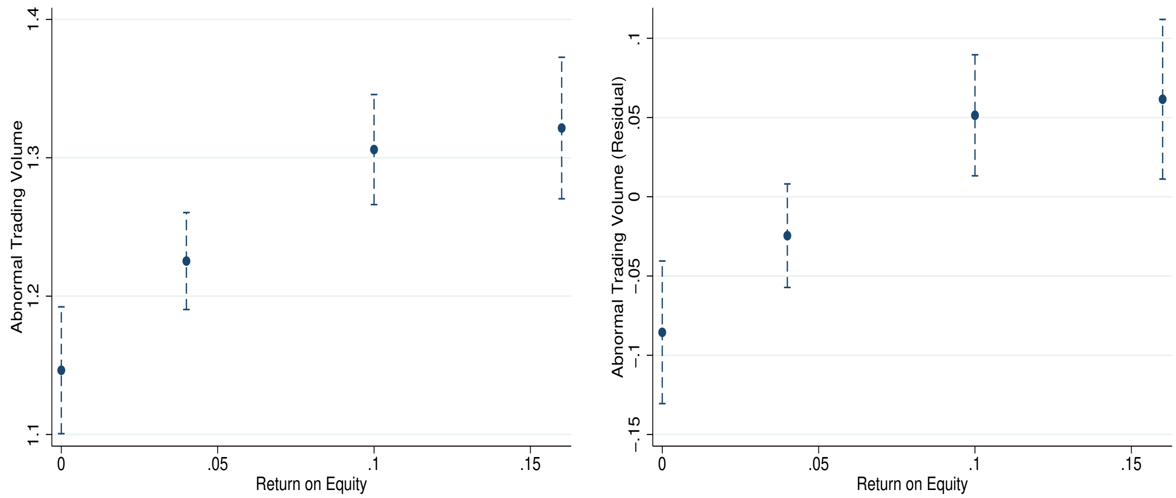
Figure 3.5. Abnormal Return Variance Around Firms' Annual Earnings Announcement: China v.s. the U.S. (2009-2016)



(a) Level of Abnormal Return Variance　　**(b)** Residual of Abnormal Return Variance

Figure 3.7. Abnormal Return Variance Around Firms' Annual Earnings Announcement: China

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



(a) Level of Abnormal Trading Volume

(b) Residual of Abnormal Trading Volume

Figure 3.9. Abnormal Trading Volume Around Firms' Annual Earnings Announcement: China

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

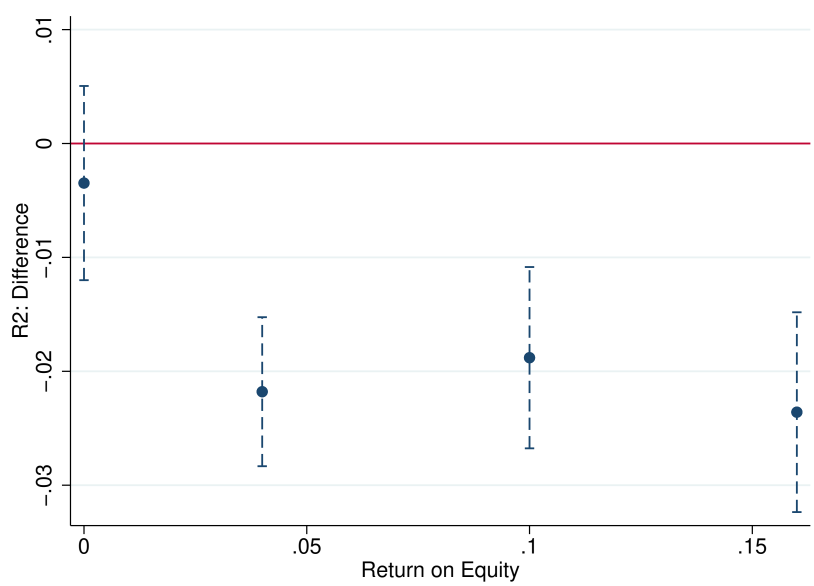
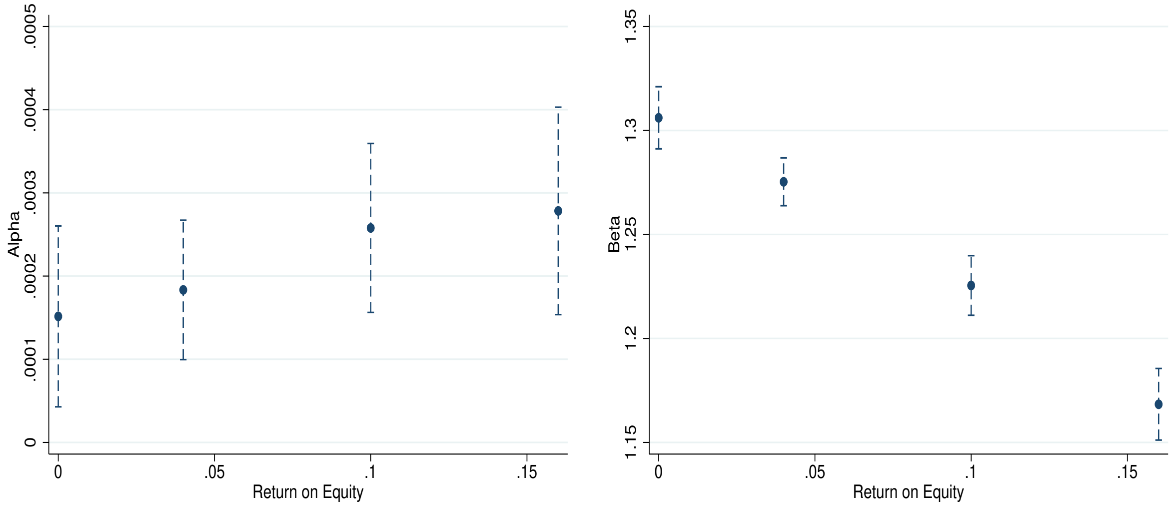


Figure 3.11. Difference in R^2: Pre- and Post- Earnings Announcement

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



(a) Alpha

(b) Beta

Figure 3.12. Alpha and Beta: Post Earnings Announcement

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

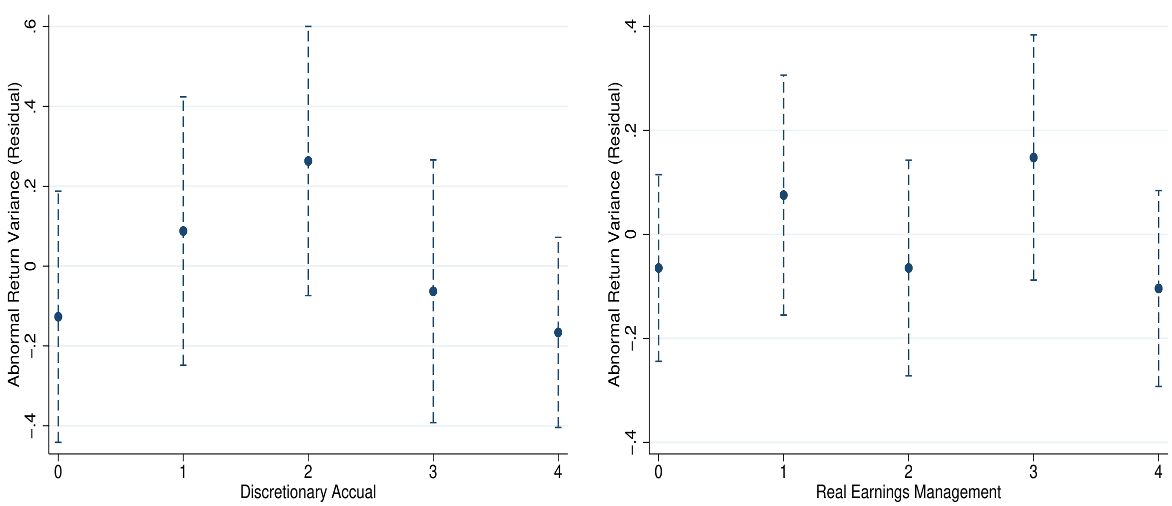


Figure 3.14. Abnormal Return Variance for Firms with Different Levels of Earnings Management in the Low Information Segment

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.

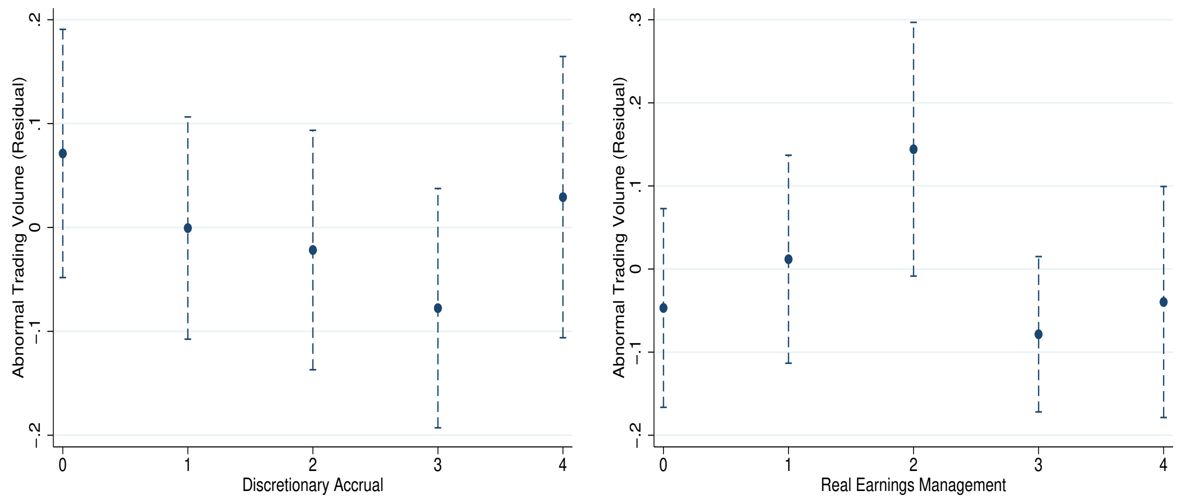


Figure 3.16. Abnormal Trading Volume for Firms with Different Levels of Earnings Management in the Low Information Segment

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

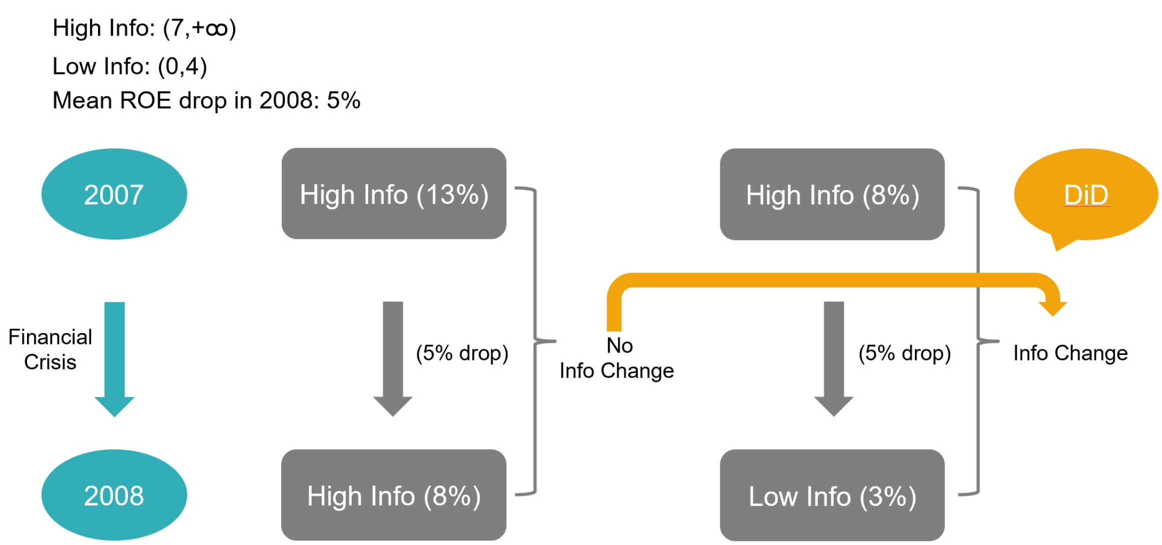
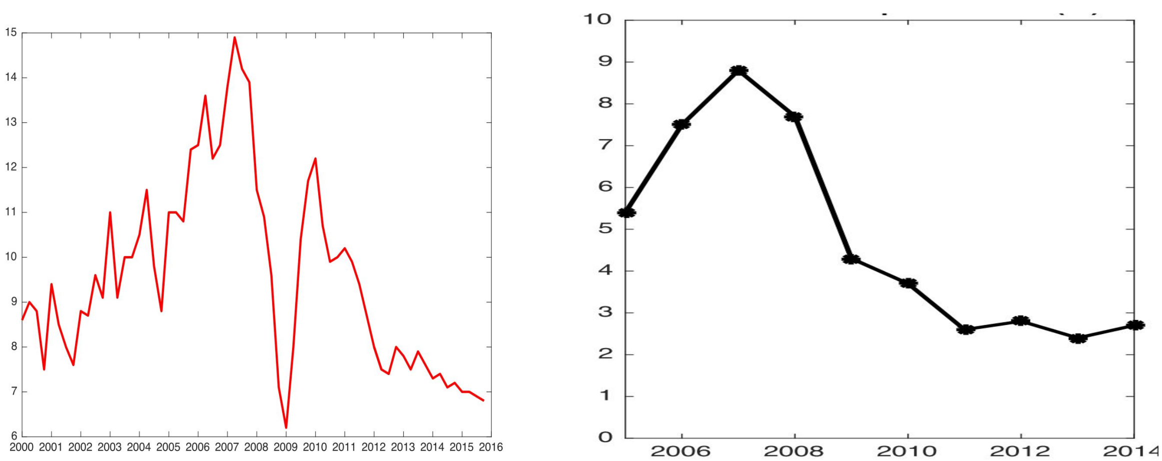


Figure 3.18. DID Design



(a) China's Quarterly GDP Growth

(b) China's Annual Export/GDP Ratio

Figure 3.19. Impact of the 2007/08 Financial Crisis on China's Economy

These two graphs are from Chong-en et al. (2016)

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

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

	(1) RM	(2) RM	(3) DA	(4) DA
$1_{ROE}\in$**(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.

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.

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).

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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.6. Second stage regression

	(1)	(2)	(3)	(4)
	CP_{t+1}	CP_{t+1}	CP_{t+1}	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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

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$\times$ 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).

Appendix A

Appendix for Chapter 1

A.1 Variable Definitions

Table A.1. Description of Main Variables

Variable	Definition
Share Turnover	*Share Turnover* is defined as the daily share volume over tradable shares outstanding (in percentage). *Abnormal Share Turnover* measures the change of *Share Turnover* before and after the implementation day using multiple event windows.
Returns	The detailed construction of stock returns is provided in the main analysis.
Δ*DISEXP*	Abnormal discretionary expenses, measured as the annual change of SG&A expenses (selling, general, and administrative) scaled by beginning-of-year total assets.
Δ*CAPEX*	Abnormal capital expenditures, measured as the annual change of capital expenditures scaled by beginning-of-year total assets.
Δ*ASSETS*	Asset growth rate, measured as the annual change in total assets scaled by beginning-of-year total assets.

Variable	Definition
ADA	Using the modified Jones model, I estimate the absolute discretionary accruals *ADA* as follows. I first run the following cross-sectional OLS regression within each industry year based on the 2012 CSRC industry classification: $$\frac{TA_{i,t}}{A_{i,t-1}} = \beta_1 \frac{1}{A_{i,t-1}} + \beta_2 \frac{\Delta REV_{i,t}}{A_{i,t-1}} + \beta_3 \frac{PPE_{i,t}}{A_{i,t-1}} + \varepsilon_{i,t}$$ where i indexes firm and t indexes year. Total accruals $TA_{i,t}$ are defined as the net income minus operating cash flows, $A_{i,t-1}$ is total assets at the beginning of year t, $\Delta REV_{i,t}$ is the change in sales revenues from the preceding year, and $PPE_{i,t}$ is property, plant, and equipment. The non-discretionary accruals are calculated as $$NDA_{i,t} = \hat{\beta}_1 \frac{1}{A_{i,t-1}} + \hat{\beta}_2 \frac{\Delta REV_{i,t} - \Delta AR_{i,t}}{A_{i,t-1}} + \hat{\beta}_3 \frac{PPE_{i,t}}{A_{i,t-1}} + \varepsilon_{i,t}$$ where ΔAR_{it} is the change in account receivables. *ADA* is defined as the absolute value of the difference between total accruals $\frac{TA_{i,t}}{A_{i,t-1}}$ and non-discretionary accruals $NDA_{i,t}$.
Δ*ADA*	Abnormal discretionary accruals, measured as the annual change of *ADA*.
Δ*DISEXPRD*	A slightly modified version of Δ*DISEXP*, measured as the annual change of SG&A and R&D expenses scaled by beginning-of-year total assets. R&D is set to zero if missing.
Δ*CAPEXRD*	A slightly modified version of Δ*CAPEX*, measured as the annual change of capital expenditures and R&D expenses scaled by beginning-of-year total assets. R&D is set to zero if missing.

Variable	Definition
$Marginable_i^k$	A dummy variable that equals one if firm i is approved for margin trading in round k.
V_i^k	The standardized forcing variable used to rank the non-marginable stocks.
$\mathbf{1}[V_i^k > 0]$	A dummy variable that equals one if stock i's ranking index is above the experiment cutoff in round k.
Total Assets	Book value of total assets.
Market Cap.	Market value of equity.
B/M	Book-to-market ratio, calculated as the book value of equity divided by the market value of equity.
ROA	Operating return on assets, calculated as the operating income over beginning-of-year total assets.
Leverage	Long-term debt plus current liabilities scaled by the book value of debt and equity.
Asset Growth	Same definition as $\Delta ASSETS$; control variable for earnings management.
Cash Flow	Cash flow from operations over beginning-of-year total assets.
Cash Flow Vol.	Cash flow volatility is estimated by the standard deviations of cash flows of a firm in the sample period, scaled by beginning-of-year total assets.
ΔNI	Earnings surprise, defined as the absolute value of the current ROA (return on assets) minus last year's ROA.
Analyst No.	Previous month's number of analysts who follow a firm and report forecasts.

Variable	Definition
Dispersion	Previous month's analyst forecast dispersion, defined as the standard deviations of earnings per share forecasts reported by analysts.
Beta	Stock-level beta in the preceding year, calculated from the CAPM model using daily returns.

A.2 Margin Account Details

This section describes the details of margin accounts in China and explains why investors can only leverage their position on marginable stocks. Let me first introduce some definitions for clarity. Since the term "margin" has been overused to denote multiple concepts in this field, I will try to use more specific terms when possible. In this Appendix, I will also use the original definition of the *initial* and *maintenance* ratios listed in the official CSRC rules. To facilitate the understanding, I have converted such ratios to the standard U.S. version in the main body. A demonstration of the conversion procedure is also provided in this Appendix.

- *Cash Deposit* (*CD*): the initial money or securities that an investor has to deposit with the broker to increase their purchasing power. For simplicity, I assume the deposit to be purely cash of 200 RMB. This can also be achieved by depositing 308 RMB worth of securities based on the 65% security discount rate or a combination between cash and securities. With 200 RMB worth of *Cash Deposit*, an investor has a purchasing power of $200/50\% = 400$ RMB that can only be spent on marginable stocks.

- *Initial Equity* (*IE*): the value of securities an investor actually purchased; can fall in between 0 and 400 RMB (*Cash Deposit*/50%). I will only consider the case that *Initial Equity* is larger than *Cash Deposit*. Otherwise, the margin account is essentially identical to a cash account.

- *Margin Debt* (*MD*): the net amount of money an investor borrows from the broker; equals *Initial Equity* − *Cash Deposit*. A term commonly used in the U.S. but not as much in China.

- *Current Equity* (*CE*): the floating equity value after the establishment of *Initial Equity*.

- *Initial Deposit Ratio* (*IR%*): the ratio of *Cash Deposit* over *Initial Equity*; cannot be smaller than 50%. The 50% *IR%* is identical to the required minimum initial margin of

50% in the United States.

- *Maintenance Deposit Ratio* (*MR%*): the ratio of (*Cash Deposit*+*Current Equity*) over *Initial Equity*, ignoring any interest rate or transaction fee. A margin call would happen when *MR%* is lower than 130%. An investor can only extract cash from the margin account if *MR%* is higher than 300% and must keep it at at least 300% afterward.

Row 1–4 of Table A.2 illustrates the initial phase, where an investor opens a margin account and deposit their 200 RMB cash. As the investor increases *Initial Equity* from 250 to 400 RMB, their *Initial Deposit Ratio* falls from 80% to the minimum requirement of 50% (in red). *Maintenance Deposit Ratio* does not come into play in the initial phase since, by design, it is higher than 130% even at a 50% *Initial Deposit Ratio*.

The rest of Table A.2 illustrates the later phase where *Initial Equity* has been established in the initial phase and remains the same after that. The stock price fluctuates over time, leading to a changing *Current Equity*. Row 5–8 considers the situation when *Maintenance Deposit Ratio* hits the lower limit of 130% for a margin call (in red). For a *Initial Deposit Ratio* of 50% (row 8), a price drop of 20% would lead to a margin call ($[200+320]/400 = 130\%$; $[320-400]/400 = -20\%$). For investors with lower leverage (say, 80% *Initial Deposit Ratio*), the price has to decrease by 50% to trigger the margin call. In the United States, *maintenance margin* is defined as the amount of equity that an investor must maintain in the margin account over the current market value of the stock held by the investor. Therefore, the minimum *maintenance margin* is met in China when an investor receives a margin call with a minimum *Initial Deposit Ratio* of 50% (row 8), which equals $(\textit{Current Equity} - \textit{Margin Debt})/\textit{Current Equity} = (320-200)/320 = 37.5\%$. This is comparable to the 30–40% minimum *maintenance margin* required by major U.S. brokerage firms.

Another key feature is the cash extraction requirement as described in Row 9–12, which is the underlying mechanism why investors cannot leverage their positions on non-marginable stocks in China. A reasonable concern is that investors may borrow 400 RMB altogether, use half of it on marginable stocks, and transfer the other half to a separate cash account eligible for non-marginable stocks. However, the CSRC requires that all traders cannot extract any cash from their margin accounts unless *Maintenance Deposit Ratio* reaches 300% (in red). For investors with a *Initial Deposit Ratio* of 50% (row 12), they can only extract money when their stock value rises to 2.5 times the purchase price ($[200+1000]/400 = 300\%$; $1000/400 = 2.5$), an event unlikely to happen to the majority of margin traders within a short period. The same dilemma even happens to investors with a much higher *Initial Deposit Ratio* of 80%, requiring *Current Equity* to be at least 2.2 times the *Initial Equity*. But perhaps more importantly, the extracted cash, even invested on non-marginable stocks through a separate cash account, is no longer used as collateral and cannot provide any leverage.

Table A.2. Initial and Maintenance Margin Requirements

Phase	Events	Row	*CD*	*IE*	*MD*	*CE*	Price%	*IR*%	*MR*%
Initial		1	200	250	50	250	0%	80%	180%
		2	200	300	100	300	0%	67%	167%
		3	200	350	150	350	0%	57%	157%
		4	200	400	200	400	0%	50%	150%
Later	Margin Call	5	200	250	50	125	-50%	80%	130%
		6	200	300	100	190	-37%	67%	130%
		7	200	350	150	255	-27%	57%	130%
		8	200	400	200	320	-20%	50%	130%
Later	Cash Extract	9	200	250	50	550	120%	80%	300%
		10	200	300	100	700	133%	67%	300%
		11	200	350	150	850	143%	57%	300%
		12	200	400	200	1000	150%	50%	300%

Appendix B

Appendix for Chapter 2

B.1 Proof of Lemma

Proof. Without loss of generality, we will prove for the case of t = 0. The price-dividend ratio Φ_0 is given by

$$\Phi_0 = \mathrm{E}_0\left[\int_0^\infty e^{\int_0^t -\mu_v dv + \left(g - \frac{1}{2}\sigma_d^2\right)dv + \sigma_d dB_v^d}\right] \tag{B.1}$$

The expected return μ_t satisfies

$$\mu_t = \bar{\mu} + (\mu_0 - \bar{\mu})\, e^{-Kt} + \int_0^t e^{-K(t-v)}\sigma dB_v \tag{B.2}$$

The cumulative expected return is

$$\int_0^t \mu_v dv = \bar{\mu}t + \frac{\mu_0 - \bar{\mu}}{K}\left(1 - e^{-Kt}\right) + \int_0^t \int_0^v e^{-K(v-u)}\sigma dB_u dv \tag{B.3}$$

Changing the integration order,

$$\int_0^t \int_0^v e^{-K(v-u)}\sigma dB_u dv = \int_0^t \int_u^t e^{-K(v-u)}\sigma dv dB_u = \frac{\sigma}{K}\int_0^t \left(1 - e^{-K(t-u)}\right) dB_u \tag{B.4}$$

So we have

$$
\begin{aligned}
&\int_0^t -\mu_v dv + g dv - \frac{1}{2}\sigma_d^2 dv + \sigma_d dB_v^d \\
&= \int_0^t -\mu_v dv + g dv - \frac{1}{2}\sigma_d^2 dv + \sigma_d(\rho dB_v + \sqrt{1-\rho^2} dB_v^{\perp}) \\
&= -(\bar{\mu} - g + \frac{1}{2}\sigma_d^2)t - \frac{\mu_0 - \bar{\mu}}{K}(1 - e^{-Kt}) + \int_0^t (-\frac{\sigma}{K}(1 - e^{-K(t-u)}) + \rho\sigma_d) dB_u \\
&\qquad + \sigma_d \sqrt{1-\rho^2} dB_u^{\perp}
\end{aligned}
\tag{B.5}
$$

This is a normal random variable. The mean is given by the first two terms. The variance is given by Ito' Isometry

$$
\begin{aligned}
&\int_0^t \left(-\frac{\sigma}{K}\left(1 - e^{-K(t-u)}\right) + \rho\sigma_d\right)^2 du + \sigma_d^2\left(1-\rho^2\right) du \\
&= \frac{\sigma^2}{K^2}\left(t - \frac{2}{K}\left(1-e^{-Kt}\right) + \frac{1}{2K}\left(1-e^{-2Kt}\right)\right) + \frac{\rho\sigma\sigma_d}{K}\frac{1}{K}\left(1-e^{-Kt}\right) - \frac{2\rho\sigma\sigma_d}{K}t + \sigma_d^2 t
\end{aligned}
\tag{B.6}
$$

Using the moment generating function of a normal random variable, we get

$$
\begin{aligned}
\Phi &= \int_0^\infty \mathrm{E}_0\left[e^{\int_0^t -\mu_v dv + \left(g - \frac{1}{2}\sigma_d^2\right)dv + \sigma_d dB_v^d}\right] \\
&= \int_0^\infty e^{-(\hat{\mu}-g)t} e^{-\frac{\mu_0-\hat{\mu}}{K}\left(1-e^{-Kt}\right) + \frac{\sigma^2}{2K^2}\left(t - \frac{2}{K}\left(1-e^{-Kt}\right) + \frac{1}{2K}\left(1-e^{-2Kt}\right)\right)} dt.
\end{aligned}
\tag{B.7}
$$

■

We need the following transversality condition

$$
\hat{\mu} - g - \frac{\sigma^2}{2K^2} > 0 \tag{B.8}
$$

The transversality condition implies that the limit either $K \to 0$ or $\sigma \to \infty$ is not appropriate. Next, implied cost of equity capital v is given by

$$
\Phi_0 = \frac{1}{v_0 - g} \tag{B.9}
$$

which gives

$$v_0 = g + \frac{1}{\Phi_0} \tag{B.10}$$

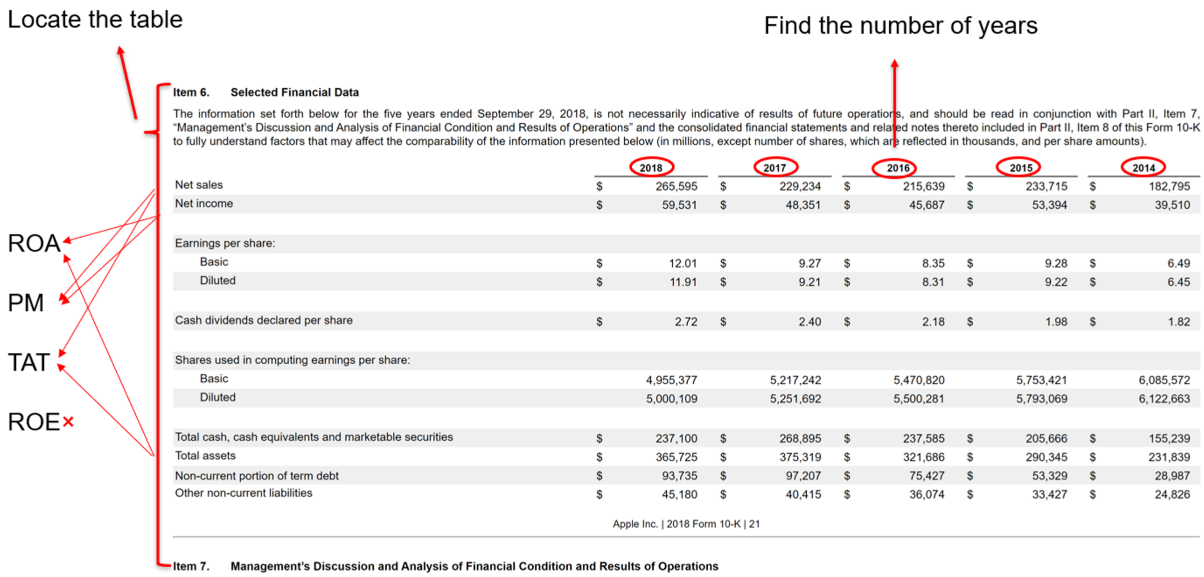
B.2 Disclosure Measure Construction

This section describes our textual analysis procedure. We mainly apply two coding schemes to replicate Botosan's disclosure index: table coding that features figure positioning and key information extraction; and text coding that features fuzzy search and number matching.

We use table coding for most items (a-e) in Category of Table 1 since the summary of historical results is typically listed in an integrated table within "Item 6. Selected Financial Data" in Form 10-K. Essentially, we want to count the number of years/quarters that a firm provides sufficient information to calculate the financial ratios, including return on assets (ROA), net profit margin (PM), asset turnover (TAT), return on equity (ROE), etc. Figure B.1 shows an example of the table coding procedure for Apple's 10-K file in fiscal 2018. Apple provides ROA information for five years since we can extract total assets and net income from 2014 to 2018 (similar results for PM and TAT). As a comparison, ROE is missing since there is no information on total shareholders' equity for the past five years.

For the table coding procedure, there are three files that may contain the desired annual report information in the EDGAR system. Most of our annual report information comes from the traditional 10-K files. However, firms may also choose to put quantitative and MDA information in the EX-13 form or only uploaded an all-in-one text file at the early stage of EDGAR implementation. Figure B.2 shows an example of Walmart's 2015 annual report page with all three different files uploaded simultaneously. For each firm in each year, we first go over the balance sheet, income statement, and cash flow statement in the original 10-K file. If the above information is missing, we would turn to the EX-13 form and the all-in-one text file and check whether that firm used an alternative reporting method at the time.

For the remaining items (I.f to IV.g) in Table 1, we mainly use the text coding method.



Item 6. Selected Financial Data

The information set forth below for the five years ended September 29, 2018, is not necessarily indicative of results of future operations, and should be read in conjunction with Part II, Item 7, "Management's Discussion and Analysis of Financial Condition and Results of Operations" and the consolidated financial statements and related notes thereto included in Part II, Item 8 of this Form 10-K to fully understand factors that may affect the comparability of the information presented below (in millions, except number of shares, which are reflected in thousands, and per share amounts).

	2018	2017	2016	2015	2014
Net sales	$ 265,595	$ 229,234	$ 215,639	$ 233,715	$ 182,795
Net income	$ 59,531	$ 48,351	$ 45,687	$ 53,394	$ 39,510
Earnings per share:					
Basic	$ 12.01	$ 9.27	$ 8.35	$ 9.28	$ 6.49
Diluted	$ 11.91	$ 9.21	$ 8.31	$ 9.22	$ 6.45
Cash dividends declared per share	$ 2.72	$ 2.40	$ 2.18	$ 1.98	$ 1.82
Shares used in computing earnings per share:					
Basic	4,955,377	5,217,242	5,470,820	5,753,421	6,085,572
Diluted	5,000,109	5,251,692	5,500,281	5,793,069	6,122,663
Total cash, cash equivalents and marketable securities	$ 237,100	$ 268,895	$ 237,585	$ 205,666	$ 155,239
Total assets	$ 365,725	$ 375,319	$ 321,686	$ 290,345	$ 231,839
Non-current portion of term debt	$ 93,735	$ 97,207	$ 75,427	$ 53,329	$ 28,987
Other non-current liabilities	$ 45,180	$ 40,415	$ 36,074	$ 33,427	$ 24,826

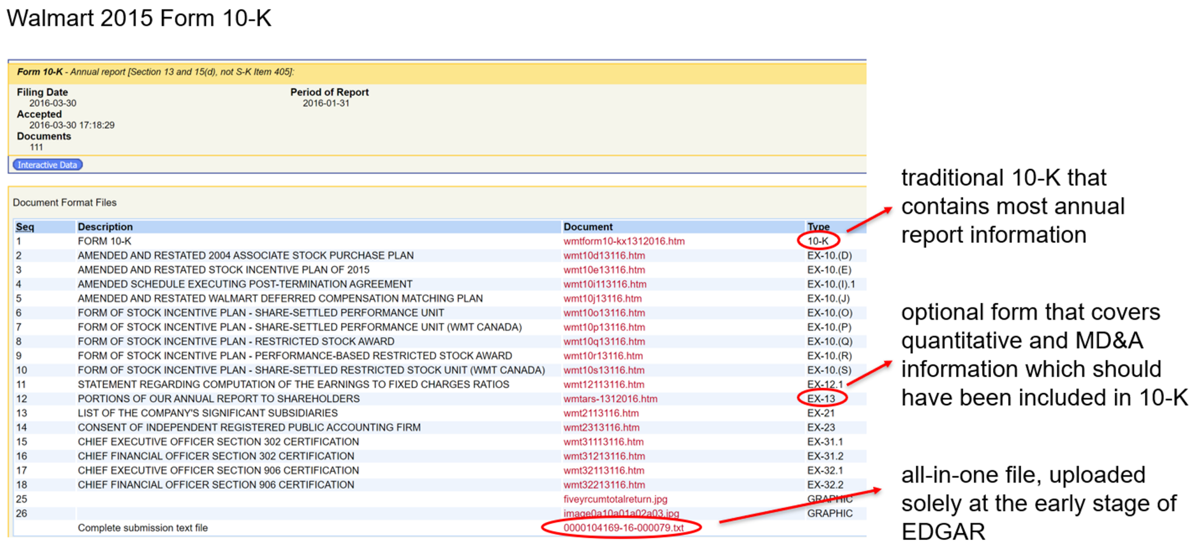
Apple Inc. | 2018 Form 10-K | 21

Item 7. Management's Discussion and Analysis of Financial Condition and Results of Operations

Figure B.1. An Example of the Table Coding

Specifically, we construct a list of conclusive keywords and collocations for each qualitative item and also require a corresponding number when necessary for each quantitative item. Figure B.3 shows an example of the text coding procedure. For instance, for Item .g "discussion of corporate strategy", we first check the standard information display format. Next, we summarize the highlighted positions where this information typically shows up (title, paragraph beginning, bolded font) and high-frequency match-up phrases (corporate/business/development strategy/plan, etc.). We require an in-sample accuracy of 90% and an out-of-sample accuracy of 85% for a fully established coding method (200 firms for each). Of the 20 individual items for the text coding procedure, we manually establish and verify the coding method and automatically calculate the final score for each firm in each year.

Our textual analysis has some superior features compared to the traditional word searching method. First, our coding scheme has loose matching tolerance. For example, if we confirm "forecasted sales" is a keyword, a sentence like "based on analyst forecasts, our sales value next year will..." is also taken into consideration in our coding scheme. Similarly, selected match-up phrases are captured as long as they are positioned in the same paragraph or within a



Seq	Description	Document	Type
1	FORM 10-K	wmtform10-kx1312016.htm	10-K
2	AMENDED AND RESTATED 2004 ASSOCIATE STOCK PURCHASE PLAN	wmt10d13116.htm	EX-10.(D)
3	AMENDED AND RESTATED STOCK INCENTIVE PLAN OF 2015	wmt10e13116.htm	EX-10.(E)
4	AMENDED SCHEDULE EXECUTING POST-TERMINATION AGREEMENT	wmt10i113116.htm	EX-10.(I).1
5	AMENDED AND RESTATED WALMART DEFERRED COMPENSATION MATCHING PLAN	wmt10j13116.htm	EX-10.(J)
6	FORM OF STOCK INCENTIVE PLAN - SHARE-SETTLED PERFORMANCE UNIT	wmt10o13116.htm	EX-10.(O)
7	FORM OF STOCK INCENTIVE PLAN - SHARE-SETTLED PERFORMANCE UNIT (WMT CANADA)	wmt10p13116.htm	EX-10.(P)
8	FORM OF STOCK INCENTIVE PLAN - RESTRICTED STOCK AWARD	wmt10q13116.htm	EX-10.(Q)
9	FORM OF STOCK INCENTIVE PLAN - PERFORMANCE-BASED RESTRICTED STOCK AWARD	wmt10r13116.htm	EX-10.(R)
10	FORM OF STOCK INCENTIVE PLAN - SHARE-SETTLED RESTRICTED STOCK UNIT (WMT CANADA)	wmt10s13116.htm	EX-10.(S)
11	STATEMENT REGARDING COMPUTATION OF THE EARNINGS TO FIXED CHARGES RATIOS	wmt12113116.htm	EX-12.1
12	PORTIONS OF OUR ANNUAL REPORT TO SHAREHOLDERS	wmtars-1312016.htm	EX-13
13	LIST OF THE COMPANY'S SIGNIFICANT SUBSIDIARIES	wmt2113116.htm	EX-21
14	CONSENT OF INDEPENDENT REGISTERED PUBLIC ACCOUNTING FIRM	wmt2313116.htm	EX-23
15	CHIEF EXECUTIVE OFFICER SECTION 302 CERTIFICATION	wmt31113116.htm	EX-31.1
16	CHIEF FINANCIAL OFFICER SECTION 302 CERTIFICATION	wmt31213116.htm	EX-31.2
17	CHIEF EXECUTIVE OFFICER SECTION 906 CERTIFICATION	wmt32113116.htm	EX-32.1
18	CHIEF FINANCIAL OFFICER SECTION 906 CERTIFICATION	wmt32213116.htm	EX-32.2
25		fiveyrcumtotalreturn.jpg	GRAPHIC
26		image0a10a01a02a03.jpg	GRAPHIC
	Complete submission text file	0000104169-16-000079.txt	

Figure B.2. An Example of the Annual Report Page in EDGAR

one-sentence gap as appropriate. Second, our coding method is time-efficient, online analytical, and easily applicable to other disclosure measures such as ESG, CSR, and innovation. We benefit from the multiprocessing frame and improve our running efficiency from 300 hours (single process) to 20 hours for the entire EDGAR universe. To avoid getting banned by EDGAR for excessive requests, we use an online real-time coding frame that can directly generate the disclosure score without downloading the 10-K file for each firm.

Third, we allow for file-sensitive and format-sensitive table coding schemes. As mentioned above, annual report information can be positioned in different files (10-K, EX-13, and all-in-one file) with different formats (HTM or TXT). Since there is a vast encoding difference between TXT and HMT and a huge layout difference between 10-K, EX-13, and the all-in-one text file, we must reconstruct our table coding scheme for each specific file/format, which ends up with five different coding methods in total. Figure B.4 illustrates how we choose the coding method under each circumstance. For each firm in each year, we first identify whether the firm uploads a 10-K file; if no, we directly go to the all-in-one text file. If a 10-K file is uploaded, we would first check the format of the 10-K file and use the corresponding method to search

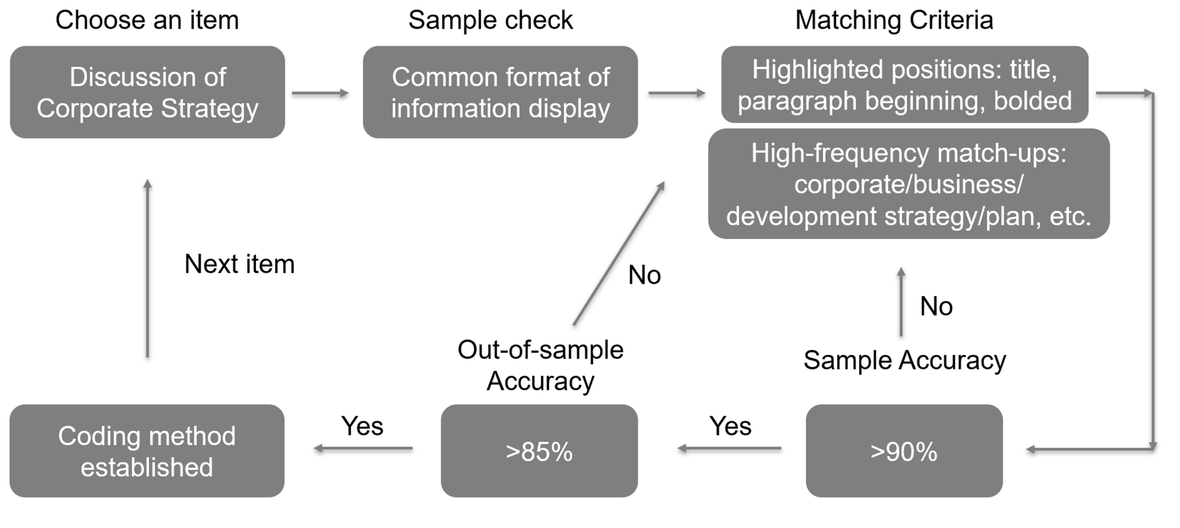


Figure B.3. An Example of the Text Coding

information encoded in HTM or TXT. If no desired information is found in the 10-K, we would explore the EX-13 form and redo the abovementioned steps. As the EDGAR developed over the years, more and more information tends to be disclosed through the file/format of 10-K HTM.

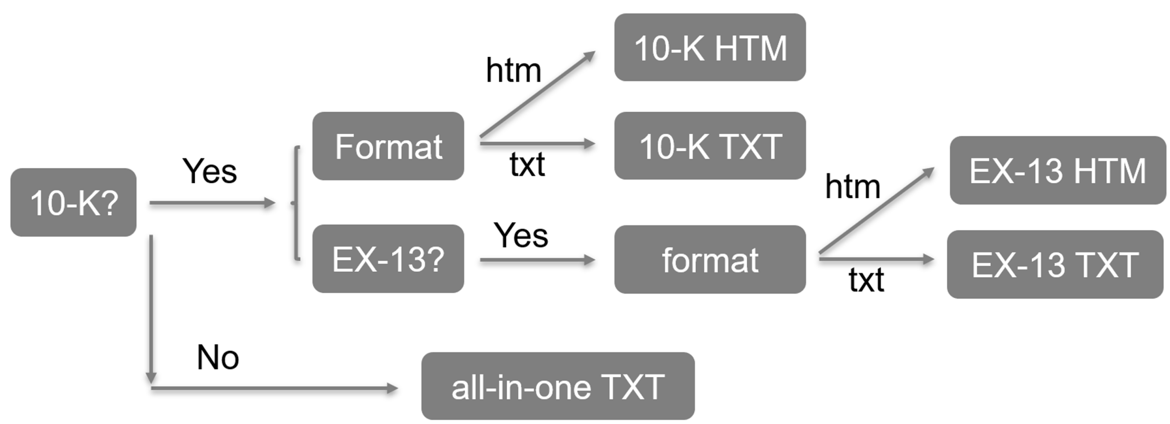


Figure B.4. Different Coding Methods for Different File/Format

In summary, our textual analysis generates a firm-year disclosure measure for each firm that ever submitted a 10-K file to the EDGAR system in that year. We end up with 144,778 firm-year observations from 1994 to 2019, of which a great portion is not included in this study since they are either not covered by Value Line or I/B/E/S or missing formula inputs for the implied cost of equity capital estimation.

Appendix C

Appendix for Chapter 3

C.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.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.

C.2.1 Discretionary Accruals (DA)

We measure each firm's DA using modified Jones model:

$$\frac{Accruals_t}{A_{t-1}} = \alpha_0 + \alpha_1 \frac{1}{A_{t-1}} + \alpha_2 \frac{\Delta S_t - \Delta AR_t}{A_{t-1}} + \alpha_3 \frac{PPE_t}{A_{t-1}} + \varepsilon_t \qquad \text{(C.1)}$$

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.

C.2.2 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., 2008; Cohen and Zarowin, 2010).

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.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.4 Delisting Threat vs. Non-delisting Threat

Table C.1 and C.2 show that investors react less to firms with delisting threat (negative earnings last year) compared with firms without delisting threat (positive earnings last year).

C.5 Bunching Estimator

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. Figure C.1 shows the counterfactual ROE distribution without the delisting policy in China.

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**
	(0.097)	(0.101)	(0.096)	(0.104)	(0.087)	(0.093)
Firm Size		-0.027		0.013		-0.017
		(0.039)		(0.042)		(0.034)
Firm Leverage		-0.039		-0.069		0.028
		(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.

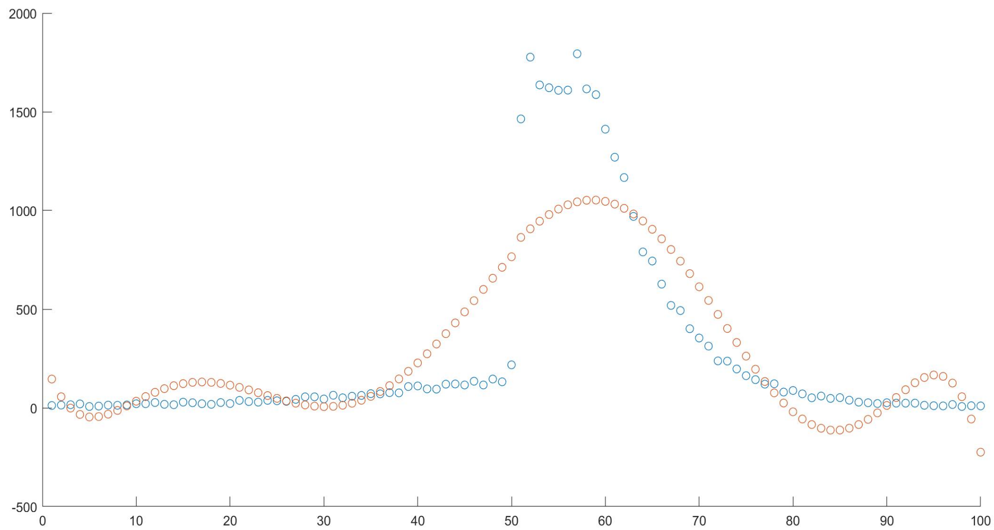


Figure C.1. Counterfactual ROE Distribution

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

Ali, A. and W. Zhang (2015). CEO tenure and earnings management. *Journal of Accounting and Economics 59*, 60–79.

Ang, A., R. Hodrick, Y. Xing, and X. Zhang (2006). The cross-section of volatility and expected returns. *Journal of Finance 61*(1), 259–299.

Appel, I. R., T. A. Gormley, and D. B. Keim (2016). Passive investors, not passive owners. *Journal of Financial Economics 121*, 111–141.

Asker, J., J. Farre-Mensa, and A. Ljungqvist (2015). Corporate investment and stock market listing: A puzzle? *Review of Financial Studies 28*, 342–390.

Badertscher, B. A. (2011). Overvaluation and the choice of alternative earnings management mechanisms. *The Accounting Review 86*, 1491–1518.

Baker, M., R. S. Ruback, and J. Wurgler (2006). Behavioral corporate finance: a survey. In *Handbook of Corporate Finance: Empirical Corporate Finance*, pp. 145–186. Elsevier.

Barber, B. M., X. Huang, K. J. Ko, and T. Odean (2020). Leveraging overconfidence. Working paper, UC Davis.

Bartov, E. (1993). The timing of asset sales and earnings manipulation. *Accounting Review*, 840–855.

Bartov, E., D. Givoly, and C. Hayn (2002). The rewards to meeting or beating earnings expectations. *Journal of accounting and economics 33*(2), 173–204.

Beatty, A., S. Liao, and J. J. Yu (2013). The spillover effect of fraudulent financial reporting on peer firms' investments. *Journal of Accounting and Economics 55*(2-3), 183–205.

Ben-David, I., J. Birru, and V. Prokopenya (2018). Uninformative feedback and risk taking: Evidence from retail forex trading. *Review of Finance 22*, 2009–2036.

Bergstresser, D. and T. Philippon (2006). CEO incentives and earnings management. *Journal of*

Financial Economics 80, 511–529.

Bessembinder, H., W. Maxwell, and K. Venkataraman (2006). Market transparency, liquidity externalities, and institutional trading costs in corporate bonds. *Journal of Financial Economics 82*(2), 251–288.

Beyer, A., D. F. Larcker, and B. Tayan (2014). Study on how investment horizon and expectations of shareholder base impact corporate decision-making. National Investor Relations Institute and Stanford University.

Bhojraj, S., P. Hribar, M. Picconi, and J. McInnis (2009). Making sense of cents: An examination of firms that marginally miss or beat analyst forecasts. *The Journal of Finance 64*(5), 2361–2388.

Bian, J., Z. Da, Z. He, D. Lou, K. Shue, and H. Zhou (2021). Margin trading and leverage management. Working paper, University of Chicago.

Bloomfield, R. and M. O'Hara (1999). Market transparency: Who wins and who loses? *The Review of Financial Studies 12*(1), 5–35.

Boehmer, E. and J. Wu (2013). Short selling and the price discovery process. *Review of Financial Studies 26*, 287–322.

Bolton, P., J. Scheinkman, and W. Xiong (2006). Executive compensation and short-termist behaviour in speculative markets. *Review of Economic Studies 73*, 577–610.

Bond, P., A. Edmans, and I. Goldstein (2012). The real effects of financial markets. *Annual Review of Financial Economics 4*, 339–360.

Botosan, C. (1997). Disclosure level and the cost of equity capital. *The Accounting Review 72*(3), 323–349.

Botosan, C. and M. Plumlee (2002). A re-examination of disclosure level and the expected cost of equity capital. *Journal of Accounting Research 40*(1), 21–40.

Botosan, C. and M. Plumlee (2005). Assessing alternative proxies for the expected risk premium. *The Accounting Review 80*(1), 21–53.

Botosan, C., M. Plumlee, and Y. Xie (2004). The role of information precision in determining the cost of equity capital. *Review of Accounting Studies 9*, 233–259.

Brav, A., R. Lehavy, and R. Michaely (2005). Using expectations to test asset pricing models. *Financial Management 34*(3), 31–64.

Brunnermeier, M. K. and L. H. Pedersen (2009). Market liquidity and funding liquidity. *Review of Financial Studies 22*, 2201–2238.

Burgstahler, D. and I. Dichev (1997). Earnings management to avoid earnings decreases and losses. *Journal of accounting and economics 24*(1), 99–126.

Bushee, B. J. (1998). The influence of institutional investors on myopic R&D investment behavior. *The Accounting Review 73*, 305–333.

Bushee, B. J. (2001). Do institutional investors prefer near-term earnings over long-run value? *Contemporary Accounting Research 18*, 207–246.

Cadman, B. and J. Sunder (2014). Investor horizon and CEO horizon incentives. *The Accounting Review 89*, 1299–1328.

Calonico, S., M. D. Cattaneo, M. H. Farrell, and R. Titiunik (2019). Regression discontinuity designs using covariates. *Review of Economics and Statistics 101*, 442–451.

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica 82*, 2295–2326.

Campbell, J. (1991). A variance decomposition for stock return. *The Economic Journal 101*, 157–179.

Cattaneo, M. D., M. Jansson, and X. Ma (2020). Simple local polynomial density estimators. *Journal of the American Statistical Association 115*, 1449–1455.

Cellini, S. R., F. Ferreira, and J. Rothstein (2010). The value of school facility investments: Evidence from a dynamic regression discontinuity design. *Quarterly Journal of Economics 125*, 215–261.

Chang, E. C., Y. Luo, and J. Ren (2014). Short-selling, margin-trading, and price efficiency: Evidence from the Chinese market. *Journal of Banking & Finance 48*, 411–424.

Chen, K. C. and H. Yuan (2004). Earnings management and capital resource allocation: Evidence from china's accounting-based regulation of rights issues. *The Accounting Review 79*(3), 645–665.

Chen, Q., I. Goldstein, and W. Jiang (2006). Price informativeness and investment sensitivity to stock price. *The Review of Financial Studies 20*(3), 619–650.

Chen, X., J. Harford, and K. Li (2007). Monitoring: Which institutions matter? *Journal of Financial Economics 86*, 279–305.

Chetty, R., J. Friedman, T. Olsen, and L. Pistaferri (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. (2013). A theory of voluntary disclosure and cost of capital. *Review of Accounting Studies 18*, 987–1020.

Chi, D. and M. Gupta (2009). Overvaluation and earnings management. *Journal of Banking & Finance 33*, 1652–1663.

Chong-en, B., H. Changtai, and S. Z. Michael (2016). The long shadow of a fiscal expansion. *Brookings Papers on Economic Activity*, 15–16.

Chordia, T., S.-W. Huh, and A. Subrahmanyam (2007). The cross-section of expected trading activity. *Review of Financial Studies 20*, 709–740.

Claus, J. and J. Thomas (2001). Equity premia as low as three percent? evidence from analysts' earnings forecasts for domestic and international stock markets. *Journal of Finance 56*(5), 1629–1666.

Cohen, D. A., A. Dey, and T. Z. Lys (2008). Real and accrual-based earnings management in the pre-and post-sarbanes-oxley periods. *The accounting review 83*(3), 757–787.

Cohen, D. A. and P. Zarowin (2010). Accrual-based and real earnings management activities around seasoned equity offerings. *Journal of accounting and Economics 50*(1), 2–19.

Collins, D. W. and S. Kothari (1989). An analysis of intertemporal and cross-sectional determinants of earnings response coefficients. *Journal of accounting and economics 11*(2-3), 143–181.

Core, J., W. Guay, and R. Verdi (2008). Is accruals quality a priced risk factor? *Journal of Accounting and Economics 46*(1), 2–22.

Cornett, M. M., A. J. Marcus, and H. Tehranian (2008). Corporate governance and pay-for-performance: The impact of earnings management. *Journal of Financial Economics 87*, 357–373.

Cremers, M., A. Pareek, and Z. Sautner (2020). Short-term investors, long-term investments, and firm value: Evidence from Russell 2000 index inclusions. *Management Science 66*, 4535–4551.

Daniel, K., M. Grinblatt, S. Titman, and R. Wermers (1997). Measuring mutual fund performance with characteristic-based benchmarks. *Journal of Finance 52*, 1035–1058.

Daske, H., L. Hail, C. Leuz, and R. Verdi (2013). Adopting a label: heterogeneity in the economic consequences around ias/ifrs adoptions. *Journal of Accounting Research 51*(3), 495–547.

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*(1), 3–42.

Dechow, P. M., R. G. Sloan, and A. P. Sweeney (1995). Detecting earnings management. *Accounting review*, 193–225.

Degeorge, F., J. Patel, and R. Zeckhauser (1999). Earnings management to exceed thresholds. *The Journal of Business 72*(1), 1–33.

Dhaliwal, D., O. Li, A. Tsang, and Y. Yang (2011). Voluntary nonfinancial disclosure and the cost of equity capital: the initiation of corporate social responsibility reporting. *The Accounting Review 86*(1), 59–100.

Durnev, A., R. Morck, B. Yeung, and P. Zarowin (2003). Does greater firm-specific return variation mean more or less informed stock pricing? *Journal of Accounting Research 41*(5), 797–836.

Easley, D. and M. O'Hara (2004). Information and the cost of capital. *The Journal of Finance 59*(4), 1553–1583.

Easton, P. (2004). Pe ratios, peg ratios, and estimating the implied expected rate of return on equity capital. *The Accounting Review 79*(1), 73–95.

Edmans, A., V. W. Fang, and K. A. Lewellen (2017). Equity vesting and investment. *Review of Financial Studies 30*, 2229–2271.

Edwards, A. K., L. Harris, and M. S. Piwowar (2004). Corporate bond market transparency and transaction costs. *Fifteenth Annual Utah Winter Finance Conference*.

Fairfield, P. M., J. S. Whisenant, and T. L. Yohn (2003). Accrued earnings and growth: Implications for future profitability and market mispricing. *The Accounting Review 78*, 353–371.

Fama, E. and K. French (1997). Industry costs of equity. *Journal of Financial Economics 43*(2), 153–193.

Fama, E. and J. MacBeth (1973). Risk, return, and equilibrium: empirical tests. *Journal of Political Economy 81*(3), 607–636.

Fama, E. F. (1998). Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics 49*(3), 283–306.

Fama, E. F. and K. R. French (2000). Forecasting profitability and earnings. *The Journal of Business 73*(2), 161–175.

Fang, V. W., A. H. Huang, and J. M. Karpoff (2016). Short selling and earnings management: A controlled experiment. *Journal of Finance 71*, 1251–1294.

Francis, J., R. LaFond, P. Olsson, and K. Schipper (2004). Costs of equity and earnings attributes. *The Accounting Review 79*(4), 967–1010.

Francis, J., R. LaFond, P. Olsson, and K. Schipper (2005). The market pricing of accruals quality. *Journal of Accounting and Economics 39*(2), 295–327.

Francis, J., D. Nanda, and P. Olsson (2008). Voluntary disclosure, earnings quality, and cost of capital. *Journal of Accounting Research 46*(1), 53–99.

Froot, K. A., A. F. Perold, and J. C. Stein (1992). Shareholder trading practices and corporate investment horizons. *Journal of Applied Corporate Finance 5*, 42–58.

Fuller, J. and M. C. Jensen (2010). Just say no to wall street: Putting a stop to the earnings game. *Journal of Applied Corporate Finance 22*, 59–63.

Garleanu, N. and L. H. Pedersen (2011). Margin-based asset pricing and deviations from the law of one price. *Review of Financial Studies 24*, 1980–2022.

Gebhardt, W., C. Lee, and B. Swaminathan (2001). Toward an implied cost of capital. *Journal of Accounting Research 39*(1), 135–176.

Gelman, A. and G. Imbens (2019). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics 37*, 447–456.

Gode, D. and P. Mohanram (2003). Inferring the cost of capital using the ohlson-juettner model. *Review of Accounting Studies 8*, 399–431.

Goldman, E. and S. L. Slezak (2006). An equilibrium model of incentive contracts in the presence of information manipulation. *Journal of Financial Economics 80*, 603–626.

Graham, J. R., C. R. Harvey, and S. Rajgopal (2005). The economic implications of corporate financial reporting. *Journal of Accounting and Economics 40*, 3–73.

Grullon, G., S. Michenaud, and J. P. Weston (2015). The real effects of short-selling constraints. *Review of Financial Studies 28*, 1737–1767.

Hail, L. and C. Leuz (2006). International differences in the cost of equity capital: Do legal institutions and securities regulation matter? *Journal of Accounting Research 44*(3), 485–531.

Hansman, C., H. Hong, W. Jiang, Y.-J. Liu, and J.-J. Meng (2021). Anticipating the direct effects of credit supply. NBER Working paper 24586.

Hardouvelis, G. A. (1990). Margin requirements, volatility, and the transitory component of stock prices. *American Economic Review 80*, 736–762.

HAW, I.-M., D. Qi, D. Wu, and W. Wu (2005). Market consequences of earnings management in response to security regulations in china. *Contemporary Accounting Research 22*(1), 95–140.

Hayn, C. (1995). The information content of losses. *Journal of accounting and economics 20*(2), 125–153.

Healy, P. M., A. P. Hutton, and K. G. Palepu (1999). Stock performance and intermediation changes surrounding sustained increases in disclosure. *Contemporary Accounting Research 16*(3), 485–520.

Heimer, R. and A. Simsek (2019). Should retail investors' leverage be limited? *Journal of Financial Economics 132*, 1–21.

Herrmann, D., T. Inoue, and W. B. Thomas (2003). The sale of assets to manage earnings in japan. *Journal of Accounting Research 41*(1), 89–108.

Hribar, P. and D. C. Nichols (2007). The use of unsigned earnings quality measures in tests of earnings management. *Journal of Accounting Research 45*, 1017–1053.

Hughes, J., J. Liu, and J. Liu (2007). Information asymmetry, diversification, and cost of capital. *The Accounting Review 82*(3), 705–729.

Hughes, J., J. Liu, and J. Liu (2009). On the relation between expected returns and implied cost of capital. *Review of Accounting Studies 14*, 246–259.

Iliev, P. (2010). The effect of SOX section 404: Costs, earnings quality, and stock prices. *Journal of Finance 65*, 1163–1196.

Imbens, G. and K. Kalyanaraman (2012). Optimal bandwidth choice for the regression discontinuity estimator. *Review of Economic Studies 79*, 933–959.

Jagannathan, R. and Z. Wang (1996). The conditional capm and the cross-section of expected returns. *Journal of Finance 51*(1), 3–53.

Jensen, M. C. (2005). Agency costs of overvalued equity. *Financial Management 34*, 5–19.

Jin, L. and S. C. Myers (2006). R2 around the world: New theory and new tests. *Journal of financial Economics 79*(2), 257–292.

Kahraman, B. and H. E. Tookes (2017). Trader leverage and liquidity. *Journal of Finance 72*, 1567–1610.

Kothari, S., X. Li, and J. Short (2009). The effect of disclosures by management, analysts, and business press on cost of capital, return volatility, and analyst forecasts: a study using content analysis. *The Accounting Review 84*(5), 1639–1670.

Kraft, A. G., R. Vashishtha, and M. Venkatachalam (2018). Frequent financial reporting and managerial myopia. *The Accounting Review 93*, 249–275.

Lambert, R. (2009). Discussion of "on the relation between expected returns and implied cost of capital". *Review of Accounting Studies 14*, 260–268.

Lambert, R., C. Leuz, and R. Verrecchia (2007). Accounting information, disclosure, and the cost of capital. *Journal of Accounting Research 45*(2), 385–420.

Lee, D. S. and T. Lemieux (2010). Regression discontinuity designs in economics. *Journal of Economic Literature 48*, 281–355.

Levine, R. and S. Zervos (1996). Stock market development and long-run growth. *The World Bank Economic Review 10*(2), 323–339.

Li, S. (2010). Does mandatory adoption of international financial reporting standards in the european union reduce the cost of equity capital? *The Accounting Review 85*(2), 607–636.

Li, V. (2015). Do false financial statements distort peer firms' decisions? *The Accounting Review 91*(1), 251–278.

Lo, A. W. and J. Wang (2000). Trading volume: Definitions, data analysis, and implications of portfolio theory. *Review of Financial Studies 13*, 257–300.

Massa, M., B. Zhang, and H. Zhang (2015). The invisible hand of short selling: Does short selling discipline earnings management? *Review of Financial Studies 28*, 1701–1736.

McCahery, J. A., Z. Sautner, and L. T. Starks (2016). Behind the scenes: The corporate governance preferences of institutional investors. *Journal of Finance 71*, 2905–2932.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics 142*, 698–714.

Merton, R. (1973). An intertemporal capital asset pricing model. *Econometrica 41*(5), 867–887.

Miller, E. M. (1977). Risk, uncertainty, and divergence of opinion. *Journal of Finance 32*, 1151–1168.

Morck, R., B. Yeung, and W. Yu (2000a). The information content of stock markets: why do emerging markets have synchronous stock price movements? *Journal of financial economics 58*(1-2), 215–260.

Morck, R., B. Yeung, and W. Yu (2000b). The information content of stock markets: why do emerging markets have synchronous stock price movements? *Journal of Financial Economics 58*(1-2), 215–260.

Narayanan, M. P. (1985). Managerial incentives for short-term results. *Journal of Finance 40*, 1469–1484.

Ohlson, J. and B. Juettner-Nauroth (2005). Expected eps and eps growth as determinants of value. *Review of Accounting Studies 10*, 349–365.

Pevzner, M., F. Xie, and X. Xin (2015a). When firms talk, do investors listen? the role of trust in stock market reactions to corporate earnings. *Journal of Financial Economics 117*(1), 190–223.

Pevzner, M., F. Xie, and X. Xin (2015b). When firms talk, do investors listen? the role of trust in stock market reactions to corporate earnings announcements. *Journal of Financial Economics 117*(1), 190–223.

Porter, M. E. (1992). Capital choices: Changing the way America invests in industry. *Journal of Applied Corporate Finance 5*, 4–16.

Roll, R. (1998). R2. *The Journal of Finance 43*(3), 541–566.

Roychowdhury, S. (2006). Earnings management through real activities manipulation. *Journal of accounting and economics 42*(3), 335–370.

Seguin, P. J. (1990). Stock volatility and margin trading. *Journal of Monetary Economics 26*, 101–121.

Sengupta, P. (1998). Corporate disclosure quality and the cost of debt. *The Accounting Review 73*(4), 459–474.

Stein, J. C. (1988). Takeover threats and managerial myopia. *Journal of Political Economy 96*, 61–80.

Stein, J. C. (1989). Efficient capital markets, inefficient firms: A model of myopic corporate behavior. *Quarterly Journal of Economics 104*, 655–669.

von Thadden, E.-L. (1995). Long-term contracts, short-term investment and monitoring. *Review of Economic Studies 62*, 557–575.

Welch, I. (2020). The wisdom of the robinhood crowd. NBER Working paper 27866.

Wu, J., L. Zhang, and X. F. Zhang (2010). The q-theory approach to understanding the accrual anomaly. *Journal of Accounting Research 48*, 177–223.

Yu, F. (2008). Analyst coverage and earnings management. *Journal of Financial Economics 88*, 245–271.

Zang, A. Y. (2011). Evidence on the trade-off between real activities manipulation and accrual-based earnings management. *The Accounting Review 87*(2), 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.

Zhang, X. F. (2007). Accruals, investment, and the accrual anomaly. *The Accounting Review 82*, 1333–1363.